Journal of Personality and Social Psychology

Men and Women Are From Earth: Examining the Latent Structure of Gender

Bobbi J. Carothers and Harry T. Reis

Online First Publication, October 22, 2012. doi: 10.1037/a0030437

Journal of Personality and Social Psychology

0022-3514/12/$12.00 DOI: 10.1037/a0030437

Men and Women Are From Earth: Examining the Latent Structure of Gender

Bobbi J. Carothers
Washington University in St. Louis

Harry T. Reis
University of Rochester

Taxometric methods enable determination of whether the latent structure of a construct is dimensional or taxonic (nonarbitrary categories). Although sex as a biological category is taxonic, psychological gender differences have not been examined in this way. The taxometric methods of mean above minus below a cut, maximum eigenvalue, and latent mode were used to investigate whether gender is taxonic or dimensional. Behavioral measures of stereotyped hobbies and physiological characteristics (physical strength, anthropometric measurements) were examined for validation purposes, and were taxonic by sex. Psychological indicators included sexuality and mating (sexual attitudes and behaviors, mate selectivity, sociosexual orientation), interpersonal orientation (empathy, relational-interdependent self-construal), gender-related dispositions (masculinity, femininity, care orientation, unmitigated communion, fear of success, science inclination, Big Five personality), and intimacy (intimacy prototypes and stages, social provisions, intimacy with best friend). Constructs were with few exceptions dimensional, speaking to Spence's (1993) gender identity theory. Average differences between men and women are not under dispute, but the dimensionality of gender indicates that these differences are inappropriate for diagnosing gender-typical psychological variables on the basis of sex.

Keywords: gender, taxometric, latent structure, personality, sex differences

Supplemental materials: http://dx.doi.org/10.1037/a0030437.supp

If you ain't wrong, you're right;
If it ain't day, it's night . . .
If it ain't dry, it's wet . . .
Gotta be this or that.

—Sunny Skylar

Sex is the most pervasive method of categorizing people. We are more likely to categorize people based on gender than race (Stangor, Lynch, Duan, & Glass, 1992). People use gender to sort individuals into categorical "natural kinds" more than they use 20 other kinds of social categories (Haslam, Rothschild, & Ernst, 2000). Parents of newborns are usually first asked the sex of their child, even before questions about the child's and mother's health (Intons-Peterson & Reddel, 1984). Because nearly all animal species come in only two forms, male and female, sex is an easy target for categorization. According to gender identity theory (Spence, 1984, 1993), from infancy onward, whether a person is male or female pervades almost all human activities and experiences. Children use information about sex to aid in normal cognitive and social functioning (Kohlberg, 1966; Martin & Halverson, 1981). Infants as young as 3 months process images of male and female faces differently (Ramsey-Rennels & Langlois, 2006). People automatically identify an unfamiliar person's sex without conscious effort (Brewer, 1988), using the distinctiveness of male and female body shapes and gaits, and can determine a person's sex seeing only how they walk in as little as 2.7 s (Barclay, Cutting, & Kozlowski, 1978; Johnson & Tassinary, 2005).

Given the obvious evolutionary significance of a person's sex, it may not be surprising that sex is used pervasively for social

Bobbi J. Carothers, Center for Public Health Systems Science, Washington University in St. Louis; Harry T. Reis, Department of Clinical and Social Sciences in Psychology, University of Rochester.

This article is based on the doctoral dissertation of the first author while at University of Rochester. Portions of this research were presented at meetings of the Society for Personality and Social Psychology and the International Association for Relationships Research. Data for this research were generously donated by Arthur Aron, Jennifer Aubé, Mark Barnett, David Buss, Mark Davis, William Graziano, Stephen Quackenbush, Harry Reis, Todd Shackelford, Brad Sheese, and Jeffry Simpson. We thank Chantal Levesque-Bristol and Bradley Bodenhamer for their valuable time in collecting the Midwestern Sample data set and John Ruscio for consultation on taxometric methods. Science inclination data were made available by the Organisation for Economic Co-operation and Development. Anthropometric and National Health and Social Life Survey (NHSLS) data were made available by the Inter-University Consortium for Political and Social Research. Data for the National Health and Nutrition Examination Survey III were originally collected and prepared by the Centers for Disease Control and Prevention. Data for the NHSLS were originally collected by the University of Chicago and the National Opinion Research Center. Neither the collectors of the original data nor the consortium bears any responsibility for the analyses or interpretations presented here.

Correspondence concerning this article should be addressed to Bobbi J. Carothers, Center for Public Health Systems Science, George Warren Brown School of Social Work, Washington University in St. Louis, Campus Box 1009, St. Louis, MO 63112. E-mail: bcarothers@wustl.edu

categorization. In this light, Hyde (2005) posed a key question: Are men and women generally the same or essentially different? In lay conceptions, men and women tend to be seen as categorically different. When asked if men and women were "basically different" in 1989, 58% of a sample agreed (40% said they were similar, and 2% had no opinion; Gallup, 1991). Moreover, lay beliefs about the characteristics of one sex tend to be negatively correlated with beliefs about the characteristics of the other sex (e.g., Foushee, Helmreich, & Spence, 1979; Kahn & Yoder, 1989). Because we categorize objects in order to simplify complex information, it has been suggested that people categorize men and women on the basis of sex to simplify a complex social world (Fiske, 2010; Taylor, Fiske, Etcoff, & Ruderman, 1978). In other words, if men and women differ, grouping people into sex categories makes sense as a simple and fast judgment rather than dealing with each individual as a unique case. That men and women do in fact differ in many important behaviors, at least some of which have meaningful effect sizes, reinforces this tendency (Eagly, 1995; Hyde, 2005).

From these and other findings, it has been argued that sex may be the strongest example of *essentialism* in lay social cognition: the belief that categories possess distinctive "deep, hidden, and unchanging properties that make their members what they are" (Prentice & Miller, 2007, p. 202; see also Haslam et al., 2000). Furthermore, lay essentialist beliefs about sex differences are often grounded in genetic explanations (e.g., Cole, Jayaratne, Cecchi, Feldbaum, & Petty, 2007), endowing them with a sense of naturalness, presumed importance, and acceptability (Haslam et al., 2000; Prentice & Miller, 2007; Wood & Eagly, 2012). It may be no surprise, then, that popular models putting men and women into fundamentally distinct categories, such as the argument that men and women come from different planets (Gray, 1992), often strike a resonant chord with lay audiences.

Interestingly, despite many thousands of studies on sex differences—in 2011 alone, PsycINFO reported 3,370 articles under the keyword *human sex differences*—psychological science has not explicitly demonstrated whether human sex differences in psychological characteristics reflect categorical differences between men and women or whether they are a matter of degree. In all likelihood, this is because research has not directly considered the question. The present research adopted an empirical perspective for determining whether gender differences are best characterized as qualitative or as a matter of degree. As we explain, the prevalence and magnitude of gender differences as they have been studied so far are uninformative with regard to this distinction, because differences per se do not tell us whether a *taxon* exists—that is, a categorical variable in which members form nonarbitrary (meaningful and naturally occurring) distinctive classes (Meehl, 1992; Meehl & Golden, 1982; Ruscio & Ruscio, 2008; Waller & Meehl, 1998). We used taxometric methods to determine whether gender differences are taxonic (i.e., representing the existence of distinct categories) or dimensional (i.e., reflecting differences of degree).

Existing Perspectives on the Taxon–Dimension Distinction

The prevailing view among scholars working in this area is that most gender differences are dimensional. Arguably the clearest examples of this belief derive from the results of meta-analyses, which are often conducted to establish the extent to which men and women differ on a given variable (e.g., Hyde, Fennema, & Lamon, 1990; Hyde & Linn, 1988; Oliver & Hyde, 1993). For example, Hyde (2005, 2007) reviewed 46 meta-analyses of sex differences spanning diverse cognitive, social, and personality domains. The observed magnitude and pattern of these results led Hyde to conclude that men and women are substantially more similar than they are different, a conclusion that implies (but does not demonstrate) dimensionality. Another dimensionalist position is that gender differences reflect underlying continuous attributes, such as personality (e.g., communion and agency; Spence & Helmreich, 1978), temperament (Brody, 2000), and hormonal differences (Thornhill & Gangestad, 2008). Researchers who emphasize the impact of culturally derived experiences similarly seem to adopt a dimensionalist orientation, by describing how specific experiences can shape the degree to which an individual develops specific traits or competencies (e.g., Eagly, 1987; Halpern, 2012; Wood & Eagly, 2002, 2012). The dimensionalist position is also popular in social neuroscience, in which arguments about the plasticity and "deep intricacy" (Eliot, 2011, p. 897) of the human brain are cited to explain why reliable sex differences in brain structure and function may have relatively little behavioral impact (e.g., Eliot, 2009, 2011; Halpern, 2012).

Nevertheless, perhaps because the taxon–dimension distinction has not received explicit empirical attention, some researchers appear to favor a more categorical interpretation of gender differences. For example, many evolutionarily oriented researchers view men's and women's social behavior as categorically distinct, reflecting the different adaptive tasks implied by biological differences between men and women (e.g., Geary, 2010). These distinctions are reflected in sexually dimorphic brain structures, such as men's larger and more lateralized brains (see Ellis et al., 2008, for recent meta-analyses) or neural regions that show differential elaboration in response to activation by sex hormones (see Geary, 2010, for a review). Likewise, researchers who theorize about sexuality and mating-related behaviors from an evolutionary perspective typically favor categorical accounts in their conceptualization of gender differences (e.g., Buss, 1995; Buss & Schmitt, 1993; Schmitt, 2002).

More generally, Zahn-Waxler and Polanichka (2004) reviewed several theoretical models positing qualitative differences in the expression and etiology of antisocial behavior in boys and girls. Another example is the "tend-and-befriend model," which proposes that men's and women's responses to stress are fundamentally different (Taylor et al., 2000), linking these behaviors to the influence of oxytocin. Indeed, the fact that researchers sometimes analyze data from men and women separately (Kashy & Kenny, 2000), and the insistence by some editors and reviewers that researchers routinely analyze their data for gender differences (see Baumeister, 1988; Eagly, 1987; and McHugh, Koeske, & Frieze, 1986, for varied positions in this debate), implies belief in at least the possibility that men's and women's behavior may be categorically distinct.

If, as mentioned above, most scholars are skeptical about the general idea that in terms of social behavior men and women represent natural kinds, why does categorical thinking persist? One reason, we propose, is that no existing research has examined explicitly the distinction between categorical and dimensional models. As described in the next section, documentation of gender differences in and of themselves is insufficient to conclusively establish that gender differences are dimensional, or whether they might be taxonic. The research described in this article was designed to provide such evidence.

Understanding Dimensions and Taxa

One reason why the underlying nature of gender differences has been difficult to address is that although biological sex is clearly a categorical variable, the variables commonly of interest to researchers and laypersons alike tend to be dimensional (e.g., masculinity, femininity, school achievement, depression, aggression), varying along a continuum. The statement that men are more aggressive than women, for example, implicitly assumes that there is one group of people who are high in aggression (men) and another group of people who are low in aggression (women). This assumption treats an observed mean difference between men and women as a special kind of category called a taxon. Examples of taxa include animal species (gophers vs. chipmunks), certain physical illnesses (e.g., one either has meningitis or not), and biological sex.

To distinguish a taxon from a mere category, Waller and Meehl (1998) used the example of students receiving an A on an exam. Though A students are in a different category from B students, they are not a taxon because the dividing line between an A and a B is determined by the grader's cutoff on a dimensional scale of performance. Knowing a person's grade in one subject is relatively uninformative about other attributes. However, when individuals are members of a taxon, they are more likely possess traits that are characteristic of that taxon than nonmembers. For example, knowing that Person A has two X chromosomes allows one to accurately conclude that Person A will develop breasts, ovulate, have little facial hair, and exhibit all other characteristics associated with being female. Like many taxa, sex has a genetic basis; with few exceptions, a person is either XX or XY. The indicators for sex (i.e., genetic makeup, anatomy, physiology) are nearly infallible. As Meehl (1995) put it, "There are gophers, there are chipmunks, but there are no gophmunks" (p. 268). When applied to sex, "between day and night there is dusk. But between male and female there is . . . essentially nothing" (Myers, 2008, p. 164).[1]

We sought to empirically determine whether standard gender differences are better conceived as taxonic or dimensional. Although men and women may differ on average in myriad ways, these differences may be dimensional, reflecting different amounts of a given attribute assessed along a single dimension, or qualitative, sorted into fundamentally distinct categories. Taxometric analysis is concerned not only with the magnitude of differences but also with the pattern and distribution of differences across multiple variables. As we will show, this difference has considerable importance for understanding the fundamental nature of gender differences.

What exactly is a taxon? In the simplest sense, a taxon means that a set of variables is essentially uncorrelated within groups due to a tendency toward restriction of range. However, because mean differences exist on each variable, correlations appear when the groups are combined in a single analysis. This is shown in Figure 1. For example, let the variables in Figure 1 be hair length and height; Group 1 is men and Group 2 is women. Within each group, there is no correlation. But when the two groups are combined in a single analysis, because the sex difference introduces additional variability for each variable, the correlation can be large (Ruscio, Haslam, & Ruscio, 2006; Ruscio & Ruscio, 2008). (This is a version of the well-known aggregation fallacy identified by Epstein, 1986.) Thus, understanding the pattern of within- and between-group correlations across multiple variables measured simultaneously in a single sample (rather than group distributions on variables from isolated samples) is necessary to examine the latent structure of gender. This is the novel contribution of the present research.

Why Distinguish Dimensions and Taxa?

There are several reasons why it is important to determine the latent structure of gender (Ruscio & Ruscio, 2008). First, whereas dimensional variations may arise from multiple additive influences, categorical differences require mechanisms capable of creating a dichotomous (on/off) structure. Second, knowledge about the underlying structure of a construct can inform classification, that is, whether individuals vary along dimensions or belong to distinct groups. Third, whereas categories imply that group comparisons are most appropriate for statistical analysis, dimensional variables are more appropriately analyzed by continuous correlational methods.

Taxometric methods have been used profitably in several areas of psychological research. For example, although the *Diagnostic and Statistical Manual of Mental Disorders* "is a categorical classification that divides mental disorders into types based on criteria sets with defining features" (American Psychiatric Association, 2000, p. xxxi), Widiger and Trull (2007) concluded that existing data better support a dimensional model of personality disorder. Other constructs sometimes thought to be categorical but found to be dimensional include attachment (Fraley & Spieker, 2003; Fraley & Waller, 1998; Roisman, Fraley, & Belsky, 2007) and depression (Ruscio & Ruscio, 2000; Whisman & Pinto, 1997). On the other hand, constructs found to be taxonic include schizophrenia (Meehl, 1962, 1990), schizotypy (Korfine & Lenzenweger, 1995; Lenzenweger & Korfine, 1992), self-monitoring (Gangestad & Snyder, 1985, 1991), and Type A behavior (Strube, 1989). In each instance, researchers have used this information to refine and advance theoretical models. Haslam and Kim (2002) provide a more extensive review of taxometric research.

Dimensions and Taxa as Applied to Gender

As noted earlier, gender research has made little or no use of taxometric methods so far. Clearly, many gender differences are present in personality and social behavior (for reviews, see Dindia & Canary, 2006; Ellis et al., 2008; Gilligan, 1982; Halpern, 2012; Maccoby & Jacklin, 1974). Hyde's 2005 (see also Hyde, 2007) review of 46 meta-analyses of gender differences revealed numerous consistent and reliable differences between men and women. Some of the personality and social variables producing larger differences included interruptions, smiling, self-disclosure, aggression, sexuality, anxiety, assertiveness, agreeableness, body-

[1] Though the existence of intersex individuals may call this into question, the rarity of their occurrence casts them as the exceptions that prove the rule.

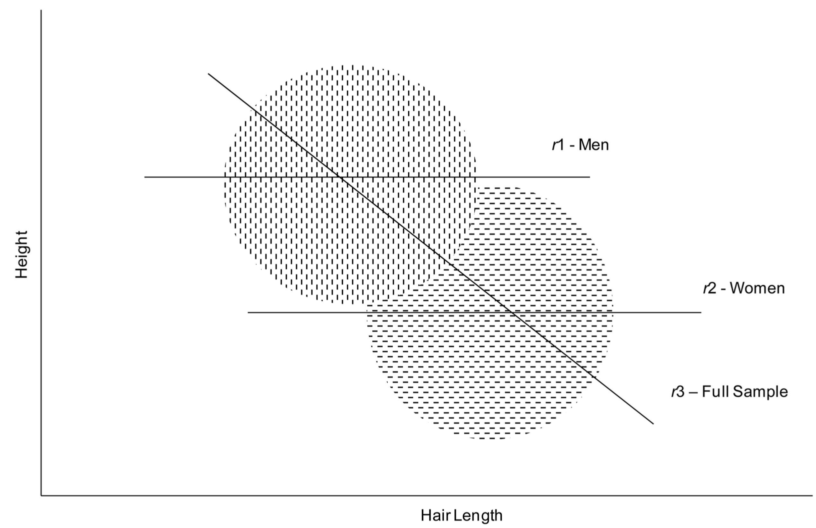

Figure 1. Within-group versus full-sample correlation patterns underlying taxometric analysis. Within-group correlations tend to be small, but can be high over the full sample. Adapted from *Multivariate Taxometric Procedures: Distinguishing Types From Continua* (p. 13, Figure 3.1), by N. G. Waller and P. E. Meehl, 1998, Thousand Oaks, CA: Sage. Copyright 1998 by Sage Publications.

esteem, and self-efficacy about computer use. Hyde's review did not include many other variables of interest to personality and social psychologists, such as sex differences in mating and relational strategies (Buss & Barnes, 1986; Buss, Larsen, Westen, & Semmelroth, 1992; Buss & Schmitt, 1993; Buss & Shackelford, 1997; Kenrick, Sadalla, Groth, & Trost, 1990) or in responses to stress (Taylor et al., 2000). (Other specific study variables are described later in this article.) What is important to note here is that each of these gender differences reflects a reliable difference between men and women. However, the existence of a gender difference in and of itself, and regardless of its magnitude, does not provide direct empirical support for its dimensionality or taxonicity. Rather, as explained above, multiple indicators, along with their intercorrelations, must be examined together in the same sample.

We sought to establish that gender differences are better represented as dimensional than as taxonic constructs in a more definitive way than prior research has done. The implications of this distinction are important. Should gender be taxonic, contrary to prevailing beliefs, men and women could legitimately be said to be qualitatively different in relevant domains. Knowing only that a person was male, we could also infer that he would be relatively aggressive, good in math, poor in verbal skills, primarily interested in short-term mating, less agreeable, and so on, to the same general degree of accuracy as inferring biological characteristics relevant to sex—for example, large waist-to-hip ratio and deeper voice. This is because if gender were taxonic, the relevant behaviors would co-occur in all members of the class. In other words, most men would score similarly to each other on all these indicators, as would women, with little overlap between the two groups. On the other hand, if gender is dimensional, as scholars commonly assume, a person's gender-appropriate behavior on one variable would not imply being high on other gender-related variables. Should our results support dimensionality, scholars would have a stronger empirical basis for their belief.

The Present Research

As discussed above, the underlying structure of gender differences is implicit in many theoretical accounts of social behavior. Our research was designed to provide explicit evidence about a wide variety of psychological traits in which men and women differ. These analyses evaluated the likelihood that existing gender differences are better represented as dimensional or taxonic. This work differs from more exploratory taxometric research because the sex of participants in the samples is already known. If a taxon appears, we can determine whether it is based on sex (in line with what Meehl, 1992, referred to as a real data pseudoproblem). In other words, participants' actual sex is the "gold standard" of prediction. Because the Bayesian distributions obtained from taxometric analyses graphically represent the probability that each individual is a member of the taxon group, one can determine whether whatever taxa emerge from these analyses accurately sort men into one group and women into the other.[2]

Our analyses were conducted on a series of domains commonly studied by researchers interested in differences between women and men. These are detailed below.

Method

Data Set Recruitment

We selected data sets on the basis of three major criteria: that they included variables that have been important in the sex-

[2] One somewhat related approach in prior research is Lippa's work on gender diagnosticity (e.g., Lippa, 1991; Lippa & Connelly, 1990). That research used Bayesian methods to try to find traits that would accurately categorize men and women. Although the idea of gender diagnosticity implies that gender is taxonic with respect to certain indicators, it does not establish whether gender is taxonic or dimensional.

differences-in-social-behavior literature, that they came as close as possible to meeting statistical criteria for taxometric analyses, and that they were available to us. We began by performing a literature search and contacted several researchers regarding data from articles that preliminarily appeared to meet the analysis requirements. In order to meet the requirements of taxometric analysis, data sets were considered eligible if there were at least 300 participants, the gender balance in the sample was no worse that 60:40, and there were at least three continuous and theoretically linked measures showing significant gender differences. Researchers who had collected these data sets were asked to share them with us. Other researchers with known experience in gender difference research were contacted and asked whether they had any data fitting the above criteria that they would be willing to share. Altogether, 31 researchers were contacted: Nine provided data; 15 reported having no data sets that fit the requirements or that requested data were no longer available (some referred us to second authors); seven did not reply. Four data sets were available online.

We also collected further data on several relational, interpersonal, and behavioral variables in one additional sample to examine the taxonicity of gender across these constructs, sometimes repeating earlier measurements with ambiguous results for replication purposes. This sample of 109 men and 167 women[3] with a mean age of 21.15 years (SD = 7.68) was drawn from an introductory-level psychology class at a large midwestern university and is periodically mentioned through the rest of the article as the Midwestern Sample.

Where a data set provided multiple scales along with the raw items for each scale, several analyses could be performed. First, analyses were performed including all mean scores of scales showing a significant sex difference. Second, when individual items were made available, analyses on the individual scales themselves could be performed.

Taxometric Procedures

Three common techniques for taxometric analyses are mean above minus below a cut (MAMBAC), maximum eigenvalue (MAXEIG), and latent mode (L-Mode; Waller & Meehl, 1998). For MAXEIG and L-Mode, at least three variables measuring the desired construct are submitted for analysis (MAMBAC can work with as few as two variables), and the shape of the obtained function is used to judge whether the construct is taxonic or dimensional. Taxonic data produce peaked MAMBAC and MAXEIG curves, and bimodal L-Mode curves. Dimensional data yield flat or U-shaped MAMBAC curves, flat MAXEIG curves, and unimodal L-Mode curves. MAXEIG results can also be used to assign a Bayesian probability that each individual belongs to the taxon group. Taxonic data produce U-shaped histograms with clumps at the tails of the x-axis. With taxonic constructs, some individuals have a very high probability of being a member of the taxon group, and everyone else has a very low probability, with few people in the middle "undecided" ranges. Dimensional constructs produce less discernible Bayesian distributions, meaning that many cases cannot be classified into one group or the other with certainty. In the event that a taxon is identified, it can be used to determine whether men and women were accurately sorted. Dimensional data can also produce U-shaped distributions (Ruscio et al., 2006), so a taxonic-looking Bayesian distribution must be interpreted within the context of the other methods. A dimensional-looking Bayesian distribution, however, makes a strong case for dimensionality.

Interpreting distributions produced by these three methods can be subjective. In the ideal case the plots obtained fit either of the alternatives mentioned above, but in practice these graphs are sometimes ambiguous. We used three methods to control for this ambiguity. First, we drew conclusions only from results that replicated consistently with all three methods. Each procedure tests for taxonicity in a different way, and because interpretation is largely visual, most researchers recommend that multiple methods be used to confirm one another (Waller & Meehl, 1998). Second, simulation methods (Ruscio et al., 2006) help determine the suitability of data for taxometric analyses. On the basis of properties of the actual data (n, number of indicators, correlations between indicators, skew), it is possible to simulate what the data would look like if they were taxonic or dimensional. Data are considered appropriate for taxometric analysis if the simulations are distinguishable. Taxonic and dimensional simulations that are too similar to be distinguished from each other indicate that the data are unsuitable for that method. Third, ambiguous results can be compared with their simulations. Comparison curve fit indices (CCFIs; Ruscio et al., 2006) provide an objective measure of whether the curve generated by the actual data is closer to its taxonic (CCFI of 1) or dimensional (CCFI of 0) simulation. Judgments based on visual inspection should supersede CCFI results when they disagree (Ruscio et al., 2006).

The CCFI for each analysis is reported in Table 1. In Figures 2–4 and A1–A22 (see Appendix, included in the supplemental materials), MAMBAC results are the upper left quadrant, MAXEIG in the upper right, L-Mode in the lower left, and Bayesian probabilities in the lower right. For MAMBAC and MAXEIG, two graphs are presented, one superimposing the actual data over simulated taxonic data (left panel) and the other superimposing the identical data over simulated dimensional data (right panel). L-Mode is presented in one panel, with actual data drawn in the heavy solid line, simulated taxonic data drawn in the lighter solid line, and simulated dimensional data drawn in the dotted line. Comparison of the degree to which actual results match these simulations was used to assign a result to each analysis. Numerical output is also given in the form of base rate estimates (i.e., the proportion of the sample assigned to the taxon) and is provided in the supplemental materials (see Supplemental Table 1). These estimates are unreliable in dimensional data for all three methods, for if there are no true taxon or complement groups, the concept of "base rate" is moot, though consistent estimates may indicate taxonicity.

[3] Seven men and five women were deleted for indicating nonheterosexuality to prevent the possibility of sexual orientation from introducing an additional taxon or contributing additional unaccounted-for variance. Results including these individuals were similar to those reported here. We do admit to a heterosexual bias in our treatment of gender indicators, and although the extent to which sexual orientation can influence the results of gender taxonicity is intriguing, it is beyond the scope of this article.

Table 1
Summary of Graphical Taxonicity and Comparison Curve Fit Indices (CCFIs)

Variable	MAMBAC (Fit)[a]	MAXEIG (Fit)	L-Mode (Fit)
Study Set 1: Validation			
Sex-stereotyped activities	Taxonic (.44)	Taxonic (.64)	Taxonic (.68)
Physical strength: NCAA field events	Taxonic (.66)	Taxonic (.67)	Taxonic (.53)
Anthropometric measurements	Taxonic (.71)	Taxonic (.58)	Dimensional (.56)
Study Set 2: Sexuality and mating			
Sexual attitudes and behaviors	Dimensional (.44)	Dimensional (.34)	Dimensional (.26)
Mate Selectivity (Run 1)	Dimensional (.24)	Unclear (.26)	Unclear (.39)
Mate Selectivity (Run 2)	Dimensional (.38)	Dimensional (.47)	Dimensional (.27)
SOI (Run 1)	Unclear (.16)	Dimensional (.52)	Dimensional (.41)
SOI (Run 2)	Dimensional (.38)	Dimensional (.32)	Dimensional (.34)
Study Set 3: Interpersonal orientation			
IRI subscales (Run 1)	Dimensional (.32)	Dimensional (.38)	Dimensional (.36)
IRI subscales (Run 2)	Dimensional (.29)	Dimensional (.52)	Dimensional (.60)
RISC	Dimensional (.26)	Dimensional (.33)	Dimensional (.30)
Study Set 4: Gender-related dispositions			
Positive Femininity, Negative Masculinity (reversed), Negative Femininity, Unmitigated Communion	Dimensional (.35)	Dimensional (.49)	Dimensional (.23)
Masculinity (reversed), Femininity, Care Orientation	Dimensional (.22)	Taxonic (.33)	Dimensional (.38)
Care Orientation items	Taxonic (.31)	Unclear (.29)	Unclear (.35)
Masculinity, Femininity, Fear of Success	Dimensional (.25)	Dimensional (.34)	Dimensional (.19)
Masculinity items	Dimensional (.23)	Dimensional (.31)	Dimensional (.25)
Femininity items	Dimensional (.30)	Dimensional (.40)	Dimensional (.46)
Fear of Success items	Dimensional (.26)	Dimensional (.49)	Dimensional (.18)
Science Inclination	Dimensional (.19)	Dimensional (.33)	Dimensional (.11)
Big Five personality traits	Dimensional (.26)	Dimensional (.33)	Dimensional (.61)
Study Set 5: Intimacy			
Intimacy Prototype items	Dimensional (.44)	Dimensional (.65)	Dimensional (.25)
Intimacy Stage items	Dimensional (.36)	Dimensional (.27)	Dimensional (.33)
PAIR-M (Best Friend)	Dimensional (.30)	Dimensional (.41)	Taxonic (.50)
SPS items	Dimensional (.35)	Dimensional (.41)	Dimensional (.48)
Study Set 6: All psychological scales (Run 2)			
IRI subscales, mean of SPS and PAIR-M, RISC, Mate Selectivity, SOI (reversed)	Dimensional (.47)	Unclear (.47)	Taxonic (.49)

Note. MAMBAC = mean above minus below a cut; MAXEIG = maximum eigenvalue; L-Mode = latent mode; NCAA = National Collegiate Athletic Association; SOI = Sociosexual Orientation Inventory; IRI = Interpersonal Reactivity Index; RISC = Relational-Interdependent Self-Construal Scale; PAIR-M = Modified Personal Assessment of Intimacy in Relationships Scale; SPS = Social Provisions Scale.
[a] CCFI = 0 (strongly dimensional), CCFI = .5 (inconclusive), CCFI = 1 (strongly taxonic).

All analyses were run in the R language with Ruscio's (2008) programming for taxometric analyses. Ten sets of both taxonic and dimensional simulations were used for MAMBAC, MAXEIG, and L-Mode. All indicators were standardized for all methods to permit comparisons of the actual curves against their simulations. Unless otherwise stated, MAMBAC and MAXEIG were always run with 10 internal replications with only the averaged curves displayed and no smoothing. MAMBAC was run with the summed indicator method, with 50 cuts along input indicators and setting aside 25 cases at each extreme; MAXEIG was run with 50 windows at a 90% overlap; and L-Mode was run with default values (though when there is a "substantial" difference in the number of people in the complement and taxon groups, one must change the default setting on the program to begin looking for the smaller mode at the trough between the two modes). The inchworm consistency test was implemented whenever a large- or small-*n* taxon was suspected from a combination of sloping MAMBAC and MAXEIG curves, along with Bayes or L-Mode outputs, in order to more closely examine otherwise dimensional-but-skewed or inconclusive MAXEIG results.

Establishing Indicator Appropriateness

Several data qualifications are preferred in order to detect taxa: a large sample size (N of at least 300), at least three indicators for most analyses, within-group correlations of .3 or less, and an effect size of at least $d = 1.25$. Taxa are also most easily detected when the sample contains an even mix of taxon and complement members, though small- or large-*n* taxa can still be detected if the previous conditions are favorable (Waller & Meehl, 1998).

Preliminary analyses were performed for each data set to ensure adequate indicator validity for taxometric analyses with regard to effect size and within-sex (nuisance) correlation. Effect sizes were generally lower than ideal, with the highest effect size for a psychological variable at $d = 1.17$ (appeal of sex with more than one partner). Within-sex correlations between indicators were also often greater than .3. The general recommendation is to combine indicators with within-group correlations greater than .3 into a composite indicator (Ruscio & Ruscio, 2004). However, this was often problematic with scale items, as one hallmark of scale reliability is high item correlation, and the composite indicators would often, in turn, correlate highly with remaining items. Given that two out of the three taxometric methods used here require a minimum of three indicators and that the statistical power of these analyses increases with the number of indicators, a significant reduction in the number of indicators was undesirable. Fortunately, another gauge of the appropriateness of an entire data set for taxometric analyses is found with data simulation techniques described above. If one can still distinguish between the taxonic and dimensional simulations derived from a given data set, those data as a whole are suitable for analysis. Lastly, curves from individual indicators provided by MAMBAC and MAXEIG can be examined for consistency. If some indicators demonstrate a pattern very different from that of the others, their inclusion in an averaged curve would mask a clear taxonic or dimensional structure, and are therefore inappropriate for inclusion in the averaged curve (Ruscio et al., 2006). An indicator was considered for removal if its pattern deviated from the others in both MAMBAC and MAXEIG analyses.

Given the low effect sizes for gender relative to what is ideal for taxometric analyses, the approach to indicator selection was one that would provide the best chances to detect sex-based taxa, lest these analyses be accused of stacking the deck toward a dimensional result. The preliminary analyses first determined which indicators demonstrated significant gender differences. Indicators that did not demonstrate a statistically significant gender difference were eliminated from further analyses. Correlations were then examined separately for each sex. Indicators with correlations above .3 but below .5 for both sexes were retained as long as the taxonic and dimensional simulations were distinguishable. In cases of indistinguishable simulations, combining highly within-sex correlated indicators was attempted. When such combinations resulted in continued high nuisance correlations, problematic indicators with the lowest effect sizes were dropped from the analysis instead. Unless otherwise stated, either pairs of indicators yielding a correlation greater than .5 were always combined or, in the case of remaining problematic correlations, the indicator with the smaller effect size for the sex difference was deleted. Individual curves for MAMBAC and MAXEIG were checked for consistency; if an indicator showed a pattern that was different from that of the other indicators for both methods, it was dropped if possible. Significance testing for gender differences, means, standard deviations, effect sizes, within-sex correlations, and indicator deletion/combination can be found in the supplemental materials (see Supplemental Table 2). Additionally, taxometric analyses cannot be performed with missing data, so only those individuals with complete data could be used.

Study Set 1: Validation With Sex-Stereotyped Activities and Physical Measures

The purpose of Study Set 1 was to demonstrate that taxometric procedures can effectively detect taxa in gender-related constructs and demonstrate that sex itself is taxonic.

Sex-Stereotyped Activities

Participants and measure. To develop a measure of sex-stereotyped activities, we asked 30 introductory-level students to record five activities they enjoyed doing in their free time, five activities they thought men typically enjoyed in their free time, and five activities they thought women typically enjoyed in their free time. These items were compiled into a 129-item list that was rated by 38 college students (13 men and 25 women) in terms of enjoyment and how often they engage in the activity. A 1–7 scale was used for each rating, randomly varying the order of the enjoyment (1 = *do not enjoy at all* to 7 = *enjoy very much*) and frequency (1 = *rarely* to 7 = *nearly every day*) sections. The 28 items that revealed significant sex differences for both ratings and no order effects were retained. The Midwestern Sample was then asked to rate how much they enjoyed doing each of these 28 activities[4] on a 1–5 scale. Only those items with a $d > 1$ were retained for our analyses (see Supplemental Table 2). These were boxing (reversed), construction (reversed), playing golf (reversed), playing videogames (reversed), watching pornography (reversed), taking a bath, talking on the phone, scrapbooking, watching talk shows, and cosmetics (beauty design, hair styling, makeup, and nail care activities combined into a single cosmetic variable due to their high correlation).

Results. As shown in Figure 2, both MAMBAC and MAXEIG curves were peaked, the L-Mode curve was bimodal, and the Bayesian distribution was U-shaped and accurately sorted men and women into separate groups. All indicate a taxonic structure based on sex. MAXEIG and L-Mode fit indices, shown in Table 1, supported taxonicity; average base rates were consistent with each other; and L-Mode's right-mode estimate was consistent with the actual sex ratio of the sample (see Supplemental Table 1). Sex-stereotyped activities were therefore indicative of a sex taxon.

Physical Strength

Participants and measures. Data were obtained from the 1998–2002 Divisions I, II, and III NCAA Championships for outdoor track decathlons and heptathlons (National Collegiate Athletic Association, 2007). Four events included in both competitions were analyzed providing distances for long jump, shot put, high jump, and javelin throw (all distances were expressed in meters). The 97 men and 81 women who completed all four events were included in these analyses. When an individual competed more than once, only the data from their last year of competition where they completed all events were included. All events demonstrated significant sex differences. Although the correlations between long jump and high jump exceeded .5 for both men ($r = .510$) and women ($r = .522$), they were not combined because

[4] The final sex-stereotyped activities measurement can be obtained from the first author.

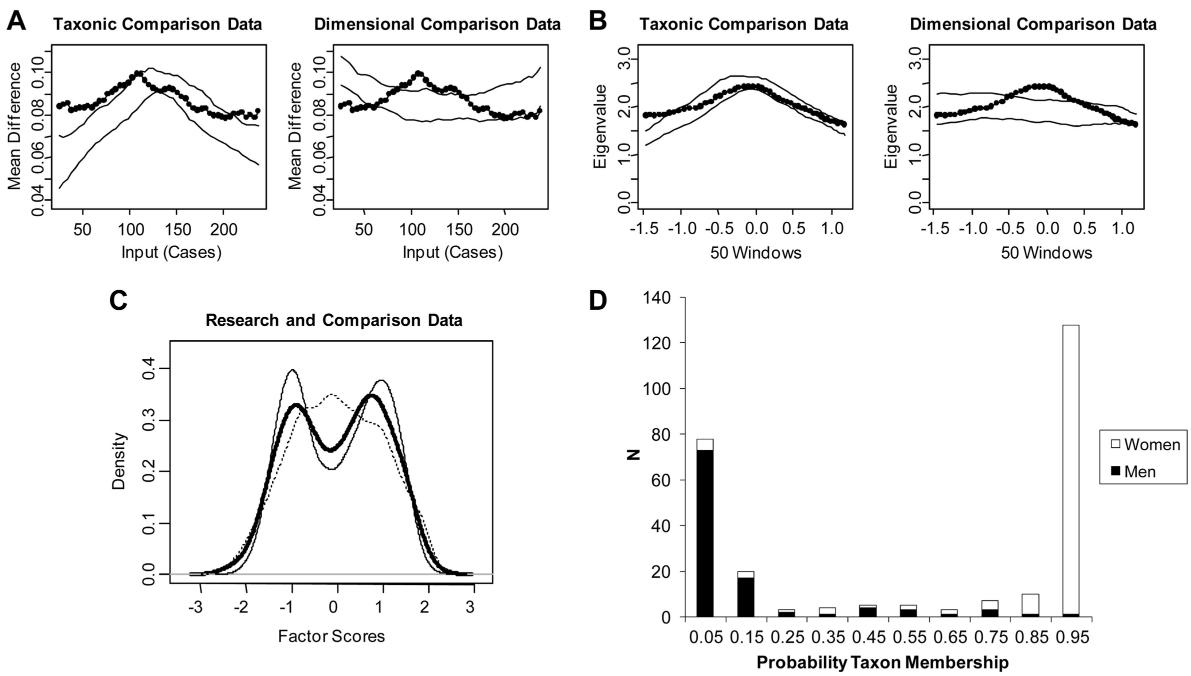

Figure 2. Sex-stereotyped activities: Self-rated enjoyment of boxing (reversed), construction (reversed), playing videogames (reversed), playing golf (reversed), watching pornography (reversed), taking a bath, talking on phone, scrapbooking, watching talk shows, and cosmetics. Mean above minus below a cut (A) and maximum eigenvalue (B) curves were peaked, latent mode (C) was bimodal, and the Bayesian distribution (D) was U-shaped, accurately sorting men and women into separate groups.

remaining correlations would still be problematic. Given that the effect sizes were so high for all variables (d = 1.01–2.60), and the fact that the taxonic and dimensional simulations were easily distinguishable for all methods, all indicators were initially retained for analysis.

Results. Upon preliminary examination of curves for individual indicators in MAMBAC and MAXEIG, shot put was removed from final analyses because it demonstrated a pattern remarkably different from that of the other three indicators (see Supplemental Figure 1), making it an inappropriate candidate to be included in the average curve.

As shown in Figure 3, MAMBAC and MAXEIG were peaked and L-Mode was bimodal. Fit indices for all showed moderate support for a closer fit to their taxonic simulations (see Table 1). Bayesian membership probabilities demonstrated a U-shaped distribution, sorting men into the taxon and women into the complement (nontaxon group). Base rate estimates for MAMBAC, MAXEIG, and L-Mode were consistent with each other as well as with the sex ratio. Physical strength as measured by long jump, high jump, and javelin throw is therefore indicative of a sex taxon.

Anthropometric Measurements

Participants and measures. Data from the National Health and Nutrition Examination Survey III (Centers for Disease Control and Prevention, 1996) conducted from 1988 to 1994 with 33,994 participants from age 2 and up was used to examine anthropometric measurements. The original data set was trimmed to a subset of 584 men and 663 women age 17 or older who provided body measurements. Several measurements were taken in the original collection, but only weight, height, shoulder breadth, arm circumference, and waist-to-hip ratio were of interest. All variables demonstrated significant sex differences. Due to the high within-sex interitem correlation, weight, shoulder breadth, and arm circumferences were combined into one "size" indicator.

Results. Upon preliminary examination of curves for individual indicators in MAMBAC and MAXEIG, the combined size indicator was removed because it demonstrated a pattern remarkably different from that of the other two indicators (see Supplemental Figure 2), making it an inappropriate candidate for averaging the final curve. The size indicator was replaced with shoulder breadth, since it had the largest effect size of all the size indicators.

As shown in Figure A1, MAMBAC was slightly peaked, MAXEIG was clearly peaked, but L-Mode was unimodal. The fit index for MAMBAC was closer to its taxonic simulation, though indices for MAXEIG and L-Mode were inconclusive (see Table 1). Bayesian probabilities did place a fair portion of the sample in the middle of the distribution, though men were placed at the high end while women were placed at the low end of the distribution. Most evidence supports a sex taxon for anthropometric measurements.

Discussion

The measurement of stereotyped activities was clearly taxonic according to each taxometric method. However, the deliberateness with which this measurement was constructed should be kept in mind. The items were selected precisely because they empirically

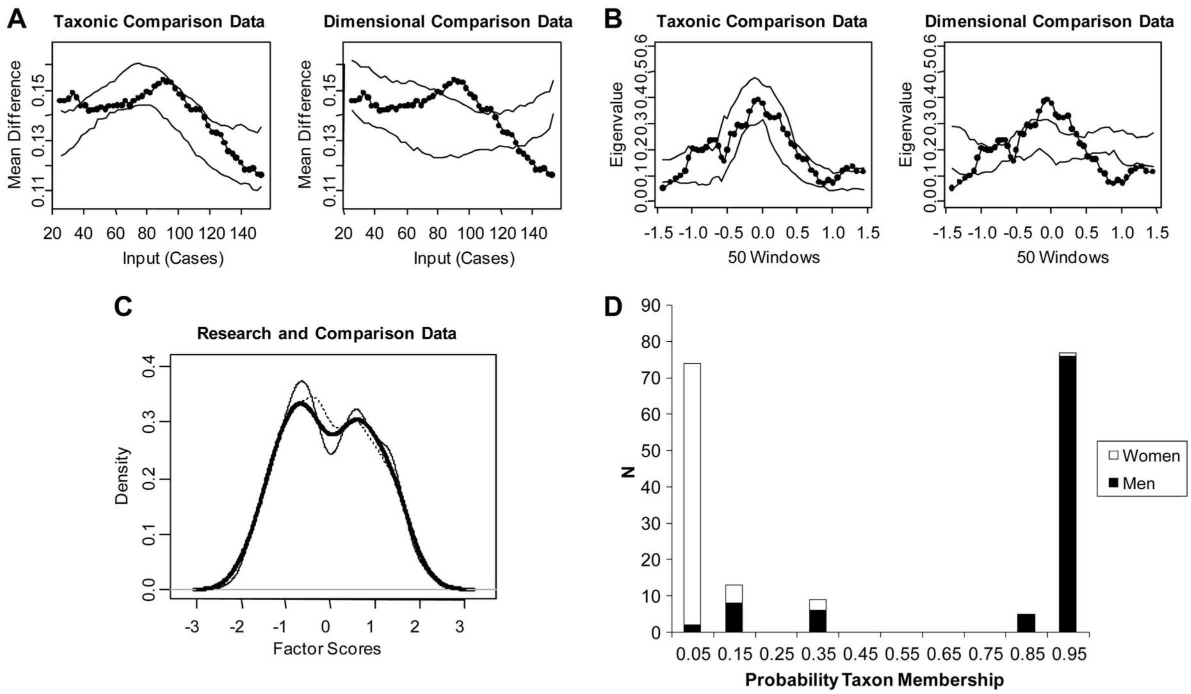

Figure 3. Physical strength: National Collegiate Athletic Association long jump, high jump, javelin throw. Mean above minus below a cut (A) and maximum eigenvalue (B) were peaked and latent mode (C) was bimodal. Bayesian probabilities (D) demonstrated a U-shaped distribution, sorting men into the taxon and women into the complement.

demonstrated gender differences, stacking the deck in favor of a taxonic result. Including a more representative (and probably more androgynous) list of activities may well have resulted in more ambiguous results. Nonetheless, the measurement is an excellent showcase of the ability of taxometrics to identify a gender taxon and accurately sort men and women.

Physical strength and anthropometric measurements were taxonic according to sex. Sex differences in strength and body morphology are readily observable ways in which men and women differ, taxa that can quite literally be identified by the naked eye. It is easy to see upon visual inspection that our bodies are built to do different things, as confirmed by the results.

Study Set 2: Sexuality and Mating

Sexual Attitudes and Behaviors

Participants and measures. Data from the National Health and Social Life Survey (Laumann, Gagnon, Michael, & Michaels, 1992), a nationwide community sample of people between the ages of 18 and 59 living in households (i.e., not including prisoners or students living away from home), were used to examine sexual attitudes and behaviors. With over 1,000 variables and 3,000 participants, the original data set was trimmed to 940 men and 1,296 women, all of whom identified themselves as heterosexual, were U.S. citizens fluent in English, and were rated as able to understand, frank, and cooperative in their responses. We selected items that revealed large gender differences and have been implicated in prior gender difference research as described earlier. These were the appeal of sex with more than one partner, the appeal of having sex with a stranger (4-point scale from 1 = *not at all appealing* to 4 = *very appealing*), unwillingness to have sex without love (reversed, 4-point scale from 4 = *strongly agree* to 1 = *strongly disagree*), how often they have an orgasm during intercourse (4-point scale from 1 = *never* to 4 = *always*), how often they think about sex (6-point scale from 1 = *never* to 6 = *several times a day*), and masturbation frequency (10-point scale from 1 = *0 times this year* to 10 = *more than once a day*). Due to high within-sex correlation, appeal of sex with more than one partner and appeal of having sex with a stranger were combined.

Results. As shown in Figure A2, MAMBAC and MAXEIG curves were nonpeaked, L-Mode was not clearly bimodal (set to search for a right mode beyond a factor score of 1), and all curves were closer to their dimensional than taxonic simulations (see Table 1). The Bayesian distribution appeared slightly U-shaped and sorted men and women with some degree of accuracy, but this was the only evidence that pointed toward taxonicity. Base rate estimates tended to underestimate the sample sex ratio. Thus the latent structure of sexual attitudes and behavior appeared to be dimensional.

Mate Selectivity (Run 1)

Participants and measures. David Buss provided the data from Buss, Shackelford, Kirkpatrick, and Larsen (2001, Study 2), from which 208 male and 365 female college undergraduates provided complete data from the Mate Selection Survey (Hill, 1945), a rating of "the importance of the following factors in choosing a mate" for 18 characteristics such as chastity and good looks on a 4-point scale (1 = *irrelevant/unimportant* to 4 = *indispensable*). Most items showed significant sex differences, though many had to be combined due to high within-sex correlations. One indicator was the mean of importance of similar education, education and intelligence, good financial prospect, ambition and industriousness, and favorable social status; the second was the mean of number of children desired, desire for home and children, preferred marriage age (reversed; all variables were standardized before averaging); the third was the mean of dependable character, emotional stability, pleasing disposition; good looks (reversed) and good cook (reversed) were left as their own indicators. The resulting within-sex correlations were all below .3.

Results. As shown in Figure A3, the MAMBAC curve was nonpeaked, and MAXEIG and L-Mode were indiscernible. Fit indices for all curves indicated that they were closer to their dimensional than taxonic simulations (see Table 1), but the MAXEIG and L-Mode simulations did not do a good job of differentiating taxonic and dimensional data. Bayesian probabilities showed a large proportion of the sample with a high probability of taxon membership made up mostly of women, but an even distribution of the sample throughout the rest of the probabilities. Base rate estimates were inconsistent with each other and the sample sex ratio. Thus the general pattern for Mate Selectivity appeared dimensional, with a cautionary note that the data might not have been appropriate for analyses, given the similarity of their taxonic and dimensional simulations.

Mate Selectivity (Run 2)

Participants and measures. The same 18-item Mate Selection Survey was collected on the Midwestern Sample, this time with a 1–5 rating. A slightly different set of items yielded the largest significant sex differences for use in the taxometric analysis: good looks (reversed), dependable character, emotional stability and maturity, desire for home and children, ambition and industriousness, and good financial prospect.

Results. As shown in Figure A4, MAMBAC and MAXEIG curves were nonpeaked, L-Mode was unimodal, and all curves were closer to their dimensional than taxonic simulations (see Table 1). Bayesian probabilities distinguished two groups, with a complement group low in mate selectivity composed mostly of men. Given the results of the Bayesian probabilities and the downward-sloping MAXEIG curve, an inchworm consistency test was performed on MAXEIG to confirm dimensionality; none of the indicator inputs nor the averaged curve demonstrated signs of taxonicity (see Supplemental Figure 3). Base rate estimates for the three methods were inconsistent with each other. Overall, Mate Selectivity appeared dimensional in this sample, as well.

Sociosexual Orientation (Run 1)

Participants and measures. Jeffry Simpson provided unpublished data from the Sociosexual Orientation Inventory (SOI; Simpson & Gangestad, 1991) completed by both members of 104 heterosexual college undergraduate couples as part of a larger study (Campbell, Simpson, Kashy, & Fletcher, 2001). The SOI measures willingness to have sex outside a committed relationship. Intraclass correlations demonstrated that three of the indicators were uncorrelated within couple. Therefore, the 208 participants could be used without concern for dependence (Kashy & Kenny, 2000). These three indicators were number of one-night stands, $\hat{\rho} = .01$, $F(103, 103) = 1.02$, $p > .2$; how often they fantasize about having sex with someone other than their current partner, $\hat{\rho} = -.03$, $F(103, 103) = 1.06$, $p > .2$; and can imagine being comfortable with casual sex with different partners, $\hat{\rho} = .07$, $F(103, 103) = 1.16$, $p > .2$. Number of one-night stands asks participants to fill in the relevant number and was capped at a value of 30 for all analyses (as per scoring instructions for college samples; Simpson & Gangestad, 1991). Fantasizing asks for ratings on a 1–8 scale (1 = *never* to 8 = *at least once a day*), and comfort with casual sex asks for ratings on a 1–9 scale (1 = *strongly disagree* to 9 = *strongly agree*). Although the correlations between fantasizing and comfort with casual sex were high for both men ($r = .505$) and women ($r = .315$), combining the items was not an option, as that would leave fewer than the three indicators needed for MAXEIG and L-Mode.

Results. MAMBAC and MAXEIG curves for the number of one-night stands indicator demonstrated patterns different from those of the other two indicators (see Supplemental Figure 4), but since MAXEIG and L-Mode require three indicators, it was necessary to include in the analysis. As shown in Figure A5, MAMBAC demonstrated a small peak at the far right, MAXEIG was nonpeaked, and L-Mode was unimodal. L-Mode's taxonic and dimensional simulations were nearly indistinguishable, indicating that the data may not be appropriate for this analysis. The fit index for MAMBAC was dimensional, but inconclusive for MAXEIG and L-Mode (see Table 1). Bayesian probabilities separated two groups, with a small-*n* high-SOI taxon consisting mostly of men. All methods demonstrated a skew indicating a greater number of people at the low end of the scale. Base rate estimates for all three methods were inconsistent with each other and the sample sex ratio. The fact that a greater proportion of men were placed in the complement group than the taxon group indicates that whatever underlies this construct is not biological sex. Given the Bayesian distribution results, the upward slopes of MAMBAC and MAXEIG, and the cusp at the end of MAMBAC, an inchworm consistency test was performed on MAXEIG to investigate the possibility of a small-*n* taxon. Of the three variables, the curve using fantasizing about having sex with someone other than their current partner as an input demonstrated small-*n* taxonicity, which was enough to also make the averaged curve cusp and appear small-*n* taxonic (see Supplemental Figure 5). In sum, sociosexuality showed no evidence of a gender taxon, but did show some signs of a small taxon of men high in SOI.

Sociosexual Orientation (Run 2)

Participants and measures. The SOI was also collected in the Midwestern Sample. Items demonstrating significant differences were number of partners expected in the next 5 years (capped at 30), number of one-night stands, fantasizing about someone other than their partner, feeling comfortable with sex without love, comfort with casual sex, and requiring closeness for sex (reversed). The last three were measured with a 1–5 rating. Because these last three items were highly correlated with each other, they were aggregated for the taxometric analysis.

Results. As shown in Figure A6, MAMBAC and MAXEIG curves were nonpeaked, and L-Mode was mostly unimodal, though did find a very small mode at the far right when set to look past a factor score of 4. All curves were closer to their dimensional than taxonic simulations, as confirmed by dimensional fit indices (see Table 1), though the taxonic and dimensional simulations for L-Mode were not easily distinguished. As in Run 1, the Bayesian probabilities distinguished two groups, with approximately half the men and a few women in the high-SOI taxon, and most of the women and nearly half the men sorted into the low-SOI complement. In this analysis, the MAMBAC and MAXEIG curves showed a substantial upward slope, indicating positive skew of the measure and a possible small-*n* taxon. To examine this possibility, we performed an inchworm consistency test to confirm the dimensionality of MAXEIG. The two curves using number of one-night stands and fantasizing about sex with someone other than their partner as inputs showed some evidence of a small-*n* taxon; there was no evidence for the other two indicators, and the averaged curve did not peak, indicating dimensionality (see Supplemental Figure 6). Base rate estimates were inconsistent with each other and the sample sex ratio. Sociosexuality showed no evidence of a taxon based on biological sex and good evidence of a dimensional structure, although, as in Run 1, there was some evidence of a small-*n* taxon.

Discussion

Constructs dealing with sexual behaviors and mating appeared generally dimensional with all three taxometric methods, although the Bayesian membership probabilities tended to sort men and women into separate groups with fair accuracy in the expected directions—the high end of sociosexuality and low end of mate selectivity were more likely to contain men, and the opposite ends more likely to contain women. Recall that a taxonic-looking Bayesian distribution does not necessarily indicate taxonicity, though the accurate sex sorting is difficult to ignore.

The meaning of dimensional sexuality is interesting when interpreted along the lines of mating behavior as explained by evolutionary psychology. Given the qualitatively different reproductive structures and functions of men and women as identified by evolutionary psychologists, qualitatively different mating strategies would seem reasonable. Indeed, the effect sizes for sex on these variables were the largest of all the variables examined in this study (with the exception of the validation variables). However, the dimensional findings suggest that those large mean differences were not consistent within each person; that is, scoring in a masculine way on one variable did not guarantee doing so on the others. Indeed, qualitative differences may apply only in minimal parental investment situations, in which the consequences of a poor mate choice are far more costly for females than for males, resulting in females being more selective than males, and where competition among males for access to females is more intense (Trivers, 1972). Of course in humans, both parents typically devote long periods to raising their offspring; thus the standards that men and women have in choosing long-term mates tend to be equally high on many characteristics (Kenrick et al., 1990). The SOI does lend itself to a great deal of skew, particularly in college populations, with most people at low to moderate levels, and a few with very high levels. Additionally, the Run 1 of the SOI should be interpreted with caution, as this was a sample of couples who in restricting their sexual activities to their current partner underrepresent those who are particularly high in sociosexuality and would not limit themselves to one person for long enough to be considered part of a couple. The addition of the Midwestern Sample did lend more credence to the possibility of a small-*n* taxon of men who were very willing to have sex without any commitment. Although taxometric analyses generally require large samples, the detection of small-*n* taxa is particularly demanding in this regard. Further study with these constructs using much larger samples would shed further light on their latent structures.

Study Set 3: Interpersonal Orientation

Empathy

Participants and measures. The Interpersonal Reactivity Index (IRI; Davis, 1980, 1983) was completed by 139 male and 184 female college students provided by Mark Davis (Run 1), as well as the Midwestern Sample (Run 2). The IRI is a multidimensional measure of empathy, encompassing Fantasy, Empathic Concern, Perspective Taking, and Personal Distress subscales. Each subscale contains seven phrases such as “I often have tender, concerned feelings for people less fortunate than me” rated on a 5-point scale from 1 = *does not describe me well* to 5 = *describes me very well*.

All four subscales showed sex differences for Run 1. The Empathic Concern and Perspective Taking subscales demonstrated high within-sex correlations, and so were combined. For Run 2, the Perspective Taking subscale did not show significant sex differences, so it was excluded from any further analyses for that sample.

Results. As shown in Figure A7 for Run 1, MAMBAC and MAXEIG were nonpeaked and closer to their dimensional simulations than their taxonic ones, as confirmed by their moderately dimensional fit index (see Table 1). L-Mode was unimodal with a moderately dimensional fit index, but its taxonic and dimensional simulations were not easily distinguishable, indicating that L-Mode may not be appropriate for these data. Bayesian probabilities presented a gap in the distribution indicating a possible taxon, though not one that sorted men and women into different groups. Base rate estimates were inconsistent with each other and the sample sex ratio. Most evidence for Run 1 of the entire IRI measurement supported a dimensional structure.

As shown in Figure A8 for Run 2, MAMBAC and MAXEIG were nonpeaked and appear closer to their dimensional than taxonic simulations, as confirmed by strong dimensional fit index for MAMBAC, though the index for MAXEIG was inconclusive (see Table 1). L-Mode was mostly unimodal, but a very small mode at the left side could be distinguished when setting it to search past a factor score of −3, though the taxonic and dimensional simulations were not easily distinguishable, and its fit index was inconclusive. Bayesian probabilities demonstrated a large-*n* taxon, though there was no sorting according to sex. Base rate estimates were inconsistent with each other and the sample sex ratio. Due to the slight cusp at the left side of the MAMBAC curve and the possible small mode on the left side of L-Mode, an inchworm consistency test was run, but only the curve with Fantasy as an input demonstrated signs of taxonicity (see Supplemental Figure 7). Run 2 therefore confirmed the dimensionality of the entire IRI.

Relational Interdependence

Participants and measures. The Midwestern Sample completed the Relational-Interdependent Self-Construal Survey (RISC; Cross, Bacon, & Morris, 2000), an 11-item measurement of how much a person's relationships make up his or her sense of self. Participants rated items such as "In general, my close relationships are an important part of my self-image" on a 5-point scale (1 = *strongly disagree* to 5 = *strongly agree*). Items 2, 3, 5, 7, and 9 (reversed) were used for this analysis due to high within-sex correlations for the other items.

Results. As shown in Figure A9, MAMBAC and MAXEIG curves were both nonpeaked and closer to their dimensional than taxonic simulations, as confirmed by their dimensional fit indices (see Table 1). L-Mode was unimodal and closer to its dimensional than taxonic simulation, also confirmed by a dimensional fit index. Bayesian probabilities demonstrated a negatively skewed distribution, with little sorting according to sex. MAMBAC and L-Mode base rates were consistent with each other but not with the sample sex ratio, which was better estimated by the MAXEIG base rate. Evidence supported a dimensional structure for the RISC.

Discussion

Measurements of interpersonal orientation were generally dimensional. Thus, contrary to the assertions of pop psychology titles like *Men Are From Mars, Women Are From Venus* and *The Rules*, it is untrue that men and women think about their relationships in qualitatively different ways. Even leading researchers in gender and stereotyping can fall into the same trap. Taylor et al. (2002) claimed that "females seek and give social support at levels that are markedly, robustly, and *qualitatively* different from those of men" (p. 752, emphasis added). Although these gender differences in behavior are no doubt marked and robust (quantitative), it appears that they are not of different (qualitative) kinds. As with the sex and mating analyses, the differences between men and women on interpersonal variables are not consistent enough on an individual basis—those who score in a stereotypical way on one measure do not necessarily do so on another. On the other hand, the attitudes people report on a paper-and-pencil measurement and their actual behavior are not entirely consistent (Mischel, 1968). Thus the dimensionality of interpersonal cognitions need not preclude the taxonicity of interpersonal behaviors. Methods that more pointedly measure interpersonal behaviors (how many birthday cards have they sent this year, how many times a month do they call a friend just to see how he or she is, etc.) may more readily reveal a gender taxon.

Study Set 4: Gender-Related Dispositions

Masculinity, Femininity, Unmitigated Communion

Participants and measures. The Extended Personal Attributes Questionnaire regarding positive (desirable) and negative (undesirable) masculinity and femininity (Spence, Helmreich, & Holahan, 1979), with 5-point ratings between extreme degrees of each trait (1 = *very arrogant* to 5 = *not at all arrogant*), was administered to 107 male and 259 female college undergraduates (Aubé, 2004). It was measured in conjunction with Helgeson's (1993) Unmitigated Communion Scale, measuring the concern for others at the expense of the self, rated on a 5-point scale for agreement with eight phrases (e.g., "I always place the needs of my family above my own"). All but positive masculinity demonstrated significant gender differences. Raw data were unavailable for each item of the scales, so one taxometric analysis was performed for the mean levels of positive femininity, negative masculinity (reversed), negative femininity, and unmitigated communion.

Results. As shown in Figure A10, MAMBAC and MAXEIG curves were both nonpeaked and closer to their dimensional than taxonic simulations, as confirmed by their moderate (MAMABC) and weakly (MAXEIG) dimensional fit indices (see Table 1). L-Mode was unimodal and closer to its dimensional than taxonic simulation as indicated by its strong dimensional fit index. Bayesian probabilities demonstrated a gap in the distribution, but with little sorting by sex. Base rate estimates were inconsistent with each other and the sample sex ratio. Most evidence therefore supported a dimensional structure for masculinity, femininity, and unmitigated communion as indicators for gender.

Masculinity, Femininity, Care Orientation

Participants and measures. Stephen Quackenbush provided data from 244 male and 272 female college undergraduates who filled out the masculinity and femininity scales of the Bern Sex-Role Inventory (Bem, 1974), as well as the Lessons Learned Questionnaire of moral orientation, from which justice and care orientation factors were extracted (Barnett, Quackenbush, & Sinisi, 1995). This measure asked people to recall which personal experience had the greatest impact on their moral development, and rate on a 7-point scale to what extent they learned a moral lesson about a particular concept. Justice items include concepts such as "the rights of others" and "the difference between right and wrong," and care items include concepts such as "intimate relationships" and "caring for others." Gender differences were significant for all but justice orientation. Item scores for the masculinity and femininity scales were no longer available, so two sets of taxometric analyses were conducted. The first used masculinity (reversed), femininity, and care orientation as indicators of gender; the second used the three care orientation items of the Lessons Learned Questionnaire that demonstrated significant sex differences (Items 12–14).

Results: Masculinity, femininity, and care orientation. As shown in Figure A11, MAMBAC was nonpeaked and closer to its dimensional than taxonic simulation, as confirmed by its strongly dimensional fit index (see Table 1). MAXEIG was peaked, though it was closer to its dimensional simulation, if only weakly indicated by its fit index. L-Mode was unimodal and closer to its dimensional than taxonic simulation, as confirmed by its moderately dimensional fit index. Bayesian probabilities separated the sample into two distinct groups with some sorting according to sex. Base rate estimates were inconsistent with each other and the sample sex ratio. The evidence demonstrated reserved support for a dimensional structure of the Bern Sex-Role Inventory and care orientation as indicators of gender.

Results: Care orientation items. As shown in Figure A12, MAMBAC showed a very small peak at the left end of the distribution, though it was closer to its dimensional than taxonic simulation, as confirmed by a moderately dimensional fit index (see Table 1). Additionally, MAMBAC's base rate estimate was far lower than what would be expected by a peak so far to the left. The MAXEIG curve showed a gentle peak, but it was to the left of where taxonic simulations predicted it to be, and the fit index was strongly dimensional. L-Mode was multimodal (setting it to search for the left mode beyond −1.5 and right mode beyond 1 standard deviation), yet was also closer to its dimensional than taxonic simulation. Bayesian probabilities were ambiguous, placing only a small portion of the sample in the lower range of the distribution, but demonstrated no sorting according to sex. All methods reflected the generally negative skew of the three items used for analysis (0.04, −1.18, and −1.16). On the basis of these mixed results, the structure of care orientation is unclear at this time.

Masculinity, Femininity, Fear of Success

Participants and measures. Harry Reis (1980) provided data from 337 high school boys and 364 girls who completed the Personal Attributes Questionnaire (PAQ; Spence, Helmreich, & Stapp, 1974) for masculinity and femininity, as well as the Fear of Success Scale (FOSS; Zuckerman & Allison, 1976). The PAQ is similar to the Extended Personal Attributes Questionnaire, except that it only taps desirable aspects of masculinity and femininity. For the FOSS, participants rated phrases such as "Often the cost of success is greater than the reward" on a 1–5 scale (1 = *disagree completely* to 5 = *agree completely*). Data for all the scale items were available, so it was possible to perform four taxometric analyses: one for the structure of gender as defined by masculinity (reversed), femininity, and fear of success as well as one for each of the individual scales. Six of the seven masculinity items (competitive, difficulty with decisions [reversed], never gives up, confident, feels superior, and stands up under pressure), all eight femininity items (emotional, devote self to others, gentle, helpful, kind, aware of others' feelings, understanding, and warm to others), and nine of the 18 FOSS items demonstrated significant sex differences, so were used for analysis of the individual scales. Item 6 of the FOSS was sex significant but deleted for individual scale analysis because it was in the opposite direction from what was expected, leaving Items 1, 4, 5, 8, 9, 11, 13, 16, and 17 for final analysis, with Items 4, 5, 8, 11, and 16 reversed so that women scored higher on all items.

Results: Masculinity, femininity, fear of success. As shown in Figure A13, MAMBAC and MAXEIG curves were nonpeaked and closer to their dimensional than taxonic simulations, as confirmed by their strong (MAMBAC) and moderate (MAXEIG) dimensional fit indices (see Table 1). L-Mode was unimodal and closer to its dimensional than taxonic simulation as confirmed by its strongly dimensional fit index, though the dimensional and taxonic simulations were difficult to distinguish. Bayesian probabilities showed a U-shaped distribution, and although there appeared to be some sorting of women into the higher end, men appeared to be more evenly distributed. Base rates for MAMBAC and L-Mode accurately predicted the sample sex ratio, but the base rate for MAXEIG overestimated it. Most of the evidence supported a dimensional structure of gender as defined by masculinity, femininity, and fear of success.

Results: Masculinity items. As shown in Figure A14, MAMBAC and MAXEIG curves were nonpeaked and closer to their dimensional than taxonic simulations, as confirmed by their strong (MAMBAC) and moderate (MAXEIG) dimensional fit indices (see Table 1). L-Mode was unimodal and closer to its dimensional than taxonic simulation, as confirmed by its strong dimensional fit index. The Bayesian distribution was somewhat U-shaped, with women more likely to be on the low end, but little sorting of men. Base rates for MAXEIG and L-Mode accurately predicted the sample sex ratio, but the base rate for MAMBAC overestimated it. The evidence supports a dimensional structure for masculinity as measured by the PAQ.

Results: Femininity items. As shown in Figure A15, MAMBAC and MAXEIG curves were nonpeaked and closer to their dimensional than taxonic simulations, as confirmed by their moderate (MAMBAC) to weak (MAXEIG) dimensional fit indices (see Table 1). L-Mode was unimodal, though did detect a very small mode at the far left when setting to look beyond a factor score of −3.5. It was closer to its dimensional than taxonic simulation, but the simulations were not easily distinguished, and its fit index was only weakly dimensional. Bayesian probabilities placed most of the sample near the center of the distribution, and showed no sorting according to sex. Base rates for MAMBAC and accurately predicted the sample sex ratio, but the base rates for MAXEIG and L-Mode overestimated it. The evidence supported a dimensional structure for femininity as measured by the PAQ.

Results: FOSS items. As shown in Figure A16, MAMBAC and MAXEIG curves were nonpeaked and closer to their dimensional than taxonic simulations, as confirmed by their strong (MAMBAC) and weakly (MAXEIG) dimensional fit indices (see Table 1). L-Mode was unimodal, closer to its dimensional than taxonic simulation, and had a strong dimensional fit index. Bayesian probabilities showed more women at the higher end of the distribution, but little sorting for men. MAMBAC and L-Mode base rates accurately estimated the sample sex ratio, but the base rate for MAXEIG underestimated it. Evidence supported a dimensional structure for fear of success as measured by the FOSS.

Science Inclination

Participants and measures. Data from the Programme for International Student Assessment, an international study of 15-year-old students, were used to examine inclination toward science subjects and careers (Organisation for Economic Co-operation and Development, 2006). The Programme for International Student Assessment conducts surveys every 3 years regarding a variety of academic subjects in order to evaluate educational systems. The 2006 collection focused on science (Organisation for Economic Co-operation and Development, 2009). With almost 500 variables and 400,000 participants, the original data set was trimmed to 2,660 girls and 2,645 boys from the United States. We selected items with 1–4 ratings that could easily be averaged into cohesive scale scores that demonstrated gender differences: enjoyment (i.e., "I like reading about <broad science>," 1 = *strongly disagree* to 4 = *strongly agree*), ease of performing science tasks (i.e., "Recognize the science question that underlies a newspaper report on a health issue," 1 = *I couldn't do this* to 4 = *I could do this easily*), value of science (i.e., "Advances in <broad science and technology> usually improve people's living conditions," 1 = *strongly disagree* to 4 = *strongly agree*), frequency of science activities (i.e., "Watch TV programs about <broad science>," 1 = *never or hardly ever* to 4 = *very often*), interest in learning (i.e., "Topics in physics," 1 = *no interest* to 4 = *high interest*), environmental awareness (i.e., "The increase of greenhouse gases in the atmosphere," 1 = *I have never heard of this* to 4 = *I am familiar with this and I would be able to explain this well*), interest in a science career (i.e., "I would like to work in a career involving <broad science>," 1 = *strongly disagree* to 4 = *strongly agree*), and ease of learning science (i.e., "Learning advanced <school science> topics would be easy for me," 1 = *strongly disagree* to 4 = *strongly agree*). Within-sex correlations among most scale averages were high, so enjoyment, value, activity frequency, interest in learning, and interest in science career were combined into a single variable; ease of learning science and ease of science tasks were combined into another variable, with environmental awareness remaining as a separate variable. Many within-sex correlations for the three resulting indicators were still above .5, but given that the taxonic and dimensional simulations were easily distinguishable for all methods, these indicators were retained for analysis.

Results. As shown in Figure A17, MAMBAC and MAXEIG curves were nonpeaked and closer to their dimensional than taxonic simulations, as confirmed by their strong dimensional fit indices (see Table 1). (MAXEIG was run with 75 windows at 50% overlap to complete the analysis.) L-Mode was unimodal, closer to its dimensional than taxonic simulation, and had a strong dimensional fit index. Bayesian probabilities placed most of the sample in the middle of the distribution. MAMBAC and L-Mode base rates accurately estimated the sample sex ratio, but the base rate for MAXEIG slightly overestimated it. Evidence supported a dimensional structure for science inclination.

Big Five Personality Traits

Participants and measures. Brad Sheese and William Graziano provided data from 393 male and 460 female college undergraduates taking an introductory psychology class on the Big Five Inventory (John, Donahue, & Kentle, 1991). Students provided 1–9 ratings (1 = *disagree strongly* to 9 = *agree strongly*) to a question stem of "I see myself as someone who . . ." and items such as "is talkative" and "does a thorough job." Significant sex differences were found for the Extraversion, Openness, Agreeableness, and Emotional Stability subscales, but not for Conscientiousness. Extraversion and Agreeableness scores were reversed so that men scored high on all scales.

Results. As shown in Figure A18, MAMBAC and MAXEIG curves were flat and closer to their dimensional than taxonic simulations, as confirmed by their strong (MAMBAC) and moderate (MAXEIG) dimensional fit indices (see Table 1). Taxonic and dimensional simulations for MAXEIG were indistinguishable, so these data may not have been appropriate for this analysis. L-Mode was unimodal with a weak taxonic fit index, though its taxonic and dimensional simulations were also indistinguishable. Bayesian probabilities demonstrated a separation of groups, with little sorting according to sex. Base rates for MAMBAC and MAXEIG were inconsistent with each other and the sample sex ratio, though the base rate for L-Mode accurately predicted it. Most evidence therefore supported a dimensional structure of personality as measured by extraversion, openness, agreeableness, and emotional stability, four of the five factors of personality, though results should be interpreted with caution given the indistinguishable nature of the MAXEIG and L-Mode simulations.

Discussion

The fact that the constructs of masculinity and femininity were dimensional is striking, given the psychological and semantic link of these traits to sex (Cantor & Mischel, 1979), and that even a subtle manipulation into categorical thinking can lead to taxonic results in the use of rating scales (Beauchaine & Waters, 2003). Particularly with the Reis (1980) data that provided item scores for masculinity and femininity, rating oneself on a series of gendered traits might cue the participants into realizing that they were rating themselves on gendered characteristics. Such a realization could easily cause their subsequent ratings to drift to the "appropriate" responses, given the importance and centrality of gender identity. Apparently, this did not happen here. Our results confirm Korfine and Lenzenweger's (1995) findings with femininity and extend them to masculinity. The dimensionality of gendered personality means that masculinity and femininity are not all-or-nothing traits, but that they are truly a continuum.

Cognitions regarding success and science inclination also demonstrated a clear pattern. Models for fear of success and science inclination were consistently dimensional, despite the data having been collected in a school setting where students' assessments of their academic performance and interest would be most salient.

Smetana (2002) demonstrated the dimensionality of personality as conceptualized by the five factor model, and those results are replicated here. Despite Dahlstrom's (1995) advocacy of personality taxonomies, the attempt to apply a typology to a nontaxonic construct does not work. One can only really discuss the degree to which one demonstrates the characteristics of extraversion, agreeableness, and so forth. Likewise, when examining the four factors that did demonstrate significant sex differences (extraversion, openness, agreeableness, and emotional stability) as a whole, the result was dimensional, thereby rendering the five factor model of personality a poor choice in accurately sorting men and women.

Study Set 5: Intimacy

Intimacy Prototype

Participants and measures. Arthur Aron provided data from 182 male and 337 female college undergraduates from Studies 1, 2, 4, and 6 of Aron and Westbay (1996) on the love prototype scale (Fehr, 1988). For a series of relationship features such as openness and caring, participants rated "How central is this feature to love?" on an 8-point scale (1 = *extremely poor feature* to 8 = *extremely good feature*). Five of the 12 items demonstrated significant sex differences: supportive, patience, caring, trust, and feel good about the self.

Results. As shown in Figure A19, MAMBAC and MAXEIG curves were nonpeaked, and MAMBAC was closer to its dimensional than taxonic simulation. (MAXEIG was run with an overlap of 75% in order to complete analyses.) The fit index for MABAC was inconclusive, however, and MAXEIG had a weak taxonic fit (see Table 1). L-Mode was unimodal and had a strong dimensional fit, but the simulations were nearly indistinguishable, even when set to capture a possible small left mode beyond a factor score of −3. Bayesian probabilities place nearly the entire sample in a taxon and a very small portion in the midrange. All curves demonstrated the shifts resulting from the consistent negative skew in the distributions of the original items (ranging from −0.89 to −3.41). Base rate estimates were consistent with each other, but overestimated the sample sex ratio. Intimacy prototype was therefore dimensional.

Intimacy Stage

Participants and measures. Arthur Aron also provided data from Aron and Westbay (1996) in which 88 men and 200 women (Studies 1 and 2) completed the Modified Eriksonian Psychosocial Stage Inventory (Darling-Fisher & Leidy, 1988) measuring identity and intimacy stage resolution. Undergraduates completed the Identity and Intimacy subscales, each containing 10 statements, five of which reflect successful crisis resolution and five reflect unsuccessful resolution, and each rated on a 5-point scale (1 = *almost always true* to 5 = *hardly ever true*). Sex differences were statistically significant for intimacy stage resolution, but not for identity stage resolution. Six out of the 10 intimacy stage items demonstrated significant sex differences and were included for analyses (Items 3, 4, 5, 8, 15, and 16). Items 5 and 15 were reverse scored so that women scored higher than men on all items.

Results. As shown in Figure A20, MAMBAC and MAXEIG curves were nonpeaked, L-Mode was unimodal, all curves were closer to their dimensional than taxonic simulations, and the Bayesian distribution demonstrated no discernibly taxonic pattern. Fit indices showed moderate (MAMBAC and L-Mode) and strong (MAXEIG) support for a dimensional structure (see Table 1). Base rate estimates were inconsistent with each other, with only MAXEIG approaching the sample sex ratio. The intimacy stage construct was therefore considered dimensional.

Intimacy With Best Friend

Participants and measures. The Midwestern Sample completed the Best Friend version of the Modified Personal Assessment of Intimacy in Relationships Scale (PAIR-M; Thériault, 1998), a 16-item measure of intimacy derived from Schaefer and Olson's (1981) Personal Assessment of Intimacy in Relationships Scale, which is oriented to marriage. The PAIR-M has three subscales tapping positive ("I listen to my best friend when he/she needs someone to talk to"), negative ("Serious discussions make me realize how few ideas I have in common with my best friend"), and social ("I prefer that my best friend and I spend time with other friends rather than just with each other") intimacy, rated on a 5-point scale (1 = *strongly disagree* to 5 = *strongly agree*). Eight of the 11 items (Items 1, 3, 7, 9, 10, 12, 14, and 16) on the positive and negative subscales showed a significant sex difference and acceptable within-sex correlations, so were used in the taxometric analyses. Scores for Items 7, 9, 12, and 14 were reversed so that women scored higher on average for all items.

Results. As shown in Figure A21, MAMBAC and MAXEIG curves were nonpeaked and closer to their dimensional simulations, as confirmed by moderate (MAMBAC) and weak (MAXEIG) dimensional fit indices (see Table 1). L-Mode was bimodal, set to search for a right mode beyond a factor score of 0.5, but its fit index was inconclusive. Bayesian probabilities placed most people at the high end of the scale, with a small group in the middle of the distribution, and an even smaller group of men at the low end. Base rate estimates were inconsistent with each other, with only that from MAMBAC approaching the sample sex ratio. Therefore, the PAIR-M was likely dimensional.

Social Provisions in Intimate Relationships

Participants and measures. The Midwestern Sample also completed the Social Provisions Scale (SPS; Cutrona & Russell, 1987), a measure assessing the provisions of intimacy within close relationships with 24 items such as "There are people I can depend on to help me if I really need it." The measure is composed of six subscales: Guidance, Reassurance of Worth, Social Integration, Attachment, Nurturance, and Reliable Alliance, each measured with four items rated on a 5-point scale (1 = *strongly disagree* to 5 = *strongly agree*). All but the Reassurance of Worth and Social Integration subscales demonstrated significant sex differences, but a taxometric analysis using the other four subscales was not possible because of large within-sex correlations (most $r > .5$). Taxometric analyses for each subscale were also not possible, because MAXEIG and L-Mode require at least three indicators, and even if three of the four subscale items demonstrated significant sex differences, they typically had within-sex correlations exceeding .5. Therefore, one analysis was run on the five items that both demonstrated significant sex differences and had acceptable within-group correlations (Items 9, 15, 17, 23, and 24).

Results. As shown in Figure A22, MAMBAC and MAXEIG were nonpeaked, as confirmed by moderate (MAMBAC) to weak (MAXEIG) dimensional fit indices (see Table 1). L-Mode was unimodal, but had an inconclusive fit index. Bayesian probabilities grouped almost the entire sample into one location on the distribution. Base rate estimates were inconsistent with each other and the sample sex ratio. Therefore, the SPS was most likely dimensional in structure.

Discussion

Measures of intimacy and social support were dimensional, indicating that being intimacy ready (Erikson, 1950) is a matter of degree, rather than simply being ready or not. The consistent negative skew for the items measuring these constructs is also noteworthy, demonstrating that most people are capable of having and being ready for intimate relationships and that most feel well integrated into their social networks.

Study Set 6: All Psychological Scales

Participants and Measures

All the psychological scales completed by the Midwestern Sample (IRI, SPS, PAIR-M, RISC, MSQ, and SOI; i.e., most of the mating, interpersonal, and intimacy scales, excluding the behavioral measurement used for validation) were examined in one taxometric analysis. For measures containing multiple subscales,

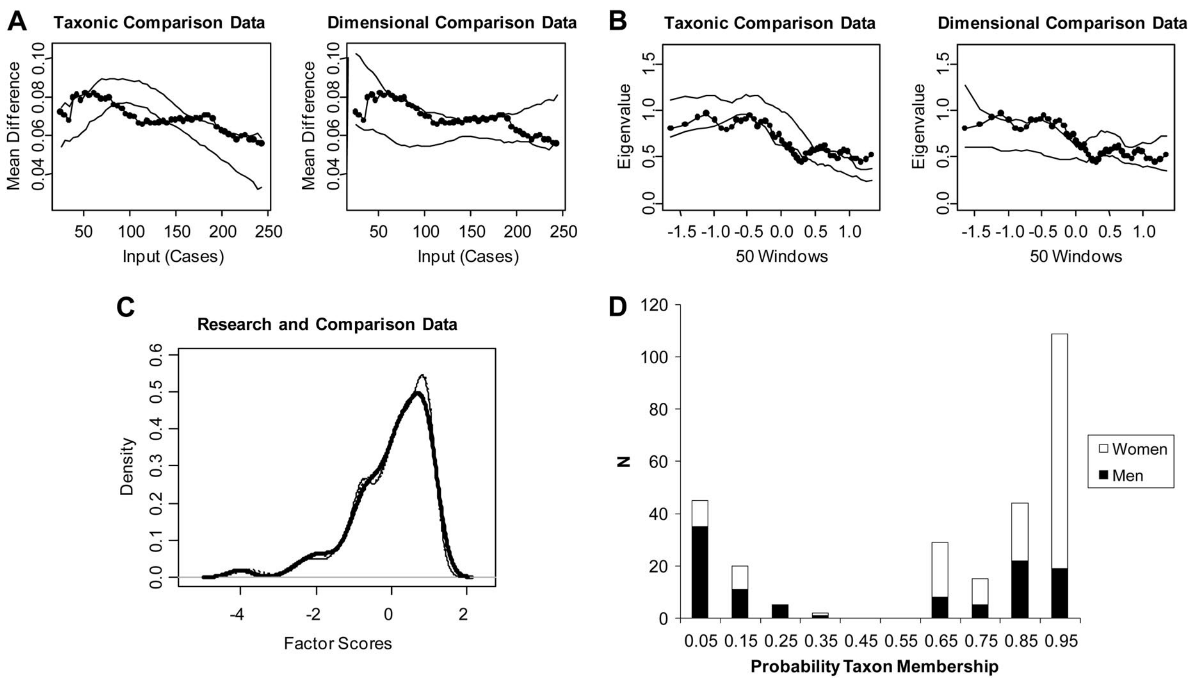

Figure 4. All psychological scales: Mean of Interpersonal Reactivity Index's Fantasy, Empathic Concern, and Personal Distress subscales; grand mean of mean of Social Provisions Scale's Guidance, Attachment, Nurturance, and Reliable Alliance subscales; mean of Modified Personal Assessment of Intimacy in Relationships Scale's Positive Intimacy and Negative Intimacy (reversed) subscales; Relational-Interdependent Self-Construal Scale; Mate Selectivity; Sociosexual Orientation Inventory (reversed). Mean above minus below a cut (A) was mostly flat, though with a peak at the far left of the curve, to the left of where the taxonic simulation predicted. Maximum eigenvalue (B) was nonpeaked but appeared to fit its taxonic and dimensional simulations equally well. Latent mode (C) demonstrated a very small secondary mode at the far left of the distribution (set to search for left mode beyond a factor score of −3). However, taxonic and dimensional simulations were not easily differentiated, indicating that these data may not be appropriate for this method. Bayesian probabilities (D) demonstrated a U-shaped distribution, with most women sorted into the high end but men distributed more evenly.

only those subscales demonstrating significant sex differences were averaged to create a single score for the whole scale (Fantasy, Empathic Concern, and Personal Distress for the IRI; Guidance, Attachment, Nurturance, and Reliable Alliance for the SPS; Positive Intimacy and Negative Intimacy [reversed] for the PAIR-M). Due to high within-sex correlations between the SPS and PAIR-M, the means of both measures were averaged into a single indicator.

Results

As shown in Figure 4, MAMBAC was mostly flat, though it did demonstrate a peak at the far left of the curve. The peak was to the left of where the taxonic simulation predicted, though its fit index was inconclusive (see Table 1) and did not correspond to the generated base rate of 0.58. MAXEIG was nonpeaked, but appeared to fit its taxonic and dimensional simulations equally well, as confirmed by its inconclusive fit index. L-Mode demonstrated a very small secondary mode at the far left of the distribution, and once the parameter of looking beyond a factor score of −3 was added, the generated base rates (0.81 and 0.93) correspond to this mode. However, its taxonic and dimensional simulations were not easily differentiated, indicating that these data may not have been appropriate for this method, as its fit index was also inconclusive. Bayesian probabilities demonstrated a U-shaped distribution, with most women sorted into the high end of the distribution, but with men distributed more equally into both ends. Given the slope of the MAMBAC and MAXEIG curves, the slight possibility of a bimodal L-Mode curve, and the U-shaped Bayesian distribution, an inchworm consistency test was performed. Taxonic curves were produced for only two (IRI and RISC) of the five input indicators, and the averaged curve was lumpy but not peaked, indicating dimensionality (see Supplemental Figure 8). All the base rates were inconsistent with each other, and only that for MAMBAC approached the sample sex ratio. Given the inconsistency of the base rates and methods, the overall analysis of gender in terms of mating and interpersonal variables was most likely dimensional.

General Discussion

We begin with a brief overview of the results. We analyzed 122 unique indicators from 13 studies comprising 13,301 individuals. Our first analysis examined leisure activities that are traditionally sex stereotyped, physical strength, and body measurements, and these gave strong evidence of a sex-specific taxon. This demonstrates that the analysis we used is capable of providing evidence of a sex taxon with both rating data and physical measurements. We then conducted a series of taxometric analyses in four specific areas. With regard to sexuality and mating, there was little evidence of a sex-related taxon for sexual attitudes and behaviors and mate preferences, though there were some signs of small-*n* taxonicity for sociosexuality. Instead, dimensional solutions were clearly favored. The next subset of analyses for interpersonal orientations included measures of empathy and relational interdependence, which were also largely dimensional. Analyses of gender-related dispositions included masculinity–femininity, fear of success, science inclination, and Big Five personality traits. Again, the evidence overwhelmingly supported dimensional solutions. The final set of analyses examined intimacy that again largely supported a dimensional model.

As further evidence that gender is generally dimensional, the fit indices found in Table 1 were submitted to single-sample *t* tests comparing the average CCFI values of MAMBAC, MAXEIG, and L-Mode for all the nonvalidation data sets against a test value of .5. Recall that CCFI values range from 0 (strong dimensional fit) to 1 (strong taxonic fit), with .5 as inconclusive. Indeed, the mean for MAMBAC ($M = .30$, $SD = .08$) was significantly lower than .5, $t(20) = -12.1$, $p < .001$, $d = 2.63$; as was the mean for MAXEIG ($M = .39$, $SD = .10$), $t(20) = -5.03$, $p < .001$, $d = 1.10$; and the mean for L-Mode ($M = .35$, $SD = .13$), $t(20) = -5.35$, $p < .001$, $d = 1.17$. All effect sizes surpassed the value of .8 needed to be considered "large."

In sum, the various attributes examined in this research overwhelmingly are described better by a dimensional than a taxonic model. In other words, our data provide clear empirical evidence to support the belief of researchers who see psychological gender differences in dimensional terms.

What Does a Dimensional Model Mean?

The traditional and easiest way to think of gender differences is in terms of a mean difference: on average, individuals of one sex score higher than individuals of the other sex. What does it mean for there to be clear differences between the mean scores of men and women, but no evidence for separation into taxon and complement groups? With a taxonic construct, knowing a few properties of an individual allows one to make a fairly accurate diagnosis of that individual to a taxon, and to make further accurate inferences of additional properties. Consider one of the variables we examined, PAQ–Masculinity. Had there been a sex taxon, knowing that an individual is high on one marker of masculinity would permit the diagnosis that this individual was male and thus also high on other markers (or "symptoms") of masculinity compared to women. However, this is not the case with dimensional constructs.

In order for PAQ–Masculinity to be taxonic by sex, not only would an average sex difference be required on all items, but a relatively large effect size and particular covariance pattern must emerge. Most men would have to score higher than most women on all items (large and consistent effect size). Additionally, though a male who scores high on one item does not necessarily score high on other items compared to other men (thus no within-group correlation), he does consistently outscore most of the women on all items (thus the correlation in the overall sample). Given the dimensional results we obtained, these data can claim only the first characteristic; on average, men scored differently from women. In other words, there are average sex differences for each "symptom" of gender, but they are not consistent or big enough to accurately diagnose group membership. This can be seen in the Bayesian distributions, which make no separation between groups. Given the multivariate nature of taxometric analyses, the focus is on consistency and covariation over a number of variables measured at the same time in the same sample, rather than measured differences of variables measured in isolation. The demonstrated lack of consistency over variables within the same sample is what differentiates this study from the historical meta-analyses of gender differences.

What does it mean to conceptualize gender-related differences as dimensional as opposed to taxonic? It means that the possession of traits associated with gender is not as simple as "this or that." This idea is consistent with the multidimensional and multifactorial nature of gender identity theory:

> However, even among those with a strong, unambiguous gender identity, men and women do not exhibit all the attributes, interests, attitudes, roles, and behaviors expected of their sex according to their society's descriptive and prescriptive stereotypes but only some of them. They may also display some of the characteristics and behaviors associated with the other sex. In more abstract terms, these various categories of gender phenomena are not only multidimensional but multifactorial, a contention well supported by the empirical data. (Spence, 1993, p. 633)

Although gender differences *on average* are not under dispute, the idea of consistently and inflexibly gender-typed individuals is. That is, there are not two distinct genders, but instead there are linear gradations of variables associated with sex, such as masculinity or intimacy, all of which are continuous (like most social, psychological, and individual difference variables). Thus, it will be important to think of these variables as continuous dimensions that people possess to some extent, and that may be related to sex, among whatever other predictors there may be. Of course, the term *sex differences* is still completely reasonable. In a dimensional model, differences between men and women reflect all the causal variables known to be associated with sex, including both nature and nurture. But at least with regard to the kinds of variables studied in this research, grouping into "male" and "female" categories indicates overlapping continuous distributions rather than natural kinds.

This research also adds further evidence to the current debate about whether it is more profitable to focus this literature on gender differences or gender similarities (Hyde, 2005). "The gender similarities hypothesis states, instead, that males and females are alike on most—but not all—psychological variables" (Hyde, 2005, p. 590). Our research shows, moreover, that even those variables on which males and females are not alike may be evidence of variations along a continuous dimension rather than categorical, and as Hyde terms them, "overinflated claims of gender differences" (Hyde, 2005, p. 590). Clearly, if differences between men and women are conceptualized as variations along a continuum, there is little reason to reify these differences with the sorts of extremities typically mentioned. Instead, these differences would be seen as reflecting all the influences that are brought to bear on an individual's growth, development, and experience, and would be relatively amenable to modification.

Our findings are silent with respect to the question of whether gender differences in the variables we studied are caused primarily by biological factors or experience (Eagly & Wood, 1999). In our view, both biological and social causes are essentially continuous, leading individuals to develop various proclivities and dispositions to one or another extent, and encouraging them to follow certain developmental pathways to a greater extent than others (Archer, 1984; Halpern, 2012; Maccoby, 2002). It is unlikely that any of these pathways are fully discrete. Although taxa such as chemical elements and animal species are often referred to as "natural kinds," this is not invariably the case, as Meehl (1992) pointed out with regard to the political taxon of Trotskyists, who presumably came to such membership via social learning and environment rather than having been born into it. If taxonicity does not necessarily mean membership is rooted in biological causes, it seems reasonable that dimensionality is not necessarily rooted in learning or environment. In short, the dimensionality of gender does not address the social or biological basis of differences between men and women.

If gender is dimensional, why do categorical stereotypes of men and women persist in everyday life? Although our research does not speak to this issue, several explanations seem relevant. One reason is that people tend to think categorically (Medin, 1989), or as Fiske (2010) put it, referring to both laypeople and researchers, "we love dichotomies" (p. 689). People use easily accessible categories to help organize the abundance of information that the social world presents, a mental shortcut that has come to be known as the "cognitive miser" hypothesis (Fiske & Taylor, 1991). Because sex is one of the most readily observed human traits, it forms an easy and common basis for categorizing other persons. As a result, because other qualities tend to be accommodated to accessible categories, and because men and women do differ in myriad ways, category-based generalizations maximize the difference between the sexes while minimizing differences within them (e.g., Fiske & Neuberg, 1990; Taylor et al., 1978). Furthermore, as Krueger, Hasman, Acevedo, and Villano (2003) showed, it may be rational to accentuate intergroup differences whenever these differences are easy to learn, fairly accurate, and helpful for action.

Another reason for the endurance of categorical beliefs about gender is that people tend to essentialize human categories when such categories are discrete, have sharp boundaries, are rooted in cultural traditions, are involuntary and immutable, and are perceived to originate in natural distinctions (i.e., are nonarbitrarily part of what makes "living kinds . . . the things they are"; Prentice & Miller, 2007, p. 205; see also Prentice & Miller, 2006). Essentialized categories are seen as having deep significance and meaningful coherence, factors that lead them to play important roles in social perception and behavior. As Prentice and Miller (2006, 2007) noted, however, psychological essentialism need not reflect actual taxonic differences. Nonetheless, "every time people invoke biology to explain gender differences, they further strengthen the view that women and men are different human kinds" (Prentice & Miller, 2007, p. 205). Of course, the widespread attention to "male" and "female" as discrete categories that is virtually endemic across public media and lay theories of gender only serves to bolster these beliefs. Popular examples include self-help gurus such as Gray (1992) and Fein and Schneider (1995), who attribute relational problems between men and women to their putatively taxonic differences in goals, behaviors, and communication styles.

It might also be asked whether our research implies that male-related qualities and female-related qualities should be considered as a single bipolar dimension. Recall that early conceptions of gender used a unidimensional model of masculinity and femininity (e.g., Hathaway & McKinley, 1943; Strong, 1936; Terman & Miles, 1936). In the 1970s, this model was replaced by a two-dimensional structure that considered masculinity and femininity as independent dimensions each running from low to high levels (Bem, 1974; Constantinople, 1973; Spence, Helmreich, & Stapp, 1975). Our results are not inconsistent with this two-dimensional framework. Taxometric analyses discriminate between taxonic and dimensional constructs, but dimensional constructs may be multidimensional. Hence, knowing that masculinity and femininity, for example, are not taxonic is not informative about whether a unidimensional or multidimensional model is preferable or, for that matter, whether multiple dimensions are correlated or orthogonal. Our results also indicate that the practice of categorizing individuals in terms of masculinity and femininity into a fourfold scheme, as was common in the early literature on androgyny, is clearly inappropriate (Coyne & Whiffen, 1995; Ruscio & Ruscio, 2008). Aside from losing statistical power, this practice misspecifies the most appropriate underling conceptual model and may lead to incorrect results (De Boeck, Wilson, & Acton, 2005; Maxwell & Delaney, 1993; Ruscio et al., 2006).

Considering Historical and Cultural Context

It may be fruitful to consider how our findings are bound to the cultural and historical context within which the data were collected. With a few exceptions, most of these data were collected from young Americans in the last quarter of the 20th century. This is a time and setting in which differences between men and women were shrinking, reflecting societal, economic, and educational circumstances that contributed to the increasing liberalization of gender roles (Brooks & Bolzendahl, 2004). Indeed, it seems likely that were we to examine new data sets collected in 2012, they would, if anything, be even more likely to be dimensional. This point suggests two important implications. First, to the extent that our data sets are outdated, they should have been more likely to reveal a taxonic structure (which they did not), making our support for dimensionality more compelling. Second, if suitable data sets can be found, historical comparisons of underlying structures may prove revealing of the impact of societal trends.

Consider the possibility that archived data sets on gender-related dispositions similar to those we analyzed might be found from the early part of the 20th century. If changes in gender norms and social roles are indeed related to changes in prescriptions for gendered behavior (Diekman & Goodfriend, 2006), perhaps a sex-related taxon would have been identified. If so, cross-temporal comparisons could be taken as evidence of the effects of societal trends on "de-taxonification."[5] In other words, taxometric methods suggest an interesting new approach to examining variations in the structure of gender-related behaviors, attitudes, beliefs, and traits across historical and cultural milieus.

[5] We thank the editor for suggesting this term and analysis.

A similar point can be made about cross-cultural comparisons. Cultures that place greater prescriptions on the behavior of women and men are likely to produce larger sex differences, but more pointedly, might also differentiate men and women more consistently, which then might be evidenced in gender taxa. It would be valuable to replicate our analyses with data collected in such cultures. Most cultures divide labor by sex, but the degree of such division, as well as which tasks are done by which sex, often varies from one culture to the next (Wood & Eagly, 2002). Cross-cultural differences in taxonicity provide a useful perspective on the role of biology and culture in shaping gender differences (and on lay assumptions of essentialism, as discussed above). As Wood and Eagly's (2002, 2012) biosocial model explains, behavioral sex differences reflect the cultural expression of biological traits. In nonindustrialized cultures, sex differences in the division of labor are driven by two primary forces: men's larger size and greater strength and women's childbearing and -rearing responsibilities. Insofar as child care does not interrupt a mother's activities and women have a means of augmenting their physical strength, they can and do participate in behaviors that are otherwise male dominated. Availability of such means varies over ecology and culture. In our own culture, hired child care is available, and where strength is required, machines and tools are often used. With regard to our research, the key point is to be able to empirically examine how the ecology and prescriptive beliefs of a given culture are reflected not only in differences between men and women, but in the way that such differences are structurally organized within persons. Taxometric methods are a useful method for conducting such analyses.

At this point, it may be useful to return to the definition of a taxon: a category in which members form nonarbitrary classes. Cultural patterns of behavior are rarely (if ever) arbitrary. They exist because they are (or at least one time were) adaptive in a particular set of conditions. Conditions in a given setting may change, of course, but beliefs may not, in which case sex-limited restrictions on behavior would still be part of socialization practices, making them a potential cause of categorical stereotyping of the behavior of women and men. The existence of a taxon implies relatively strict and pervasive differentiations between categories: A person displays not just one but all attributes of his or her category. Dimensions, in contrast, imply considerably more flexibility, whereby individuals may display one sex-related behavior but not others.

Limitations and Future Directions

Several limitations apply to this research. First, our findings are necessarily limited to the particular domains we examined; it is possible that other domains might reveal taxa. With the exception of the anthropomorphic and physical strength measurements, our analyses were limited to self-report measures. It is conceivable that the dimensional model fits how men and women see themselves but that analyses of actual behaviors would be more likely to reveal taxa (as our analyses of anthropometric and athletic behavior did). In our defense, we note that we deliberately chose variables that in the social and personality psychology literature have consistently been identified as differing between among men and women—sexuality, masculinity–femininity, and intimacy, for example, are staples in the gender differences literature. Nevertheless, extending taxometric analysis to behavioral data represents a potentially valuable next step in this line of research. If behavioral measures do eventually reveal gender taxons, the fact that in everyday interaction lay perceivers can observe behavior more readily than self-perceptions can help to explain the gap between the gender-similarities stance of the scholarly community and the gender-differences stance of the lay population.

A second limitation is based on score distributions. Several constructs examined here (sexual attitudes and behaviors, mate selectivity Run 2, both runs of the SOI, PAIR-M, intimacy prototype, SPS, and care orientation) demonstrated consistent skew throughout their indicators. Aron and Westbay (1996) transformed the Intimacy Prototype scale to correct for the skew in their analyses. With the exception of the SOI, none of the skew values were extreme. Rather, as is particularly important for the multivariate nature of taxometric analyses, repeated instances of skew in one direction, rather than the average skew value, are most likely responsible for the shape of the curves that emerged. This begs the question of why we did not correct for skew, for which there is a reasonable answer: Skew may indicate taxonicity! Skew is the result of distributional "oddballs" pulling the curve, and it is precisely these oddballs that taxonomists are interested in. People who skew the curve are often deleted because something is assumed to make them qualitatively different from the rest of the sample. To delete them in taxometric analyses or to transform the curve as a correction would be to eliminate the very kinds of people being sought in very small- or large-*n* taxa (Lenzenweger, 2004). For this reason, we thought it appropriate to examine gendered variables for taxonicity, even though the effect size for sex differences was usually smaller than the ideal of 1.25. Possibly, a small single-sex taxon of particularly stereotypical males or females could have emerged, even with small *d* values for the whole sample. It was also possible that a construct might have been taxonic, but without accurately distinguishing men and women. We found possible hints of a small-*n* taxon with the SOI; examination with a much larger sample is recommended to further explore this possibility.

A final limitation concerns the taxometric procedure itself, which relies, in part, on subjective judgments. We attempted to deal with this limitation in three ways. First, we relied on three parallel procedures—MAMBAC, MAXEIG, and L-Mode—for all analyses, which was intended to obviate problems particular to any one of them. Second, in all cases, we compared our plots to simulated taxonic and dimensional plots based on categorical and dimensional simulations with the same distributional characteristics as our actual data. Third, we used fit indices (Ruscio et al., 2006), a mathematically precise index of fit. Most published taxometric research uses no more than one of these criteria. Nonetheless, because of this subjectivity, psychometricians have developed more mathematically precise methods for detecting and analyzing latent structure (such as finite mixture modeling and Dimcat; De Boeck et al., 2005; McLachlan & Peel, 2000). In a recent analysis, McGrath and Walters (2012) compared taxometric procedures to finite mixture modeling using 50,000 Monte Carlo data sets. They concluded that taxometric procedures were "superior to finite mixture modeling for distinguishing between dimensional and categorical models" (p. 284), although not necessarily so for identifying the number of classes in models found to be categorical. Because our purpose was exactly to determine whether a dimensional or categorical model provided a better account of observed sex differences, it is appropriate to conclude that the greater mathematical precision of finite mixture modeling would not be preferable to the admittedly somewhat subjective taxometric methods that we used.

Conclusion

For some time, there has been a striking difference in the way that most scholars and the lay public conceptualize sex differences. Whereas most researchers, with a few noteworthy exceptions, have conceived of psychological sex differences as dimensional constructs, laypersons were more likely to view these differences as fundamentally taxonic. We conducted our analyses with the goal of making explicit the mathematical properties that follow from these distinctive positions and then testing their relevance for a diverse set of measures. In all instances the dimensional approach prevailed. At least with regard to the measures we examined, therefore, it can be concluded that they unambiguously represent exemplars of the same underlying attributes rather than qualitatively distinct categories of human characteristics.

References

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: American Psychiatric Association.

Archer, J. (1984). Gender roles as developmental pathways. *British Journal of Social Psychology, 23,* 245–256. doi:10.1111/j.2044-8309.1984.tb00635.x

Aron, A., & Westbay, L. (1996). Dimensions of the prototype of love. *Journal of Personality and Social Psychology, 70,* 535–551. doi:10.1037/0022-3514.70.3.535

Aubé, J. (2004). Extended Personal Attributes Questionnaire [Data file].

Barclay, C. D., Cutting, J. E., & Kozlowski, L. T. (1978). Temporal and spatial factors in gait perception that influence gender recognition. *Perception & Psychophysics, 23,* 145–152. doi:10.3758/BF03208295

Barnett, M. A., Quackenbush, S. W., & Sinisi, C. S. (1995). The role of critical experiences in moral development: Implications for justice and care orientations. *Basic and Applied Social Psychology, 17,* 137–152. doi:10.1080/01973533.1995.9646136

Baumeister, R. F. (1988). Should we stop studying sex differences altogether? *American Psychologist, 43,* 1092–1095. doi:10.1037/0003-066X.43.12.1092.b

Beauchaine, T. P., & Waters, E. (2003). Pseudotaxonicity in MAMBAC and MAXCOV analyses of rating-scale data: Turning continua into classes by manipulating observer's expectations. *Psychological Methods, 8,* 3–15. doi:10.1037/1082-989X.8.1.3

Bem, S. L. (1974). The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology, 42,* 155–162. doi:10.1037/h0036215

Brewer, M. B. (1988). A dual process model of impression formation. In T. K. Scrull & R. S. Wyer Jr. (Eds.), *Advances in social cognition* (Vol. 1, pp. 1–36). Hillsdale, NJ: Erlbaum.

Brody, L. R. (2000). The socialization of gender differences in emotional expression: Display rules, infant temperament, and differentiation. In A. H. Fischer (Ed.), *Gender and emotion: Social psychological perspectives* (pp. 24–47). New York, NY: Cambridge University Press. doi:10.1017/CBO9780511628191.003

Brooks, C., & Bolzendahl, C. (2004). The transformation of US gender role attitudes: Cohort replacement, social-structural change, and ideological learning. *Social Science Research, 33,* 106–133. doi:10.1016/S0049-089X(03)00041-3

Buss, D. M. (1995). Psychological sex differences: Origins through sexual selection. *American Psychologist, 50,* 164–168. doi:10.1037/0003-066X.50.3.164

Buss, D. M., & Barnes, M. (1986). Preferences in human mate selection. *Journal of Personality and Social Psychology, 50,* 559–570. doi:10.1037/0022-3514.50.3.559

Buss, D. M., Larsen, R. J., Westen, D., & Semmelroth, J. (1992). Sex differences in jealousy: Evolution, physiology, and psychology. *Psychological Science, 3,* 251–255. doi:10.1111/j.1467-9280.1992.tb00038.x

Buss, D. M., & Schmitt, D. P. (1993). Sexual strategies theory: An evolutionary perspective on human mating. *Psychological Review, 100,* 204–232. doi:10.1037/0033-295X.100.2.204

Buss, D. M., & Shackelford, T. K. (1997). From vigilance to violence: Mate retention tactics in married couples. *Journal of Personality and Social Psychology, 72,* 346–361. doi:10.1037/0022-3514.72.2.346

Buss, D. M., Shackelford, T. K., Kirkpatrick, L. A., & Larsen, R. J. (2001). A half century of mate preferences: The cultural evolution of values. *Journal of Marriage and Family, 63,* 491–503. doi:10.1111/j.1741-3737.2001.00491.x

Campbell, L., Simpson, J. A., Kashy, D. A., & Fletcher, G. J. O. (2001). Ideal standards, the self, and flexibility of ideals in close relationships. *Personality and Social Psychology Bulletin, 27,* 447–462. doi:10.1177/0146167201274006

Cantor, N., & Mischel, W. (1979). Prototypes in person perception. *Advances in Experimental Social Psychology, 12,* 3–52. doi:10.1016/S0065-2601(08)60258-0

Centers for Disease Control and Prevention. (1996). National Health and Nutrition Examination Survey [Data files]. Retrieved from http://www.cdc.gov/nchs/nhanes/nh3data.htm

Cole, E. R., Jayarante, T. E., Cecchi, L. A., Feldbaum, M., & Petty, E. M. (2007). Vive la difference? Genetic explanations for perceived gender differences in nurturance. *Sex Roles, 57,* 211–222. doi:10.1007/s11199-007-9248-7

Constantinople, A. (1973). Masculinity–femininity: An exception to a famous dictum. *Psychological Bulletin, 80,* 389–407. doi:10.1037/h0035334

Coyne, J. C., & Whiffen, V. E. (1995). Issues in personality as diathesis for depression: The cause of sociotropy–dependency and autonomy–self-criticism. *Psychological Bulletin, 118,* 358–378. doi:10.1037/0033-2909.118.3.358

Cross, S. E., Bacon, P. L., & Morris, M. L. (2000). The relational-interdependent self-construal and relationships. *Journal of Personality and Social Psychology, 78,* 791–808. doi:10.1037/0022-3514.78.4.791

Cutrona, C. E., & Russell, D. (1987). The provisions of social relationships and adaptation to stress. In W. H. Jones & D. Perlman (Eds.), *Advances in personal relationships* (Vol. 1, pp. 37–68). Greenwich, CT: JAI Press.

Dahlstrom, W. G. (1995). Pigeons, people, and pigeon-holes. *Journal of Personality Assessment, 64,* 2–20. doi:10.1207/s15327752jpa6401_1

Darling-Fisher, C. S., & Leidy, N. K. (1988). Measuring Eriksonian development in the adult: The modified Erikson Psychosocial Stage Inventory. *Psychological Reports, 62,* 747–754. doi:10.2466/pr0.1988.62.3.747

Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catalog of Selected Documents in Psychology, 10,* 85.

Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology, 44,* 113–126. doi:10.1037/0022-3514.44.1.113

De Boeck, P., Wilson, M., & Acton, G. S. (2005). A conceptual and psychometric framework for distinguishing categories and dimensions. *Psychological Review, 112,* 129–158. doi:10.1037/0033-295X.112.1.129

Diekman, A. B., & Goodfriend, W. (2006). Rolling with the changes: A role congruity perspective on gender norms. *Psychology of Women Quarterly, 30,* 369–383. doi:10.1111/j.1471-6402.2006.00312.x

Dindia, K., & Canary, D. J. (Eds.). (2006). *Sex differences and similarities in communication* (2nd ed.). Mahwah, NJ: Erlbaum.

Eagly, A. H. (1987). Reporting sex differences. *American Psychologist, 42,* 756–757. doi:10.1037/0003-066X.42.7.755

Eagly, A. H. (1995). The science and politics of comparing women and men. *American Psychologist, 50,* 145–158. doi:10.1037/0003-066X.50.3.145

Eagly, A. H., & Wood, W. (1999). The origins of sex differences in human behavior: Evolved dispositions versus social roles. *American Psychologist, 54,* 408–423. doi:10.1037/0003-066X.54.6.408

Eliot, L. (2009). *Pink brain, blue brain: How small differences grow into troublesome gaps—and what we can do about it*. New York, NY: Houghton Mifflin Harcourt.

Eliot, L. (2011). The trouble with sex differences. *Neuron, 72,* 895–898. doi:10.1016/j.neuron.2011.12.001

Ellis, L., Hershberger, S., Field, E., Wersinger, S., Pellis, S., Geary, D., . . . Karadi, K. (2008). *Sex differences: Summarizing more than a century of scientific research*. New York, NY: Psychology Press.

Epstein, S. (1986). Does aggregation produce spuriously high estimates of behavior stability? *Journal of Personality and Social Psychology, 50,* 1199–1210. doi:10.1037/0022-3514.50.6.1199

Erikson, E. (1950). *Childhood and society*. New York, NY: Norton.

Fehr, B. (1988). Prototype analysis of the concepts of love and commitment. *Journal of Personality and Social Psychology, 55,* 557–579. doi:10.1037/0022-3514.55.4.557

Fein, E., & Schneider, S. (1995). *The rules: Time-tested secrets for capturing the heart of Mr. Right*. New York, NY: Warner Books.

Fiske, S. T. (2010). Venus and Mars or down to Earth: Stereotypes and realities of gender differences. *Perspectives on Psychological Science, 5,* 688–692. doi:10.1177/1745691610388768

Fiske, S. T., & Neuberg, S. L. (1990). A continuum of impression formation, from category-based to individuating processes: Influences of information and motivation on attention and interpretation. *Advances in Experimental Social Psychology, 23,* 1–74. doi:10.1016/S0065-2601(08)60317-2

Fiske, S. T., & Taylor, S. E. (1991). *Social cognition* (2nd ed.). New York, NY: McGraw-Hill.

Foushee, H. C., Helmreich, R. L., & Spence, J. T. (1979). Implicit theories of masculinity and femininity: Dualistic or bipolar? *Psychology of Women Quarterly, 3,* 259–269. doi:10.1111/j.1471-6402.1979.tb00544.x

Fraley, R. C., & Spieker, S. J. (2003). Are infant attachment patterns continuously or categorically distributed? A taxometric analysis of strange situation behavior. *Developmental Psychology, 39,* 387–404. doi:10.1037/0012-1649.39.3.387

Fraley, R. C., & Waller, N. G. (1998). Adult attachment patterns: A test of the typological model. In J. A. Simpson & W. S. Rholes (Eds.), *Attachment theory and close relationships* (pp. 77–114). New York, NY: Guilford Press.

Gallup, G. (1991). *The Gallup poll: Public opinion 1990*. Wilmington, DE: Scholarly Resources.

Gangestad, S., & Snyder, M. (1985). "To carve nature at its joints": On the existence of discrete classes in personality. *Psychological Review, 92,* 317–349. doi:10.1037/0033-295X.92.3.317

Gangestad, S. W., & Snyder, M. (1991). Taxonomic analysis redux: Some statistical considerations for testing a latent class model. *Journal of Personality and Social Psychology, 61,* 141–146. doi:10.1037/0022-3514.61.1.141

Geary, D. C. (2010). *Male, female: The evolution of human sex differences* (2nd ed.). Washington, DC: American Psychological Association. doi:10.1037/12072-000

Gilligan, C. (1982). *In a different voice*. Cambridge, MA: Harvard University Press.

Gray, J. (1992). *Men are from Mars, women are from Venus*. New York, NY: Harper Collins.

Halpern, D. F. (2012). *Sex differences in cognitive abilities* (4th ed.). New York, NY: Psychology Press.

Haslam, N., & Kim, H. C. (2002). Categories and continua: A review of taxometric research. *Genetic, Social, and General Psychology Monographs, 128,* 271–320.

Haslam, N., Rothschild, L., & Ernst, D. (2000). Essentialist beliefs about social categories. *British Journal of Social Psychology, 39,* 113–127. doi:10.1348/014466600164363

Hathaway, S. R., & McKinley, J. C. (1943). *The Minnesota Multiphasic Personality Inventory*. New York, NY: Psychological Corporation.

Helgeson, V. S. (1993). Implications of agency and communion for patient and spouse adjustment to a first coronary event. *Journal of Personality and Social Psychology, 64,* 807–816. doi:10.1037/0022-3514.64.5.807

Hill, R. (1945). Campus values in mate selection. *Journal of Home Economics, 37,* 554–558.

Hyde, J. S. (2005). The gender similarities hypothesis. *American Psychologist, 60,* 581–592. doi:10.1037/0003-066X.60.6.581

Hyde, J. S. (2007). New directions in the study of gender similarities and differences. *Current Directions in Psychological Science, 16,* 259–263. doi:10.1111/j.1467-8721.2007.00516.x

Hyde, J. S., Fennema, E., & Lamon, S. J. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin, 107,* 139–155. doi:10.1037/0033-2909.107.2.139

Hyde, J. S., & Linn, M. C. (1988). Gender differences in verbal ability: A meta-analysis. *Psychological Bulletin, 104,* 53–69. doi:10.1037/0033-2909.104.1.53

Intons-Peterson, M. J., & Reddel, M. (1984). What do people ask about a neonate? *Developmental Psychology, 20,* 358–359. doi:10.1037/0012-1649.20.3.358

John, O. P., Donahue, E. M., & Kentle, R. L. (1991). *The Big Five Inventory–Versions 4a and 54*. Berkeley: University of California, Berkeley, Institute of Personality and Social Research.

Johnson, K. L., & Tassinary, L. G. (2005). Perceiving sex directly and indirectly: Meaning in motion and morphology. *Psychological Science, 16,* 890–897. doi:10.1111/j.1467-9280.2005.01633.x

Kahn, A. S., & Yoder, J. D. (1989). The psychology of women and conservatism: Rediscovering social change. *Psychology of Women Quarterly, 13,* 417–432. doi:10.1111/j.1471-6402.1989.tb01011.x

Kashy, D. A., & Kenny, D. A. (2000). The analysis of data from dyads and groups. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (pp. 451–477). New York, NY: Cambridge University Press.

Kenrick, D. T., Sadalla, E. K., Groth, G., & Trost, M. R. (1990). Evolution, traits, and the stages of human courtship: Qualifying the parental investment model. *Journal of Personality, 58,* 97–116. doi:10.1111/j.1467-6494.1990.tb00909.x

Kohlberg, L. (1966). A cognitive-developmental analysis of children's sex-role concepts and attitudes. In E. E. Maccoby (Ed.), *The development of sex differences* (pp. 82–173). Stanford, CA: Stanford University Press.

Korfine, L., & Lenzenweger, M. F. (1995). The taxonicity of schizotypy: A replication. *Journal of Abnormal Psychology, 104,* 26–31. doi:10.1037/0021-843X.104.1.26

Krueger, J. I., Hasman, J. F., Acevedo, M., & Villano, P. (2003). Perceptions of trait typicality in gender stereotypes: Examining the role of attribution and categorization processes. *Personality and Social Psychology Bulletin, 29,* 108–116. doi:10.1177/0146167202238376

Laumann, E. O., Gagnon, J. H., Michael, R. T., & Michaels, S. (1992). National Health and Social Life Survey [Computer file]. Chicago, IL: University of Chicago and National Opinion Research Center [Producer]. Ann Arbor, MI: Inter-University Consortium for Political and Social Research [Distributor].

Lenzenweger, M. F. (2004). Consideration of the challenges, complications, and pitfalls of taxometric analysis. *Journal of Abnormal Psychology, 113,* 10–23. doi:10.1037/0021-843X.113.1.10

Lenzenweger, M. F., & Korfine, L. (1992). Confirming the latent structure and base rate of schizotypy: A taxometric analysis. *Journal of Abnormal Psychology, 101,* 567–571. doi:10.1037/0021-843X.101.3.567

Lippa, R. (1991). Some psychometric characteristics of gender diagnosticity measures: Reliability, validity, consistency across domains, and relationship to the Big Five. *Journal of Personality and Social Psychology, 61,* 1000–1011. doi:10.1037/0022-3514.61.6.1000

Lippa, R., & Connelly, S. (1990). Gender diagnosticity: A new Bayesian approach to gender-related individual differences. *Journal of Personality and Social Psychology, 59,* 1051–1065. doi:10.1037/0022-3514.59.5.1051

Maccoby, E. E. (2002). Gender and social exchange: A developmental perspective. *New Directions for Child and Adolescent Development,* 2002(95), 87–105. doi:10.1002/cd.39

Maccoby, E. E., & Jacklin, C. N. (1974). *The psychology of sex differences*. Stanford, CA: Stanford University Press.

Martin, C. L., & Halverson, C. F. (1981). A schematic processing model of sex typing and stereotyping. *Child Development, 52,* 1119–1134. doi:10.2307/1129498

Maxwell, S. E., & Delaney, H. D. (1993). Bivariate median splits and spurious statistical significance. *Psychological Bulletin, 113,* 181–190. doi:10.1037/0033-2909.113.1.181

McGrath, R. E., & Walters, G. D. (2012). Taxometric analysis as a general strategy for distinguishing categorical from dimensional latent structure. *Psychological Methods, 17,* 284–293. doi:10.1037/a0026973

McHugh, M. C., Koeske, R. D., & Frieze, I. H. (1986). Issues to consider in conducting nonsexist psychological research: A guide for researchers. *American Psychologist, 41,* 879–890. doi:10.1037/0003-066X.41.8.879

McLachlan, G., & Peel, D. (2000). *Finite mixture models*. New York, NY: Wiley.

Medin, D. L. (1989). Concepts and conceptual structure. *American Psychologist, 44,* 1469–1481. doi:10.1037/0003-066X.44.12.1469

Meehl, P. E. (1962). Schizotaxia, schizotypy, schizophrenia. *American Psychologist, 17,* 827–838. doi:10.1037/h0041029

Meehl, P. E. (1990). Toward an integrated theory of schizotaxia, schizotypy, and schizophrenia. *Journal of Personality Disorders, 4,* 1–99. doi:10.1521/pedi.1990.4.1.1

Meehl, P. E. (1992). Factors and taxa, traits and types, differences of degree and differences in kind. *Journal of Personality, 60,* 117–174. doi:10.1111/j.1467-6494.1992.tb00269.x

Meehl, P. E. (1995). Bootstraps taxometrics: Solving the classification problem in psychopathology. *American Psychologist, 50,* 266–275. doi:10.1037/0003-066X.50.4.266

Meehl, P. E., & Golden, R. R. (1982). Taxometric methods. In P. Kendall & J. Butcher (Eds.), *Handbook of research methods in clinical psychology* (pp. 127–181). New York, NY: Wiley.

Mischel, W. (1968). *Personality and assessment*. New York, NY: Wiley.

Myers, D. G. (2008). *Social psychology* (9th ed.). Boston, MA: McGraw-Hill.

National Collegiate Athletic Association. (2007). [Championship information]. Retrieved from http://www.ncaasports.com

Oliver, M. B., & Hyde, J. S. (1993). Gender differences in sexuality: A meta-analysis. *Psychological Bulletin, 114,* 29–51. doi:10.1037/0033-2909.114.1.29

Organisation for Economic Co-operation and Development. (2006). *PISA 2006 Student Questionnaire* [Data file and codebook]. Available from http://pisa2006.acer.edu.au/downloads.php

Organisation for Economic Co-operation and Development. (2009). *PISA 2006 Technical Report*. Retrieved from http://browse.oecdbookshop.org/oecd/pdfs/free/9809041e.pdf

Prentice, D. A., & Miller, D. T. (2006). Essentializing differences between women and men. *Psychological Science, 17,* 129–135. doi:10.1111/j.1467-9280.2006.01675.x

Prentice, D. A., & Miller, D. T. (2007). Psychological essentialism of human categories. *Current Directions in Psychological Science, 16,* 202–206. doi:10.1111/j.1467-8721.2007.00504.x

Ramsey-Rennels, J. L., & Langlois, J. H. (2006). Infant's differential processing of female and male faces. *Current Directions in Psychological Science, 15,* 59–62. doi:10.1111/j.0963-7214.2006.00407.x

Reis, H. T. (1980). *A longitudinal study of quantitative-verbal interests*. Unpublished manuscript, University of Rochester, Rochester, NY.

Roisman, G. I., Fraley, C. R., & Belsky, J. (2007). A taxometric study of the Adult Attachment Interview. *Developmental Psychology, 43,* 675–686. doi:10.1037/0012-1649.43.3.675

Ruscio, J. (2008). Taxometric program code [Data file]. Retrieved from http://www.tcnj.edu/~ruscio/taxometrics.html

Ruscio, J., Haslam, N., & Ruscio, A. M. (2006). *Introduction to the taxometric method: A practical guide*. Mahwah, NJ: Erlbaum.

Ruscio, J., & Ruscio, A. M. (2000). Informing the continuity controversy: A taxometric analysis of depression. *Journal of Abnormal Psychology, 109,* 473–487. doi:10.1037/0021-843X.109.3.473

Ruscio, J., & Ruscio, A. M. (2004). A nontechnical introduction to the taxometric method. *Understanding Statistics, 3,* 151–194. doi:10.1207/s15328031us0303_2

Ruscio, J., & Ruscio, A. M. (2008). Categories and dimensions: Advancing psychological science through the study of latent structure. *Current Directions in Psychological Science, 17,* 203–207. doi:10.1111/j.1467-8721.2008.00575.x

Schaefer, M. T., & Olson, D. H. (1981). Assessing intimacy: The PAIR Inventory. *Journal of Marital and Family Therapy, 7,* 47–60. doi:10.1111/j.1752-0606.1981.tb01351.x

Schmitt, D. P. (2002). A meta-analysis of sex differences in romantic attraction: Do rating contexts moderate tactic effectiveness judgments? *British Journal of Social Psychology, 41,* 387–402. doi:10.1348/014466602760344278

Simpson, J. A., & Gangestad, S. W. (1991). Individual differences in sociosexuality: Evidence for convergent and discriminant validity. *Journal of Personality and Social Psychology, 60,* 870–883. doi:10.1037/0022-3514.60.6.870

Smetana, D. K. (2002). The five factor model: Dimensions or categories? *Dissertation Abstracts International: Section B. Sciences and Engineering, 62*(11), 5420.

Spence, J. T. (1984). Gender identity theory and its implications for the concepts of masculinity and femininity. *Nebraska Symposium on Motivation, 32,* 59–95.

Spence, J. T. (1993). Gender-related traits and gender ideology: Evidence for a multifactorial theory. *Journal of Personality and Social Psychology, 64,* 624–635. doi:10.1037/0022-3514.64.4.624

Spence, J. T., & Helmreich, R. L. (1978). *Masculinity and femininity: Their psychological dimensions, correlates and antecedents*. Austin: University of Texas Press.

Spence, J. T., Helmreich, R., & Holahan, C. K. (1979). Negative and positive components of psychological masculinity and femininity and their relationship to self-reports of neurotic and acting out behaviors. *Journal of Personality and Social Psychology, 37,* 1673–1682. doi:10.1037/0022-3514.37.10.1673

Spence, J. T., Helmreich, R., & Stapp, J. (1974). The Personal Attributes Questionnaire: A measure of sex role stereotypes and masculinity–femininity. *JSAS Catalog of Selected Documents in Psychology, 4,* 43.

Spence, J. T., Helmreich, R., & Stapp, J. (1975). Ratings of self and peers on sex role attributes and their relation to self-esteem and conceptions of masculinity and femininity. *Journal of Personality and Social Psychology, 32,* 29–39. doi:10.1037/h0076857

Stangor, C., Lynch, L., Duan, C., & Glass, B. (1992). Categorization of individuals on the basis of multiple social features. *Journal of Personality and Social Psychology, 62,* 207–218. doi:10.1037/0022-3514.62.2.207

Strong, E. K. (1936). Interests of men and women. *Journal of Social Psychology, 7,* 49–67. doi:10.1080/00224545.1936.9921647

Strube, M. J. (1989). Evidence for the type in Type A behavior: A taxometric analysis. *Journal of Personality and Social Psychology, 56,* 972–987. doi:10.1037/0022-3514.56.6.972

Taylor, S. E., Fiske, S. T., Etcoff, N. L., & Ruderman, A. J. (1978). Categorical and contextual bases of person memory and stereotyping. *Journal of Personality and Social Psychology, 36,* 778–793. doi:10.1037/0022-3514.36.7.778

Taylor, S. E., Klein, L. C., Lewis, B. P., Gruenewald, T. L., Gurung, R. A. R., & Updegraff, J. A. (2000). Biobehavioral responses to stress in females: Tend-and-befriend, not fight-or-flight. *Psychological Review, 107,* 411–429. doi:10.1037/0033-295X.107.3.411

Taylor, S. E., Lewis, B. P., Gruenewald, T. L., Gurung, R. A. R., Updegraff, J. A., & Klein, L. C. (2002). Sex differences in biobehavioral responses to threat: Reply to Geary and Flinn (2002). *Psychological Review, 109,* 751–753. doi:10.1037/0033-295X.109.4.751

Terman, L., & Miles, C. C. (1936). *Sex and personality*. New York, NY: McGraw-Hill.

Thériault, J. (1998). Assessing intimacy with the best friend and the sexual partner during adolescence: The PAIR-M Inventory. *Journal of Psychology, 132,* 493–506. doi:10.1080/00223989809599282

Thornhill, R., & Gangestad, S. W. (2008). *The evolutionary biology of human female sexuality*. New York, NY: Oxford University Press.

Trivers, R. L. (1972). Parental investment and sexual selection. In B. Campbell (Ed.), *Sexual selection and the descent of man* (pp. 136–179). Chicago, IL: Aldine.

Waller, N. G., & Meehl, P. E. (1998). *Multivariate taxometric procedures: Distinguishing types from continua*. Thousand Oaks, CA: Sage.

Whisman, M. A., & Pinto, A. (1997). Hopelessness depression in depressed inpatient adolescents. *Cognitive Therapy and Research, 21,* 345–358. doi:10.1023/A:1021830600946

Widiger, T. A., & Trull, T. J. (2007). Plate tectonics in the classification of personality disorder: Shifting to a dimensional model. *American Psychologist, 62,* 71–83. doi:10.1037/0003-066X.62.2.71

Wood, W., & Eagly, A. H. (2002). A cross-cultural analysis of the behavior of women and men: Implications for the origins of sex differences. *Psychological Bulletin, 128,* 699–727. doi:10.1037/0033-2909.128.5.699

Wood, W., & Eagly, A. H. (2012). Biological construction of sex differences and similarities in behavior. In M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 46, pp. 55–124). San Diego, CA: Academic Press.

Zahn-Waxler, C., & Polanichka, N. (2004). All things interpersonal: Socialization and female aggression. In M. Putallaz & K. L. Bierman (Eds.), *Aggression, antisocial behavior, and violence among girls: A developmental perspective* (pp. 48–68). New York, NY: Guilford Press.

Zuckerman, M., & Allison, S. N. (1976). An objective measure of fear of success: Construction and validation. *Journal of Personality Assessment, 40,* 422–430. doi:10.1207/s15327752jpa4004_12a

Received November 21, 2011
Revision received September 12, 2012
Accepted September 14, 2012 ■