American Psychological Association

American Psychologist

© 2018 American Psychological Association
0003-066X/19/$12.00

2019, Vol. 74, No. 2, 171–193
http://dx.doi.org/10.1037/amp0000307

The Future of Sex and Gender in Psychology: Five Challenges to the Gender Binary

Janet Shibley Hyde
University of Wisconsin—Madison

Rebecca S. Bigler
University of Texas at Austin

Daphna Joel
Tel-Aviv University

Charlotte Chucky Tate
San Francisco State University

Sari M. van Anders
University of Michigan

The view that humans comprise only two types of beings, women and men, a framework that is sometimes referred to as the "gender binary," played a profound role in shaping the history of psychological science. In recent years, serious challenges to the gender binary have arisen from both academic research and social activism. This review describes 5 sets of empirical findings, spanning multiple disciplines, that fundamentally undermine the gender binary. These sources of evidence include neuroscience findings that refute sexual dimorphism of the human brain; behavioral neuroendocrinology findings that challenge the notion of genetically fixed, nonoverlapping, sexually dimorphic hormonal systems; psychological findings that highlight the similarities between men and women; psychological research on transgender and nonbinary individuals' identities and experiences; and developmental research suggesting that the tendency to view gender/sex as a meaningful, binary category is culturally determined and malleable. Costs associated with reliance on the gender binary and recommendations for future research, as well as clinical practice, are outlined.

Keywords: gender, sex differences, transgender, neuroscience, social neuroendocrinology

From its beginnings in the 1800s, psychological research and practice firmly espoused the assumption that there are two and only two categories of people: women and men. By the early part of the 20th century, Woolley (1910) had written a review of psychological research on the differences between women and men. In the 1930s, psychologists developed the concept of psychological masculinity-femininity (Terman & Miles, 1936) and argued that masculinity was necessary for good adjustment for men, as was femininity for women (Pleck, 1981). These approaches are based on what is referred to as the *gender binary*. In addition to the core belief that there are two discrete categories into which all individuals can be sorted, the gender binary system also typically assumes that one's category membership is biologically determined, apparent at birth, stable over time, salient and meaningful to the self, and a powerful predictor of a host of psychological variables.

Over the past two decades, however, a confluence of forces has challenged psychology's assumption of the gender binary. These forces range from the transgender activist movement (Martinez-San Miguel & Tobias, 2016; Stryker, 2008) and the intersex activist movement (Dreger & Herndon, 2009; Reis, 2007) to research in neuroscience and psychological science. This article synthesizes research that

This article was published Online First July 19, 2018.

Janet Shibley Hyde, Department of Psychology and Department of Gender & Women's Studies, University of Wisconsin—Madison; Rebecca S. Bigler, Department of Psychology, University of Texas at Austin; Daphna Joel, School of Psychological Sciences and Sagol School of Neuroscience, Tel-Aviv University; Charlotte Chucky Tate, Department of Psychology, San Francisco State University; Sari M. van Anders, Departments of Psychology, Women's Studies, and Neuroscience, University of Michigan; as of July 2018, Departments of Psychology, Gender Studies, and Neuroscience, Queen's University.

Parts of this research were supported by the Israel Science Foundation (Grant 217/16) to Daphna Joel, and by the National Science Foundation (REC 0635444) to Janet Hyde. The authors thank Kelley Kidwell for creating Figure 1. Coauthors after Janet Shibley Hyde are listed alphabetically.

Correspondence concerning this article should be addressed to Janet Shibley Hyde, Department of Psychology, University of Wisconsin—Madison, 1202 West Johnson Street, Madison, WI 53706. E-mail: jshyde@wisc.edu

Janet Shibley Hyde

challenges the gender binary from multiple perspectives, focusing especially on neuroscience, behavioral neuroendocrinology, research on gender similarities and differences, research on the experiences of transgender individuals, and the developmental psychology underlying the psychological process of categorizing by gender. Costs associated with the gender binary framework are considered, as are implications for future research and clinical practice of moving beyond a gender binary framework.

Terminology in this area is complex and controversial.[1] Some authors have argued that *sex* should be used for biologically based differences between males and females, whereas *gender* should be used for differences between women and men that are produced socioculturally (e.g., Muehlenhard & Peterson, 2011; Unger, 1979; West & Zimmerman, 1987). Others have argued that biological and sociocultural factors are typically intertwined, and thus the distinction between the terms *sex* and *gender* should be abandoned (Yoder, 2003). In this article, the term *gender/sex* is frequently used, to recognize that the biological and the sociocultural are typically inseparable (van Anders, 2015; van Anders & Dunn, 2009). The term *sex* is used here to refer to biological systems involving the X and Y chromosomes, pre- and postnatal sexual differentiation, and hormones that influence sexual differentiation of the external genitals, which, in turn, serve as the basis for sex assignment at birth. Individuals with statistically atypical genitals or internal reproductive structures can be termed *intersex* or *sex diverse*, and constitute roughly 1% to 2% of the population (at least among White people; Blackless et al., 2000; Lee et al., 2016). The term *gender* is used here to refer to sociocultural systems that include norms and expectations for males and females, which vary as a function of intersections with other factors (Cole, 2009; Else-Quest & Hyde, 2016), as well as psychological processes such as identity, femininity, masculinity, and gender-conformity and nonconformity. The term *transgender* is used in this article as an umbrella term for individuals who self-label differently than their birth-assigned category (for a list of terms concerning gender that are transgender inclusive, see American Psychological Association [APA], 2015), and *cisgender* is used to refer to individuals whose self-labeling is the same as their birth-assigned category. Finally, the term *nonbinary* is used here for individuals who self-identify in ways outside the two categories of female and male (e.g., agender, genderfluid, bigender; Bornstein, 1994; Brooks, 2017; Tate, Youssef, & Bettergarcia, 2014).

The Challenge From Neuroscience

The gender/sex binary is salient in neuroscience, and especially in its popularized versions. For example, in her book *The Female Brain*, Brizendine (2006) claimed that

> scientists have documented an astonishing array of structural, chemical, genetic, hormonal, and functional brain differences between women and men. We've learned that men and women have different brain sensitivities to stress and conflict. . . . Women may remember the smallest details of their first date, and their biggest fights, while their husbands barely remember that these things happened. Brain structure and chemistry have everything to do with why this is so. (p. 4)

There are indeed average differences between women and men in brain structure and function (Lenroot & Giedd, 2010; Ruigrok et al., 2014). These differences are, however, often misinterpreted as innate or preprogrammed, context independent, and stable over time (e.g., Joel & McCarthy, 2017). Furthermore, it is often implicitly assumed that these differences add up to create two types of brains, one typical of females and the other typical of males. For this assumption to be accurate, differences between females and males in the structure of specific brain regions should be both highly dimorphic in the population and internally consistent in the individual. It turns out that they are neither.

Sex and Biological Dimorphism

Two fundamental assumptions underlie current thinking about sex as a biological system and about its relations with other systems: (a) that sex is a *dimorphic* system (i.e., a

[1] The five authors have produced, independently, five lines of systematic empirical work that challenge the gender/sex binary. As is true of the field of gender/sex research, we—as a group—do not agree completely on the terminology used to refer to men and women (i.e., sex versus gender) or some of the tenets of the development of men and women. Interested readers should consult the authors' respective works for more detailed discussions.

Rebecca S. Bigler

system that can take one of only two forms), and (b) that the effects of sex on other systems (e.g., the brain, gender identity) are characterized by a dimorphic outcome (e.g., male vs. female brain, male vs. female gender identity). Dimorphism can be observed at the level of a single element of a system—for example, the gonads are generally dimorphic: in most cases, they differentiate into just two forms, ovaries or testes. Dimorphism can also characterize systems that comprise several dimorphic elements—for example, the internal genitals are generally a dimorphic system because, in most cases, humans are born with either a uterus, cervix, fallopian tubes, and vagina, or with seminal vesicles, vas deferens, epididymis, and prostate. For a system to show sexual dimorphism, each of its elements should be dimorphic, that is, should exist in only two different forms or categories, one typical of females and the other typical of males, and all the elements within an individual should be *internally consistent*, that is, either all in the form typical of females or all in the form typical of males (Joel, 2011, 2012, 2014).

The Human Brain and the Gender Binary

Although many studies have reported differences between women and men in brain structure (e.g., Lenroot & Giedd, 2010; Ruigrok et al., 2014), these differences are not sexually dimorphic; rather, there is considerable overlap between the distributions of women and men. This is true even for regions showing the largest sex differences known to date. For example, the intermediate nucleus of the human hypothalamus is about twice as large, on average, in men compared with women, yet in approximately 30% of men, the size of this nucleus falls in the female-typical range (Garcia-Falgueras, Ligtenberg, Kruijver, & Swaab, 2011).

In terms of internal consistency, data on sex differences in the rodent brain suggest that internal consistency is rare. This is because sex differences in specific brain features can be different, and even opposite, under different environmental conditions, and because these sex-by-environment interactions vary across brain features. Consider, for example, sex differences in the density of cannabinoid receptors in the rat hippocampus. Under typical laboratory conditions, the density of these receptors is higher in male than female rats. However, following 3 weeks of mild stress, the sex difference in the dorsal part of the hippocampus is reversed; the density of receptors in females is the same as that seen in nonstressed males, and the density of receptors in stressed males is the same as that seen in nonstressed females (Reich, Taylor, & McCarthy, 2009). In other words, the brain sex difference is context dependent. The experience of stress did not, however, reverse sex differences in the density of cannabinoid receptors in the entire brain. Different patterns were found, for example, in the ventral hippocampus. This example demonstrates that complex sex-by-environment interactions produce a brain structure that is multimorphic rather than dimorphic.

Similar sex-by-environment interactions have been demonstrated in rodent research for many brain regions (e.g., cortex, amygdala, cerebellum), for many brain features (e.g., neuronal density, dendritic morphology, neurotransmitter systems), and following many different types of environmental conditions (e.g., rearing conditions, exposure to drugs) at different points throughout development (from in utero to adulthood, reviewed by Joel, 2011, 2012). Thus, it is unlikely that brains are internally consistent and dimorphic; rather, each brain comprises a unique mosaic of features, some more common in females and others more common in males (Joel, 2011, 2012).

A recent study that assessed, for the first time, sexual dimorphism in the human brain in terms of the degree of internal consistency found it to be rare. Specifically, Joel and colleagues (2015) analyzed different structural measures, such as volume, cortical thickness, and connectivity, using MRI of over 1,400 human brains from four data sets. In each dataset, they assessed internal consistency in 7 to 12 features chosen because they showed the largest differences (i.e., least overlap) between women and men. For each such feature, the researchers defined the forms that were more common in women compared with men (female-end form), the forms that were more common in men compared with women (male-end form), and the forms that were similarly common in both women and men (intermediate form). For example, for a brain region that was larger, on average, in women, the female-end form was defined as volumes in the top 33% for women, and the male-end form was defined as volumes in the lowest 33% for men. Next, the researchers

Daphna Joel

tested, for each brain, whether it was internally consistent—that is, whether all elements had the female-end form, or all had the male-end form, or all had the intermediate form—and contrasted internal consistency with *mosaicism*, that is, having at least one element with the female-end form and at least one element with the male-end form. Regardless of the sample, age, type of imaging, and method of analysis, mosaicism was much more common than internal consistency; mosaicism was seen in 23% to 53% of the brains and internal consistency in 0.7% to 10.4% of brains, depending on the sample and specific brain measure. (The remaining brains had either combined female-end and intermediate features or male-end and intermediate features.) Accordingly, sex differences in the human brain do not add up to create two types of brain, a male brain and a female brain. Instead, most brains are gender/sex mosaics.

Summary

The division of humans into two categories, females and males, on the basis of the form of their genitalia is often accompanied by the assumption that males and females belong to two distinct categories in other domains; however, current scientific evidence refutes this assumption for the brain. The distributions for men and women on different brain features are overlapping, and internal consistency across features within individuals is rare. Thus, human brains are not internally consistent for male-typical and female-typical features. Instead, most human brains are a mosaic of these features.

The Challenge From Behavioral Neuroendocrinology

Belief in a gender binary is found not only in neuroscience but also in behavioral neuroendocrinology. This belief involves two assumptions: (1) that gonadal hormones are dimorphic (i.e., that there are "female" hormones, such as estrogen and progesterone, and "male" hormones, such as testosterone), and (2) that levels of these hormones are genetically determined and fixed. Current research in behavioral neuroendocrinology and, in particular, social neuroendocrinology, challenges both of these assumptions.

Androgens, Estrogens, and the Gender Binary

Many people mistakenly assume that there are male hormones and female hormones, but this idea is challenged by the presence of estrogens (e.g., estradiol) and androgens (e.g., testosterone) in both women and men as well as in gender-diverse (e.g., nonbinary) people, because these hormones, as well as progesterone, are produced by both ovaries and testes as well as the adrenal glands and through peripheral conversion in fatty tissue; these sources are present in all bodies (for a review, see Gillies & McArthur, 2010). Another common misunderstanding is that these hormones circulate at sexually dimorphic or nonoverlapping levels. In actuality, average levels of estradiol and progesterone do not differ between women and men (Liening, Stanton, Saini, & Schultheiss, 2010; van Anders, 2010). Moreover, the changes in steroid levels that accompany reproductive phases (e.g., pregnancy, ovulation) highlight the breakdown of the gender binary; for example, nonpregnant women have estradiol and progesterone levels more similar to men than to pregnant women (Tulchinsky, Hobel, Yeager, & Marshall, 1972). Classification on the basis of steroid levels (estradiol and progesterone) would make for a very different kind of binary, one between pregnant women and everyone else (nonpregnant women, men, and gender-diverse individuals).

Differences in the levels of these hormones vary across the life span, with no differences during the prenatal period except for one brief span corresponding to genital sexual differentiation, and no differences from birth to adolescence except one period during the first year of life. In short, fetuses and prepubertal children cannot be categorized into a gender binary on the basis of androgens and estrogens. During adolescence, testosterone levels increase in both boys and girls, but at a much higher average rate for boys (Gillies & McArthur, 2010). However, the size of this difference has been mischaracterized; although testosterone levels are higher in men than women, on average, the difference is much smaller than widely believed and the distributions show considerable overlap (Granger, Shirtcliff, Booth, Kivlighan, & Schwartz, 2004; Liening et al., 2010; Overpeck, Colson, Hohmann, Applestine, & Reilly, 1978;

Charlotte Chucky Tate

Sapienza, Zingales, & Maestripieri, 2009; van Anders, 2010). In fact, the unquestioned belief in the gender binary has hampered the study of these hormones because many researchers have studied "male" hormones (e.g., testosterone) only in men and "female" hormones (e.g., estradiol and progesterone) only in women (van Anders, 2013). It is only recently that researchers have turned to closely examining actual empirical data on gender/sex variation and overlap in adult testosterone levels.

Androgens and Estrogens: Genetically Determined and Fixed?

In addition to gonadal hormones being viewed as dimorphic, this presumed dimorphism is generally assumed to be genetically determined and fixed. Innate fixedness is an important feature of people's beliefs in the gender binary, and gonadal hormones are presumed to underlie the gender binary and contribute to its stability. Research, however, demonstrates that malleability is an important characteristic of biological phenomena, as exemplified by research on neuroplasticity and epigenetics (Pittenger & Duman, 2008; Weaver et al., 2004). However, when it comes to gender/sex and hormones, plasticity and malleability are still largely ignored or presented in ways that fix biology in new critical periods or early life programming (Pitts-Taylor, 2010; Richardson et al., 2014).

The assumption that levels of gonadal hormones are innate and fixed is challenged by evidence that their levels vary widely *within* individuals and are socially modulated (e.g., Nyby, 2008; van Anders, Goldey, & Bell, 2014; van Anders & Watson, 2006). That is, hormone levels are not a fixed characteristic of individuals—even though they may show some trait-like patterns—but instead are a set of changing and interdependent parameters (Wagner, 2006). Evidence for this assertion comes from a large body of research, including research described in the next section.

Hormones such as estradiol and progesterone are already understood to show variability within the context of reproductive phases in women (e.g., menstrual cycle, pregnancy, menopause). Research on the social modulation of these hormones demonstrates that they vary in response to social context and behaviors as well, often in ways that counter gendered stereotypes of these hormones. For example, engaging in dominance contests can increase estradiol and progesterone (Oxford, Tiedtke, Ossmann, Özbe, & Schultheiss, 2017; Stanton & Schultheiss, 2007). Social closeness also increases progesterone (S. L. Brown et al., 2009), as might social rejection (Duffy, Harris, Chartrand, & Stanton, 2017).

Because testosterone is often seen as a biological *account* of differences between men and women, understanding whether it is innate and fixed, or socially influenced and malleable, is important. Testosterone is understood in many ways to be the biological essence of sex, accounting for female–male differences, in general, and maleness, in particular (Fine, 2017; van Anders, 2013). As such, when testosterone is studied, it is generally studied as the sole cause of gender/sex differences, and social factors are excluded. In addition, it has traditionally been studied only as a cause and not as an outcome of behavior. These biologically deterministic approaches run counter to scientific evidence concerning testosterone, as described next.

Is testosterone genetically determined? Testosterone has a relatively high heritability, implying that genetics account for a relatively high proportion of variability in testosterone levels (Harris, Vernon, & Boomsma, 1998; Kuijper et al., 2007). But heritability estimates of testosterone also demonstrate a relatively large role for nongenetic factors, including the environment (e.g., time of day) and social factors (Harris et al., 1998; Kuijper et al., 2007; van Anders, 2013). Moreover, the genetic contribution is likely overestimated because it reflects samples taken from people in the same social and behavioral contexts; that is, if people were sampled across a wider range of environmental circumstances, environment would likely account for more variance, and heritability estimates would be lower (van Anders et al., 2014). Thus, the idea that testosterone research supports an innate gender binary because of its heritability is challenged by evidence showing that testosterone levels are influenced to a considerable degree by nongenetic factors, and by more recent research demonstrating that they dynamically respond to social context.

Is testosterone fixed? Research on hormones challenges the notion of the gender binary as fixed by showing that social factors influence testosterone. Social neuroendo-

Sari M. van Anders

crinology attends to this "reverse relationship" (effects of behaviors on hormones) and also focuses on recursive links between hormones and social behaviors, all while attending to social context (van Anders, Goldey, & Kuo, 2011; van Anders & Watson, 2006). A growing number of researchers use social neuroendocrine framings that take social context into account (e.g., Gettler, McDade, Feranil, & Kuzawa, 2011; Hamilton, Carré, Mehta, Olmstead, & Whitaker, 2015). For example, sexual thoughts increase testosterone levels in women (Goldey & van Anders, 2011); testosterone responses to sexual thoughts are correlated with the type of fantasy content in men (Goldey, Avery, & van Anders, 2014); parenting behaviors decrease testosterone but only when they are nurturant (van Anders, Tolman, & Volling, 2012); and relationship transitions are recursively linked with testosterone (Dibble, Goldey, & van Anders, 2017).

A large and growing body of research documents the ways that social behavioral contexts modulate testosterone levels. For example, recent research with large samples over multiple waves shows strong evidence for sexual and relational modulation of testosterone and very little for androgenic modulation of these phenomena (Das & Sawin, 2016). Other research shows evidence for both directions, as when the presence of many family members in one's social network is associated with lower testosterone at later points, and testosterone negatively influences perceived later social support (Das, 2017). Some research shows iterative associations. For example, many fathers to-be show decreases in testosterone over the duration of their partners' pregnancies, and these decreases in testosterone can predict parenting behaviors (Gettler et al., 2011; Saxbe et al., 2017). But nurturant parenting behaviors, in particular, seem to decrease testosterone in men (van Anders et al., 2012). This research not only counters the notion of testosterone as fixed but also expands notions of which behavioral contexts are meaningful for investigations with androgens and challenges their assumed link to masculinity.

Indeed, testosterone responses can be parsed more meaningfully into decreases related to nurturance (involving warm, close, supportive, and/or loving contact) and increases related to competition (involving acquisition of resources, broadly defined), providing a theoretical basis that is grounded in both sociocultural and evolutionary understandings as well as other endocrine systems (via the steroid/peptide theory of social bonds; van Anders et al., 2011). This matters because testosterone—like other hormones—has evolved to respond to some, but not other, behavioral contexts. The empirical evidence that breaks down fixed views of testosterone shows that it is not only socially modulated but also socially situated, with social modulation of testosterone itself grounded in sociocultural processes and experiences.

Indeed, some social neuroendocrine research has begun to examine how gendered experiences themselves might modulate hormones, with evidence suggesting that testosterone can respond to gender socialization and gender norms (van Anders, Steiger, & Goldey, 2015). Thus, hormone research is showing the ways that gendered expectations and lived experiences can actually shape the very hormones thought to underlie the essence of femaleness and maleness, again challenging basic tenets of the gender binary.

Summary

Social neuroendocrine research challenges the gender binary in multiple ways. Androgens and estrogens are not two distinct sets of sex hormones—one set for women and one set for men—but rather hormones that are found in all humans. That is, human bodies produce hormones like estradiol, testosterone, and progesterone regardless of gender/sex, and levels of estradiol and progesterone are similar in men and women. Moreover, levels of these hormones are not fixed, but are dynamic and can be influenced by gendered social experiences. In humans, for example, testosterone baseline levels across gender/sex change as a function of multiple environmental factors (e.g., time of day, season) and social factors (e.g., relationship status, mood; van Anders, et al., 2014). Thus, the idea that testosterone—or any hormone—is the biological basis of the gender binary is belied by the scientific research, and challenges to the biology of the gender binary are challenges to the gender binary itself.

The Challenge From Psychological Research: Gender Differences and Similarities

As noted earlier, for more than a century, psychologists have devoted themselves to research on psychological gender/sex differences. That research rests on an assumption that there are

just two categories of people: females and males. Often, results are translated into statements about psychological dimorphisms, as if there was no overlap in the female and male distributions for the behavior being measured. For example, findings of (often small) average gender differences are translated into statements such as "Girls and boys play differently. They learn differently. They fight differently . . . They hear differently" (Sax, 2005, p. 28). In the sections that follow, results based on two different methods for synthesizing data on psychological gender differences are presented. One is a "mosaic" analysis like the one described in the section on neuroscience; the other uses meta-analysis.

The Psychological Gender Mosaic

Do the few behavioral and psychological variables that show large gender/sex differences add up to create two genders, each with its own dimorphic set of psychological and behavioral characteristics? This question can be answered by applying an analysis of internal consistency to these variables, just as was done to brain regions showing large gender/sex differences in the neuroscience section of this article. Joel and colleagues (2015) did exactly that and found that internal consistency in personality traits, attitudes, interests, and behaviors is extremely rare. In contrast, most humans possess both feminine (i.e., more common in women than men) and masculine (i.e., more common in men than women) psychological characteristics (see also Koestner & Aube, 1995; Spence, 1993). Of special interest was a dataset of 10 highly gender-stereotyped behaviors in U.S. college students (boxing, construction, playing golf, playing video games, scrapbooking, taking a bath, talking on the phone, watching porn, watching talk shows, and using cosmetics). Even for these highly gendered behaviors, which showed very large gender differences ($1.0 < d < 2.03$), less than 1% of the students exhibited only feminine or only masculine behaviors, whereas over 55% showed some combination of both feminine and masculine behaviors (*feminine* and *masculine* were defined here as the scores characteristic of the most extreme 33% of women and men, respectively). Thus, although stereotypes of women and men clearly exist, individuals who consistently match these stereotypes are rare.

Meta-Analyses of Psychological Gender Differences

The statistical technique of meta-analysis allows a much more powerful and nuanced analysis of research on gender differences than individual studies (for an overview of methods, see Borenstein, Hedges, Higgins, & Rothstein, 2009). Meta-analysis can synthesize dozens or hundreds of studies and tell us how large a gender difference is, how much overlap there is between male and female distributions, and whether the magnitude of the gender difference varies according to factors such as age, ethnicity, and nationality. Because of the dominance of the gender binary in psychological research, meta-analyses to date have synthesized research examining differences between just two gender categories: women and men.

The magnitude of a gender difference is typically assessed using the statistic $d = (M_M - M_F)/s_w$, where M_M is the mean score for males, M_F is the mean score for females, and s_w is the pooled within-groups standard deviation. According to this formula, a positive value of d means that males scored higher on the measure, and a negative value indicates that females scored higher. With meta-analysis, a d value, or effect size, is computed for each study and then a weighted average d is computed across all studies. According to conventions established by Cohen (1988), $d = 0.20$ is a small difference, $d = 0.50$ is a moderate difference, and $d = 0.80$ is a large difference. Hyde (2005) added the interpretation that a d value ≤ 0.10 is trivial.

Figure 1 shows four possible alternatives for the distribution of males' and females' scores on a trait, which could be anything from hippocampus size to mathematics performance. Panel A shows a very large gender/sex difference ($d = 5.0$)—a dimorphism, in the language of biology, because this difference is so large that there is virtually no overlap between the two distributions. Panels B, C, and D show the overlap of male and female distributions when $d = 0.20$, 0.50, and 0.80, respectively. Even for $d = 0.80$, there is still substantial overlap in distributions, that is, the trait is not dimorphic.

Today, numerous meta-analyses of research on psychological gender differences are available, and many of them produce surprising results. The three meta-analyses reviewed here are illustrative.

Mathematics performance. Mathematics is stereotyped as an area of male superiority, and the implicit association test shows that people associate males and math more closely than they do females and math (Nosek, Banaji, & Greenwald, 2002). Meta-analyses, however, challenge this stereotype.

One meta-analysis synthesized data from state assessments of U.S. children's math performance from Grades 2 through 11, based on the testing of more than 7 million children (Hyde, Lindberg, Linn, Ellis, & Williams, 2008). Across grades, d ranged between -0.02 and 0.06. That is, all the differences were trivial or nonexistent, and the authors concluded that girls had reached parity with boys in mathematics. A second meta-analysis accumulated data from 242 studies, representing the testing of more than 1.2 million people (Lindberg, Hyde, Petersen, & Linn, 2010). Overall, $d = 0.05$, again indicating no gender difference.

Depression. Depression is stereotyped as a female disorder, and the expression of depressive symptoms, such as sadness and tearfulness, violates male gender role stereo-

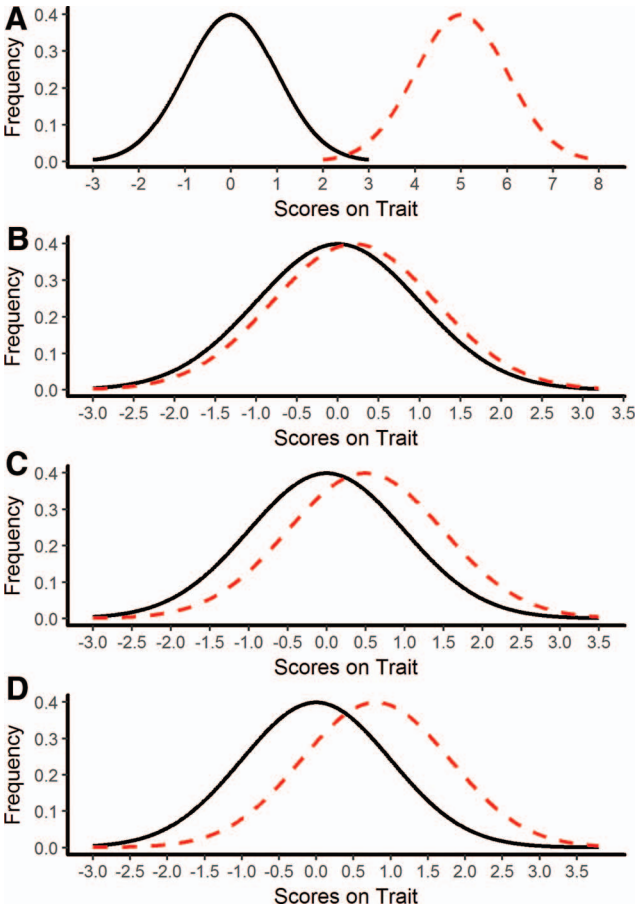

Figure 1. Two distributions of scores with different effect sizes for gender/sex differences. In these hypothetical cases, the dashed (red) distribution can be thought of as the distribution for females, and the solid (black) as the distribution for males. The top panel (A) shows distributions reflecting gender/sex dimorphism, with $d = 5.0$. The second panel (B) shows $d = 0.20$, the next (C) shows $d = 0.50$, and the bottom panel (D) shows $d = 0.80$. See the online article for the color version of this figure.

types (Tran, Cole, & Weiss, 2012). Scientists generally find that the gender ratio (women: men) for depression is 2:1 (e.g., Andrade et al., 2003; Weissman & Klerman, 1977). The disorder is also believed to show a particular developmental pattern, with no gender difference in childhood, and a difference emerging between 13 and 15 years of age (Hankin et al., 1998).

A meta-analysis identified data based on nationally representative samples, yielding diagnostic data from more than 1.7 million persons in 90 nations (Salk, Hyde, & Abramson, 2017). Overall, the odds ratio (*OR*; female:male) was 1.95, which is close to the generally believed value of 2.0. That single value, however, masks considerable variability in *OR*s across factors such as age and nationality. The results indicated that even by age 12, $OR = 2.37$; that is, the gender difference emerges earlier than previously thought. The *OR* peaked at 3.01 in the 13- to 15-years age group, and then declined into the 20s and remained relatively stable around 1.8 after that. Even with such a gender disparity, though, an *OR* of approximately 2.0 means that roughly one third of all depressed people are males, a point that is taken up again in a later section.

Sexuality. Evolutionary psychologists have argued that there are large gender differences in sexual attitudes and behaviors, resulting from sexual selection, and that the differences are universal across cultures (e.g., Buss, 1989). Meta-analysis tells a different tale.

One meta-analysis synthesized research on 30 different sexual attitudes and behaviors, based on 834 distinct samples and seven large national data sets (Petersen & Hyde, 2010). Most gender differences were small. Examples for behavior include same-gender sexual experience ($d = -0.05$) and age at first sex ($d = 0.20$). Examples for attitudes include general sexual permissiveness ($d = 0.21$), attitudes about masturbation ($d = 0.02$), and attitudes about homosexuality (the term used in most questionnaires; $d = -0.13$). Some exceptions were found; gender differences were larger for use of pornography ($d = 0.63$), masturbation ($d = 0.53$), and attitudes about casual sex ($d = 0.45$).

Another finding from this meta-analysis, which has also been documented in other meta-analyses (e.g., Else-Quest, Hyde, & Linn, 2010), is that the magnitude of gender differences can vary depending on the situation and, in particular, on the sociocultural context. Cross-national analyses probed whether the magnitude of the gender difference correlated with empowerment of women within each nation (Else-Quest & Grabe, 2012). The prediction from sociocultural theory is that gender gaps in sexuality will be smaller in nations in which women are more empowered (Eagly & Wood, 1999). That is exactly what the results indicated for some measures. For example, the gender gap in masturbation was negatively correlated, across nations, with gender empowerment. The same was true for casual sex experience. Thus, gender differences in various aspects of sexuality are not universal across cultures. Instead, they vary across cultures in predictable and theoretically meaningful ways.

Gender Similarities and the Gender Binary

By 2005, enough meta-analyses of gender differences had appeared to make an overall pattern clear: Most psychological gender differences are small or trivial. Hyde (2005) therefore proposed the *gender similarities hypothesis*, which states that men and women are very similar on most, but not all, psychological variables. The evidence came from a review of 46 meta-analyses of gender differences, extracting effect sizes from each. These meta-analyses all focused on behaviors believed to show gender differences, and covered a wide array of psychological domains, including cognitive abilities, communication, social and person-

ality variables (e.g., aggression, helping behavior, leadership), and psychological well-being. Based on 124 effect sizes, 30% were trivial ($d \leq 0.10$) and an additional 48% were small, falling in an interval around $d = 0.20$. That is, a total of 78%, or fully three quarters, of gender differences were small or trivial. There were a few exceptions, especially in the areas of aggression and sexuality. The science has continued to support the gender similarities hypothesis, with an independent replication coming 10 years later (Zell, Krizan, & Teeter, 2015).

Even for research conducted under the assumption of the gender binary, the results indicate that males and females are not dimorphic psychologically. In fact, as captured in the gender similarities hypothesis, psychological research on gender differences provides massive evidence of similarities, which challenges the gender binary.

Summary

Analyses using one set of statistical methods indicate that, even for variables showing relatively large differences between women and men, humans are characterized psychologically by a gender mosaic. Meta-analytic findings point to the gender similarities hypothesis, which asserts that females and males are quite similar on most, but not all, psychological variables. The small effect sizes for gender differences found in meta-analyses imply large overlap in the distribution of scores for men and women, which challenges the gender binary.

The Challenge From Psychological Research With Transgender and Nonbinary Individuals

In another instance reflecting a belief in the gender binary, most psychologists have implicitly assumed that everyone will adopt self-labels that are consistent with the category (female or male) that they were assigned at birth (based on the form of the genitals), and then conform to the gender stereotypes, roles, and expectations associated with that gender category. Although most people in the United States—and possibly worldwide—are cisgender, individuals whose identities do not conform to the gender binary are found throughout recorded history and across diverse cultures (e.g., Devor, 1997; Herdt, 1993). Some cultures have been marked by recognition and acceptance of such individuals. For example, many indigenous nations of the Americas had more than two gender categories and many alternate gender expressions (Tompkins, 2015). European colonizers brought with them a firm belief in the gender binary and criminalized or pathologized such practices (Tompkins, 2015). In the United States today, the visibility of individuals who do not conform to the gender binary is increasing. Some of these individuals are fighting for, and winning, the right to be formally recognized as having an identity that is neither male nor female (see, e.g., the case of Jamie Shupe; Foden-Vencil, 2016).

Today, *transgender* or *trans* refers to a spectrum of people and experiences, often collected under the term the *transgender umbrella* (e.g., Stryker, 2008). This umbrella can include those with a gender identity that does not conform to the birth-assigned category as well as those who practice nonconformity to social expectations in gender expression that cater to cisgender, heterosexual individuals. One surveillance report concluded that 0.5% of the U.S. adult population between the ages of 18 and 64 are transgender (Conron, Scott, Stowell, & Landers, 2012), although this is likely an underestimate. Using Centers for Disease Control and Prevention data, a more recently estimated prevalence in the United States was 0.6%, with a range across states from 0.3% in North Dakota to 0.8% in Hawaii (Flores, Herman, Gates, & Brown, 2016). The 0.6% prevalence translates to 1.4 million adults, which is not a small number.

Transgender and nonbinary individuals have largely been ignored in psychological research or have been separated out as disordered, dysphoric, or otherwise outside mainstream considerations of gender. The failure to integrate transgender and nonbinary experiences into traditional models of (cis)gender means that psychologists will consistently miss important aspects of how gender organizes and functions within people's lives (e.g., Tate, Youssef, & Bettergarcia, 2014; van Anders, 2015). The integration of these groups into research should help researchers uncover new experiences and highlight existing ones that have been obscured when only cisgender experience is the focus.

Psychological research on transgender and nonbinary individuals poses three key challenges to the gender binary. First, transgender and nonbinary individuals show that birth-assigned categories are imperfect for predicting how a person will self-label their gender identity, thereby undermining a key assumption of the gender binary. Second, although many cis and trans women and men experience their identities as one category and not the other, others experience their identity as nonbinary (Galupo, Pulice-Farrow, & Ramirez, in press; Joel, Tarrasch, Berman, Mukamel, & Ziv, 2014; Tate et al., 2014), challenging the assumption that gender/sex comprises only the dichotomous categories of male and female. Third, some transgender experiences highlight the fact that self-labeling with respect to gender ("being gender") is separable from the enactment of traits, roles, and behaviors ("doing gender"). Each challenge is described below.

Assignment at Birth ≠ Gender Identity

The existence and lived experiences of transgender and nonbinary individuals pose a direct challenge to the gender binary by providing evidence that sex assignment at birth

does not invariably predict individuals' felt gender identity. Research on self-concepts in childhood also supports this statement. Olson, Key, and Eaton (2015) examined cis and trans children's self-concepts as "girl" or "boy" using an implicit association test. They found differences between the two groups who shared the same birth-assigned sex category (i.e., between trans girls and cis boys, and between trans boys and cis girls) but similarities between the groups who shared the same gender identity. That is, trans girls and cis girls were statistically indistinguishable from each other on this task, as were trans boys and cis boys. This study is one of the first quantitative demonstrations of similar psychological processes at work for those with the same gender identity, irrespective of birth-assigned category. Thus, gender identity is not invariably linked to sex category at birth, and gender identity is a stronger predictor of experience than sex category assigned at birth.

Nonbinary Gender Identities

As noted above, some individuals do not self-categorize as exclusively female or male but as either the combination of the two (concurrently or fluidly) or as something else (Galupo et al., in press; Schudson, Dibble, & van Anders, 2017; Tate et al., 2014). These individuals are usually described as *nonbinary* and, indeed, their identities fundamentally challenge the gender binary. One study examined the narrative reports of those identifying specifically as *gender variant* and *agender* for those who resided in the United States and other countries where English was a primary language (Galupo et al., in press). The agender respondents in that sample provided poignant insights into their sense of self-categorization in a manner that requires more than two gender groups. For instance, one agender respondent noted, "I do not have an internal gender. It is not androgynous; it is not fluid; it is non-existent." Another agender respondent described their experience in mathematically graphable terms: "If there are two axes on a graph, one for how female a person feels, and another for how male they feel, I am very close to the coordinate (0,0)." In another study, participants reported that being able to mark fluidity—change over time and place—in one's gender/sex is key (Schudson et al., 2017). Other studies have found nonbinary gender experiences in cisgender individuals (Joel et al., 2014; Martin et al., 2017; Olson et al., 2015), demonstrating that studies of small and neglected groups can open up new questions, which are relevant to all and not just to the original group. Accordingly, fully integrating nonbinary experiences into psychology's views of gender identity requires that scholars recognize that *male* and *female* are insufficient for capturing the full range of identities and acknowledge that gender/sex may be irrelevant to individuals' sense of who they are.

Being Versus Doing Gender

Traditional psychological research conceptualized gender as a state of being (e.g., "I am a woman"). In contrast, sociologists West and Zimmerman (1987, 2009) introduced the concept of "doing gender," which philosopher Butler (1990, 1993) expanded upon as "gender performativity." West and Zimmerman (1987) argued, "We conceive of gender as an emergent feature of social situations: both as an outcome of and a rationale for various social arrangements and as a means of generating one of the most fundamental divisions of society" (p. 126). This concept of "doing gender" is closely linked to the view that gender is socially constructed at many levels, including interpersonal interactions and cultural messages (e.g., Marecek, Crawford, & Popp, 2004). For example, people in the United States "do gender" by engaging in gender displays through their choices concerning clothing, accessories, and grooming (e.g., wearing skirts and high heels, and shaving legs) and their adoption of roles and activities (e.g., working in computer science and repairing cars). Doing gender also occurs in interpersonal interactions, as, for example, when women adopt passive roles and men adopt active roles. This "doing gender" concept was so powerful that some scholars wondered whether gender was nothing more than a series of stylized acts—nothing more than performance that becomes felt as natural inclination (Butler, 1993).

Yet transgender and nonbinary individuals remind scholars that the "being gender" concept (e.g., "I am a woman") also has a place in considering gender as a whole. Moreover, being and doing gender should be considered in an integrated way for all individuals, whether cisgender, transgender, or nonbinary. To facilitate this integrated view, Tate and colleagues (2014) introduced an analytic model of gender as a bundle of interrelated, though separate, constructs, depicted in Figure 2. The bundle of gender con-

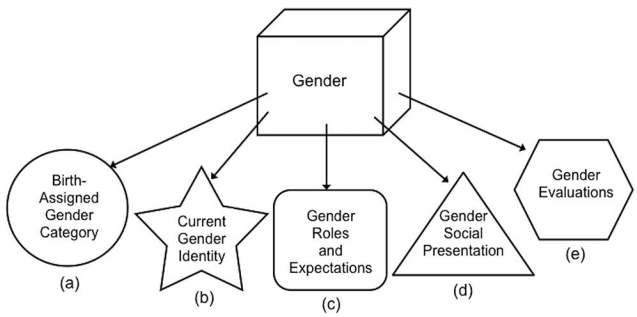

Figure 2. The facets within the bundle of constructs studied as gender within psychological science. Facets are numbered using letters (a through e) as a quick reference within the text. As indicated by their different shapes, the facets are presumed to function as separable or distinguishable from other facets. The box shape above all facets indicates that each of the facets is presumed to fit into the larger category called "gender" (Tate et al., 2014).

structs is meant to explicitly remind researchers that there are at least five separable components of *gender*. These components or facets are organizable around the intra- and interpersonal ways in which gender is described in the social world. As Figure 2 notes, the facets are (a) birth-assigned category (which is commonly determined by a physician based on genital appearance); (b) current identity (which is the individual's sense of what gender category [or categories] is an authentic statement of who they are); (c) gender roles, ideologies, and expectations (which is the set of social expectations that the individual finds self-relevant [or irrelevant] based on one's self-assigned categorization); (d) social presentation of gender identity (which is the set of interpersonal signals that convey one's self-assigned categorization, such as apparel, vocal modulation, and name use); and (e) evaluations of the social world based on gender (which includes gender bias [e.g., sexism] as well as self–other comparisons and other ingroup–outgroup attitudes, beliefs, and behaviors). In Figure 2, "doing gender" is depicted by two facets: (c) gender roles, expectations, and ideologies, and (d) social presentation of gender. Only one facet from Figure 2 is clearly about being gender: (b) current identity.

Research based on Tate et al.'s (2014) and other models of gender as a bundle of separable constructs (e.g., Huston, 1983; Spence & Helmreich, 1978) provides insight into the process of gender/sex development. For example, among those children whose self-assigned gender/sex category is important and directs their doing of gender, not everyone in the same self-assigned category will do gender similarly. There is substantial variability in agreement with stereotypes and engagement in particular gendered behaviors *within* gender groups (Liben & Bigler, 2002), some of which is based on intersections with sexuality (e.g., Lucal, 1999) as well as ethnicity and social class (e.g., Cole, 2009).

The study of these constructs among transgender individuals is likely to lead to further insights into how the being and doing of gender are connected. For example, the narratives of trans women and men (e.g., Bono, 2011; Boylan, 2013; Green, 2004; Khosla, 2015; Mock, 2014; Morris, 1974) suggest that "being gender" could be a precondition for "doing gender" in a way that is not tied to one's birth-assigned sex category. Such narratives often feature "turning points," such that once the individual realized that she was a girl (see Boylan, 2013; Mock, 2014; Morris, 1974) or that he was a boy (see Bono, 2011; Green, 2004; Khosla, 2015), they paid attention to the relevant stereotypes for the self-labeled category (not the birth-assigned category). Additionally, as their sense of self developed within the self-assigned label, they found that the relevant stereotypes for that category affected them more strongly—*even if they did not agree with the stereotypes*. That is, many narratives of trans women and men show that it is one's *self*-assigned gender identity that can focus a person's attention on how gender is expected to be done within that category.

Summary

Transgender and nonbinary individuals' experiences pose several serious challenges to the gender binary. First, transgender and nonbinary individuals show that birth-assigned categories are imperfect at predicting how a person self-categorizes with respect to gender/sex, undermining a key assumption of the gender binary. Second, although cis and trans women and men may experience their identities as one category and not the other, other individuals experience gender/sex as continuous, or even irrelevant to the self (Galupo et al., in press; Joel et al., 2014; Tate et al., 2014; van Anders, 2015). Third, transgender and nonbinary individuals highlight the fact that gender category self-labeling ("being gender") and gender roles and expectations ("doing gender") are both necessary for understanding how individuals psychologically process the systems of socialization around them in ways that are not linked to birth-assigned categories.

The Challenge From Developmental Psychology

Gender/sex is typically accorded special status by psychologists as a social category that is hardwired to emerge in the human psyche. Thus, psychologists' view of gender/sex contrasts with that of most other social categories (e.g., groups based on race, ethnicity, nationality, and social class), whose use and meaning are often considered context dependent and variable across individuals, time, and situations (e.g., Smedley & Smedley, 2005). Developmental research leads to questions about the special status afforded to gender/sex.

The last 20 years has seen a proliferation of research related to the question of why any particular attribute—including gender/sex—becomes a salient basis for categorization among children. When integrated with studies of conceptual and language development, social categorization research suggests that gender/sex emerges as a psychologically salient and meaningful dimension of human variation during childhood, not as the inevitable result of an innate mechanism, but instead as the result of societal practices that guarantee that children (over)learn to categorize the self and others into the binary categories of male and female.

Practices That Establish Gender/Sex as a Salient and Binary Category

If attention to gender/sex as a meaningful, binary social category is not automatic, what conditions promote it? The question is difficult to answer empirically because there are practical limits to the ability to experimentally manipulate children's gender environments. Researchers have therefore used novel-group para-

digms to study the conditions that cause children to attend to and categorize others as a function of their social group membership (Bigler, 2013). In such studies, children are typically assigned to experimentally created novel social groups, often marked by different colored t-shirts. Characteristics of the groups and their treatment within the environment (e.g., laboratory, classroom) are then manipulated, and children's views of the novel groups are assessed. Results from such studies are summarized in a theoretical account of the formation of social stereotypes and prejudices, titled *developmental intergroup theory* (Bigler & Liben, 2006, 2007), which posits that children's endogenous qualities, including their cognitive capacities (e.g., classification skill, perspective taking) and idiosyncratic characteristics (traits, interests), interact in dynamic ways with their environmental context to induce children to attend to, categorize others, and develop stereotypes and prejudice concerning gender/sex. Here, the environmental contributions to these processes are highlighted.

Heightening the perceptual discriminability of gender/sex. Inducing children to attend to and sort people routinely on the basis of attributes that they are unable to perceptually detect (e.g., many religious groups and nationalities) is difficult. When given sorting tasks, children often group by gender/sex, race, or clothing, but they rarely sort by imperceptible traits (e.g., political party membership; Bigler & Liben, 2006). Gender is highly perceptually salient; however, this is true largely because of social conventions. In fact, children are unable to detect the gender of other children when they appear without culturally stereotypic markers of their gender (e.g., hair styles, makeup, and clothing; Wild et al., 2000). Throughout much of U.S. history, gender-differentiated dress was legally mandated in most public settings, including schools and workplaces (Bartlett, 1994), further suggesting that gender/sex is perceptually marked so that it becomes psychologically salient, when it is not naturally so. Strong cultural norms continue to dictate that men and women differ in their (a) use of cosmetics and accessories (jewelry, hats, purses, belts, and shoes), (b) treatment of body hair, and (c) sculpting of their body shape via weight lifting, dieting, surgery, and so on.

The claim is not that children and adults are unable to detect adults' gender/sex without social markers. Instead, based on the scientific evidence, the claim is that (a) gender/sex can be made more or less perceptually salient, and (b) children would be unlikely to habitually sort their peers into binary gender/sex categories (i.e., *boys* and *girls*) were the cultural cues (e.g., hairstyles, clothing) that are currently linked to gender/sex to disappear.

Perceptual distinctiveness is not, however, sufficient to explain the psychological salience of social groups, including gender. That is, the mere presence of a discriminable property does not necessarily lead children to use it regularly as a basis of social categorization (A. S. Baron, Dunham, Banaji, & Carey, 2014; Bigler & Liben, 2006). This is true of biologically based attributes (e.g., skin color or height) as well as markers of novel groups (colored t-shirts). Infants and children are fully able to perceptually discriminate people in terms of many attributes that do not become the basis for routine and meaningful categorization (e.g., straight or curly hair, or right- or left-handedness). Additional causal mechanisms are required to account for children's chronic use of gender/sex to sort the self and others.

Linguistic labeling of gender/sex. It is well known that language affects children's conceptualization of inanimate objects and animate beings. As R. Brown (1958) argued, words serve as invitations to form categories. Infants and preschoolers readily accept such invitations. Waxman and her colleagues have shown that infants (3- to 24-month-olds) respond differently to identical stimuli as a function of whether they are given a noun label or not (Waxman, 2013; Waxman & Booth, 2001, 2003). For example, infants who hear novel objects labeled with the same noun are more likely to respond to the perceptual commonalities among objects than are infants who do not hear labels. In the case of gender, infants exposed to gendered nouns (e.g., "That man was nice") will think of and respond to people differently than infants who do not hear such nouns (e.g., "That person was nice").

Categorization, in turn, leads children to make inferences about category members. Specifically, categorization induces essentialist reasoning about categories, or the belief that things have *natures or underlying essences* that make them what they are. Category labels lead children to infer that the members of a category share deep, inherent, meaningful commonalities, even in the absence of perceptual or conceptual clues concerning such similarities (Gelman, 2003). This principle operates for social and nonsocial stimuli (e.g., Diesendruck, & Deblinger-Tangi, 2014). For example, a study of Jewish Israeli parents and children found that the single parental variable that most reliably predicted children's ethnic essentialism was parents' linguistic marking (via use of labels and generics) of Arabs and Jews (Segall, Birnbaum, Deeb, & Diesendruck, 2015).

English is characterized by a moderate level of gender marking. That is, some languages contain a greater number of linguistic forms that mark gender (e.g., Spanish, French), and some languages contain fewer (e.g., Hungarian, Turkish). In English, gendered language includes nouns (e.g., *girl*), honorific titles (e.g., *Miss*), occupational titles (e.g., *actress*), and pronouns (e.g., *she*, *his*). Words that denote gender in English use a binary system with two dichotomous markings: male and female. Although nongendered pronouns were introduced in English in the 1800s (D. Baron, 1986), there has been, and continues to be, no widely accepted linguistic convention for marking individuals who fall outside of or between the categories of male and female, or who would prefer that their gender not be marked. As a

consequence, children raised within English-speaking (and other gendered-language) environments are literally forced by language to attend to gender and view it as a binary category.

Explicit and implicit use of gender/sex for sorting. In addition to labeling, exposing children to environments that are characterized by gender/sex-based sorting contributes to children's attention to and sorting of the self and others by gender/sex. Although one might argue that the presence of the sorting is evidence of the individuals' hardwired attention to gender/sex, experimental work indicates otherwise. Exposure to the use of even novel, meaningless attributes by authority figures induces children to attend to and sort others into ingroups and outgroups, just as they do with gender/sex.

Bigler and Liben (2006, 2007) described two forms of category use: explicit and implicit. Explicit use makes category use clear to children via linguistic labels. For example, research has examined children's intergroup attitudes as a function of their teachers' use of novel color groups to organize classroom spaces and activities. In experimental classrooms, teachers used color groups to organize classroom desks, bulletin boards, and activities. In control classrooms, teachers ignored the color groups. After several weeks, children completed measures about the groups, which revealed that ingroup biases developed only in experimental classrooms, where teachers used the color groups rather than ignoring them. That is, even when groups were perceptually salient, explicit sorting by teachers was required for children to develop ingroup favoritism. The findings suggest that the mere presence of perceptual markers of gender/sex is insufficient to cause gender biases; instead, adults' use of the gender/sex binary to explicitly sort individuals (e.g., "boys" and "girls" basketball teams and bathrooms) causes gender/sex biases.

Sometimes, however, the sorting of people or the organization of the environment by gender (or other dimensions) is never articulated and instead is implicit (e.g., de facto social segregation). Experimental and correlational work indicates that children who are exposed to contexts (e.g., a room, a school, a neighborhood) in which the people are sorted into different spaces on the basis of some trait (e.g., race, shirt color) will show increased categorization and stereotyping on that basis (Bigler & Liben, 2006). Children are embedded in contexts in which the gender/sex binary is used routinely to sort people implicitly into romantic relationships, occupations, and social roles, and thus they come to attend to gender/sex and view it as a meaningful and binary social category and, without intervention, develop gender/sex stereotypes and prejudices.

Summary

Multiple causal mechanisms are involved in ensuring that gender/sex categorization is overlearned early in life. Research suggests that three societal practices—the exaggeration of the perceptual discriminability of gender/sex, routine linguistic labeling of individuals by gender/sex, and explicit and implicit sorting of individuals by gender/sex—causally contribute to children's tendency to categorize the self and others into the categories of male and female, and develop gender/sex stereotypes and prejudices. Reductions in these gender-binary-grounded practices will reduce, and possibly eliminate, children's tendency to view the world through gender binary lenses.

Costs of the Gender Binary

For decades, psychological scientists have relied on the assumption that all people can be sorted into just two categories, women and men, each with its own set of brain features, hormones, psychological characteristics, and gender identity. Furthermore, they assumed that the categorization of the self and others into the binary categories of male and female was natural, inevitable, and psychologically beneficial. Yet, as shown in earlier sections, the gender binary fundamentally misrepresents human biological and psychological states and processes. The drawbacks of the gender binary are numerous and collectively produce what is likely to be an enormous cost to human societies. These costs are especially unfortunate in that they are likely to be nearly entirely avoidable. Our review of such costs is not intended to be exhaustive but instead to be illustrative of rationales for changes to psychological, as well as broader societal, policies and practices.

First, reliance on the gender binary in research, despite evidence of its inadequacy, is an obstacle to scientific progress. Scientifically, the use of a binary variable (female/male, or woman/man) as a proxy for the many variables that are included under gender/sex is an obstacle to understanding the contribution of the different components of gender/sex to a multitude of phenomena. For example, an average difference between women and men (say, in the prevalence of depression) is often attributed to the effects of gender/sex, although it gives no clue as to the gender/sex-related mechanism(s) involved (e.g., hormone levels, gender role expectations, socioeconomic status).

Reliance on the gender binary leads researchers to treat some variables as related to one category and not the other (e.g., studying "male" hormones only in men, or nurturing behaviors only in women). This constitutes an obstacle to scientific progress on a broad array of topics, including the study of hormones and research on gender identity.

Moreover, the binary view of gender/sex overlooks the dynamic, reciprocally causal interactions between biologically based and environmentally based mechanisms, as well

as the overlap in distributions and mosaicism in hormones, brains, and behaviors. Relatedly, the gender binary fails to capture that what it means to be a woman, a man, or a nonbinary person depends on a host of other identities, experiences, and power structures, as intersectionality highlights (Cole, 2009).

Second, the gender binary denies and denigrates the existence of individuals whose bodies or identities fall outside of or between the categories of males and females. For example, intersex individuals have long fought for, and have been denied, the right to self-determination of their gender/sex, including the ability to claim an identity outside of the gender binary (Dreger & Herndon, 2009; Reis, 2007). Scholarship shows the negative outcomes on bodily integrity, autonomy, function, and trauma from surgery intended not to fix medical issues but instead to fit bodies into the gender binary (Chase, 1998; Dreger, 1999). Similarly, more than two social categories are needed to describe people's gender identities. Some people identify as women (cis or trans), some as men (cis or trans), and some as neither or nonbinary. Narratives suggest that forcing individuals into a gender category that is unwanted can have enormous psychological costs (e.g., Beemyn & Rankin, 2011).

Third, although gender similarities are the rule for most psychological attributes, many cultures hold on to beliefs that women and men are very different from each other, and these nonscientific views may (a) dictate individuals' treatment of others (e.g., gender discrimination), (b) shape individuals' conceptions of themselves (e.g., self-stereotypes), (c) undermine individuals' performance in stereotyped domains (e.g., stereotype threat), and (d) shape legal and social policies (e.g., institutional gender biases). There are many possible illustrations for such effects, and just three are presented here, beginning with academics.

In the realm of educational accomplishments, research suggests that gender stereotypes impede children's achievement in domains culturally viewed as inappropriate for their gender. As early as second grade, children hold the implicit stereotype that math is for boys (Cvencek, Meltzoff, & Greenwald, 2011), despite meta-analytic findings that girls do as well as boys in math. The worry is that, as children and adolescents develop their intentions about a future career, girls will not consider careers in math-intensive fields such as engineering and physics because they have absorbed cultural messages that math is for boys, not girls (e.g., Eccles, 1994). In this way, beliefs in the gender binary disempower people and limit human potential.

In the domain of educational policy, purported differences between male and female brains have been used as a rationale for single-sex schooling. One advocate claimed, "By the mid-1980s it was clear that the hemispheric compartmentalization of function that is so obvious in men's brains—left brain verbal, right brain spatial—applies less well or not at all to female brains" (Sax, 2005, p. 12). In fact, a meta-analysis of studies of brain lateralization showed that the gender difference is trivial, $d = 0.06$ (Voyer, 1996). Findings that the human brain is a gender/sex mosaic further suggest that assigning children to educational settings based on gender/sex runs counter to scientific evidence. Moreover, research shows that single-sex schooling produces no better outcomes for students than existing coeducational schooling does, for outcomes such as math performance, science performance, and educational aspirations (Pahlke, Hyde, & Allison, 2014). The costs to these misapplications of gender-binary-based neuroscience and psychological research are enormous. Hundreds of school districts in the United States have launched single-sex classrooms or schools, paid for by taxpayers, ostensibly supported by science, but instead contradicted by science. Costs are not only economic (e.g., training of staff and funding new buildings) but also societal, as single-sex education intensifies beliefs in the gender binary.

Sports policy provides a third example. The lack of scientific support for the hormonal or neural bases of gender binary categorization has not stopped policymakers from using them to make exclusionary policies (van Anders et al., 2017). For example, the International Association of Athletics Federations (IAAF) instituted a policy that excluded women from competing if they had high androgen levels—"hyperandrogenicity"—despite no evidentiary basis (IAAF, 2006). The international Court of Arbitration for Sport (2015) recently struck it down, although only after one banned athlete brought a case forward, and after a number of women—with and without intersex conditions or experiences—were subjected to invasive scrutiny and questioning. Indeed, many women were prevented from competing unless they agreed to coercive, unnecessary biomedical alterations, including genital surgery. The enforcement of the gender binary in sports policy on the basis of biological sex has thus had a host of negative implications for women's lives.

Toward a New View of Gender/Sex

Although it is premature to fully specify a replacement to the gender binary, the data reviewed here provide a preliminary set of findings that must underlie any new, expansive conceptualization of gender/sex:

1. Gender and sex are closely intertwined such that sex cannot be studied without consideration of gender, and studies of gender can often benefit from considering sex as well. Use of the term *gender/sex* will be helpful in expressing this close interconnection and overcoming a tendency to use gender and sex as binary categories.

2. Gender/sex is multidimensional and each component is dynamic and responsive, to both internal

forces (biological, cognitive) and external forces (social interactions, culture).

3. Individuals show variability across the different components of gender/sex, presenting a mosaic of biological and psychological characteristics that may not all align in a single category of the gender binary.

4. Thinking of oneself and others in terms of gender/sex is not inevitable.

In sum, the multidimensional, complex, interactive, and dynamic nature of gender/sex cannot be captured by a categorical variable, much less by a categorical variable with only two categories. One of the most important tasks facing psychologists is to create a better, more accurate conceptualization of gender/sex for use in education, research, and practice, and some have begun to do so (e.g., APA, 2015; Schiebinger, 2016).

Recommendations for Research and Practice

Given the evidence challenging its validity and the costs associated with the gender binary, its use by laypersons and researchers alike should be replaced. In many instances, the gender binary should be supplemented by additional categories (e.g., genderqueer). In some instances, it should be replaced or supplemented with variables that correlate with gender/sex categories (e.g., hormones, muscle mass, traits, socioeconomic status). In yet other instances, gender/sex should be ignored altogether. This section gives examples of how the emerging new understanding of gender/sex can be used to guide gender-related research and practice. It should be acknowledged, however, that the gender binary will continue to shape individuals' thinking and behavior, even as scientific understanding and practice evolve, and thus researchers will necessarily continue to use gender binary frameworks in some instances, including efforts to investigate gender-binary-based discrimination.

Implications for Research Methods

There is a distinction between the question of whether and how to treat gender/sex as a variable in a study and the question of who should be included in a study. Samples should be diverse in terms of gender, sexuality, socioeconomic status, and so on, to capture the wide variability of human experience. Whether any of these categories should be used in analyzing the results is a different question and depends on the study's aims. With regard to gender/sex, the answer could range between measuring it only to make sure a representative sample has been gathered without using it as a variable in analyzing the results, to measuring some or most of its many different biological and psychological components.

Recent calls from funding agencies to attend to sex and to include women and girls—and female nonhuman animals—represent progress in terms of who is included in a study but have also led to a limited approach that focuses merely on between-sex differences when analyzing the results. A focus on difference ignores the variability described in this article and limits the scientific potential for asking rich questions that get closer to empirical reality. Moreover, it is questionable whether this approach, which yields lists of differences between females and males, is scientifically useful, as different lists are created under different genetic, developmental, and environmental conditions (Fine, Jordan-Young, Kaiser, & Rippon, 2013; Joel & Yankelevitch-Yahav, 2014; Jordan-Young & Rumiati, 2012). This is especially relevant in humans, in whom intersectionality highlights that women and men are not homogeneous categories (Cole, 2009). Therefore much of this article has included a call to move "beyond" a female–male differences approach and instead to develop new research methods to study the rich complexity of gender/sex.

Conceptualizing and measuring gender/sex. It is becoming increasingly clear that psychology's conceptualization and measurement of gender/sex is inadequate. For example, psychologists typically report the numbers of male and female participants in the study without explaining the basis of the measurement, making clear the assumptions inherent in the measurement, offering participants alternative categories to choose, or assessing whether the individuals' gender/sex identity is personally meaningful. New methods must be adopted.

One necessary methodological change is to allow individuals to report their gender identity and the gender/sex of others in nonbinary ways. One possibility is to provide the options "female," "male," "transgender female," "transgender male," "genderqueer," and "other (specify)." Some researchers, including van Anders, use open-ended responses, asking, "What is your gender?" to avoid having participants select from options that might not reflect their own language. As another alternative, Tate, Ledbetter, and Youssef (2013) ask participants, "How do you currently identify?" to emphasize the importance of self-assigned labels. The main response options are "female," "male," "transgender female," "transgender male," "genderqueer (click for more options)," and "intersex." If participants indicate a genderqueer self-label, multiple specific labels appear (e.g., "agender," "bigender," and other self-labels that are currently used by nonbinary and gender nonconforming persons in the United States), from which the respondent chooses as many as are relevant.

The proliferation of discrete gender/sex categories may be perceived by some as untenable (e.g., on the basis of logistical or statistical grounds), and it seems possible that the gender binary will morph into a *gender trinary*, defined as men, women, and "everybody else." In our view, this strat-

egy, like the gender binary, is insufficient for achieving a full scientific understanding of gender/sex. Instead of grouping participants into cisgender, transgender, and nonbinary participants, one solution is to use continuous measures of gender/sex, described below, which can be used across the entire sample with no loss of statistical power.

In addition to the self, study participants are routinely asked to identify their parents, children, or friends using the traditional gender binary categories (e.g., mom or dad, daughter or son), a practice that reifies the gender binary and thus should be supplemented with neutral or open categories (e.g., parent, child, sibling).

A second methodological advance is to inquire about both birth-assigned and self-assigned gender/sex identities. One group recommends a question about birth assignment, current gender identity, and gender as currently lived (Bauer, Braimoh, Scheim, & Dharma, 2017). Another group follows the question about self-labeling with the question "What category were you assigned to at birth?" with the options "female," "male," and "intersex" (Tate et al., 2013). If the participant chooses "intersex" as a birth-assigned category, they answer an additional question about the gender that they were raised as (female or male).

With respect to developmental *stability*, researchers must acknowledge and address the fact that many gender/sex constructs, not just identity, may change over time for an individual (e.g., Schudson et al., 2017). Thus, researchers may want to examine how the self-relevance of gender norms, same-gender favoritism, gender stereotyping, feelings of gender typicality, and so on, change for individuals over time.

A third methodological change involves treating all gender/sex constructs as multidimensional and continuous, rather than unidimensional and dichotomous. With respect to *multidimensionality*, progress in understanding gender will require scholars to consider gender/sex to be a bundle of interrelated but separate constructs, including the "being" and the "doing" of gender (Tate et al., 2014). One can examine gender/sex identities, stereotypes, and behaviors via a variety of measures, including self-report, observational, and implicit measures (e.g., Olson et al., 2015). Moreover, researchers can use these measures and conceptualizations for cisgender, transgender, and nonbinary participants to understand where similarities and differences may lie. In addition to social–cognitive methods (e.g., reaction-time studies), researchers can explore the full range of self-report methods from the personality and individual differences literature (Tate, 2014; Tate et al., 2014). Furthermore, to the extent that gender/sex is a bundle of interrelated constructs, then researchers should use multivariate methods and statistics that respect this interrelationship while estimating the effects of each phenomenon.

With respect to the dimensional or *continuous* nature of gender/sex, it might, at least in some contexts, be thought of as lying along one or more continua instead of placement within discrete categories. For gender identity, researchers should assume that many individuals feel varying degrees of similarity or belongingness to both the categories "male" and "female." For example, with one measure, participants rate their felt similarity to both women and men, using questions such as "In the past 12 months, have you thought of yourself as a woman?" and "In the past 12 months, have you thought of yourself as a man?" with response options of "always," "often," "sometimes," "rarely," and "never" (Joel et al., 2014). In a community sample of 2,155 Israeli participants (mean age = 36), 35% felt to some extent like the other gender (for a similar approach with children, see Martin, Andrews, England, Zosuls, & Ruble, 2017).

Another possible approach includes conceptualizing gender/sex and its multiple facets as "branched" or "coincident" as opposed to aligned, concordant, or consistent (van Anders, 2015). Conceptualizing gender/sex as multifactorial requires new theories of gender/sex and new measurement approaches. One of these is sexual configurations theory (SCT; van Anders, 2015), which, in both theory and measurement, allows people to articulate being agender, as well as their binary or nonbinary gender/sex, and those that challenge gender norms. Newer measures like SCT, built with insights from a more diverse set of gender identities and experiences, have proven useful for exploring gender among gender-diverse individuals that include cisgender, transgender, gender nonconforming, agender, and gender nonbinary (Schudson et al., 2017).

A final recommendation concerns taking seriously the notion that asking participants about their own and others' gender/sex identities, even with the use of nonbinary categories, implicitly conveys to participants that gender/sex is important. In some cases, simple solutions can be applied, such as asking adults about their gender identity at the end of a study so that it does not influence responding. In other cases, solutions are not immediately obvious. Although it remains important to ask about gender/sex given how important it is in stratifying people's lives, it needs to be situated within an array of demographic questions and not positioned as the only or the most important way to define people. The issue is perhaps especially serious in developmental research. Children are unlikely to be familiar with various nonbinary terms, and thus researchers who assess children's gender identities, stereotypes, and beliefs are likely to continue to rely on gender binary frameworks. The field will need to develop methods of assessing gender/sex (e.g., individuals' attention to it, categorization of others by it, and attitudes concerning it) that do not inherently cause attention to gender/sex and its conceptualization as a binary, thereby influencing participants' psychological conceptions of gender/sex in the process of measurement. SCT is one of these measures (Schudson et al., 2017; van Anders, 2015). Researchers will need to consider what the goal of address-

ing gender/sex is for their research and consider questions accordingly.

Gender/sex for a nonbinary, biosocial world. *Gender/sex* is a research concept that helps to overcome the gender binary (Jordan-Young & Rumiati, 2012; van Anders, 2015). It reminds us that when researchers study human phenomena, they study phenomena that exist in women, men, and gender-diverse people: whole people who are embodied in ways that reflect both evolution and social context.

The concept and term *gender/sex* highlights the extent to which evolved sex and sociocultural gender are intertwined. For example, genitals and genes are physical, biological, and evolved, and are thus studied as part of sex. But scholars have demonstrated that these aspects of sex reflect gender (culture), too. For example, some people label a genital tubercle of a certain size a *clitoris*, whereas others label it a *penis*, in part related to cultural notions of how these genitals will be used (Dreger, 1998; Fausto-Sterling, 2000; Kessler, 1998). Another example is that the X and Y chromosomes are called *sex chromosomes*, with the X being associated with femaleness and the Y with maleness, and yet the human X chromosome contains multiple genes for sperm (Fujimura, 2006; Richardson, 2013). In both examples, what scientists call *sex*—genitals, genes—reflects how biological sex is socially interpreted and constructed.

The concept of gender/sex has been useful within social neuroendocrinology; it provides a framework for dynamically incorporating evolutionary and social context into research with hormones and humans (van Anders, 2013; van Anders et al., 2011), for example, by exploring how gendered social experiences influence what are typically seen as bodily markers of sex, like testosterone (van Anders et al., 2015). The concept of gender/sex is likely to prove useful in other fields as well. For example, gender-differentiated exposure to media over time may shape neural structures in ways that might otherwise be interpreted as innate sex differences, as with the claim that "men are more visual." Others have theorized that feminine norms limit girls' exposure to dirt, impacting immune physiology and having far-reaching effects on gendered patterns of autoimmune diseases (Clough, 2011). Gender/sex as a conceptual tool allows researchers to develop integrated research questions and paradigms that simultaneously explore biological and sociocultural influences that relate to gender/sex.

Writing for scientific audiences. Just as the APA issued guidelines in its publication manual prohibiting the use of the masculine generic (American Psychological Association, 2010, p. 73), the APA as well as researchers, reviewers, and editors will need to be more open to the changing landscape of gender, which will need to be reflected in language and terminology. One impetus for the change concerns the need to represent linguistically the growing number of individuals who claim an identity that falls outside of male and female. A second impetus for change is the need to embrace the use of gender-neutral pronouns to describe researchers and participants alike. This should be done not only for people who request it but also as a general policy, in view of the growing conviction that elimination of routine, non-necessary gender labeling is an essential component of efforts to prevent and reduce gender stereotyping and prejudice among youth (Bigler & Leaper, 2015). Although there is not widespread agreement on any single convention, several different systems of gender-neutral language have been proposed, including pronouns (e.g., *ze* in English, singular *they*), titles (e.g., *Mx*), and nouns (*partner*, *spouse*; Tamburin, 2015).

Implications for Clinical Practice

The new understanding of gender/sex that is emerging from research represents a serious challenge for clinical practice. For centuries, the treatment of individuals seeking assistance for mental health problems has been shaped by physicians' and therapists' belief in the gender binary (Brabender & Mihura, 2016). Treatment practices that are rooted in the gender binary persist, but they are under increasing scrutiny, as evidenced, for example, by the issuing of new gender/sex-related APA standards and guidelines (APA, 2007, 2015; discussed below). Space limitations do not permit coverage of the full range of clinical implications of new approaches to gender/sex here, but instead two points are highlighted: the role of the gender binary in (a) producing gender stereotypes that lead to bias in mental health diagnoses and treatment, and (b) shaping the diagnosis and treatment of adults and children who seek treatment for issues related specifically to gender (e.g., gender identity disorder, gender dysphoria).

Gender stereotypes and mental health. As described above, belief in the existence of two distinct, mutually exclusive gender/sex categories (i.e., the gender binary) promotes gender stereotyping. For example, the existence of gender/sex differences in means or in probabilities of some variable (e.g., aggression, disordered eating) is often translated to a categorical difference (women are like this, and men are like that); belief in such categorical differences can be harmful in clinical practice.

Consider the case of depression. There is a gender/sex difference in rates of depression; women are diagnosed with the disorder more often than men. Categorical views of gender/sex have contributed to the stereotyping of depression as a female disorder. Nonetheless, roughly one third of those with diagnosable depression are men (Salk et al., 2017). Imposing a categorical view that "depression = women" can lead to serious underdiagnosis for men and overdiagnosis for women; misdiagnosis, in turn, leads to treatment that is not optimal. Indeed, research indicates that even when they exhibit the same symptoms, men are less likely than women to be diagnosed with depression (Calla-

han et al., 1997; Stoppe, Sandholzer, Huppertz, Duwe, & Staedt, 1999).

Belief in the existence of two distinct, mutually exclusive gender/sex categories also encourages gender essentialist views, or beliefs that males and females are different in ways that are innate, biologically based, and inevitable. Such views are likely to affect theories of the causes of psychological disorders and their treatment. In particular, the search for environmental sources of disorders may be neglected because of strong gender essentialist views. In the case of depression, for example, estrogens have been implicated in the disorder (Hyde, Mezulis, & Abramson, 2008), but the field neglected the possibility that hormonal changes could be a result of exogenous, rather than endogenous, factors. As an example of such exogenous factors, a nationwide prospective cohort study of more than 1 million women living in Denmark reported that the first diagnosis of depression (and use of an antidepressant) was linked to onset of hormonal contraception, with the highest rates among adolescents (Skovlund, Mørch, Kessing, & Lidegaard, 2016).

In sum, psychologists need to be sensitized to the ways in which gender stereotypes of disorders may obscure understanding of etiological factors as well as affect their choices of diagnosis and treatment. Stereotypes can lead to overdiagnosis (of those who are stereotype-consistent) and underdiagnosis (of those who are stereotype-inconsistent) in ways that have major implications for health and access to treatment.

Gender-related mental health issues. The gender binary may have especially serious consequences for the diagnosis and treatment of individuals who seek assistance for mental health issues related to gender/sex. Belief in the gender binary dictates that children should be encouraged to develop identities that fall into one of the two gender/sex categories: male or female. The matter is a highly contentious one in the United States; the last 10 years have seen divisive, protracted debates over diagnoses such as gender identity disorder and gender dysphoria, and their treatment (e.g., Davy, 2015; Lawrence, 2014).

In 2015, the APA issued guidelines for psychological practice with transgender and gender-nonconforming people (APA, 2015). The overarching goal is practices that are affirming for transgender, nonbinary, and gender-nonconforming individuals. Many of the guidelines are consistent with the research reviewed in this article. For example, one of the guidelines asserts that gender is a nonbinary construct and another recognizes fluidity in gender identity over development. The report, however, did not take a stand about the preferred treatment of prepubertal children who present as transgender or gender nonconforming (TGNC):

> Due to the evidence that not all children persist in a TGNC identity into adolescence or adulthood, and because no approach to working with TGNC children has been adequately, empirically validated, consensus does not exist regarding best practice with prepubertal children. (APA, 2015, p. 842)

Despite entrenched beliefs in the gender binary, many therapists now advocate for clients' right to change genders, and many are forging ahead to probe the implications of nonbinary views. For example, a 2015 special issue of the journal *Psychology of Sexual Orientation and Gender Diversity* was dedicated to trans- and queer-inclusive theory and practice in therapy, focusing mainly on resilience (Meyer, 2015). As a second example, transgender and cisgender clinicians contributed to an edited book on affirmative therapy with transgender clients (Singh & dickey, 2016).

Following from the expansive view of gender/sex, everyone, including children, should be able to express their felt gender identity. Furthermore, this approach should apply to individuals, including children, whose identities and behavior combine elements traditionally associated with masculinity and femininity as well as those who completely reject gender/sex as an organizing self-construct. For this to occur, it is important to view all individuals as belonging to a common human group that varies in quantitative ways along various gender-related dimensions (rather than as dichotomous groups that vary qualitatively from each other) and to advocate for societal changes aimed at expanding our views of gender/sex, to make space for all identities, expressions, and behaviors.

Conclusion

For more than a century, psychological scientists have relied on the gender binary in research. Recent empirical findings stemming from multiple disciplines provide fundamental challenges to the notion that humans can be understood as belonging to only two categories: women and men. This evidence includes neuroscience findings that refute gender/sex dimorphism of the human brain; behavioral endocrinology findings that challenge the notion of biologically fixed, gender dimorphic hormonal systems; psychological findings that highlight the similarities between males and females; psychological research on transgender and nonbinary individuals' identities and experiences; and developmental research suggesting that the tendency to view gender/sex as a meaningful, binary category is not innate but instead is culturally determined and malleable. Furthermore, societal changes in the ways that laypersons think and feel about gender/sex are increasingly incompatible with the gender binary.

In addition to the challenges presented by science, there are enormous costs to maintaining the gender binary. These costs include the myriad negative consequences of gender stereotyping and prejudice. For these reasons, the gender binary should be replaced by a conception of gender/sex

that stresses multiplicity and diversity, including a multiple-category (rather than binary) system, whose categories are not mutually exclusive (one can identify as more than one), fluid (one's identity can change across time), and allow for the possibility that gender is viewed as irrelevant to the self. Following the recommendations for research and practice outlined here should facilitate scientific progress and promote positive developmental outcomes among all individuals.

References

American Psychological Association (APA). (2007). Guidelines for psychological practice with girls and women. *American Psychologist, 62,* 949–979. http://dx.doi.org/10.1037/0003-066X.62.9.949

American Psychological Association (APA). (2010). *Publication manual of the American Psychological Association* (6th ed.). Washington, DC: Author.

American Psychological Association (APA). (2015). Guidelines for psychological practice with transgender and gender nonconforming people. *American Psychologist, 70,* 832–864. http://dx.doi.org/10.1037/a0039906

Andrade, L., Caraveo-Anduaga, J. J., Berglund, P., Bijl, R. V., De Graaf, R., Vollebergh, W., . . . Wittchen, H. U. (2003). The epidemiology of major depressive episodes: Results from the International Consortium of Psychiatric Epidemiology (ICPE) Surveys. *International Journal of Methods in Psychiatric Research, 12,* 3–21. http://dx.doi.org/10.1002/mpr.138

Baron, A. S., Dunham, Y., Banaji, M., & Carey, S. (2014). Constraints on the acquisition of social category concepts. *Journal of Cognition and Development, 15,* 238–268. http://dx.doi.org/10.1080/15248372.2012.742902

Baron, D. (1986). *Grammar and gender*. New Haven, Connecticut: Yale University Press.

Bartlett, K. T. (1994). Only girls wear barrettes: Dress and appearance standards, community norms, and workplace equality. *Michigan Law Review, 92,* 2541–2582. http://dx.doi.org/10.2307/1290002

Bauer, G. R., Braimoh, J., Scheim, A. I., & Dharma, C. (2017). Transgender-inclusive measures of sex/gender for population surveys: Mixed-methods evaluation and recommendations. *PLoS ONE, 12*(5), e0178043. http://dx.doi.org/10.1371/journal.pone.0178043

Beemyn, G., & Rankin, S. (2011). *The lives of transgender people*. New York, NY: Columbia University Press.

Bigler, R. S. (2013). Understanding and reducing social stereotyping and prejudice among children. In M. R. Banaji & S. A. Gelman (Eds.), *Navigating the social world: What infants, children, and other species can teach us* (pp. 327–331). Oxford, UK: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780199890712.003.0060

Bigler, R. S., & Leaper, C. (2015). Gendered language: Psychological principles, evolving practices, and inclusive policies. *Policy Insights from the Behavioral and Brain Sciences, 2,* 187–194. http://dx.doi.org/10.1177/2372732215600452

Bigler, R. S., & Liben, L. S. (2006). A developmental intergroup theory of social stereotypes and prejudice. In R. V. Kail (Ed.), *Advances in child development and behavior* (Vol. 34, pp. 39–89). San Diego, CA: Elsevier. http://dx.doi.org/10.1016/S0065-2407(06)80004-2

Bigler, R. S., & Liben, L. S. (2007). Developmental intergroup theory: Explaining and reducing children's social stereotyping and prejudice. *Current Directions in Psychological Science, 16,* 162–166. http://dx.doi.org/10.1111/j.1467-8721.2007.00496.x

Blackless, M., Charuvastra, A., Derryck, A., Fausto-Sterling, A., Lauzanne, K., & Lee, E. (2000). How sexually dimorphic are we? Review and synthesis. *American Journal of Human Biology, 12,* 151–166.

Bono, C. (2011). *Transition: How I became a man*. New York, NY: Penguin.

Borenstein, M., Hedges, L. V., Higgins, J., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. New York, NY: Wiley. http://dx.doi.org/10.1002/9780470743386

Bornstein, K. (1994). *Gender outlaw: On men, women, and the rest of us*. New York, NY: Routledge.

Boylan, J. F. (2013). *She's not there: A life in two genders*. New York, NY: Broadway Paperbacks.

Brabender, V. M., & Mihura, J. L. (Eds.). (2016). *Handbook of gender and sexuality in psychological assessment*. New York, NY: Routledge.

Brizendine, L. (2006). *The female brain*. New York, NY: Morgan Road Books.

Brooks, J. (February, 2017). *Boy? Girl? Both? Neither? A new generation overthrows gender*. Retrieved from https://ww2.kqed.org/futureofyou/2017/02/15/boy-girl-both-neither-a-new-generation-overthrows-gender/

Brown, R. (1958). *Words and things: An introduction to language*. New York, NY: Macmillan.

Brown, S. L., Fredrickson, B. L., Wirth, M. M., Poulin, M. J., Meier, E. A., Heaphy, E. D., . . . Schultheiss, O. C. (2009). Social closeness increases salivary progesterone in humans. *Hormones and Behavior, 56,* 108–111. http://dx.doi.org/10.1016/j.yhbeh.2009.03.022

Buss, D. M. (1989). Sex differences in human mate preferences: Evolutionary hypotheses tested in 37 cultures. *Behavioral and Brain Sciences, 12,* 1–14. http://dx.doi.org/10.1017/S0140525X00023992

Butler, J. (1990). *Gender trouble: Feminism and the subversion of identity*. New York, NY: Routledge.

Butler, J. (1993). *Bodies that matter: On the discursive limits of sex*. New York, NY: Routledge.

Callahan, E. J., Bertakis, K. D., Azari, R., Helms, L. J., Robbins, J., & Miller, J. (1997). Depression in primary care: Patient factors that influence recognition. *Family Medicine, 29,* 172–176.

Chase, C. (1998). Surgical progress is not the answer to intersexuality. *The Journal of Clinical Ethics, 9,* 385–392.

Clough, S. (2011). Gender and the hygiene hypothesis. *Social Science & Medicine, 72,* 486–493. http://dx.doi.org/10.1016/j.socscimed.2010.11.021

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.

Cole, E. R. (2009). Intersectionality and research in psychology. *American Psychologist, 64,* 170–180. http://dx.doi.org/10.1037/a0014564

Conron, K. J., Scott, G., Stowell, G. S., & Landers, S. J. (2012). Transgender health in Massachusetts: Results from a household probability sample of adults. *American Journal of Public Health, 102,* 118–122. http://dx.doi.org/10.2105/AJPH.2011.300315

Court of Arbitration for Sport. (2015). *Interim arbitral award (Dutee Chand v. Athletics Federation of India (AFI) & International Association of Athletics Federations (IAAF), CAS 2014/A/3759)*. Retrieved from http://www.tas-cas.org/fileadmin/user_upload/award_internet.pdf

Cvencek, D., Meltzoff, A. N., & Greenwald, A. G. (2011). Math-gender stereotypes in elementary school children. *Child Development, 82,* 766–779. http://dx.doi.org/10.1111/j.1467-8624.2010.01529.x

Das, A. (2017). Network connections and salivary testosterone among older U.S. women: Social modulation or hormonal causation? *The Journals of Gerontology: Series B: Psychological Sciences and Social Sciences*. Advanced online publication. http://dx.doi.org/10.1093/geronb/gbx111

Das, A., & Sawin, N. (2016). Social modulation or hormonal causation? Linkages of testosterone with sexual activity and relationship quality in a nationally representative longitudinal sample of older adults. *Archives of Sexual Behavior, 45,* 2101–2115. http://dx.doi.org/10.1007/s10508-016-0815-2

Davy, Z. (2015). The DSM–5 and the politics of diagnosing transpeople. *Archives of Sexual Behavior, 44,* 1165–1176. http://dx.doi.org/10.1007/s10508-015-0573-6

Devor, H. (1997). *FTM: Female-to-male transsexuals in society*. Bloomington, IN: Indiana University Press.

Dibble, E. R., Goldey, K. L., & van Anders, S. M. (2017). Pair bonding and testosterone in men: Longitudinal evidence for trait and dynamic associations. *Adaptive Human Behavior and Physiology, 3,* 71–90. http://dx.doi.org/10.1007/s40750-016-0054-8

Diesendruck, G., & Deblinger-Tangi, R. (2014). The linguistic construction of social categories in toddlers. *Child Development, 85,* 114–123. http://dx.doi.org/10.1111/cdev.12130

Dreger, A. D. (1998). *Hermaphrodites and the medical invention of sex*. Cambridge, MA: Harvard University Press.

Dreger, A. D. (1999). *Intersex in the age of ethics*. Hagerstown, MD: University Publishing Group.

Dreger, A. D., & Herndon, A. M. (2009). Progress and politics in the intersex rights movement: Feminist theory in action. *GLQ: A Journal of Lesbian and Gay Studies, 15,* 199–224. http://dx.doi.org/10.1215/10642684-2008-134

Duffy, K. A., Harris, L. T., Chartrand, T. L., & Stanton, S. J. (2017). Women recovering from social rejection: The effect of the person and the situation on a hormonal mechanism of affiliation. *Psychoneuroendocrinology, 76,* 174–182. http://dx.doi.org/10.1016/j.psyneuen.2016.11.017

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

Eccles, J. S. (1994). Understanding women's educational and occupational choices: Applying the Eccles et al. model of achievement-related choices. *Psychology of Women Quarterly, 18,* 585–609. http://dx.doi.org/10.1111/j.1471-6402.1994.tb01049.x

Else-Quest, N. M., & Grabe, S. (2012). The political is personal: Measurement and application of nation-level indicators of gender equity in psychological research. *Psychology of Women Quarterly, 36,* 131–144. http://dx.doi.org/10.1177/0361684312441592

Else-Quest, N. M., & Hyde, J. S. (2016). Intersectionality in quantitative psychological research: I. Theoretical and epistemological issues. *Psychology of Women Quarterly, 40,* 155–170. http://dx.doi.org/10.1177/0361684316629797

Else-Quest, N. M., Hyde, J. S., & Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: A meta-analysis. *Psychological Bulletin, 136,* 103–127. http://dx.doi.org/10.1037/a0018053

Fausto-Sterling, A. (2000). *Sexing the body: Gender politics and the construction of sexuality*. New York, NY: Basic Books.

Fine, C. (2017). *Testosterone Rex: Myths of sex, science, and society*. New York, NY: Norton.

Fine, C., Jordan-Young, R., Kaiser, A., & Rippon, G. (2013). Plasticity, plasticity, plasticity . . . and the rigid problem of sex. *Trends in Cognitive Sciences, 17,* 550–551. http://dx.doi.org/10.1016/j.tics.2013.08.010

Flores, A. R., Herman, J. L., Gates, G. J., & Brown, T. N. T. (2016). *How many adults identify as transgender in the United States?* Los Angeles, CA: Williams Institute.

Foden-Vencil, K. (2016, June 17). *Neither male nor female: Oregon resident legally recognized as third gender*. Retrieved from http://www.npr.org/2016/06/17/482480188/neither-male-nor-female-oregon-resident-legally-recognized-as-third-gender

Fujimura, J. H. (2006). Sex genes: A critical sociomaterial approach to the politics and molecular genetics of sex determination. *Signs, 32,* 49–82. http://dx.doi.org/10.1086/505612

Galupo, P., Pulice-Farrow, L. K., & Ramirez, J. L. (in press). "Like a constantly flowing river": Gender identity flexibility among non-binary transgender individuals. In J. D. Sinnott (Ed.), *Identity flexibility during adulthood: Perspectives in adult development* (pp. 163–177). New York, NY: Springer. http://dx.doi.org/10.1007/978-3-319-55658-1_10

Garcia-Falgueras, A., Ligtenberg, L., Kruijver, F. P. M., & Swaab, D. F. (2011). Galanin neurons in the intermediate nucleus (InM) of the human hypothalamus in relation to sex, age, and gender identity. *The Journal of Comparative Neurology, 519,* 3061–3084. http://dx.doi.org/10.1002/cne.22666

Gelman, S. A. (2003). *The essential child*. New York, NY: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780195154061.001.0001

Gettler, L. T., McDade, T. W., Feranil, A. B., & Kuzawa, C. W. (2011). Longitudinal evidence that fatherhood decreases testosterone in human males. *Proceedings of the National Academy of Sciences of the United States of America, 108,* 16194–16199. http://dx.doi.org/10.1073/pnas.1105403108

Gillies, G. E., & McArthur, S. (2010). Estrogen actions in the brain and the basis for differential action in men and women: A case for sex-specific medicines. *Pharmacological Reviews, 62,* 155–198. http://dx.doi.org/10.1124/pr.109.002071

Goldey, K. L., Avery, L. R., & van Anders, S. M. (2014). Sexual fantasies and gender/sex: A multimethod approach with quantitative content analysis and hormonal responses. *Journal of Sex Research, 51,* 917–931. http://dx.doi.org/10.1080/00224499.2013.798611

Goldey, K. L., & van Anders, S. M. (2011). Sexy thoughts: Effects of sexual cognitions on testosterone, cortisol, and arousal in women. *Hormones and Behavior, 59,* 754–764. http://dx.doi.org/10.1016/j.yhbeh.2010.12.005

Granger, D. A., Shirtcliff, E. A., Booth, A., Kivlighan, K. T., & Schwartz, E. B. (2004). The "trouble" with salivary testosterone. *Psychoneuroendocrinology, 29,* 1229–1240. http://dx.doi.org/10.1016/j.psyneuen.2004.02.005

Green, J. (2004). *Becoming a visible man*. Nashville, TN: Vanderbilt University Press.

Hamilton, L. D., Carré, J. M., Mehta, P. H., Olmstead, N., & Whitaker, J. D. (2015). Social neuroendocrinology of status: A review and future directions. *Adaptive Human Behavior and Physiology, 1,* 202–230. http://dx.doi.org/10.1007/s40750-015-0025-5

Hankin, B. L., Abramson, L. Y., Moffitt, T. E., Silva, P. A., McGee, R., & Angell, K. E. (1998). Development of depression from preadolescence to young adulthood: Emerging gender differences in a 10-year longitudinal study. *Journal of Abnormal Psychology, 107,* 128–140. http://dx.doi.org/10.1037/0021-843X.107.1.128

Harris, J. A., Vernon, P. A., & Boomsma, D. I. (1998). The heritability of testosterone: A study of Dutch adolescent twins and their parents. *Behavior Genetics, 28,* 165–171. http://dx.doi.org/10.1023/A:1021466929053

Herdt, G. (Ed.). (1993). *Third sex, third gender: Beyond sexual dimorphism in culture and history*. New York, NY: Zone Books.

Huston, A. C. (1983). Sex-typing. In E. M. Hetherington (Ed.) & P. H. Mussen (Series Ed.), *Handbook of child psychology: Vol. 4. Socialization, personality, and social development* (pp. 387–486). New York, NY: Wiley.

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

Hyde, J. S., Lindberg, S. M., Linn, M. C., Ellis, A. B., & Williams, C. C. (2008). Gender similarities characterize math performance. *Science, 321,* 494–495. http://dx.doi.org/10.1126/science.1160364

Hyde, J. S., Mezulis, A. H., & Abramson, L. Y. (2008). The ABCs of depression: Integrating affective, biological, and cognitive models to explain the emergence of the gender difference in depression. *Psychological Review, 115,* 291–313. http://dx.doi.org/10.1037/0033-295X.115.2.291

International Association of Athletics Federations. (2006). *IAAF policy on gender verification*. Retrieved from https://oii.org.au/wpcontent/uploads/2009/01/iaaf_policy_on_gender_verification.pdf

Joel, D. (2011). Male or female? Brains are intersex. *Frontiers in Integrative Neuroscience, 5,* 57. http://dx.doi.org/10.3389/fnint.2011.00057

Joel, D. (2012). Genetic-gonadal-genitals sex (3G-sex) and the misconception of brain and gender, or, why 3G-males and 3G-females have intersex brain and intersex gender. *Biology of Sex Differences, 3,* 27. http://dx.doi.org/10.1186/2042-6410-3-27

Joel, D. (2014). Sex, gender, and brain—A problem of conceptualization. In S. Schmitz & G. Höppner (Eds.), *Gendered neurocultures: Feminist and queer perspectives on current brain discourses* (pp. 169–186). Vienna, Austria: Zaglossus. e.U.

Joel, D., Berman, Z., Tavor, I., Wexler, N., Gaber, O., Stein, Y., . . . Assaf, Y. (2015). Sex beyond the genitalia: The human brain mosaic. *Proceedings of the National Academy of Sciences of the United States of America, 112,* 15468–15473. http://dx.doi.org/10.1073/pnas.1509654112

Joel, D., & McCarthy, M. M. (2017). Incorporating sex as a biological variable in neuropsychiatric research: Where are we now and where should we be? *Neuropsychopharmacology, 42,* 379–385. http://dx.doi.org/10.1038/npp.2016.79

Joel, D., Tarrasch, R., Berman, Z., Mukamel, M., & Ziv, E. (2014). Queering gender: Studying gender identity in "normative" individuals. *Psychology & Sexuality, 5,* 291–321. http://dx.doi.org/10.1080/19419899.2013.830640

Joel, D., & Yankelevitch-Yahav, R. (2014). Reconceptualizing sex, brain and psychopathology: Interaction, interaction, interaction. *British Journal of Pharmacology, 171,* 4620–4635. http://dx.doi.org/10.1111/bph.12732

Jordan-Young, R., & Rumiati, R. I. (2012). Hardwired for sexism? Approaches to sex/gender in neuroscience. *Neuroethics, 5,* 305–315. http://dx.doi.org/10.1007/s12152-011-9134-4

Kessler, S. J. (1998). *Lessons from the intersexed*. New Brunswick, NJ: Rutgers University Press.

Khosla, D. (2015). *Both sides now: One man's journey through womanhood*. New York, NY: Penguin.

Koestner, R., & Aube, J. (1995). A multifactorial approach to the study of gender characteristics. *Journal of Personality, 63,* 681–710. http://dx.doi.org/10.1111/j.1467-6494.1995.tb00510.x

Kuijper, E. A. M., Lambalk, C. B., Boomsma, D. I., van der Sluis, S., Blankenstein, M. A., de Geus, E. J. C., & Posthuma, D. (2007). Heritability of reproductive hormones in adult male twins. *Human Reproduction, 22,* 2153–2159. http://dx.doi.org/10.1093/humrep/dem145

Lawrence, A. A. (2014). Gender assignment dysphoria in the. *DSM–5. Archives of Sexual Behavior, 43,* 1263–1266. http://dx.doi.org/10.1007/s10508-013-0249-z

Lee, P. A., Nordenström, A., Houk, C. P., Ahmed, S. F., Auchus, R., Baratz, A., . . . Global DSD Update Consortium. (2016). Global disorders of sex development update since 2006: Perceptions, approach and care. *Hormone Research in Paediatrics, 85,* 158–180. http://dx.doi.org/10.1159/000442975

Lenroot, R. K., & Giedd, J. N. (2010). Sex differences in the adolescent brain. *Brain and Cognition, 72,* 46–55. http://dx.doi.org/10.1016/j.bandc.2009.10.008

Liben, L. S., & Bigler, R. S. (2002). The developmental course of gender differentiation: Conceptualizing, measuring, and evaluating constructs and pathways. *Monographs of the Society for Research in Child Development, 67*(2), vii-147. http://dx.doi.org/10.1111/1540-5834.t01-1-00187

Liening, S. H., Stanton, S. J., Saini, E. K., & Schultheiss, O. C. (2010). Salivary testosterone, cortisol, and progesterone: Two-week stability, interhormone correlations, and effects of time of day, menstrual cycle, and oral contraceptive use on steroid hormone levels. *Physiology & Behavior, 99,* 8–16. http://dx.doi.org/10.1016/j.physbeh.2009.10.001

Lindberg, S. M., Hyde, J. S., Petersen, J. L., & Linn, M. C. (2010). New trends in gender and mathematics performance: A meta-analysis. *Psychological Bulletin, 136,* 1123–1135. http://dx.doi.org/10.1037/a0021276

Lucal, B. (1999). What it means to be gendered me: Life on the boundaries of a dichotomous gender system. *Gender and Society, 13,* 781–797. http://dx.doi.org/10.1177/089124399013006006

Marecek, J., Crawford, M., & Popp, D. (2004). On the construction of gender, sex, and sexualities. In A. H. Eagly, A. E. Beall, & R. J. Sternberg (Eds.), *The psychology of gender* (2nd ed., pp. 192–216). New York, NY: Guilford Press.

Martin, C. L., Andrews, N. C. Z., England, D. E., Zosuls, K., & Ruble, D. N. (2017). A dual identity approach for conceptualizing and measuring children's gender identity. *Child Development, 88,* 167–182. http://dx.doi.org/10.1111/cdev.12568

Martinez-San Miguel, Y., & Tobias, S. (Eds.). (2016). *Trans studies: The challenge to hetero- and homo normativities*. New Brunswick, NJ: Rutgers University Press.

Meyer, I. H. (2015). Resilience in the study of minority stress and health of sexual and gender minorities. *Psychology of Sexual Orientation and Gender Diversity, 2,* 209–213. http://dx.doi.org/10.1037/sgd0000132

Mock, J. (2014). *Redefining realness: My path to womanhood, identity, love, and so much more*. New York, NY: Atria Books.

Morris, J. (1974). *Conundrum*. New York, NY: Harcourt Brace Jovanovich.

Muehlenhard, C. L., & Peterson, Z. D. (2011). Distinguishing between sex and gender: History, current conceptualizations, and implications. *Sex Roles, 64,* 791–803. http://dx.doi.org/10.1007/s11199-011-9932-5

Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Math = male, me = female, therefore math not = me. *Journal of Personality and Social Psychology, 83,* 44–59. http://dx.doi.org/10.1037/0022-3514.83.1.44

Nyby, J. G. (2008). Reflexive testosterone release: A model system for studying the nongenomic effects of testosterone upon male behavior. *Frontiers in Neuroendocrinology, 29,* 199–210. http://dx.doi.org/10.1016/j.yfrne.2007.09.001

Olson, K. R., Key, A. C., & Eaton, N. R. (2015). Gender cognition in transgender children. *Psychological Science, 26,* 467–474. http://dx.doi.org/10.1177/0956797614568156

Overpeck, J. G., Colson, S. H., Hohmann, J. R., Applestine, M. S., & Reilly, J. F. (1978). Concentrations of circulating steroids in normal prepubertal and adult male and female humans, chimpanzees, rhesus monkeys, rats, mice, and hamsters: A literature survey. *Journal of Toxicology and Environmental Health, 4,* 785–803. http://dx.doi.org/10.1080/15287397809529700

Oxford, J. K., Tiedtke, J. M., Ossmann, A., Özbe, D., & Schultheiss, O. C. (2017). Endocrine and aggressive responses to competition are moderated by contest outcome, gender, individual versus team competition, and implicit motives. *PLoS ONE, 12*(7), e0181610. http://dx.doi.org/10.1371/journal.pone.0181610

Pahlke, E., Hyde, J. S., & Allison, C. M. (2014). The effects of single-sex compared with coeducational schooling on students' performance and attitudes: A meta-analysis. *Psychological Bulletin, 140,* 1042–1072. http://dx.doi.org/10.1037/a0035740

Petersen, J. L., & Hyde, J. S. (2010). A meta-analytic review of research on gender differences in sexuality, 1993–2007. *Psychological Bulletin, 136,* 21–38. http://dx.doi.org/10.1037/a0017504

Pittenger, C., & Duman, R. S. (2008). Stress, depression, and neuroplasticity: A convergence of mechanisms. *Neuropsychopharmacology, 33,* 88–109. http://dx.doi.org/10.1038/sj.npp.1301574

Pitts-Taylor, V. (2010). The plastic brain: Neoliberalism and the neuronal self. *Health: An Interdisciplinary Journal for the Social Study of Health,*

Illness and Medicine, 14, 635–652. http://dx.doi.org/10.1177/1363459309360796

Pleck, J. H. (1981). *The myth of masculinity*. Cambridge, MA: MIT Press.

Reich, C. G., Taylor, M. E., & McCarthy, M. M. (2009). Differential effects of chronic unpredictable stress on hippocampal CB1 receptors in male and female rats. *Behavioural Brain Research, 203,* 264–269. http://dx.doi.org/10.1016/j.bbr.2009.05.013

Reis, E. (2007). Divergence or disorder? The politics of naming intersex. *Perspectives in Biology and Medicine, 50,* 535–543. http://dx.doi.org/10.1353/pbm.2007.0054

Richardson, S. S. (2013). *Sex itself: The search for male and female in the human genome*. Chicago, IL: University of Chicago Press. http://dx.doi.org/10.7208/chicago/9780226084718.001.0001

Richardson, S. S., Daniels, C. R., Gillman, M. W., Golden, J., Kukla, R., Kuzawa, C., & Rich-Edwards, J. (2014). Society: Don't blame the mothers. *Nature, 512,* 131–132. http://dx.doi.org/10.1038/512131a

Ruigrok, A. N., Salimi-Khorshidi, G., Lai, M. C., Baron-Cohen, S., Lombardo, M. V., Tait, R. J., & Suckling, J. (2014). A meta-analysis of sex differences in human brain structure. *Neuroscience and Biobehavioral Reviews, 39,* 34–50. http://dx.doi.org/10.1016/j.neubiorev.2013.12.004

Salk, R. H., Hyde, J. S., & Abramson, L. Y. (2017). Gender differences in depression in representative national samples: Meta-analyses of diagnoses and symptoms. *Psychological Bulletin, 143,* 783–822. http://dx.doi.org/10.1037/bul0000102

Sapienza, P., Zingales, L., & Maestripieri, D. (2009). Gender differences in financial risk aversion and career choices are affected by testosterone. *Proceedings of the National Academy of Science of the United States of America, 106,* 15268–15273. http://dx.doi.org/10.1073/pnas.0907352106

Sax, L. (2005). *Why gender matters: What parents and teachers need to know about the emerging science of sex differences*. New York, NY: Broadway Books.

Saxbe, D. E., Edelstein, R. S., Lyden, H. M., Wardecker, B. M., Chopik, W. J., & Moors, A. C. (2017). Fathers' decline in testosterone and synchrony with partner testosterone during pregnancy predicts greater postpartum relationship investment. *Hormones and Behavior, 90,* 39–47. http://dx.doi.org/10.1016/j.yhbeh.2016.07.005

Schiebinger, L. (2016). *Gender variables in health research*. Retrieved from https://stanfordmedicine.app.box.com/s/fc1lv3gy62k05g4r7twvemgcty2wx1er

Schudson, Z. C., Dibble, E. R., & van Anders, S. M. (2017). Gender/sex and sexual diversity via sexual configurations theory: Insights from a qualitative study with gender and sexual minorities. *Psychology of Sexual Orientation and Gender Diversity, 4,* 422–437. http://dx.doi.org/10.1037/sgd0000241

Segall, G., Birnbaum, D., Deeb, I., & Diesendruck, G. (2015). The intergenerational transmission of ethnic essentialism: How parents talk counts the most. *Developmental Science, 18,* 543–555. http://dx.doi.org/10.1111/desc.12235

Singh, A. A. & Dickey, L. M. (Eds.). (2016). *Affirmative counseling and psychological practice with transgender and gender nonconforming clients*. Washington, DC: American Psychological Association.

Skovlund, C. W., Mørch, L. S., Kessing, L. V., & Lidegaard, Ø. (2016). Association of hormonal contraception with depression. *Journal of the American Medical Association Psychiatry, 73,* 1154–1162. http://dx.doi.org/10.1001/jamapsychiatry.2016.2387

Smedley, A., & Smedley, B. D. (2005). Race as biology is fiction, racism as a social problem is real: Anthropological and historical perspectives on the social construction of race. *American Psychologist, 60,* 16–26. http://dx.doi.org/10.1037/0003-066X.60.1.16

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

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

Stanton, S. J., & Schultheiss, O. C. (2007). Basal and dynamic relationships between implicit power motivation and estradiol in women. *Hormones and Behavior, 52,* 571–580.

Stoppe, G., Sandholzer, H., Huppertz, C., Duwe, H., & Staedt, J. (1999). Gender differences in the recognition of depression in old age. *Maturitas, 32,* 205–212. http://dx.doi.org/10.1016/S0378-5122(99)00024-9

Stryker, S. (2008). *Transgender history*. Berkeley, CA: Seal Press.

Tamburin, A. (2015, September 5). Colleges trend toward gender-neutral pronouns. *USA Today*. Retrieved from http://www.usatoday.com/story/news/nation/2015/09/05/colleges-trend-toward-gender-neutral-pronouns/71780214/

Tate, C. (2014). Gender identity as a personality process. In B. L. Miller (Ed.), *Gender identity: Disorders, developmental perspectives and social implications* (pp. 1–22). Hauppauge, NY: Nova Science Publishers.

Tate, C. C., Ledbetter, J. N., & Youssef, C. P. (2013). A two-question method for assessing gender categories in the social and medical sciences. *Journal of Sex Research, 50,* 767–776. http://dx.doi.org/10.1080/00224499.2012.690110

Tate, C. C., Youssef, C. P., & Bettergarcia, J. N. (2014). Integrating the study of transgender spectrum and cisgender experiences of self-categorization from a personality perspective. *Review of General Psychology, 18,* 302–312. http://dx.doi.org/10.1037/gpr0000019

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

Tompkins, K. W. (2015). Intersections of race, gender, and sexuality: Queer of color critique. In S. Herring (Ed.), *Cambridge companion to gay and lesbian literature* (pp. 173–189). New York, NY: Cambridge University Press. http://dx.doi.org/10.1017/CCO9781107110250.013

Tran, C. V., Cole, D. A., & Weiss, B. (2012). Testing reciprocal longitudinal relations between peer victimization and depressive symptoms in young adolescents. *Journal of Clinical Child and Adolescent Psychiatry, 41,* 353–360. http://dx.doi.org/10.1080/15374416.2012.662674

Tulchinsky, D., Hobel, C. J., Yeager, E., & Marshall, J. R. (1972). Plasma estrone, estradiol, estriol, progesterone, and 17-hydroxyprogesterone in human pregnancy. I. Normal pregnancy. *American Journal of Obstetrics and Gynecology, 112,* 1095–1100. http://dx.doi.org/10.1016/0002-9378(72)90185-8

Unger, R. K. (1979). Toward a redefinition of sex and gender. *American Psychologist, 34,* 1085–1094. http://dx.doi.org/10.1037/0003-066X.34.11.1085

van Anders, S. M. (2010). Chewing gum has large effects on salivary testosterone, estradiol, and secretory immunoglobulin A assays in women and men. *Psychoneuroendocrinology, 35,* 305–309. http://dx.doi.org/10.1016/j.psyneuen.2009.06.009

van Anders, S. M. (2013). Beyond masculinity: Testosterone, gender/sex, and human social behavior in a comparative context. *Frontiers in Neuroendocrinology, 34,* 198–210. http://dx.doi.org/10.1016/j.yfrne.2013.07.001

van Anders, S. M. (2015). Beyond sexual orientation: Integrating gender/sex and diverse sexualities via sexual configurations theory. *Archives of Sexual Behavior, 44,* 1177–1213. http://dx.doi.org/10.1007/s10508-015-0490-8

van Anders, S. M., & Dunn, E. J. (2009). Are gonadal steroids linked with orgasm perceptions and sexual assertiveness in women and men? *Hormones and Behavior, 56,* 206–213. http://dx.doi.org/10.1016/j.yhbeh.2009.04.007

van Anders, S. M., Goldey, K. L., & Bell, S. N. (2014). Measurement of testosterone in human sexuality research: Methodological considerations. *Archives of Sexual Behavior, 43,* 231–250. http://dx.doi.org/10.1007/s10508-013-0123-z

van Anders, S. M., Goldey, K. L., & Kuo, P. X. (2011). The steroid/peptide theory of social bonds: Integrating testosterone and peptide responses for classifying social behavioral contexts. *Psychoneuroendocrinology, 36,* 1265–1275. http://dx.doi.org/10.1016/j.psyneuen.2011.06.001

van Anders, S. M., Schudson, Z. C., Abed, E. C., Beischel, W. J., Dibble, E. R., Gunther, O. D., . . . Silver, E. R. (2017). Biological sex, gender, and public policy. *Policy Insights from the Behavioral and Brain Sciences, 4,* 194–201. http://dx.doi.org/10.1177/2372732217720700

van Anders, S. M., Steiger, J., & Goldey, K. L. (2015). Effects of gendered behavior on testosterone in women and men. *Proceedings of the National Academy of Sciences of the United States of America, 112,* 13805–13810. http://dx.doi.org/10.1073/pnas.1509591112

van Anders, S. M., Tolman, R. M., & Volling, B. L. (2012). Baby cries and nurturance affect testosterone in men. *Hormones and Behavior, 61,* 31–36. http://dx.doi.org/10.1016/j.yhbeh.2011.09.012

van Anders, S. M., & Watson, N. V. (2006). Social neuroendocrinology: Effects of social contexts and behaviors on sex steroids in humans. *Human Nature, 17,* 212–237. http://dx.doi.org/10.1007/s12110-006-1018-7

Voyer, D. (1996). On the magnitude of laterality effects and sex differences in functional lateralities. *Laterality: Asymmetries of Body, Brain and Cognition, 1,* 51–84. http://dx.doi.org/10.1080/713754209

Wagner, C. K. (2006). The many faces of progesterone: A role in adult and developing male brain. *Frontiers in Neuroendocrinology, 27,* 340–359. http://dx.doi.org/10.1016/j.yfrne.2006.07.003

Waxman, S. R. (2013). Building a better bridge. In M. Banaji, S. Gelman, & S. Lehr (Eds.), *Navigating the social world: The early years* (pp. 292–296). New York, NY: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780199890712.003.0053

Waxman, S. R., & Booth, A. E. (2001). Seeing pink elephants: Fourteen-month-olds' interpretations of novel nouns and adjectives. *Cognitive Psychology, 43,* 217–242. http://dx.doi.org/10.1006/cogp.2001.0764

Waxman, S., & Booth, A. (2003). The origins and evolution of links between word learning and conceptual organization: New evidence from 11-month-olds. *Developmental Science, 6,* 128–135. http://dx.doi.org/10.1111/1467-7687.00262

Weaver, I. C. G., Cervoni, N., Champagne, F. A., D'Alessio, A. C., Sharma, S., Seckl, J. R., . . . Meaney, M. J. (2004). Epigenetic programming by maternal behavior. *Nature Neuroscience, 7,* 847–854. http://dx.doi.org/10.1038/nn1276

Weissman, M. M., & Klerman, G. L. (1977). Sex differences and the epidemiology of depression. *Archives of General Psychiatry, 34,* 98–111. http://dx.doi.org/10.1001/archpsyc.1977.01770130100011

West, C., & Zimmerman, D. H. (1987). Doing gender. *Gender & Society, 1,* 125–151. http://dx.doi.org/10.1177/0891243287001002002

West, C., & Zimmerman, D. H. (2009). Accounting for doing gender. *Gender & Society, 23,* 112–122. http://dx.doi.org/10.1177/0891243208326529

Wild, H. A., Barrett, S. E., Spence, M. J., O'Toole, A. J., Cheng, Y. D., & Brooke, J. (2000). Recognition and sex categorization of adults' and children's faces: Examining performance in the absence of sex-stereotyped cues. *Journal of Experimental Child Psychology, 77,* 269–291. http://dx.doi.org/10.1006/jecp.1999.2554

Woolley, H. T. (1910). A review of the recent literature on the psychology of sex. *Psychological Bulletin, 7,* 335–342. http://dx.doi.org/10.1037/h0066338

Yoder, J. (2003). *Women and gender: Transforming psychology* (2nd ed.). Upper Saddle River, NJ: Pearson/Prentice Hall.

Zell, E., Krizan, Z., & Teeter, S. R. (2015). Evaluating gender similarities and differences using metasynthesis. *American Psychologist, 70,* 10–20. http://dx.doi.org/10.1037/a0038208

Received July 8, 2017
Revision received January 22, 2018
Accepted January 22, 2018 ■