

He is a Stud, She is a Slut! A Meta-Analysis on the Continued Existence of Sexual Double Standards

Joyce J. Endendijk¹ , Anneloes L. van Baar¹, and Maja Deković¹

Personality and Social Psychology Review
2020, Vol. 24(2) 163–190
© 2019 by the Society for Personality
and Social Psychology, Inc.



Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/1088868319891310
pspr.sagepub.com



Abstract

(Hetero)sexual double standards (SDS) entail that different sexual behaviors are appropriate for men and women. This meta-analysis ($k = 99$; $N = 123,343$) tested predictions of evolutionary and biosocial theories regarding the existence of SDS in social cognitions. Databases were searched for studies examining attitudes or stereotypes regarding the sexual behaviors of men versus women. Studies assessing differences in evaluations, or expectations, of men's and women's sexual behavior yielded evidence for traditional SDS ($d = 0.25$). For men, frequent sexual activity was more expected, and evaluated more positively, than for women. Studies using Likert-type-scale questionnaires did not yield evidence of SDS (combined $M = -0.09$). Effects were moderated by level of gender equality in the country in which the study was conducted, SDS-operationalization (attitudes vs. stereotypes), questionnaire type, and sexual behavior type. Results are consistent with a hybrid model incorporating both evolutionary and sociocultural factors contributing to SDS.

Keywords

sexual double standards, meta-analysis, gender, sexuality, social cognitions

Women and men often are held to different standards of appropriate behavior (Foschi, 2000; Prentice & Carranza, 2002). For example, women are penalized more than men for self-promoting behavior (Rudman, 1998) and for speaking in a direct and dominant manner (Carli et al., 1995). In addition, compared with women, men are penalized for passiveness (Costrich et al., 1975) and modest behavior (Moss-Racusin et al., 2010). Such “backlash” effects (Rudman & Fairchild, 2004), whereby men and women receive sanctions for violating social standards for their behavior, can have far-ranging negative consequences for individuals and society.

The present analysis focuses on (hetero)sexual double standards (SDS), in which different sexual behaviors are expected of, and valued for, men and women (Emmerink, Vanwesenbeeck, et al., 2016; Zaikman & Marks, 2017). Traditionally, men/boys are expected to be sexually active, dominant, and the initiator of (hetero)sexual activity, whereas women/girls are expected to be sexually reactive, submissive, and passive. Moreover, traditionally men are granted more sexual freedom than women. As a consequence, men and women can be treated differently for the same sexual behaviors. For example, slut-shaming is experienced by 50% of girls, compared with 20% of boys (Hill & Kearl, 2011).

Furthermore, traditional SDS have been associated with gender differences in sexual coercion and violence (Shen et al., 2012), as well as in sexual pleasure and achieving orgasms (Kiefer et al., 2006; Sanchez et al., 2012). SDS have also been associated with gender differences in sexual risk

behavior, specifically with more sexual partners for men, and more reluctance to request or insist on condom use for women (Lefkowitz et al., 2014). Traditional SDS have further been related to other societal problems, such as homophobia, sexism, and gender inequality (Zaikman & Marks, 2014; Zaikman, Marks, et al., 2016).

However, research is inconsistent about the continued existence and extent of SDS (for narrative reviews, see Bordini & Sperb, 2013; Crawford & Popp, 2003; Zaikman & Marks, 2017), which might be due to, among other reasons, differences between the studies in the conceptualization and measurement of SDS. Because of the negative implications of SDS for men and women, it is important to illuminate whether SDS still exist. This undertaking necessitates probing whether the conclusions about its existence depend on the sexual behavior type assessed, or on how SDS are measured or conceptualized. Therefore, we conducted a meta-analysis to examine whether SDS are present in society and which measures and conceptualizations yield evidence for the existence of SDS, and which do not. We define existence of SDS as the degree to which people have internalized SDS

¹Utrecht University, The Netherlands

Corresponding Author:

Joyce J. Endendijk, Child and Adolescent Studies, Utrecht University,
Heidelberglaan 1, P.O. Box 80140, 3508 TC Utrecht, The Netherlands.
Email: J.J.Endendijk@uu.nl

in their own social cognitions (i.e., stereotypes, attitudes; Greenwald et al., 2002).

Theoretical Perspectives on SDS

In our work, we draw on several distinct yet sometimes overlapping theoretical frameworks to make predictions about the existence and moderators of SDS. Our goal was to provide a broad theoretical overview of the conditions under which SDS would be present. First, evolutionary theory (Buss & Schmitt, 1993; Trivers, 1972) and biosocial theory (Wood & Eagly, 2002, 2012) assume the existence of SDS and make predictions about moderators of the strength of SDS, specifically with regard to the behavioral specificity of SDS, cultural differences, and historical change. Second, we use the theoretical framework of male and female control theory, which builds on premises of evolutionary and biosocial theory, to make predictions about gender differences in SDS. Third, we employ the gender-intensification hypothesis, which is similar to biosocial theory in its focus on gender roles, to predict age differences in SDS. The specific predictions we derived from each perspective are primarily based on our interpretation of the theories. The original theorists did not necessarily specify these concrete predictions, but we believe that they logically follow from their core propositions.

Theoretical perspectives on the existence of SDS

Evolutionary theory. Evolutionary theories, and especially gender differences in parental investment and reproductive strategies, provide rationales for the differential expectations and evaluations of men's and women's sexual behavior. Regarding parental investment, women biologically invest more in their children than men (e.g., egg cells are more precious than sperm cells, 9-month pregnancy, delivery; Trivers, 1972). Due to the lower parental investment of men compared with women, there is a high degree of competition among males for female mates. In this context, being highly dominant and assertive sexually is likely to increase mating success for men (Buss & Schmitt, 1993). In addition, men benefit more than women from having frequent sex with multiple partners, as this increases the likelihood of passing their genes on to a next generation. In contrast, the higher parental investment by women makes them more selective with regard to choosing mates or withholding sex, until they may be sure that the partner can provide resources for their children (Oliver & Hyde, 1993) and is willing to assist in raising their children (Wiederman & Allgeier, 1992). These evolutionary processes are supposed to unconsciously influence how we view sexual behavior of others and ourselves. More specifically, it has been suggested that physical or behavioral traits that indicate reproductive fitness elicit positive evaluations and traits that indicate a lack of fitness elicit negative evaluations (Milhausen & Herold, 2001).

Biosocial theory. According to biosocial theory (Wood & Eagly, 2002, 2012), different norms for the behavior of men and women arise from societies' division in gender roles: the female role of homemaker and the male role of economic provider. These different roles emerged, among others, from biological differences between men and women, with men being physically stronger, and women investing more in childbearing and nursing. As such, this theory integrates evolutionary processes related to parental investment and sexual strategies, although the division of gender roles is viewed as the most proximal cause of gender differences. Traditionally, the male role is characterized by competence, independence, assertiveness, power, and leadership, whereas the female role is characterized by submissiveness, kindness, consideration, helping, nurturing, and caring. People are expected to behave according to their gender roles and behavior that adheres to gender roles elicits positive evaluations, whereas behavior that violates gender roles elicits negative evaluations (Gaunt, 2012). Because gender roles are social constructions, socialization processes such as observational learning, reward, and punishment are important for learning what constitutes appropriate sexual behavior for men and women (Bussey & Bandura, 1999). These processes influence men's and women's own sexual behavior, but also their cognitions about SDS.

Applied to sexual behavior, the power difference in gender roles means that society expects men to be sexually agentic, that is, dominant, powerful, and assertive, and rewards men for such behaviors. In contrast, society expects women to be sexually communal, that is, submissive, passive, and reactive in sexual relationships, and accordingly rewards women for such behaviors. Most previous research on gender differences in sexual agency has focused on initiation patterns, with men initiating sex more frequently than women (Dworkin & O'Sullivan, 2005). Other examples of sexually agentic behaviors are perpetration of interpersonal sexual violence and willingness to engage in casual sexual relations (Mosher & Danoff-Burg, 2005), or being the director of sex as well as the teacher and expert in sex (Schwartz & Rutter, 2000). Examples of women's more sexually communal behaviors are associating sex with complying and submitting, not communicating one's own desires (Fetterolf & Sanchez, 2015; Kiefer et al., 2006), and consent to unwanted sexual activities in relationships (O'Sullivan & Allgeier, 1998).

Theoretical perspectives on moderators of SDS

Gender differences in SDS. Other theories that provide relevant predictions with regard to SDS are male and female control theory (Baumeister & Twenge, 2002). This theoretical framework integrates both evolutionary and sociocultural (i.e., feminist) perspectives to explain suppression of female sexuality in general, and more specifically to explain gender differences in SDS. According to male control the-

ory, SDS can be viewed as a male privilege that men want to keep in place. Advantages for men that arise from SDS, and the associated sexual dominance of men and suppression of female sexuality, are improved certainty about paternity (Buss, 1994), reduced male insecurity, and prevention of social chaos caused by widespread, indiscriminate sex (Baumeister & Twenge, 2002; Hyde & DeLamater, 1997). As such, SDS are part of a patriarchal system that is created by and for men, and suppresses women. Men are invested in patriarchy more than women, and therefore men might also be more supportive of SDS (Rudman et al., 2013). In addition, women are less accepting than men of social hierarchies that subordinate women (Lee et al., 2011), such as traditional gender roles.

In contrast, female control theory (Baumeister & Twenge, 2002) proposes that female sexuality is a more valuable resource than male sexuality, because gender differences in sexual desire leads to a higher demand for female sexuality. As a consequence, SDS in which female sexuality is suppressed have advantages for women, because they can trade highly valued sexual favors for lower valued favors from men, such as economic provision, monogamous relationships, and parental investment. There is, however, less empirical support for female control theory than there is for male control theory (Petersen & Hyde, 2010; Rudman et al., 2013). Because male control theory proposes that SDS constitute a form of male privilege that men want to control and keep in place, we predicted that men would be more likely than women to hold traditional SDS.

Behavioral specificity of SDS. An important question is whether the existence of SDS is behavior specific. In the current meta-analysis, we defined sexual behavior as any behavior involving oral sex and/or sexual intercourse (vaginal/anal), because too few studies examined SDS in other behaviors, such as petting, kissing, or manual stimulation of the genitalia. Previous research examined the existence of SDS in a myriad of sexual behaviors, ranging from premarital sex in committed relationships (e.g., Sprecher et al., 2013), having sex outside committed relationships (i.e., casual sex; e.g., Penhollow et al., 2017), having sex for the first time at the age of 16 or younger (i.e., early sexual debut; e.g., Sprecher et al., 1987), sexual infidelity or having a sexual affair (e.g., Haavio-Mannila & Kontula, 2003), being highly sexually active (e.g., many sexual partners versus a few sexual partners, engaging in threesomes, multiple sexual partners at the same time) versus being less sexually active (e.g., Sprecher et al., 1991), to sexual coercion between persons that are in a power or age hierarchy, for example a teacher and a student (e.g., Dollar et al., 2004).

Even though evolutionary theory and biosocial theory have often been pitted against each other in the literature, there is accumulating evidence for hybrid models explaining gender differences in sexuality from the interplay between evolutionary predispositions and sociocultural

pressures (Lippa, 2009). For example, the relative power of evolutionary and biosocial theory to explain gender differences may vary depending on the behavior under consideration (Cross et al., 2013; Lippa, 2009). In terms of reproductive fitness, men would benefit more than women from having frequent casual sex with many partners, having an early sexual debut, and having sex with other person's during a committed relationship (i.e., sexual infidelity; Buss & Schmitt, 1993; Petersen & Hyde, 2010). For men engaging in these behaviors is likely to increase the success of passing genes on to the next generation, whereas for women refraining or postponing these behaviors is likely to be a more successful reproductive strategy because of their higher parental investment. Therefore, based on evolutionary theory we expect SDS to be most prevalent for these specific behaviors and less for other sexual behaviors, such as premarital sex in committed relationships, or sexual coercion. It is important to realize that sexual behaviors are not limited to the context of reproduction, because use of condoms and contraceptives can prevent actual impregnation, but rather evolutionary processes can be viewed as general tendencies underlying sexual behavior (Zaikman & Marks, 2017).

Biosocial theory (Wood & Eagly, 2002, 2012) would predict that SDS are most prevalent in sexual encounters where there is a power/status difference between men and women, and less so in the other sexual behaviors (e.g., early sexual debut, casual sex, sexual infidelity, premarital sex in committed relationship, high sexual activity level; see also Zaikman & Marks, 2017). In patriarchal societies, men hold more power than women, which is supposed to underlie gender differences in sexual behavior and attitudes (Fugère et al., 2008; Petersen & Hyde, 2010). In such a context, society affords sexual agency more to men than to women. More specifically, SDS may serve to justify male sexual coercion toward women (Krahé et al., 2000; Warner, 2000), possibly because men exerting power over women in a sexual context fit with their agentic role. In addition, male victims of sexual coercion or rape might be perceived as powerless and not willing to have sex, which violates men's (hetero)sexual agentic gender role (Weis, 2010). Double standards with regard to sexual coercion in power/age hierarchies have most often been studied in person perception studies. In such studies, participants evaluate a perpetrator (e.g., higher status or older) and/or a victim (e.g., lower status or younger) in a hypothetical sexual context. Based on biosocial theory we expect the strongest SDS in the context of sexual coercion/power hierarchy, with male perpetrators being evaluated less negatively or penalized less (e.g., deserving less punishment, judged to be less exploitative) than female perpetrators, and male victims being evaluated more negatively (e.g., condemned more, more damaged reputation) than female victims.

Yet, an alternative hypothesis is also possible on the basis of the societal norm that men need to protect women, because

women are more vulnerable. It has been argued that the traditional male gender role therefore also encompasses chivalry norms (Eagly, 1987). In the context of sexual coercion in power/age hierarchies, male perpetrators violate the chivalry aspect of their gender role. Therefore, people might penalize male perpetrators more than female perpetrators. This expectation is consistent with research on violence in general, showing that people evaluated violence from a man to a woman more negatively than violence from a woman to a man (e.g., Felson & Feld, 2009).

Cross-cultural differences and changes over time. Evolutionary theories, and specifically the perspective of obligate sex differences (i.e., persistent and relatively uniform psychological sex differences across cultures due to hormonal or genetic effects), would predict that there are no differences in SDS between countries. If these double standards evolved from adaptive gender differences in reproductive strategies, they would be universal and should be visible in all countries (Schmitt, 2015). However, according to the emergently moderated perspective cross-cultural variation in sex differences is the result of moderating factors in the local ecology, like religion and gender equality. Level of gender equality is particularly relevant in the context of SDS. Yet, the emergently moderated perspective does not yield a testable hypothesis about whether increasing levels of gender equality would suppress or accentuate gender differences in the norms for sexual behavior. Therefore, we will only test the prediction from the obligate sex difference perspective.

With regard to cross-cultural differences, biosocial theory would argue that gender roles are a product of culture (Wood & Eagly, 2002, 2012) and thus SDS might differ between countries. To quantify cross-cultural differences in gender equality, two measures have been developed that assess the level of gender equality in countries across the world: the gender inequality index (United Nations Development Program, 2017) and the global gender gap score (World Economic Forum, 2017). Data from these measures showed that Scandinavian and Western European countries generally have the smallest gender gap in the world and that North American countries have a somewhat bigger gender gap. Latin-American and Asian societies have intermediate levels of gender inequality. The largest gender inequality can be found in Middle East and North African societies. Biosocial theory would predict that lower gender equality scores of countries on these measures are associated with more traditional SDS.

Regarding changes over time, evolutionary theory would not predict changes in SDS over the last 60 years, because evolutionary changes are generally slow. However, biosocial theory would predict that SDS would be less traditional in recent studies compared with older studies. In recent decades, the division of gender roles has become less strict in most modern Western societies (Eagly & Wood, 1999), which according to biosocial theory would lead to less

differentiation in the norms for the sexual behavior of men and women (Wood & Eagly, 2002, 2012). Moreover, gender equality has increased in most Western societies over the decades (Inglehart & Norris, 2003).

Age differences in SDS. Neither evolutionary theory nor biosocial theory makes direct predictions with regard to age differences in the existence of SDS. Yet, pressures to conform to gender roles increase with child age and might be highest in adolescence (Basow & Rubin, 1999). According to the gender-intensification hypothesis, this might be because adolescence is a period in which boys and girls become increasingly different as a result of the convergence of biological, social, and cognitive changes (Hill & Lynch, 1983). Also, major developments in sexuality take place in adolescence (DeLamater & Friedrich, 2002), making sexuality a highly salient issue on which adolescents evaluate each other (Kreager et al., 2016). Therefore, we expected SDS to be more prevalent in adolescent samples than in adult samples.

SDS: Previous Findings

Reiss (1964) conducted the first systematic study of SDS in the 1960s, indicating that more sexual permissiveness was granted to men than to women. Since then, dozens of studies have been published on SDS, albeit with inconsistent results. Several studies did not find clear evidence of SDS (e.g., Gentry, 1998; O'Sullivan, 1995), whereas others clearly demonstrated the existence of traditional SDS (e.g., Marks, 2008; Marks & Fraley, 2007). Some recent studies even found evidence for a reversed double standard (e.g., Howell et al., 2011; Zaikman, Vogel, et al., 2016), in which women were evaluated more positively than men for high sexual activity (Milhausen & Herold, 1999).

Several narrative reviews tried to summarize the inconsistent body of research (Bordini & Sperb, 2013; Crawford & Popp, 2003; Fugère et al., 2008). Crawford and Popp (2003) concluded that traditional double standards for some sexual behaviors still exist, for example, for initiating sex, casual sex, sex at an early age, and having many sexual partners, but that for other sexual behaviors a double standard is no longer present, for example, for sex before marriage. Fugère et al. (2008) concluded that in some studies men held more traditional SDS than women and that SDS might be more traditional in non-U.S. samples (Russian, Japanese) compared with U.S. samples. Similar to Crawford and Popp (2003), Bordini and Sperb (2013) concluded that premarital sex and casual sex are accepted for both men and women in Western cultures, whereas a double standard still exists for other sexual behaviors, such as being highly sexually active or having a high number of sexual partners. From these reviews, we can conclude that the following moderators appear to play a role in the existence of SDS: sexual behavior type, gender, and cultural background.

In addition, two meta-analyses examined gender differences in sexuality, encompassing 30 specific sexual behaviors and attitudes (Oliver & Hyde, 1993; Petersen & Hyde, 2010). These meta-analyses also assessed a *gender difference* in SDS attitudes, but did not report the overall existence of SDS across men and women. Men reported more traditional SDS than women in the meta-analysis on studies conducted between 1993 and 2007 (Petersen & Hyde, 2010), whereas women reported more traditional SDS than men in the meta-analysis on studies conducted between 1974 and 1993 (Oliver & Hyde, 1993). Both meta-analyses each only included seven studies about the SDS, which might be because the search terms were not specific enough for SDS. Moreover, a small number of included studies assessing SDS precluded robust examination of moderators.

Finally, recently Zaikman and Marks (2017) conducted a theory-based narrative review on SDS, describing evidence for hypotheses based on evolutionary theory, biosocial theory, and cognitive social learning theory. First, they did not identify any evidence for evolutionary theory's proposition that SDS would be most evident for sexual behaviors that could lead to reproduction. This hypothesis could not be studied in the current meta-analysis, because too few studies examined SDS in sexual behaviors that cannot lead to reproduction, such as petting or kissing. Second, they presented evidence for the prediction of biosocial theory that SDS are more evident when there are power differences between men and women. Third, they identified evidence for SDS being more prevalent in cultures characterized by higher levels of gender inequality and for SDS becoming more egalitarian over time. These findings were in line with biosocial theory, but not with evolutionary theory. Regarding predictions of cognitive social learning theory, they identified evidence for the role of traditional gender-role socialization and a high level of sexual experience in the existence of SDS.

Conceptualization and Measurement of SDS

Inconsistencies in previous research on SDS could be due to differences in conceptualization, measurement, and study design.

Conceptualization. Most previous research has conceptualized SDS in terms of attitudes: people's differential evaluation of sexual behaviors of men versus women. Yet, research using this conceptualization has provided inconsistent evidence of SDS. Therefore, Milhausen and Herold (2001) proposed a reconceptualization of SDS. They distinguished between people's personal acceptance of SDS (e.g., "I would think badly of a man/woman who had protected sexual intercourse with a woman/man he or she was not emotionally committed to") and people's knowledge of the existence of SDS in society (e.g., "Who do you think has more sexual freedom today?"). Using this reconceptualization, they found that most people still believe SDS exist at

a societal level, but on a personal level most people held egalitarian standards. The distinction between personal attitudes and more generally shared social expectations is similar to the common distinction in social psychology between knowledge of cultural stereotypes (i.e., socially shared set of expectations about a certain social group) and personal attitudes (i.e., negative or positive evaluation of certain social group or behavior of this group) (Greenwald et al., 2002).

Because SDS are grounded in negative evaluations of those who behave in ways that violate social sexual standards as well as stereotyped beliefs about gender (Lai & Hynie, 2011), in the current meta-analysis we assessed SDS in terms of individuals' attitudes, that is, their personal evaluation of sexual behavior of men versus women, and in terms of their stereotypes, that is, their expectations about the sexual behavior of men and women. Although, as explained below, we were not able to distinguish between personal stereotypical beliefs and knowledge of cultural stereotypes in our analysis, we nonetheless expected that the broad distinction between attitudes and stereotypes might be important. Consider, for example, that findings indicate that in children, as well as adults, content of gender stereotypes has not changed over time, whereas gender attitudes have become more egalitarian (Ruble, 1983; Signorella et al., 1993). Thus, studies conceptualizing SDS in terms of stereotypes (e.g., participants' responses to questions such as "Who has more sexual freedom?") might be more likely to yield evidence for a double standard, compared with studies using the attitude conceptualization (e.g., participants' differential evaluation of either a male or female target in a friends-with-[sexual] benefits scenario) or studies conceptualizing SDS as a combination of stereotypes and attitudes (e.g., aggregating participants' agreement with attitude statements such as "It is worse for a woman to sleep around than it is for a man" with stereotype statements such as "It is expected that a woman is less sexually experienced than her partner.").

Measurement type and study design. Dual-process models of cognition propose that social cognitions can be present both at an explicit and an implicit level (Gawronski & Creighton, 2013). Explicit cognitions are overtly expressed ideas that are under conscious control and, therefore, are especially prone to social-desirable responding (Greenwald et al., 2009). Self-report questionnaires of stereotypes and attitudes tap into explicit cognitions. Implicit cognitions, on the other hand, are supposedly relatively inaccessible to conscious awareness, are elicited unintentionally, require little amounts of cognitive resources, and cannot be stopped voluntarily (Gawronski & Bodenhausen, 2006). They are most often assessed with implicit association tests (IATs) that measure the strength of automatic associations between concepts (e.g., men, women, Black people, gay people) and attributes (e.g., career, family, angry, good, bad). An important implication of dual-process models for social cognitions about

controversial subjects, such as gender, is that people might report more egalitarian cognitions on explicit self-report measures than are suggested by their responses on implicit measures. The explicit format enables them to better control the expression of potentially socially undesirable cognitions (Gawronski & Bodenhausen, 2006). Research has indeed shown that implicitly assessed gender stereotypes were more traditional than explicitly assessed gender stereotypes (Endendijk et al., 2013).

In the SDS literature, studies often used self-report questionnaires resulting in a composite score indicating explicit SDS-cognitions (e.g., Caron et al., 1993). Questionnaires differ, however, in whether they assess participants' agreement with statements that are indicative of a traditional double standard (e.g., "It is up to the man to initiate sex."; double standard scale [DSS] of Caron et al., 1993), or participants' evaluation of parallel items about men's and women's sexual behavior (e.g., "A [girl/boy] who has sex on the first date is easy"; SDS scale of Muehlenhard & Quackenbush, 1998; premarital sexual permissiveness scale by Reiss, 1964). Implicit measures have hardly been used to assess SDS (see, for example, Marks, 2008; Sakaluk & Milhausen, 2012), and findings regarding differences in implicit and explicit SDS have been inconsistent. For example, Marks (2008) found evidence of double standards for person evaluation under divided attention (i.e., implicit condition), but not for person evaluation under full attention (i.e., explicit condition). In contrast, Sakaluk and Milhausen (2012) found that men held more traditional SDS than women on an explicit self-report questionnaire, but men held egalitarian standards on an IAT, whereas women demonstrated reversed double standards on the IAT.

Another measurement type that is often used, and could be considered as a relatively implicit measure of SDS, is an experimental task in which participants have to evaluate a vignette or scenario that describes the sexual behavior of a hypothetical male and/or female (e.g., Barron, 2010). Such tasks are considered less explicit methodologies for measuring SDS (Jonason & Marks, 2009; Reid et al., 2011; Weaver et al., 2013). For instance, in a between-subject design, researchers randomly assign vignettes/scenarios to participants, who are generally unaware of the presence of other vignettes presented to other participants. Or in a within-subject design, researchers administer vignettes/scenarios in a counter-balanced way to participants. Such procedures make it less explicit that the participant's SDS-cognitions are assessed than with self-report questionnaires. Because of the less explicit nature of the assessment, social desirability may play a less important role in experimental designs using vignettes than in studies using self-report questionnaires (Greenwald et al., 2009).

In addition, between-subjects designs are likely more implicit in nature than within-subjects designs. In between-subjects designs, participants evaluate the sexual behavior of either a man or a woman, whereas in within-subjects designs,

participants evaluate the sexual behavior of both genders, which likely makes the focus on gender differences more explicit. In previous studies, scholars also suspected that social desirability and demand characteristics (whereby participants form an interpretation of the experiment's purpose and unconsciously change their responses to fit that interpretation) play a larger role in within-subject research on SDS than in between-subject research (Marks & Fraley, 2005; Milhausen & Herold, 2001). Thus, evidence of SDS is less likely to be found in studies using relatively explicit measures or designs, than in studies using relatively implicit measures or designs. More specifically, studies using Likert-type-scale questionnaires are less likely to yield evidence for SDS than studies assessing the differential evaluation of men and women engaging in the same sexual behavior, or studies using IATs or similar reaction-time measures. Similarly, between-subjects designs are more likely to yield evidence for SDS than within-subjects designs.

Current Study

The current meta-analysis of SDS tested the following grand hypotheses based on evolutionary theory (Buss & Schmitt, 1993; Trivers, 1972) and biosocial theory (Wood & Eagly, 2002, 2012): (a) people expect behaviors associated with high sexual activity more from men than from women, and behaviors associated with low sexual activity more from women than from men; (b) people evaluate highly sexually active men more positively (or less negatively) than highly sexually active women, and low sexually active women more positively (or less negatively) than low sexually active men.

In addition, we tested several hypotheses regarding specific moderators. Both evolutionary theory and biosocial theory propose that SDS are behavior specific, with evolutionary theory predicting SDS to be most evident for having frequent casual sex with many partners, having an early sexual debut, and sexual infidelity, whereas biosocial theory predicts SDS to be most evident for sexual coercion in a power/age hierarchy. In addition, biosocial theory predicts that there are differences in SDS between countries and changes over time, but evolutionary theory does not. The hypotheses that were based on evolutionary theory and biosocial theory are similar to the hypotheses presented in a recent narrative review of the theories relevant to the study of SDS (Zaikman & Marks, 2017).

Regarding the demographic moderator age, the gender-intensification hypothesis predicts that studies with adolescent samples are more likely to yield evidence for SDS than studies with adult samples. In terms of gender differences, male control theory proposes that men would be more likely than women to hold traditional SDS-cognitions. With regard to measurement moderators of SDS, dual-process models of social cognition predict that studies using relatively explicit measures or designs (e.g., self-reports, within-subjects

designs) yield less evidence for SDS than studies using relatively implicit measures or designs (e.g., IATs, vignettes, between-subjects designs). Last, studies conceptualizing SDS as stereotypes may be more likely to yield evidence for SDS than studies conceptualizing SDS as personal attitudes, or as a combination of stereotypes and personal attitudes.

The current meta-analysis extends previous narrative and meta-analytic reviews in the following ways: (a) examination of the presence and strength of SDS from 1960 until 2019; (b) examination of the effect of moderators related to sample, sexual behavior type, measurement and design, and publication year; (c) a theory-based meta-analytic approach. Such a meta-analysis is currently lacking, but is essential for better and more rigorous experimental design of future SDS research and has the potential to enhance understanding of etiology and functioning of SDS (Zaikman & Marks, 2017).

Method

Literature Search

We used the PRISMA guidelines for conducting and reporting the current meta-analysis (Moher et al., 2009; see Supplemental Appendix A). The current meta-analysis was preregistered (see Supplemental Appendix B). Via three search methods, we identified eligible studies until March 1, 2019. First, we searched the electronic databases of Scopus, ERIC, PsycINFO, Online Contents, and PiCarta for empirical, peer-reviewed articles using the following search terms: ((premarital AND sex* AND standard*) OR (sex* AND double AND standard*) OR (sex* AND permissive* AND gender AND attitude*)). These search terms are similar to the search terms used in previous systematic narrative reviews on SDS (Bordini & Sperb, 2013; Crawford & Popp, 2003). We checked whether the search terms yielded all articles included in previous narrative reviews on SDS. This was indeed the case. Second, we searched the reference lists of relevant narrative reviews on SDS and meta-analyses on gender differences in sexuality. Third, we searched the reference lists of the articles and dissertations that met our inclusion criteria for eligible studies. We applied a very broad strategy with this reference search, including all articles that mentioned any of our search terms in the title terms. The database search and reference list search together yielded 1,364 hits. Figure 1 depicts the flowchart of the literature search.

Inclusion Criteria

To be included in the meta-analysis, studies had to assess participants' cognitions about SDS in terms of attitudes (one's personal evaluation of sexual behavior of men versus women) and/or stereotypes (a socially shared set of expectations about the sexual behavior of men and women). Examples of personal attitude conceptualizations of SDS

are (a) participants' differential evaluation of either a male or female target in a friends-with-(sexual)benefits scenario; or (b) participants' differential agreement with statements about men and women's sexual behavior ("A [girl/guy] who has sex on the first date is 'easy'"). Examples of stereotype conceptualizations of SDS are (a) participants' differential expectations about the percentage of males and the percentage of females who have had sex by the age of 18; (b) participants' level of agreement with statements such as "A girl is usually looking for a man to marry, but a boy is usually looking for sex" or "Women who are sexually experienced with multiple partners are usually not respected as much as men who are sexually experienced"; and (c) participants' responses to questions such as "Who generally has more sexual freedom?" Studies examining differences in the sexual *behavior* of men and women were, thus, not included. Importantly, half of the studies assessing stereotype aspects of SDS used a conceptualization of stereotypes as personal beliefs (e.g., personal expectations about the percentage of males and females engaging in a certain sexual behavior). The other half of the studies conceptualized SDS stereotypes as collective beliefs (e.g., "Who generally has more sexual freedom?"). There were too few studies examining stereotypes to investigate these two stereotype aspects as separate moderator categories. Therefore, they were grouped together.

Following recommendations by Crawford and Popp (2003), we excluded studies from this meta-analysis when they employed a design that confounded participant gender and target gender (e.g., participants only evaluating the sexual behavior of same- or opposite-gender targets). Such a design makes it impossible to draw conclusions about the existence of SDS. We did not set any restrictions with regard to the language of the paper, as long as an English abstract was available for screening purposes. During the full-text screening phase, papers written in languages other than English (eight Spanish, two Portuguese) were translated by native Spanish- or Portuguese speakers. Of the included publications, two were published in Spanish and one in Portuguese.

When a study lacked sufficient information to be included ($k = 11$), we contacted the 11 corresponding authors by email with a request for additional information (e.g., sample information, information about measures, effect size). However, none of these studies could be included, because the contacted authors did not have access to the data anymore, or the authors never responded.

We determined level of agreement between the first author and a research assistant on the inclusion of studies on a random subset of 100 studies, oversampling included studies. Agreement between the coders was satisfactory (agreement 94%, kappa = .84). In case of disagreements between the coders, the coders discussed the disagreement until they reached consensus. After the reliability assessment, the first author screened the remainder of the articles.

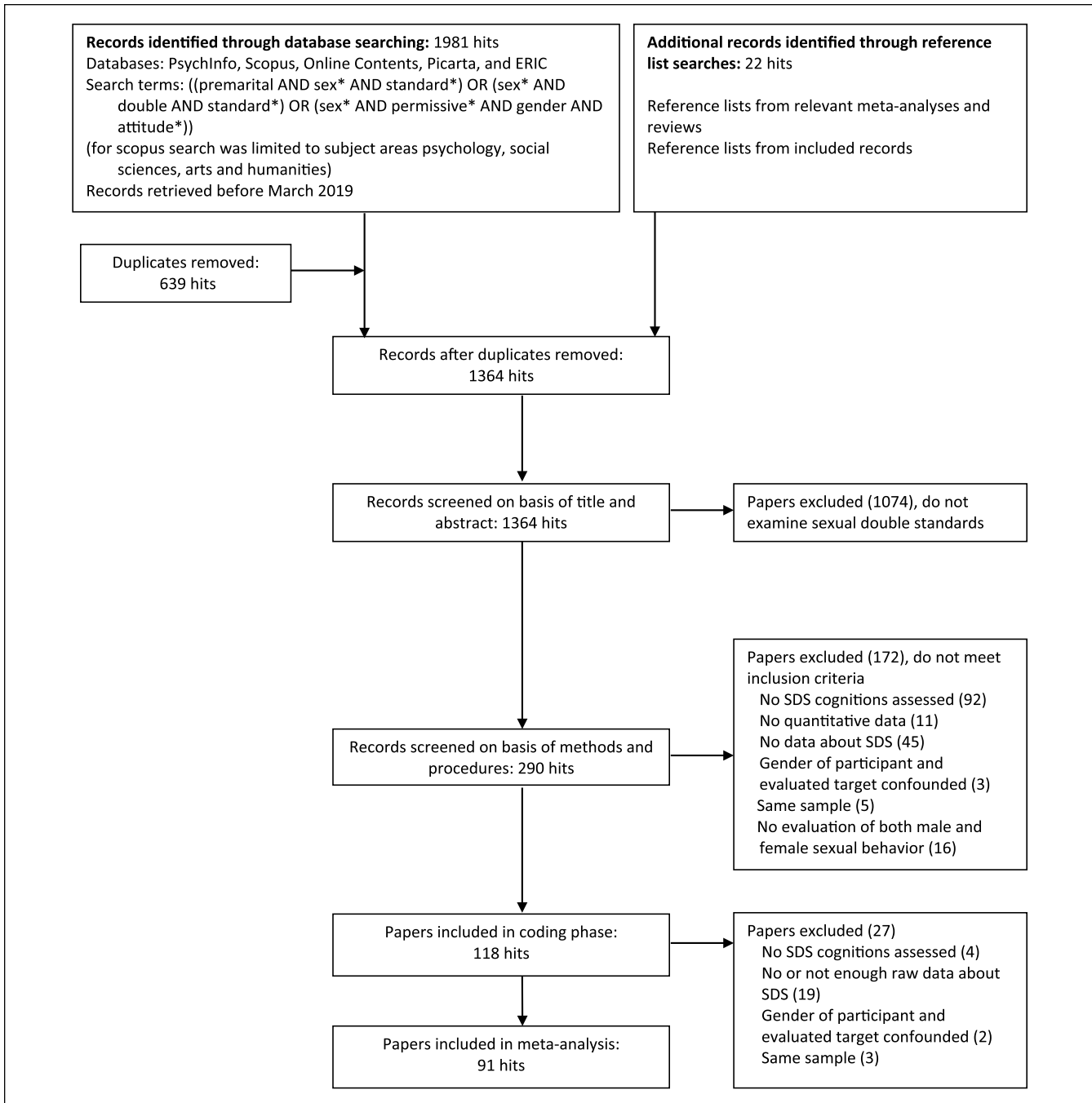


Figure 1. Flow chart of literature search process.

The studies that we included in the meta-analyses are presented in Tables 1 and 2.

Data Extraction

We coded the following moderators (see Supplemental Appendix C for further explanations of the categories): regarding measurement and design moderators, we coded SDS conceptualization (stereotype, attitude, combination),

the measurement type (questionnaire [i.e., explicit], vignettes/scenarios [i.e., relatively implicit], other), questionnaire type (DSS of Caron et al., 1993; SDS scale of Muehlenhard & Quackenbush, 1998; premarital sexual permissiveness scale by Reiss, 1964; other), design of study (cross-sectional, within-subjects design, between-subjects design), and sexual behavior type assessed (being a perpetrator of sex in power/age hierarchy, being a victim of sex in power/age hierarchy, casual premarital sex, infidelity/affair,

Table 1. Studies Included in Meta-Analysis on Differential Evaluation/Expectation of Men's and Women's Sexual Behavior.

Study	Sample	N	Age	Education	Gender equality	country	Ethnicity	Measure	Questionnaire	Conceptualization	Sexual behavior	Year	Publication	Design
Allgeier & Fogel (1978)	♂	60	19–25	H	0.7645			V		AT	MO	1978	J	BS
	♀	59	19–25	H	0.7645			V		AT	MO	1978	J	BS
Axinn et al. (2011)	♂♀	904	31	M	0.7645			Q	0	AT	ESD	2011	J	WS
Barron (2010)	♂	150	21	H	0.7645		W	V		AT	CP	2011	D	BS
	♀	150	21	H	0.7645		W	V		AT	CP	2011	D	BS
Bauman & Wilson (1976)	♂68	98		H	0.7645		W	Q	5	AT	CS, PSLA	1968	J	WS
	♀68	88		H	0.7645		W	Q	5	AT	CS, PSLA	1968	J	WS
	♂72	107		H	0.7645		W	Q	5	AT	CS, PSLA	1972	J	WS
	♀72	68		H	0.7645		W	Q	5	AT	CS, PSLA	1972	J	WS
Boyer & Galupo (2015)	♂♀	348	20		0.7645		M	V		AT	CS	2015	J	BS
Castaneda & Collins (1998)	♂US	35	22	H	0.7645		W	V		AT	MO	1998	J	BS
	♀US	71	22	H	0.7645		W	V		AT	MO	1998	J	BS
	♂Mex	56	21	H	0.7645		LH	V		AT	MO	1998	J	BS
Conley et al. (2013)	♀Mex	74	21	H	0.7645		LH	V		AT	MO	1998	J	BS
	1a:♂♀	195	22		0.7645		M	V		S	CS	2013	J	BS
	1b:♂♀	87	24		0.7645		M	V		S	CS accept	2013	J	BS
		87	24		0.7645		M	V		S	CS refuse	2013	J	BS
	2a:♂♀	1029	22		0.7645		W	V		S	CS accept	2013	J	BS
	2b:♂♀	1030	22		0.7645		W	V		S	CS refuse	2013	J	BS
Do & Fu (2010)	♂	6,643		M	0.697		EA	Q	0	AT	PSU	2010	J	WS
	♀	6,973		M	0.697		EA	Q	0	AT	PSU	2010	J	WS
Dollar et al. (2004)	♂♀	240	19–25	H	0.7645		W	V		AT	CP	2004	J	BS
Donald (1983)	♂	100	19–25	H	0.7645		W	V		AT	MO	1983	D	WS
	♀	100	19–25	H	0.7645		W	V		AT	MO	1983	D	WS
Eisenman & Dantzer (2006)	♂	128	19–25		0.7645		LH	Q	0	AT	MO	2006	J	WS
	♀	199	19–25		0.7645		LH	Q	0	AT	MO	2006	J	WS
Ford et al. (2003)	♂	597	>25	M	0.6385		LH	Q	0	AT	CS, IA, PSU	2003	J	WS
Garcia (1982)	♂♀	80	19–25		0.7645			V		AT	MO	1982	J	BS
	♂♀	79	19–25		0.7645			V		AT	MO	1982	J	BS
Guo (2019)	♂W	182	33	M	0.7645		W	Q	0	AT	CS, ACT, PSU	2019	J	WS
	♀W	152	34	M	0.7645		W	Q	0	AT	CS, ACT, PSU	2019	J	WS
	♂A	103	30	M	0.7645		O	Q	0	AT	CS, ACT, PSU	2019	J	WS
	♀A	69	31	M	0.7645		O	Q	0	AT	CS, ACT, PSU	2019	J	WS
Haavio-Mannila & Kontula (2003)	♂F71	1081	>25		0.8825			Q	0	AT	IA	1971	J	WS
	♀F71	1044	>25		0.8825			Q	0	AT	IA	1971	J	WS
	♂F92	1100	>25		0.8825			Q	0	AT	IA	1992	J	WS
	♀F92	1142	>25		0.8825			Q	0	AT	IA	1992	J	WS
	♂F99	693	>25		0.8825			Q	0	AT	IA	1999	J	WS
	♀F99	721	>25		0.8825			Q	0	AT	IA	1999	J	WS
	♂Est1	297	>25		0.8045			Q	0	AT	IA	2000	J	WS
	♀Est1	338	>25		0.8045			Q	0	AT	IA	2000	J	WS
	♂Est2	147	>25		0.8045			Q	0	AT	IA	2000	J	WS
	♀Est2	196	>25		0.8045			Q	0	AT	IA	2000	J	WS
Hackathorn & Harvey (2011)	♂R	795	>25		0.7195			Q	0	AT	IA	1996	J	WS
	♀R	1138	>25		0.7195			Q	0	AT	IA	1996	J	WS
Jonason & Marks (2009)	♂	53	19	H	0.7645		M	V		AT	IA	2011	J	BS
	♀	62	19	H	0.7645		M	V		AT	IA	2011	J	BS
Kaats & Davis (1970)	1:♂♀	120	20	H	0.7645			V		AT	MO	2009	J	BS
	2:♂♀	105	21	H	0.7645			V		AT	MO	2009	J	BS
Kleinfelter (1999)	1:♂	155	20	H	0.7645			Q	5	AT	PSU	1970	J	WS
	1:♀	222	20	H	0.7645			Q	5	AT	PSU	1970	J	WS
	2:♂	84	20	H	0.7645			Q	5	AT	PSU	1970	J	WS
	2:♀	97	20	H	0.7645			Q	5	AT	PSU	1970	J	WS
Koon-Magnin & Ruback (2012)	♂	69	20	H	0.7645		M	V		S, AT	CS, ACT, ESD, MO	1999	D	WS
	♀	150	20	H	0.7645		M	V		S, AT	CS, ACT, ESD, MO	1999	D	WS
Manalastas & David (2018)	♂♀	485	19	H	0.7645		W	V		AT	CP	2012	J	BS
	♂	3,467	22	M	0.6815		O	Q	0	AT	PSU	2018	J	WS
Marks & Fraley (2005)	♀	3,540	22	M	0.6815		O	Q	0	AT	PSU	2018	J	WS
	1:♂♀	8,080	23		0.7645		M	V		AT	ACT	2005	J	BS
Marks & Fraley (2007)	2:♂♀	144	19	H	0.7645		M	V		AT	ACT	2005	J	BS
	♂♀	468	19	H	0.7645		M	V		AT	ACT	2007	J	BS
Marks (2008)	1:♂♀	12	19	H	0.7645		M	V		AT	ACT (17p)	2008	J	BS
	2:♂♀	12	19	H	0.7645		M	V		AT	ACT (1p)	2008	J	BS
	3:♂♀	12	19	H	0.7645		M	O		AT	ACT (17p)	2008	J	BS
	4:♂♀	12	19	H	0.7645		M	O		AT	ACT (1p)	2008	J	BS
Marks et al. (2019)	♂♀	4,455	31		0.7645		O			AT	ACT	2018	J	BS
Miller (2012)	♂♀	1186	20	H	0.7645		W	Q	2	AT	MO	2012	D	WS
Murstein & Mercy (1994)	♂	98	20	H	0.7645		W	Q	0	AT	CS	1994	J	WS
	♀	148	20	H	0.7645		W	Q	0	AT	CS	1994	J	WS
Papp et al. (2015)	♂	123	20	H	0.7645		W	V		AT	MO	2015	J	BS
	♀	185	20	H	0.7645		W	V		AT	MO	2015	J	BS

(continued)

Table 1. (continued)

Study	Sample	N	Age	Education	Gender equality	country	Ethnicity	Measure	Questionnaire	Conceptualization	Sexual behavior	Year	Publication	Design
Reiss (1964)	B12-18	143	12-18		0.7645		B	Q	5	AT	CS, PSLA, PSE	1964	J	WS
	W12-18	56	12-18		0.7645		W	Q	5	AT	CS, PSLA, PSE	1964	J	WS
	B19-25	444	19-25	H	0.7645		B	Q	5	AT	CS, PSLA, PSE	1964	J	WS
	W19-25	160	19-25	H	0.7645		W	Q	5	AT	CS, PSLA, PSE	1964	J	WS
	B>25	1,251	>25		0.7645		B	Q	5	AT	CS, PSLA, PSE	1964	J	WS
	W>25	139	>25		0.7645		W	Q	5	AT	CS, PSLA, PSE	1964	J	WS
Robinson & Jedlicka (1982)	♂65	129	19-25	H	0.7645			Q	O	AT	ACT	1965	J	WS
	♀65	115	19-25	H	0.7645			Q	O	AT	ACT	1965	J	WS
	♂70	137	19-25	H	0.7645			Q	O	AT	ACT	1970	J	WS
	♀70	158	19-25	H	0.7645			Q	O	AT	ACT	1970	J	WS
	♂75	138	19-25	H	0.7645			Q	O	AT	ACT	1975	J	WS
	♀75	298	19-25	H	0.7645			Q	O	AT	ACT	1975	J	WS
	♂80	169	19-25	H	0.7645			Q	O	AT	ACT	1980	J	WS
	♀80	230	19-25	H	0.7645			Q	O	AT	ACT	1980	J	WS
Sahl & Keene (2010)	♂♀	2871	19-25	H	0.7645		M	V		AT	CP	2010	J	BS
Sheeran et al. (1996)	♂	277	16		0.8270			Q	O	S, AT	ACT	1996	J	WS
	♀	405	16		0.8270			Q	O	S, AT	ACT	1996	J	WS
Smith et al. (2008)	♂♀	151	18	H	0.7645			V		AT	MO	2008	J	BS
Soller & Haynie (2017)	♂♀	8,458	16	M	0.7645		M	Q	O	AT	ESD	2017	J	WS
Spreadbury (1982)	♂	38	19-25	H	0.7645		W	Q	O	AT	CS, PSLA, PSE	1982	J	WS
	♀	129	19-25	H	0.7645		W	Q	O	AT	CS, PSLA, PSE	1982	J	WS
Sprecher et al. (1987)	♂♀	553	21	H	0.7645			V		AT	ESD	1987	J	BS
Sprecher (1989)	♂♀	666	19	H	0.7645			V		AT	MO	1989	J	BS
Thompson et al. (2018)	♂♀	793	27		0.7645		W	V		AT	MO	2018	J	BS
Ward & Rivadeneyra (1999)	♂♀	314	19	H	0.7645		M	Q	O	S	ACT	1999	J	WS
Ward (2002)	♂	97	20	H	0.7645		M	Q	O	S	ACT	2002	J	WS
	♀	172	20	H	0.7645		M	Q	O	S	ACT	2002	J	WS
Weaver et al. (2013)	♂♀	404	19	H	0.8385			V		S, AT	CS	2013	J	BS
Young et al. (2010)	♂	206	19-25	H	0.7645		W	V		AT	ACT	2010	J	BS
	♀	294	19-25	H	0.7645		W	V		AT	ACT	2010	J	BS
Young et al. (2016)	1:♂	44	12-18		0.7645		M	V		AT	ACT	2016	J	BS
	1:♀	44	12-18		0.7645		M	V		AT	ACT	2016	J	BS
	2:♂	44	12-18		0.7645		M	V		AT	ESD	2016	J	BS
	2:♀	43	12-18		0.7645		M	V		AT	ESD	2016	J	BS
Zaikman & Marks (2014)	1:♂	28	19	H	0.7645		M	V		AT	ACT low	2014	J	BS
	1:♀	54	19	H	0.7645		M	V		AT	ACT low	2014	J	BS
	2:♂	20	19	H	0.7645		M	V		AT	ACT high	2014	J	BS
	2:♀	46	19	H	0.7645		M	V		AT	ACT high	2014	J	BS
Zaikman, Marks, et al. (2016)	1:♂	58	33		0.7645		M	V		AT	ACT low	2016	J	BS
	1:♀	57	33		0.7645		M	V		AT	ACT low	2016	J	BS
	2:♂	65	33		0.7645		M	V		AT	ACT high	2016	J	BS
	2:♀	60	33		0.7645		M	V		AT	ACT high	2016	J	BS

Note. Numbers under questionnaire refer to double standard scale of Caron (1), sexual double standard scale of Muehlenhard (2), personal acceptance of double standard scale by Milhausen (3), scale for the assessment of sexual double standards in youth by Emmerink (4), premarital sexual standards scale by Reiss (5). H = high educational level; M = mixed; O = other; MO = mixed/other; A = Asia(n); F = Finnish; R = Russian; Est = Estonian; Mex = Mexican; US = United States; W = White/Caucasian; B = Black; LH = Latino/a/Hispanic; H = Hispanic; EA = East-Asian; Q = questionnaire; V = vignettes/scenarios; S = stereotypes; AT = attitudes; PSLA = premarital sex in love/with affection; PSE = premarital sex engaged; PSU = premarital sex unspecified; CS = casual sex; IA = infidelity/affair; CP = coercion/sex in power hierarchy; ACT = sexual activity; ESD = early sexual debut; J = journal publication; D = dissertation; C = correlational; L = longitudinal; WS = within-subjects design; BS = between-subjects design.

Table 2. Studies Included in Meta-Analysis on SDS-Cognitions Assessed With Likert-type-Scale Questionnaires.

Study	Sample	N	Age	Education	Gender equality	country	Ethnicity	Measure	Questionnaire	Conceptualization	Sexual behavior	Year	Publication	Design
Ali-Faisal (2014)	♂♀	403	25	H	0.8015		M	Q	1	MO	MO	2014	D	C
Alves et al. (2008)	♂	140	21	H	0.8230		LH	Q	2, 3	S, MO	MO	2008	J	C
	♀	168	21	H	0.8230		LH	Q	2, 3	S, MO	MO	2008	J	C
Askun & Ataca (2007)	♂♀	540	21	H	0.6540			Q	1	MO	MO	2007	J	C
Balanko (2002)	♀	21	20	H	0.8385		M	Q	2	MO	MO	2002	D	O
Bareket et al. (2018)	♂	108	28		0.8115		O	Q	O	AT	CS	2018	J	WS
Bay-Cheng & Zucker (2007)	♀	342	19	H	0.7645		W	Q	2	MO	MO	2007	J	C
Bliss (2013)	♂	71	19	H	0.7645		W	Q	1	MO	MO	2013	D	C
	♀	82	19	H	0.7645		W	Q	1	MO	MO	2013	D	C
Caron et al. (1993)	♂	131	18	H	0.7645			Q	1	MO	MO	1993	J	C
	♀	199	18	H	0.7645			Q	1	MO	MO	1993	J	C
Clarke et al. (2015)	♂	140	29	H	0.8110			Q	2	MO	MO	2015	J	C
Danube et al. (2016)	♀	360	25		0.7645		M	Q	1	MO	MO	2016	J	O
Diéguez et al. (2003)	♂	2,246	20	H	0.8330		LH	Q	2	MO	MO	2003	J	C
	♀	3,368	20	H	0.8330		LH	Q	2	MO	MO	2003	J	C
Doan (2006)	♀	87	30	H	0.7645		M	Q	2	MO	MO	2006	D	C

(continued)

Table 2. (continued)

Study	Sample	N	Age	Education	Gender equality	country	Ethnicity	Measure	Questionnaire	Conceptualization	Sexual behavior	Year	Publication	Design
Eaton & Matamala (2014)	♂	263	20	H	0.7645		M	Q	2	MO	MO	2014	J	C
	♀	292	20	H	0.7645		M	Q	2	MO	MO	2014	J	C
Emmerink, Vanwesenbeeck, et al. (2016)	♂	210	18	M	0.8465		M	Q	4	AT	MO	2016	J	C
	♀	255	18	M	0.8465		M	Q	4	AT	MO	2016	J	C
Emmerink, Van den Eijnden, et al. (2016)	♂	126	21	M	0.8465		W	Q	4	AT	MO	2016	J	C
	♀	267	21	M	0.8465		W	Q	4	AT	MO	2016	J	C
Emmerink et al. (2017)	♂	340	20	M	0.8465		W	Q	4	AT	MO	2017	J	L
	♀	478	20	M	0.8465		W	Q	4	AT	MO	2017	J	L
Eriksson & Humphreys (2014)	♂	39	19	H	0.8385			Q	1	MO	MO	2014	J	C
	♀	184	19	H	0.8385			Q	1	MO	MO	2014	J	C
Glass (1972)	♂	69	12–18		0.7645			Q	5	AT	MO	1972	J	WS
	♀	211	12–18		0.7645			Q	5	AT	MO	1972	J	WS
Greene & Faulkner (2005)	♂	608	21	M	0.7645		W	Q	1	MO	MO	2005	J	C
	♀	608	21	M	0.7645		W	Q	1	MO	MO	2005	J	C
Hanassab & Tidwell (1993)	♀US	81	23	M	0.7645		A	Q	O	S, AT	PSU, MO	1993	J	C
	♀Iran	193	21		0.5610		A	Q	O	S, AT	PSU, MO	1985	J	C
Hearn (2001)	♂♀	96	21	M	0.7645		B	Q	1	MO	MO	2001	D	C
Hong (2013)	♂♀	165	52	H	0.7645		M	Q	1	MO	MO	2013	D	C
Isarabhakdi (1999)	♂	577	18		0.6505		EA	Q	5	AT	MO	1999	J	WS
	♀	517	18		0.6505		EA	Q	5	AT	MO	1999	J	WS
Javier (2018)	♀	197	19–25	H	0.7645		O	Q	2	MO	MO	2018	D	C
Jozkowski & Ekbia (2015)	♂♀	141	19–25	H	0.7645		M	Q	2	MO	MO	2015	J	O
Kettrey (2016)	♂	3,436	19–25	H	0.7645		M	Q	O	AT	CS	2016	J	WS
	♀	7,641	19–25	H	0.7645		M	Q	O	AT	CS	2016	J	WS
Kistler (2011)	>25	163	>25	M	0.7645		W	Q	2	MO	MO	2011	D	C
	19–25	301	19–25	M	0.7645		W	Q	2	MO	MO	2011	D	C
Lefkowitz et al. (2014)	♂	208	18	H	0.7645		M	Q	2	MO	MO	2014	J	C
	♀	226	18	H	0.7645		M	Q	2	MO	MO	2014	J	C
Levin (2010)	♂	144	19	H	0.7645		M	Q	O	S	MO	2010	D	C
	♀	191	19	H	0.7645		M	Q	O	S	MO	2010	D	C
Milhausen & Herold (2001)	♂	167	21		0.8385			Q	3, O	S, AT	ACT, MO	2001	J	C
	♀	246	21		0.8385			Q	3, O	S, AT	ACT, MO	2001	J	C
Miller (2012)	♂♀	1186	20	H	0.7645		W	Q	2	MO	MO	2012	D	C
Moyano et al. (2017)	♂	448	16		0.6755		LH	Q	1	MO	MO	2017	J	C
Ortiz (2012)	♀	162	20	H	0.7645		W	Q	2	MO	MO	2012	D	C
Ortiz et al. (2016)	♂♀	313	20	H	0.7645			Q	1	MO	MO	2016	J	C
Quinn-Nilas & Kennett (2018)	♀	222	21	H	0.8385		W	Q	1	MO	MO	2018	J	C
Rudman et al. (2013)	♂	153	19	H	0.7645		M	Q	O	S, AT	MO	2013	J	C
	♀	350	19	H	0.7645		M	Q	O	S, AT	MO	2013	J	C
Sakaluk & Milhausen (2012)	♂	15	19	H	0.8385			Q, O	2, O	AT, MO	MO	2012	J	C
	♀	88	19	H	0.8385			Q, O	2, O	AT, MO	MO	2012	J	C
Shapurian & Hojat (1985)	♂Iran	199	22	H	0.5610		A	Q	O	S, AT	PSU, MO	1985	J	C
	♂UK	780	12–18		0.8270		W	Q	O	S, AT	PSU, MO	1985	J	C
	♀UK	761	12–18		0.8270		W	Q	O	S, AT	PSU, MO	1985	J	C
Shin et al. (2011)	♂	3,609	21	H	0.7935		EA	Q	O	AT	MO	2011	J	C
	♀	2180	21	H	0.7935		EA	Q	O	AT	MO	2011	J	C
Sierra et al. (2007)	♂	151	23	H	0.8330			Q	1	MO	MO	2007	J	C
	♀	249	23	H	0.8330			Q	1	MO	MO	2007	J	C
Sierra et al. (2009)	♂	700	22	H	0.6565		LH	Q	1	MO	MO	2009	J	C
Sierra et al. (2017)	♂	402	41	M	0.8330		LH	Q	2	MO	MO	2017	J	C
	♀	402	41	M	0.8330		LH	Q	2	MO	MO	2017	J	C
Sprecher et al. (2013)	♂A	468	19	H	0.7645		EA	Q	O	AT	CS, PSLA	2013	J	WS
	♀A	271	19	H	0.7645		EA	Q	O	AT	CS, PSLA	2013	J	WS
	♂B	39	19	H	0.7645		B	Q	O	AT	CS, PSLA	2013	J	WS
	♀B	66	19	H	0.7645		B	Q	O	AT	CS, PSLA	2013	J	WS
	♂H	126	19	H	0.7645		LH	Q	O	AT	CS, PSLA	2013	J	WS
	♀H	73	19	H	0.7645		LH	Q	O	AT	CS, PSLA	2013	J	WS
	♂W	4,182	19	H	0.7645		W	Q	O	AT	CS, PSLA	2013	J	WS
	♀W	2425	19	H	0.7645		W	Q	O	AT	CS, PSLA	2013	J	WS
	Ubillos et al. (2016)	♂	1341	14	M	0.8330		LH	Q	1	MO	MO	2016	J
Walfield (2018)	♀	1578	14	M	0.8330		LH	Q	1	MO	MO	2016	J	C
	♂	617	35	M	0.7645		W	Q	1	MO	MO	2018	J	C
Zhang et al. (2008)	♀	603	35	M	0.7645		W	Q	1	MO	MO	2018	J	C
	♂	112	18	H	0.7645		W	Q	1	MO	MO	2008	J	C
Zurbriggen & Morgan (2006)	♀	154	18	H	0.7645		W	Q	1	MO	MO	2008	J	C
	♂	109	19	H	0.7645		M	Q	O	AT	MO	2006	J	C
	♀	225	19	H	0.7645		M	Q	O	AT	MO	2006	J	C

Note. Numbers under questionnaire refer to double standard scale of Caron (1), sexual double standard scale of Muehlenhard (2), personal acceptance of double standard scale by Milhausen (3), scale for the assessment of sexual double standards in youth by Emmerink (4), premarital sexual standards scale by Reiss (5). H = high educational level; M = mixed; O = other; MO = mixed/other; A = Asia(n); UK = United Kingdom; US = United States; W = White/Caucasian; B = Black; LH = Latino/a/Hispanic; H = Hispanic; EA = East-Asian; Q = questionnaire; S = stereotypes; AT = attitudes; PSLA = premarital sex in love/with affection; PSU = premarital sex unspecified; CS = casual sex; ACT = sexual activity; J = journal publication; D = dissertation; C = correlational; L = longitudinal; WS = within-subjects design.

level of sexual activity, premarital sex when in love or with affection, premarital sex when engaged, premarital sex unspecified, early sexual debut, other/mixed). Questionnaires, study designs, or sexual behavior types other than the ones mentioned above were too uncommon to form a separate category for moderator analyses.

Regarding sample characteristics, we coded participant's gender, the level of gender equality in the country in which the study was conducted, and participant's mean age at the time of the assessment (categorical; adolescence 12–18 years, college aged/emerging adults 19–25 years, adults > 25 years). Regarding the level of gender equality, we averaged the country's most recent score on two global gender equality measures: the gender inequality index (reverse-coded; United Nations Development Program, 2017) and the global gender gap score (World Economic Forum, 2017). Because these indices only exist for about 10 years, we could not compute the country's level of gender equality at the year each study was conducted. We coded age in categories and not continuously, because more studies provided data on this categorical measure of age than of the continuous measure. We also coded sample size, in order to be able to assign weight to the effect sizes. As the majority of the studies examined White/Caucasian college samples with a primarily heterosexual orientation (see Tables 1 and 2), ethnicity, educational level, and sexual orientation could not be taken into account as moderators. Last, we coded year of publication (continuous).

As a check for coder reliability, the first author and a research assistant each coded the same set with 20 publications. Agreement between the coders was satisfactory for both the moderator and outcome variables (kappas for categorical variables between .71 and 1.00, average .91; and agreement between 80% and 100%, average 93%; intraclass correlations for continuous variables between .97 and 1.00, average .99). The coders resolved any disagreements by discussion. After the reliability assessment, the first author coded the remainder of the articles, but consulted one or more of the other authors in cases of doubt.

Meta-Analytic Procedures

We conducted two separate meta-analyses, one meta-analysis for studies that presented SDS as two different scores for the evaluation/expectation of the sexual behavior of men versus women, and a second meta-analysis for studies that examined SDS with Likert-type-scale questionnaires. The two types of studies could not be combined in one meta-analysis, because the point estimates derived from these studies (i.e., standardized mean difference in the first meta-analysis, versus a mean of one group at one timepoint in the second meta-analysis) cannot be analyzed together (Borenstein et al., 2009).

For the first meta-analysis, we calculated the standardized mean difference (d) for each study. We included data in the following forms in hierarchical order: (a) mean and standard

deviation for participants' expectations/evaluations of the sexual behavior of men and women (or mean difference together with p or t value); (b) correlations between target gender and participants' evaluation/expectation of a targets sexual behavior; (c) p values. We gave a positive sign to effect sizes indicating a difference in evaluation of the sexual behavior of men and women that was in line with traditional SDS (e.g., more positive or less negative evaluation of males compared with females for high sexual activity), and a negative sign to differences that were not in line with traditional SDS (e.g., more positive or less negative evaluation of females compared with males for high sexual activity). According to Cohen (1977), effect sizes of $d = 0.20$ are small, $d = 0.50$ is a medium-sized effect, and $d = 0.80$ is a large effect.

For the second meta-analysis, we used rescaled group means and standard deviations of each study to test whether the combined mean was different from zero in either the positive or negative direction. Negative scores represented reversed SDS (e.g., expecting high sexual activity more for women than for men), scores around zero represented egalitarian sexual standards (e.g., no difference in expectation or evaluation of the sexual activity of men and women), and positive scores represented traditional SDS (e.g., expecting high sexual activity more for men than for women). Because the included studies used many different scales, we rescaled each group mean to the same scale ranging from -1 to $+1$, using min-max normalization which is a form of standardization (Han & Kamber, 2006; see for applications of min-max normalization, Bandura, 2008; van Zanten et al., 2014). The following formulas were used:

$$\bar{x}' = \frac{\bar{x} - \text{Min}}{\text{Max} - \text{Min}} \times (\text{Max}' - \text{Min}') + \text{Min}'$$

and

$$s' = \left| \frac{s}{\text{Max} - \text{Min}} \right| \times (\text{Max}' - \text{Min}')$$

where \bar{x}' = normalized mean; \bar{x} = original mean; s' = normalized standard deviation; s = original standard deviation; Min, Max = minimum and maximum of original scale (Min always has to correspond with the counterstereotypical end of the scale, and Max with the traditional end of the scale); Min', Max' = rescaled minimum and maximum = -1 , 1 .

In case of an unequal spread of a scale between positive and negative numbers (i.e., SDS scale [SDSS], -30 to 48), the above solution may inadvertently revert the sign of the group mean (i.e., a group mean of 0 will become -0.23). To honor the original spread of positive and negative values, we adapted the above function for rescaling SDSS scores as follows:

$$\bar{x}' = \frac{\bar{x} - (-\text{Max})}{\text{Max} - (-\text{Max})} \times (\text{Max}' - \text{Min}') + \text{Min}'$$

$$s' = \left| \frac{s}{\text{Max} - (-\text{Max})} \right| \times (\text{Max}' - \text{Min}')$$

Using this function a score of -30 will become -0.625 , a score of 0 will still be 0 , and a score of $+48$ will be $+1$.

Statistical Analyses

As there was a considerable number of studies that reported on multiple samples, or multiple SDS-aspects (see Tables 1 and 2), we applied a multilevel random effects model (Cheung, 2014; Van den Noortgate et al., 2013). See Figure 2 for the multilevel hierarchical structure employed in these meta-analyses. Such a model accounts for the dependency between effect sizes that come from the same study, because it partitions sources of variance: variance between studies, variance between samples from the same study, variance between effect sizes from the same study, and sampling variance (Cheung, 2014; Van den Noortgate et al., 2013). Furthermore, multilevel models can model all available effect sizes within studies together which maximizes statistical power. In addition, these models can test moderators of differences in outcomes at both the within-study level and the between-study or between-sample level. We performed the analyses with the metafor package (Viechtbauer, 2015) for the R environment (Version 3.5.2; R Development Core Team, 2013), using a step-wise procedure described by Assink and Wibbelink (2016). We used the restricted maximum likelihood procedure to estimate the parameters. Significance tests and moderator analyses were performed through random-effect models, which are more conservative than fixed-effect models when there may be different effect sizes underlying different studies (Borenstein et al., 2009). Single effect sizes within each study were computed using the Comprehensive Meta-Analysis (CMA, v3) program (Borenstein et al., 2005).

For the moderator analyses, we centered each continuous variable around its mean and converted each categorical variable with k categories to $k - 1$ dummy variables through binary coding. We tested single moderators first, followed by a multiple-moderator model, including all significant single moderators. We checked for outlying effect sizes and sample sizes separately for the two meta-analyses. Z -values below 3.29 or greater than 3.29 were considered outliers (Tabachnick & Fidell, 2012).

In case of nonsignificant overall effects, we used the two one-sided test (TOST) procedure for equivalence testing (Lakens, 2017), with equivalence bounds of d $[-0.10, 0.10]$, as an indication for trivial gender differences (Hyde, 2005). In the TOST procedure a significant p value is indicative of statistical equivalence. We used likelihood ratio tests to test for the presence of significant between-study and within-study heterogeneity (σ^2) (Raudenbush & Bryk, 2002). Also, we computed 95% credibility intervals (CVs) for the overall effect as an indication of the range of true population effects

(Borenstein et al., 2009; see Supplemental Appendix D Table 4 and 5 for individual CVs in moderator analyses). CVs also provide an additional metric of heterogeneity in effect sizes. In addition, I^2 statistics give an indication of the proportion of the variation in observed effects that is due to variation in true effects (Borenstein et al., 2017). If there was evidence for heterogeneity in effect sizes (at one or more of the three levels), we conducted moderator analyses, but only conducted when at least two of the moderator categories consisted of at least four samples each (Bakermans-Kranenburg et al., 2003). For moderator models, an omnibus test of the fixed model parameters tests the null hypothesis that the group mean effect sizes are equal. To control for Type I error rates, we applied the Knapp and Hartung (2003) adjustment. To further control for Type I errors associated with multiple-moderator tests, we included the significant moderators in one multiple-moderator model. Multiple-moderator models also test the relative importance of each significant moderator and take into account possible correlation between moderators. As a proxy of R^2 change, we computed the proportional reduction in the variance components between the model without moderators and the multiple-moderator model.

Publication Bias

Smaller studies with nonsignificant results or with effect sizes in the nonhypothesized direction are less likely to be published, whereas for large studies, publication of small or nonsignificant effect sizes or effect sizes in the nonhypothesized direction is more likely, because large studies are generally deemed more trustworthy. This problem is also known as publication bias (Rosenthal, 1995). Publication bias is problematic for meta-analyses, because it could lead to an overestimation of the true effect size (Borenstein et al., 2009). A method to test for publication bias is by testing for asymmetries in funnel plots. A funnel plot is a plot of each study's effect size against its standard error (usually plotted as $1/SE$, or precision). We examined funnel plot asymmetry with Egger's regression test, which regresses the standard normal estimate on the estimate's precision (Egger et al., 1997). When this test was statistically significant, we performed the trim and fill method, which estimates the number of studies which have no symmetric counterpart on the other side of the funnel (Duval & Tweedie, 2000a, 2000b).

Results

The meta-analysis conducted on studies assessing SDS from differential evaluation or expectation of men's and women's sexual behavior included 52 studies, reporting on 116 independent samples, and 277 effect sizes. The studies reported on a total of $N = 71,442$ participants. The meta-analysis conducted on studies assessing SDS with Likert-type scale questionnaires included 47 studies, reporting on 85 independent

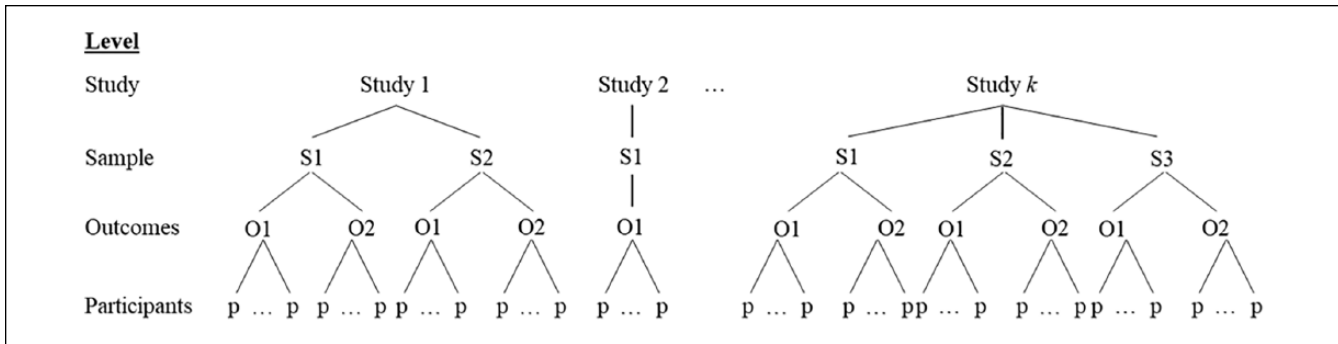


Figure 2. Multilevel hierarchical structure employed in the meta-analyses.

samples, and 129 effect sizes. The studies reported on a total of $N = 51,901$ participants.

Overall Effects

The combined effect size for the difference in evaluation or expectation of men's and women's sexual behavior was significant, but small ($d = 0.25$, 95% confidence interval [CI] [0.18, 0.33], 95% CV [-0.42, 0.92], $p < .001$, $k = 52$, independent samples = 116, effect sizes = 277). The effect size was positive, indicating that behaviors associated with high sexual activity were expected more and evaluated more positively (or less negatively) in men compared with women, and behaviors associated with sexual passivity were expected more and evaluated more positively (or less negatively) in women compared with men. We found significant variation between studies, Level 3: $\sigma^2 = 0.046$, $\chi^2(1) = 11.33$, $p < .001$; $I^2 = 38.52$, as well as significant variation between effect sizes within studies, Level 1: $\sigma^2 = 0.068$, $\chi^2(1) = 1,334.88$, $p < .0001$; $I^2 = 57.16$, but no variation between independent samples from the same studies, Level 2: $\sigma^2 < 0.001$, $\chi^2(1) = 0.00$, $p = 1.00$; $I^2 < 0.001$.

The combined mean for SDS assessed with Likert-type-scale questionnaires was not significantly different from 0 ($M = -0.09$, $SE = 0.05$, 95% CI [-0.19, 0.01], 95% CV [-0.84, 0.67], $p = .084$, $k = 47$, independent samples = 85, effect sizes = 129). The effect was also not statistically equivalent ($Z = 0.20$, $p = .42$, 90% CI [-0.18, -0.01]). This indicated that on Likert-type-scale questionnaires participants demonstrated no evidence of SDS. We found significant variation between studies, Level 3: $\sigma^2 = 0.081$, $\chi^2(1) = 29.99$, $p < .0001$; $I^2 = 56.47$, as well as significant variation between effect sizes within studies, Level 1: $\sigma^2 = 0.063$, $\chi^2(1) = 11,134.09$, $p < .0001$; $I^2 = 43.50$, but no variation between independent samples from the same studies, Level 2: $\sigma^2 < 0.001$, $\chi^2(1) = 0.00$, $p = 1.00$; $I^2 < 0.001$.

For both meta-analyses, Egger's regression test indicated that there was no evidence of publication bias (meta-analysis on differences scores: estimate = 0.743, $SE = 0.763$, 95% CI [-0.76, 2.25], $t(275) = 0.97$, $p = .33$; meta-analysis on Likert-type-scale questionnaires: estimate

= 9.47, $SE = 5.34$, 95% CI [-1.09, 20.03], $t(128) = 1.78$, $p = .078$).

Single Moderator Analyses

Tables 3 and 4 display the results of the single moderator analyses for respectively studies examining the difference in evaluation or expectation of men's and women's sexual behavior, and studies using Likert-type-scale questionnaires to assess SDS.

For the meta-analysis on studies examining the difference in evaluation or expectation of men's and women's sexual behavior, SDS conceptualization was a significant moderator. SDS were more traditional (i.e., significantly higher effect size with positive sign) in studies that assessed stereotypes than in studies that assessed attitudes. Sexual behavior type was also a significant moderator. Pairwise comparisons showed that the highest effect size (i.e., most traditional SDS) was found for being a victim of sexual coercion, followed by casual sex, and having an early sexual debut. The effect sizes for these behaviors were small to moderate. We found significantly lower combined effect sizes for sexual infidelity, level of sexual activity, other/mixed sexual behavior types, premarital sex (when in love or with affection, when engaged, or relationship status unspecified), and being a perpetrator of sexual coercion. The effect sizes for these behaviors were negligible to small. The following moderators were not significant: study design, measurement type, participant gender, participant age, level of gender equality in the country where the study was conducted, publication year.

For the meta-analysis on studies using Likert-type-scale questionnaires to assess SDS, SDS conceptualization was a significant moderator. SDS were more traditional (i.e., significantly higher combined mean with positive sign) in studies that assessed stereotypes than in studies that assessed attitudes or a combination of cognitions. For studies conceptualizing the SDS as a combination of stereotypes and attitudes, a significant reversed double standard was found as indicated by a negative combined mean. Questionnaire type was also a significant moderator.

Table 3. Meta-Analytic Results of Moderators of the Differential Evaluation and Expectation of the Sexual Behavior of Men and Women.

Moderator	#k	#IS	#ES	β_0, d^f (95% CI)	β_1 (95% CI)	F test	p	Variance level ^b			
								σ^2	I^2	σ^2	
								1	2	3	
Measurement and design											
SDS conceptualization						$F(1,275) = 17.60$	<.001	0.065**	60.99	0.036**	34.14
Stereotype	9	14	54	0.536 [0.386, 0.687]**							
Attitude	45	106	223	0.194 [0.120, 0.502]**	-0.342 [-0.502, -0.181]**						
Measurement type						$F(1,274) = 0.370$.544	0.068**	55.65	0.049**	40.14
Questionnaire	20	59	138	0.227 [0.110, 0.345]**							
Vignette/scenario	31	56	138	0.276 [0.172, 0.379]**	0.048 [-0.108, 0.205]						
Study design						$F(1,275) = 0.400$.527	0.068**	56.74	0.047**	38.98
Within subjects	22	63	150	0.228 [0.118, 0.338]**							
Between subjects	30	53	127	0.277 [0.172, 0.382]**	0.049 [-0.103, 0.201]						
Sexual behaviors ^c						$F(9,267) = 6.878$	<.001	0.048**	42.19	0.061**	53.28
Coercion (perpetrator)	3	4	7	-0.168 [-0.471, 0.113]							
Coercion (victim)	3	3	7	0.797 [0.487, 1.106]**	1.172 [0.840, 1.504]**						
Casual sex	15	29	70	0.328 [0.203, 0.452]**	0.496 [0.167, 0.826]**						
Sexual infidelity	3	15	15	0.263 [-0.020, 0.546] [†]	0.426 [-0.015, 0.867] [†]						
Sexual activity level	16	43	86	0.280 [0.162, 0.399]**	0.449 [0.123, 0.776]**						
Premarital sex: love/affection	4	12	24	0.134 [-0.043, 0.311]	0.289 [-0.067, 0.644]						
Premarital sex: engaged	2	8	8	0.129 [-0.096, 0.354]	0.273 [-0.115, 0.661]						
Premarital sex: not specified	4	8	8	0.224 [-0.035, 0.482] [†]	0.377 [-0.042, 0.796] [†]						
Early sexual debut	5	6	11	0.402 [0.181, 0.624]**	0.587 [0.201, 0.973]**						
Other/mix	13	23	39	0.165 [0.035, 0.295]*	0.328 [-0.003, 0.660] [†]						
Sample						$F(1,143) = 0.511$.467	0.043**	58.94	0.025*	34.03
Gender											
Male	25	39	74	0.175 [0.084, 0.266]**							
Female	25	39	71	0.145 [0.054, 0.236]**	-0.030 [-0.113, 0.053]						
Gender inequality country	52	116	277	0.252 [0.174, 0.329]**	-0.595 [-2.416, 1.226]						
Age						$F(1,275) = 0.414$.521	0.067**	55.47	0.049*	40.30
Adolescents (12-18)	5	12	36	0.209 [0.016, 0.401]*							
College aged (19-25)	36	69	177	0.285 [0.195, 0.374]**	0.076 [-0.125, 0.276]						
Adults (>25)	10	29	50	0.208 [0.057, 0.360]**	0.000 [-0.218, 0.217]						
Publication						$F(2,260) = 0.534$.587	0.070**	56.76	0.049**	39.09
Publication year	52	116	277	0.251 [0.174, 0.328]**	0.001 [-0.003, 0.006]						
						$F(1,275) = 0.229$.632	0.068**	56.40	0.047**	39.33

Note. Moderators or moderator-categories other than the one's presented above were not examined, because less than two of the moderator categories consisted of at least four studies each.

#k = number of independent samples; #IS = number of independent samples; #ES = number of effect sizes; d = standardized mean difference; CI = 95% confidence interval; SDS = sexual double standard.

^aFor continuous predictors, the mean effect size d indicates the mean effect size of a participant with an average value on the corresponding predictor. ^bVariance was examined at the following levels:

1 = variance within samples, that is, between effect sizes from the same sample, 2 = variance within studies, that is, between samples from the same study, 3 = variance between studies. As there was zero variance at the second level in the overall model, this level was not presented in this table. ^cReference category is coercion perpetrator, results of pairwise comparisons between other sexual behavior types can be found in Supplemental Appendix D (Table 4).

[†] $p < .10$. * $p < .05$. ** $p < .01$.

Table 4. Meta-Analytic Results of Moderators of Endorsement of Sexual Double Standards Assessed With Likert-Type-Scale Questionnaires.

Moderator	#k	#IS	#ES	$\beta_{\sigma} d^a$ (95% CI)	β_1 (95% CI)	F test	p	Variance level ^b				
								σ^2	I^2	I^2		
Measurement and design												
SDS conceptualization						$F(2, 126) = 3.15$.046	0.061**	44.11	0.077**	55.86	
Stereotype	6	13	23	0.109 [-0.075, 0.293]								
Attitude	15	36	51	-0.061 [-0.213, 0.091]	-0.170 [-0.329, -0.011]** ^c							
Combination	32	50	55	-0.128 [-0.245, 0.012]**	-0.237 [-0.444, -0.030]**							
Questionnaire type						$F(2, 126) = 8.19$	<.001	0.063**	54.42	0.055**	45.54	
DSS	17	26	26	-0.330 [-0.480, -0.180]**								
SDSS	15	24	29	0.070 [-0.080, 0.288]	0.400 [0.188, 0.611]** ^d							
Other	17	40	74	0.016 [-0.114, 0.146]	0.346 [0.148, 0.545]**							
Sample						$F(1, 110) = 2.05$.155	0.060**	41.40	0.086**	58.56	
Gender												
Male	29	33	50	-0.011 [0.136, 0.114]								
Female	34	41	62	-0.085 [-0.205, 0.035]	-0.074 [-0.177, 0.028]							
Gender equality	47	86	129	-0.078 [-0.174, 0.019]	-1.120 [-1.887, 0.353]**	$F(1, 127) = 8.35$.005	0.059**	43.15	0.078**	56.81	
Age						$F(2, 126) = 0.22$.805	0.063**	42.55	0.085**	57.42	
Adolescents (12-18)	9	18	26	-0.134 [-0.314, 0.045]								
College aged (19-25)	33	60	95	-0.084 [-0.199, 0.031]	0.052 [-0.133, 0.237]							
Adults (>25)	8	8	9	-0.062 [-0.336, 0.212]	0.085 [-0.218, 0.388]							
Publication						$F(1, 127) = 5.66$.019	0.062**	45.09	0.075**	54.87	
Publication year	47	86	129	-0.043 [-0.145, 0.059]	-0.012 [-0.021, -0.002]**							

Note. Moderators or moderator-categories other than the one's presented above were not examined, because less than two of the moderator categories consisted of at least four studies each.

#k = number of studies; #IS = number of independent samples; #ES = number of effect sizes; d = standardized mean difference; CI = 95% confidence interval; SDS = sexual double standard(s); DSS = double standard scale; SDSS = sexual double standard scale.

^aFor continuous predictors, the mean effect size d indicates the mean effect size of a participant with an average value on the corresponding predictor. ^bVariance was examined at the following levels: 1 = variance within samples, that is, between effect sizes from the same sample, 2 = variance within studies, that is, between samples from the same study, 3 = variance between studies. As there was zero variance at the second level in the overall model, this level was not presented in this table. ^cAttitudes and other cognitions do not differ significantly from each other: $\beta = -0.067$, $SE = 0.09$, $t = -0.729$, 95% CI [-0.249; 0.115], $p = .468$. ^dSDSS questionnaire and other questionnaires do not differ significantly from each other: $\beta = -0.053$, $SE = 0.09$, $t = -0.568$, 95% CI [-0.239; 0.133], $p = .571$.

* $p < .05$. ** $p < .01$.

Pairwise comparisons showed that studies using the DSS of Caron yielded a significant reversed double standard (i.e., negative combined mean). However, we found significantly more traditional SDS in studies using the SDS scale of Muehlenhard or in studies using other questionnaires. Level of gender equality in the country in which the study was conducted was a significant moderator, indicating that higher levels of gender equality were associated with less traditional SDS in the direction of a reversed double standard. Publication year was a significant moderator, indicating that a more recent publication year was associated with less traditional SDS in the direction of a reversed double standard. The following moderators were not significant: participant gender and participant age.

Multiple-Moderator Analyses Including All Significant Moderators

Tables 5 and 6 display the results for the multiple-moderator model for respectively the evaluation and expectation of the sexual behavior of men and women, and for studies examining SDS with Likert-type-scale questionnaires. SDS conceptualization was still a significant moderator in both meta-analyses. Sexual behavior type remained a significant moderator in the meta-analysis on the evaluation and expectation of the sexual behavior of men and women. Questionnaire type and gender equality in the country in which the study was conducted were still significant moderators in the meta-analysis on studies using Likert-type-scale questionnaires, but publication year was no longer significant. Effects were in the same direction as the single moderator analyses. In both multiple-moderator models, there was still significant variance left to explain at both the between-study and within-study level. Yet, in the model for studies examining differential evaluation of men and women, adding the moderators lead to a 32% reduction in the variance within samples. In the model for studies using Likert-type-scale questionnaires, adding the moderators lead to a 10% reduction in the variance within samples, and a 49% reduction in the variance between studies.

Outliers

In the first meta-analysis, we detected five outlying (four positive, one negative) effect sizes (Conley et al., 2013, Study 1b and Study 2b; Marks, 2008, implicit condition 17 partners; Weaver et al., 2013) and three studies with outlying sample sizes (Do & Fu, 2010; Marks & Fraley, 2005; Soller & Haynie, 2017). In the second meta-analysis, we did not detect outlying effect sizes, but one study had an outlying sample size (Kettrey, 2016; female sample). We conducted analyses with and without studies with outlying effect sizes. The outliers with regard to sample size were winsorized (highest nonoutlying number + difference between highest nonoutlying number and before highest nonoutlying

number). We found no differences in results for analyses including or excluding outlying effect sizes or winsorized sample sizes.

Discussion

In line with evolutionary theory (Buss & Schmitt, 1993; Trivers, 1972) and biosocial theory (Wood & Eagly, 2002, 2012), this meta-analysis demonstrated clear evidence for traditional SDS in studies assessing differences in people's evaluation, or expectation, of men's and women's sexual behavior, although the effect was small. People expected behaviors associated with high sexual activity more from men than from women, and behaviors associated with low sexual activity more from women than from men. Similarly, people evaluated highly sexually active men more positively (or less negatively) than highly sexually active women, and low sexually active women more positively (or less negatively) than low sexually active men. In contrast, the overall set of studies using Likert-type-scale questionnaires for assessing SDS did not yield evidence of SDS.

We found some significant moderator effects in one or both sets of studies. First, existence of traditional SDS was behavior specific. Second, stereotypes about SDS were more traditional than attitudes about SDS. Third, studies using the "sexual double standard scale" (SDSS; Muehlenhard & Quackenbush, 1998) reported more traditional SDS than studies using the "double standard scale" (DSS; Caron et al., 1993) which demonstrated reversed SDS. Fourth, higher levels of gender equality in a country were associated with less traditional SDS. Participant gender and age, publication year, and study design were not significant moderators.

Behavioral Specificity of SDS

Regarding sexual behavior types, we found strongest evidence of SDS for being a victim of sexual coercion, followed by casual sex, and having an early sexual debut. SDS were less evident for sexual infidelity, level of sexual activity, other/mixed sexual behavior types, premarital sex, and being a perpetrator of sexual coercion. The findings for coercion and sexual encounters within a power or age hierarchy were partly in line with the predictions from biosocial theory that SDS would be most prevalent in sexual encounters where there is a power/status difference between men and women (Wood & Eagly, 2002, 2012). However, we only found double standards for victims of sexual coercion, and not for perpetrators. That we did not find differences in the evaluation of male and female perpetrators might be because both male and female perpetrators violate gender role expectations, with men violating chivalry norms, and females violating communal characteristics. Thus, male and female perpetrators might have been evaluated equally negative for their gender-role inconsistent behavior. Moreover, the double standards for victims of sexual coercion, found in person

Table 5. Results for the Multiple-Moderator Model for Studies Examining Differential Evaluation and Expectation of the Sexual Behavior of Men and Women (#k = 52, #IS = 116, #ES = 277).

Moderator	β (SE)	95% CI	t	p
Intercept	0.172 (0.165)	[-0.153, 0.498]	1.041	.299
SDS conceptualization (reference category stereotype)				
Attitude	-0.335 (0.084)**	[-0.500, -0.170]	-3.995	<.001
Sexual behaviors (reference category coercion perpetrator)				
Coercion: victim	1.164 (0.165)**	[0.839, 1.488]	7.067	<.001
Casual sex	0.381 (0.157)*	[0.071, 0.691]	2.423	.016
Sexual infidelity	0.385 (0.207) [†]	[-0.023, 0.793]	1.856	.065
Sexual activity level	0.378 (0.154)*	[0.074, 0.682]	2.447	.015
Premarital sex: love/affection	0.201 (0.169)	[-0.133, 0.534]	1.186	.237
Premarital sex: engaged	0.182 (0.186)	[-0.185, 0.549]	0.978	.329
Premarital sex: not specified	0.372 (0.195) [†]	[-0.013, 0.756]	1.903	.058
Early sexual debut	0.528 (0.183)**	[0.167, 0.890]	2.880	.004
Other/mix	0.287 (0.156) [†]	[-0.021, 0.595]	1.834	.068
Omnibus test	F(10,266)=8.071**			<.001
	σ^2	I^2	ΔR^2	
Variance level 1 ^a	0.046**	47.32	0.32	
Variance level 3	0.046**	47.40	0.00	

Note. #k = number of studies; #IS = number of independent samples; #ES = number of effect sizes; CI = confidence interval; SDS = sexual double standard.

^aVariance was examined at the following levels: 1 = variance within samples, that is, between effect sizes from the same sample, 2 = variance within studies, that is, between samples from the same study, 3 = variance between studies. As there was zero variance at the second level in the overall model, this level was not presented in this table.

[†]p < .10. *p < .05. **p < .01.

Table 6. Results for the Multiple-Moderator Model for Studies Examining Endorsement of Sexual Double Standards With Likert-Type-Scale Questionnaires (#k = 47, #IS = 85, #ES = 129).

Moderator	β (SE)	95% CI	t	p
Intercept	0.155 (0.081)	[-0.004, 0.314]	1.925	.314
SDS conceptualization (reference category stereotype)				
Attitude	-0.178 (0.055)*	[-0.328, -0.028]	-2.346	.021
Other	-0.452 (0.110)**	[-0.669, -0.236]	-4.130	<.001
Questionnaire type (reference category DSS) ^a				
SDSS	0.427 (0.098)**	[0.234, 0.621]	4.367	<.001
Gender equality country	-1.118 (0.380)*	[-1.870, -0.366]	-2.942	.004
Publication year	-0.006 (0.004)	[-0.014, 0.002]	-1.400	.164
Omnibus test	F(5,123) = 7.998**			<.001
	σ^2	I^2	ΔR^2	
Variance level 1 ^b	0.057**	57.89	0.10	
Variance level 3	0.041**	42.06	0.49	

Note. #k = number of studies, #IS = number of independent samples; #ES = number of effect sizes; CI = confidence interval; SDS = sexual double standard; DSS = double standard scale.

^a“Other questionnaire” was a redundant predictor ($r = -1.0$ with “social cognition other”) and therefore dropped from the model. ^bVariance was examined at the following levels: 1 = variance within samples, that is, between effect sizes from the same sample, 2 = variance within studies, that is, between samples from the same study, 3 = variance between studies. As there was zero variance at the second level in the overall model, this level was not presented in this table.

*p < .05. **p < .01.

perception studies, indicate that sexual behavior within the context of a power hierarchy is evaluated more negatively (or less positive) for female victims (e.g., more condemned,

more perceived damage to reputation) than for male victims (e.g., “positive” experience that will be evaluated by peers as cool; Zaikman & Marks, 2017). Thus, girls might be blamed

for being a victim of sexual coercion (Weis, 2009), whereas boys' experiences of sexual coercion might be trivialized (Weis, 2010). This is inconsistent with the idea that male victims of sexual coercion or rape might be perceived as powerless and not willing to have sex, which violates men's (hetero)sexual agentic gender role (Weis, 2010). Because only a few studies examined the evaluation of both perpetrator and victim, replication of the existence and direction of SDS for coercion victims is necessary in future studies.

The finding that engaging in casual sex and having an early sexual debut were more expected and rewarded in men than in women, fits partly with predictions from evolutionary theory. In terms of reproductive fitness, men would benefit more than women from having casual sex and by having sex at an early age (Buss & Schmitt, 1993; Petersen & Hyde, 2010). However, similar beneficial effects would have been expected for sexual infidelity and high sexual activity with numerous partners, but for those sexual behaviors less traditional SDS were applied. The same was true for other sexual behaviors, such as premarital sex when engaged or when in love. Our findings are in line with previous narrative reviews concluding that premarital sex in particular has become accepted for both men and women (Bordini & Sperb, 2013; Crawford & Popp, 2003).

Cross-Cultural Differences in SDS

In line with predictions from biosocial theory (Wood & Eagly, 2002, 2012), and not with evolutionary theory's perspective of obligate sex differences, SDS were less traditional in countries with higher levels of gender equality. According to biosocial theory, in cultures with bigger differences in the gender roles of men and women, men have more power than women, which translates in traditional SDS (Wood & Eagly, 2002, 2012). However, level of gender equality was only a significant moderator in the meta-analysis conducted on studies using Likert-type-scale questionnaires, and not in the meta-analysis on differential evaluation and expectation of the sexual behavior of men and women. This might be because there was less variation in level of gender equality in the latter meta-analysis, as most studies in that meta-analyses were conducted in the United States. The direction of effect, albeit nonsignificant, was in the same direction as the effect from the meta-analysis on Likert-type scales.

Changes in SDS Over Time

In line with evolutionary theory, and not with biosocial theory, time period in which the study was conducted was no longer significant when controlling for other moderators, which indicated that traditional SDS have existed for decades and are still present. This finding could indicate that stable gender differences in reproductive strategies are underlying SDS. Also, it appears that even though gender roles have

become less strict in most modern Western societies (Eagly & Wood, 1999), this did not lead to less differentiation in the norms for the sexual behavior of men and women (Wood & Eagly, 2002, 2012). Possibly, it takes more time for egalitarian gender roles to permeate into the bedroom, than in other domains of life such as the work field, because sexuality is very much a private issue. Furthermore, the content of SDS may have changed over time, because most older studies focused on double standards in premarital sex in different relationship types, whereas newer studies more often focused on double standards in casual sex. Thus, changes in gender roles over time might only be reflected in changes in the behavior specificity of SDS.

Gender Differences in SDS

Regarding gender, we did not find differences between men and women in their cognitions about SDS. In light of male control theory and female control theory (Baumeister & Twenge, 2002), these findings could indicate that both male control and female control contribute equally to the existence of SDS. This means that SDS might provide evolutionary and sociocultural advantages for both genders that they would like to control. Advantages for men that arise from SDS could be improved certainty about paternity (Buss, 1994), patriarchal power over women, prevention of sexual chaos, and reduced male insecurity (Hyde & DeLamater, 1997). The advantages of SDS for women are the high value of sexual favors that they can trade for lower valued favors from men, such as economic provision, monogamous relationships, and parental investment.

Age Differences in SDS

Regarding participants' age, we did not find support for the predictions of the gender-intensification hypothesis (Hill & Lynch, 1983). It appears that adolescence is not necessarily a period that is characterized by increased gender role pressure and intensification of people's social cognitions about gender. However, it should be mentioned that most studies were conducted with high-educated college samples, mostly including emerging adults. It may be possible that the relatively small number of studies conducted with adolescents and adults, decreased the power to detect effects of age on SDS.

Implications for Evolutionary Theory and Biosocial Theory

In sum, some of the above findings are in line with evolutionary theory (Buss & Schmitt, 1993; Trivers, 1972) whereas others are in line with biosocial theory (Wood & Eagly, 2002, 2012). This converges with the findings of a recent theory-based narrative review, which demonstrated some support for predictions of both evolutionary theory and biosocial

theory about the behavioral specificity of SDS, and for predictions of biosocial theory about cultural differences in SDS (Zaikman & Marks, 2017). Each theory suggests a different mechanism that underlies SDS, but these mechanisms might be intertwined. We therefore propose that a hybrid model explaining SDS from the interplay between biological predispositions and sociocultural pressures is most appropriate (Lippa, 2009). According to biosocial theory, different norms for the behavior of men and women may have arisen from societies' division in gender roles that expects men to be assertive, dominant, and powerful, and women to be submissive, caring, and kind (Wood & Eagly, 2002, 2012). However, the division in gender roles may have a biological or evolutionary origin (Wood & Eagly, 2012), because there are gender differences in adaptive reproductive strategies leading people to view (sexual) behaviors in men and women differently. Also, the predictive power of evolutionary and sociocultural gender role pressures to explain SDS appears to depend on the sexual behavior or context under consideration. Gender roles may have more predictive power in a sexual context characterized by power/status differences. Yet, evolutionary processes might play a larger role in sexual behaviors that increase successful reproduction.

Conceptualization and Measurement of SDS

We also looked at the effects of moderators related to conceptualization and measurement of SDS. Regarding SDS conceptualization, effect sizes were significant for both stereotypes and personal attitudes. This suggests that both stereotyped beliefs about the sexual behavior of men and women and people's personal attitudes in response to sexual behavior that violates expectancies are underlying SDS. Yet, traditional SDS were more prevalent in collective or personal expectations about the sexual behavior of men and women (i.e., stereotypes) than in people's personal evaluation of the sexual behavior of men and women (i.e., attitudes). This finding is in line with the idea that people can have knowledge of collectively shared stereotypes with regard to SDS or personal stereotypical expectations about the sexual behavior of men and women, although they do not apply these stereotypes personally when evaluating other people's sexual behavior (Milhausen & Herold, 2001; Signorella et al., 1993). It has been argued that knowledge of collective stereotypes is strong, stable, and does not depend on one's experience with other people, but on culturally shared and generalized social beliefs (López-Sáez & Lisbona, 2009). Indeed, research in children as well as adults showed that content of collective gender stereotypes has not changed over time, whereas gender attitudes did become more egalitarian (e.g., Ruble, 1983; Signorella et al., 1993).

However, our findings with regard to social cognition type need to be interpreted with caution, because the vast majority of studies examined personal SDS attitudes or a mix of stereotypes and attitudes. In studies examining a

combination of stereotypes and attitudes, evidence for a reversed double standard was found, a finding that is difficult to disentangle because of the muddled operationalizations of SDS in these studies. Furthermore, in the small number of studies examining stereotypes, it was not possible to distinguish between descriptive and prescriptive aspects, or between personal stereotypes and knowledge of collective stereotypes. Yet, these distinctions are important for future research. For example, knowledge of collectively shared stereotypes is less predictive of one's own behavior toward men and women than personal stereotypes (Stangor & Schaller, 1996). Furthermore, prescriptive stereotypes (e.g., perceptions of how men and women *should* behave sexually) might be particularly relevant in the context of SDS as they have been associated with negative evaluations and backlash for people who behave in stereotype-inconsistent ways (Burgess & Borgida, 1999). Indeed, gender stereotypes in general are highly prescriptive in nature (Prentice & Carranza, 2002) and more predictive of people's personal evaluation of men and women (i.e., attitudes) than descriptive stereotypes (Gill, 2004).

As expected from dual-process models of social cognition (Gawronski & Creighton, 2013), studies using explicit Likert-type-scale questionnaires did not yield evidence for traditional SDS. Yet, studies using more implicit within- or between-subjects designs did yield evidence for SDS. The Likert-type-scale questionnaires often include items such as "It's worse for a woman to sleep around than it is for a man" in which male and female sexual behavior is explicitly contrasted to each other (Muehlenhard & Quackenbush, 1998). Therefore, in studies using such questionnaires it might have been more clear to participants that personal cognitions about SDS were assessed, leading to social-desirable responding (Greenwald et al., 2009). In between- and within-subjects designs, the focus on SDS is more implicit than in explicit self-report questionnaires. This is because in a between-subject design researchers assessed cognitions about women's and men's sexual behavior with separate items or vignettes that they randomly assign to participants, who are generally unaware of the presence of other vignettes presented to other participants. Or in a within-subject design researchers administered separate vignettes or items about women's and men's sexual behavior in a counter-balanced way to participants (Jonason & Marks, 2009; Reid et al., 2011; Weaver et al., 2013). Thus, this finding suggests that traditional SDS might only be present at a more implicit level. Previous research indeed showed that implicit assessments are less prone to social-desirable responding (Gawronski & Bodenhausen, 2006) and more likely to suggest existence of traditional gendered cognitions (Endendijk et al., 2013).

However, SDS were not different between studies using between- or within-subjects designs, or between studies using extensive vignettes/scenarios versus studies using questionnaires with different items about the sexual

behavior of men and women. This indicates that social desirability and demand characteristics might not necessarily play a larger role in within-subject research on SDS than in between-subject research (Marks & Fraley, 2005; Milhausen & Herold, 2001). Also, this finding suggests that study designs that have only a slightly less explicit focus on SDS (i.e., not contrasting male and female sexual behavior in the same items) can yield evidence for the existence of traditional SDS. This argument is consistent with one study that specifically examined differences in implicit (i.e., under divided attention) and explicit (i.e., under full attention) SDS-cognitions, showing that traditional SDS were only present at an implicit level (Marks, 2008). However, between-subjects designs, like the study by Marks (2008), have been criticized for measuring single standards (because there is no comparison with how an individual would rate another target) instead of double standards (i.e., contrasting evaluation of male vs. female target; Crawford & Popp, 2003). Therefore, using IATs might be a fruitful direction to take to examine SDS at an implicit within-subjects level (see, for example, Sakaluk & Milhausen, 2012).

Our findings regarding questionnaire type indicated that questionnaires differ in the extent to which they yield evidence for SDS, which might also explain the nonequivalent findings in studies using these methods. Studies using the DSS (Caron et al., 1993) reported reversed double standards, whereas studies using the SDSS (Muehlenhard & Quackenbush, 1998) reported more traditional double standards, which might be explained by differences in content and scoring of the questionnaires. In the DSS all but one items are formulated in the direction of a traditional double standard (e.g., "It is up to the man to initiate sex.") and participants answer the items on a scale ranging from strongly agree to strongly disagree. Such a questionnaire design cannot distinguish between people with reversed and egalitarian sexual standards, because both groups of people will (strongly) disagree with the traditional items. Therefore, we cannot be completely sure that the negative combined mean found in studies using the DSS actually reflects reversed double standards, or an egalitarian view about the sexual behavior of men and women instead. In contrast, the SDSS consists of 20 items occurring in pairs, with parallel items about men's and women's sexual behavior (e.g., "A [girl/boy] who has sex on the first date is easy"). In addition, six items contrast men's and women's sexual behavior, with some items formulated in the direction of traditional SDS (e.g., "A man should be more sexually experienced than his wife.") and others formulated in an egalitarian way (e.g., "A woman's having casual sex is just as acceptable to me as a man's having casual sex."). Participants answer all items on a scale ranging from disagree strongly to agree strongly. Difference scores are computed between the 10 male and female items and the six individual item-scores are added to these difference scores. The design of the SDSS makes it possible to assess a more complete range of reversed to

traditional double standards than with the DSS. However, the SDSS score range is asymmetrical (-30 to 48). Thus, the more traditional double standards appearing in studies using the SDSS might have been an artifact of the possible range of scores.

Limitations and Future Directions

Some limitations of this meta-analytic study need to be addressed. First, the available body of quantitative research on SDS is highly homogeneous in terms of participant age, ethnicity, and educational level. According to biosocial theory, these factors are important in the social construction of gender roles, and more specifically for the social construction of SDS (Wood & Eagly, 2002, 2012). Therefore, future studies should examine SDS in more diverse samples in terms of ethnicity, age, and educational level.

Second, almost all studies included in this meta-analysis measured SDS in a relatively explicit way, by using self-report questionnaires, even though implicit measures, such as IATs or priming tasks, are less prone to social-desirable responding than explicit measures of stereotypes, and are often better predictors of behavior (Gawronski & Bodenhausen, 2006). Thus, researchers should make use of more implicit tasks to assess SDS. Relatedly, previous research has used many different conceptualizations of SDS, sometimes combining attitudinal aspects with stereotypical aspects within one questionnaire. We advise future researchers to be more theory-driven in their conceptualization, operationalization, and predictions regarding SDS. For example, dual-process models (Gawronski & Creighton, 2013) or social cognition frameworks (e.g., Greenwald et al., 2002) could be used to further conceptualize different aspects of people's SDS-cognitions, that is, implicit, explicit, attitudes, stereotypes, knowledge of stereotypes, prescriptive versus descriptive aspects, and personal versus collective aspects. New measures need to be developed and validated before we can examine the interplay between different double standard components.

Furthermore, studies assessing SDS via questionnaires sometimes used questionnaires that did not distinguish between people with reversed and egalitarian sexual standards. With such questionnaires, it is impossible to study predictors of individual differences in SDS-cognitions. When researchers would like to use a questionnaire in future studies on SDS, they should use questionnaires with symmetrical scales to assess the complete range of SDS from reversed to traditional (e.g., 20 item-pairs of the SDSS; Muehlenhard & Quackenbush, 1998) or develop new questionnaires that can assess the complete range.

Last, most studies included in this meta-analysis focused on SDS in behaviors associated with high sexual activity and only a few studies have been conducted specifically on behaviors associated with low sexual activity. However, further study of differences in the strength of

traditional SDS between behaviors associated with high sexual activity (more male-typical) and behaviors associated with low sexual activity (more female-typical) is important. Such research can test whether boundaries for male-typical (sexual) behavior are more strict than for female-typical behavior (Hort et al., 1990). Also, research on how people acquire traditional SDS-cognitions now is essential for designing future interventions that foster egalitarian sexual standards and sexual equality for men and women.

Conclusion

The current meta-analysis demonstrated that people on average still clearly have traditional cognitions about SDS, in particular with regard to men and women having casual sex, having sex for the first time at an early age, and general sexual activity level. We also found clear evidence of traditional SDS in within- or between-subject experimental studies assessing differences in the evaluation, or expectation, of men's and women's sexual behavior. Nevertheless, SDS were less traditional in countries with higher levels of gender equality. This meta-analysis further demonstrated that both evolutionary theory and biosocial theory provide relevant and testable predictions with regard to the existence of SDS. It appears that a hybrid model including both evolutionary processes related to gender differences in parental investment and sexual strategies, as well as the societal division in gender roles can best explain double standards for the sexual behavior of men and women. This meta-analysis also demonstrated the relevance of dual-process models of social cognition (Gawronski & Creighton, 2013) for the measurement and conceptualization of SDS. Therefore, we call for more research on the interplay between evolutionary and sociocultural processes underlying SDS, using implicit as well as explicit conceptualizations and measures that are able to assess the entire range of double standards, from reversed to traditional.

Acknowledgments

The authors would like to thank Joyce ter Heide and Anne Smit for their assistance with screening and coding the articles for this meta-analysis.


Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Joyce J. Endendijk  <https://orcid.org/0000-0001-8149-912X>

Supplemental Material

Supplemental material is available online with this article.

References

References marked with an asterisk indicate studies included in the meta-analyses.

- *Ali-Faisal, S. F. (2014). *Crossing sexual barriers: The influence of background factors and personal attitudes on sexual guilt and sexual anxiety among Canadian and American Muslim women and men* [Doctoral dissertation]. <https://scholar.uwindsor.ca/etd/5051/>
- *Allgeier, E. R., & Fogel, A. F. (1978). Coital position and sex roles: Responses to cross-sex behavior in bed. *Journal of Consulting and Clinical Psychology, 46*, 588–589. <https://doi.org/10.1037/0022-006X.46.3.588>
- *Alves, M. P., Amancio, L., & Alferes, V. R. (2008). Gender and social representations: Dual sexual pattern as a function of religion and political opinion. *Psicologia: Revista da Associação Portuguesa Psicologia, 22*(2), 139–160. <https://doi.org/10.17575/rpsicol.v22i2.351>
- *Askun, D., & Ataca, B. (2007). Sexuality related attitudes and behaviors of Turkish university students. *Archives of Sexual Behavior, 36*, 741–752. <https://doi.org/10.1007/s10508-007-9186-z>
- Assink, M., & Wibbelink, C. J. (2016). Fitting three-level meta-analytic models in R: A step-by-step tutorial. *The Quantitative Methods for Psychology, 12*(3), 154–174. <https://doi.org/10.20982/tqmp.12.3.p154>
- *Axinn, W. G., Young-DeMarco, L., & Ro, M. C. (2011). Gender double standards in parenting attitudes. *Social Science Research, 40*, 417–432. <https://doi.org/10.1016/j.ssresearch.2010.08.010>
- Bakermans-Kranenburg, M. J., van IJzendoorn, M. H., & Juffer, F. (2003). Less is more: Meta-analysis of sensitivity and attachment interventions in early childhood. *Psychological Bulletin, 129*, 195–215. <https://doi.org/10.1037/0033-2909.129.2.195>
- *Balanko, S. L. (2002). *Emancipatory sexuality education: An exploration of young women's perceptions of their sexuality and sexuality education influences* [Doctoral dissertation]. <https://scholar.uwindsor.ca/etd/1915/>
- Bandura, R. (2008). *A survey of composite indices measuring country performance*. United Nations Development Programme.
- *Bareket, O., Kahalon, R., Shnabel, N., & Glick, P. (2018). The Madonna-whore dichotomy: Men who perceive women's nurturance and sexuality as mutually exclusive endorse patriarchy and show lower relationship satisfaction. *Sex Roles, 79*, 519–532. <https://doi.org/10.1007/s11199-018-0895-7>
- *Barron, S. R. B. (2010). *Perceptions of sexual coercion: Understanding the role of gender, intimacy, and alcohol use* [Doctoral dissertation]. ProQuest Dissertations Publishing. (Order No. 3427331).
- Basow, S. A., & Rubin, L. R. (1999). Gender influences on adolescent development. In N. G. Johnson, M. C. Roberts, & J. Worell (Eds.), *Beyond appearance: A new look at adolescent girls* (pp. 25–52). American Psychological Association.
- *Bauman, K. E., & Wilson, R. R. (1976). Premarital sexual attitudes of unmarried university students: 1968 vs. 1972. *Archives of Sexual Behavior, 5*, 29–37. <https://doi.org/10.1007/BF01542238>

- Baumeister, R. F., & Twenge, J. M. (2002). Cultural suppression of female sexuality. *Review of General Psychology, 6*(2), 166–203. <https://doi.org/10.1037/1089-2680.6.2.166>
- *Bay-Cheng, L. Y., & Zucker, A. N. (2007). Feminism between the sheets: Sexual attitudes among feminists, nonfeminists, and egalitarians. *Psychology of Women Quarterly, 31*, 157–163. <https://doi.org/10.1111/j.1471-6402.2007.00349.x>
- *Bliss, B. A. (2013). *Is it still a sexual offense if society doesn't find it offensive? Cultural constructions and rape proclivity* [Doctoral dissertation]. <https://knowledge.library.iup.edu/etd/74>
- Bordini, G. S., & Sperb, T. M. (2013). Sexual double standard: A review of the literature between 2001 and 2010. *Sexuality & Culture, 17*, 686–704. <https://doi.org/10.1007/s12119-012-9163-0>
- Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2005). *Comprehensive meta-analysis* (Version 3). Biostat.
- Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2009). *Introduction to meta-analysis*. Wiley.
- Borenstein, M., Higgins, J. P., Hedges, L. V., & Rothstein, H. R. (2017). Basics of meta-analysis: I2 is not an absolute measure of heterogeneity. *Research Synthesis Methods, 8*, 5–18. <https://doi.org/10.1002/jrsm.1230>
- *Boyer, C. R., & Galupo, M. P. (2015). Attitudes toward individuals in same-sex and cross-sex encounters: Implications for bisexuality. *Journal of Bisexuality, 15*, 57–68. <https://doi.org/10.1080/15299716.2014.984371>
- Burgess, D., & Borgida, E. (1999). Who women are, who women should be: Descriptive and prescriptive gender stereotyping in sex discrimination. *Psychology, Public Policy, and Law, 5*, 665–692. <https://doi.org/10.1037/1076-8971.5.3.665>
- Buss, D. M. (1994). The strategies of human mating. *American Scientist, 82*, 238–249.
- Buss, D. M., & Schmitt, D. P. (1993). Sexual strategies theory: An evolutionary perspective on human mating. *Psychological Review, 100*, 204–232. <https://doi.org/10.1037/0033-295X.100.2.204>
- Bussey, K., & Bandura, A. (1999). Social cognitive theory of gender development and differentiation. *Psychological Review, 106*, 676–713. <https://doi.org/10.1037/0033-295X.106.4.676>
- Carli, L. L., LaFleur, S. J., & Loeber, C. C. (1995). Nonverbal behavior, gender, and influence. *Journal of Personality and Social Psychology, 68*, 1030–1041. <https://doi.org/10.1037/0022-3514.68.6.1030>
- *Caron, S. L., Davis, C. M., Halteman, W. A., & Stickle, M. (1993). Predictors of condom-related behaviors among first-year college students. *The Journal of Sex Research, 30*, 252–259. <https://doi.org/10.1080/00224499309551709>
- *Castaneda, D. M., & Collins, B. E. (1998). The effects of gender, ethnicity, and a close relationship theme on perceptions of persons introducing a condom. *Sex Roles, 39*, 369–390. <https://doi.org/10.1023/A:1018819025311>
- Cheung, M. W. L. (2014). Modeling dependent effect sizes with three-level meta-analyses: A structural equation modeling approach. *Psychological Methods, 19*(2), 211–229. <https://doi.org/10.1037/a0032968>
- *Clarke, M. J., Marks, A. D. G., & Lykins, A. D. (2015). Effect of normative masculinity on males' dysfunctional sexual beliefs, sexual attitudes, and perceptions of sexual functioning. *The Journal of Sex Research, 52*, 327–337. <https://doi.org/10.1080/00224499.2013.860072>
- Cohen, J. (1977). *Statistical power analysis for the behavioral sciences*. Academic Press.
- *Conley, T. D., Ziegler, A., & Moors, A. C. (2013). Backlash from the bedroom: Stigma mediates gender differences in acceptance of casual sex offers. *Psychology of Women Quarterly, 37*, 392–407. <https://doi.org/10.1177/0361684312467169>
- Costrich, N., Feinstein, J., Kidder, L., Marecek, J., & Pascale, L. (1975). When stereotypes hurt: Three studies of penalties for sex-role reversals. *Journal of Experimental Social Psychology, 11*, 520–530. [https://doi.org/10.1016/0022-1031\(75\)90003-7](https://doi.org/10.1016/0022-1031(75)90003-7)
- Crawford, M., & Popp, D. (2003). Sexual double standards: A review and methodological critique of two decades of research. *The Journal of Sex Research, 40*, 13–26. <https://doi.org/10.1080/00224490309552163>
- Cross, C. P., Cyrenne, D. L. M., & Brown, G. R. (2013). Sex differences in sensation-seeking: A meta-analysis. *Scientific Reports, 3*, Article 2486. <https://doi.org/10.1038/srep02486>
- *Danube, C. L., Norris, J., Stappenbeck, C. A., Cue Davis, K., George, W. H., Zawacki, T., . . . Abdallah, D. A. (2016). Partner type, sexual double standard endorsement, and ambivalence predict abdication and unprotected sex intentions in a community sample of young women. *The Journal of Sex Research, 53*, 601–613. <https://doi.org/10.1080/00224499.2015.1061631>
- DeLamater, J., & Friedrich, W. N. (2002). Human sexual development. *The Journal of Sex Research, 39*, 10–14. <https://doi.org/10.1080/00224490209552113>
- *Diéguez, J. L., Sueiro, E., & López, F. (2003). The sexual double standard y variables relacionadas. *Cuadernos de Medicina Psicosomática y Psiquiatría de Enlace, 67/68*, 79–88. <http://www.editorialmedica.com/download.php?idart=164>
- *Do, M., & Fu, H. (2010). Attitudes toward premarital sex in contemporary Vietnam: Findings from a national survey. *International Journal of Sexual Health, 22*, 103–118. <https://doi.org/10.1080/19317610903498487>
- *Doan, D. J. (2006). *Couples' sexuality: Relation to gender role attitudes and division of household labor* [Doctoral dissertation]. <http://library.alliant.edu/record=b1204263~S6>
- *Dollar, K. M., Perry, A. R., Fromuth, M. E., & Holt, A. R. (2004). Influence of gender roles on perceptions of teacher/adolescent student sexual relations. *Sex Roles, 50*, 91–101. <https://doi.org/10.1023/B:SERS.0000011075.91908.98>
- *Donald, A. R. (1983). Validation of the sexual double standard construct and a subtle double standard scale. *Dissertation Abstracts International, 46*(3), 957.
- Duval, S., & Tweedie, R. (2000a). A nonparametric “trim and fill” method for accounting for publication bias in meta-analysis. *Journal of the American Statistical Association, 95*, 89–98. <https://doi.org/10.1080/01621459.2000.10473905>
- Duval, S., & Tweedie, R. (2000b). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics, 56*, 455–463. <https://doi.org/10.1111/j.0006-341X.2000.00455.x>
- Dworkin, S. L., & O'Sullivan, L. (2005). Actual versus desired initiation patterns among a sample of college men: Tapping disjunctures within traditional male sexual scripts. *The Journal of Sex Research, 42*(2), 150–158. <https://doi.org/10.1080/00224490509552268>

- Eagly, A. H. (1987). *Sex differences in social behavior: A social-role interpretation*. Lawrence Erlbaum.
- Eagly, A. H., & Wood, W. (1999). The origins of sex differences in human behavior: Evolved dispositions versus social roles. *American Psychologist, 54*, 408–423. <https://doi.org/10.1037/0003-066X.54.6.408>
- *Eaton, A. A., & Matamala, A. (2014). The relationship between heteronormative beliefs and verbal sexual coercion in college students. *Archives of Sexual Behavior, 43*, 1443–1457. <https://doi.org/10.1007/s10508-014-0284-4>
- Egger, M., Davey-Smith, G., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal, 315*, 629–634. <https://doi.org/10.1136/bmj.315.7109.629>
- *Eisenman, R., & Dantzer, M. L. (2006). Gender and ethnic differences in sexual attitudes at a Hispanic-serving university. *Journal of General Psychology, 133*, 153–162. <https://doi.org/10.3200/GENP.133.2.153-162>
- *Emmerink, P. M. J., Van den Eijnden, R. J. J. M., Ter Bogt, T. F. M., & Vanwesenbeeck, I. (2017). A scale for the assessment of sexual standards among youth: Psychometric properties. *Archives of Sexual Behavior, 46*, 1699–1709. <https://doi.org/10.1007/s10508-017-1001-x>
- *Emmerink, P. M. J., Van den Eijnden, R. J. J. M., Vanwesenbeeck, I., & Ter Bogt, T. F. M. (2016). The relationship between endorsement of the sexual double standard and sexual cognitions and emotions. *Sex Roles, 75*, 363–376. <https://doi.org/10.1007/s11199-016-0616-z>
- *Emmerink, P. M. J., Vanwesenbeeck, I., Van den Eijnden, R. J. J. M., & Ter Bogt, T. F. M. (2016). Psychosexual correlates of sexual double standard endorsement in adolescent sexuality. *The Journal of Sex Research, 53*, 286–297. <https://doi.org/10.1080/00224499.2015.1030720>
- Endendijk, J. J., Groeneveld, M. G., Van Berkel, S. R., Hallers-Haalboom, E. T., Mesman, J., & Bakermans-Kranenburg, M. J. (2013). Gender stereotypes in the family context: Mothers, fathers, and siblings. *Sex Roles, 68*, 577–590. <https://doi.org/10.1007/s11199-013-0265-4>
- *Eriksson, J., & Humphreys, T. P. (2014). Development of the Virginity Beliefs Scale. *The Journal of Sex Research, 51*, 107–120. <https://doi.org/10.1080/00224499.2012.724475>
- Felson, R. B., & Feld, S. L. (2009). When a man hits a woman: Moral evaluations and reporting violence to the police. *Aggressive Behavior, 35*, 477–488. <https://doi.org/10.1002/ab.20323>
- Fetterolf, J. C., & Sanchez, D. T. (2015). The costs and benefits of perceived sexual agency for men and women. *Archives of Sexual Behavior, 44*, 961–970. <https://doi.org/10.1007/s10508-014-0408-x>
- *Ford, N. J., Vieira, E. M., & Villela, W. V. (2003). Beyond stereotypes of Brazilian male sexuality: Qualitative and quantitative findings from Sao Paulo, Brazil. *Culture, Health & Sexuality, 5*, 53–69. <https://doi.org/10.1080/713804640>
- Foschi, M. (2000). Double standards for competence: Theory and research. *Annual Review of Sociology, 26*, 21–42. <https://doi.org/10.1146/annurev.soc.26.1.21>
- Fugère, M. A., Escoto, C., Cousins, A. J., Riggs, M. L., & Haerich, P. (2008). Sexual attitudes and double standard: A literature review focusing on participant gender and ethnic background. *Sexuality & Culture, 12*, 169–182. <https://doi.org/10.1007/s12119-008-9029-7>
- *Garcia, L. T. (1982). Sex-role orientation and stereotypes about male-female sexuality. *Sex Roles, 8*, 863–876. <https://doi.org/10.1007/BF00287856>
- Gaunt, R. (2012). “Blessed is he who has not made me a woman”: Ambivalent sexism and Jewish religiosity. *Sex Roles, 67*, 477–487. <https://doi.org/10.1007/s11199-012-0185-8>
- Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychological Bulletin, 132*, 629–731. <https://doi.org/10.1037/0033-2909.132.5.692>
- Gawronski, B., & Creighton, L. A. (2013). Dual process theories. In D. E. Carlston (Ed.), *The Oxford handbook of social cognition* (pp. 282–312). American Psychologist Association. <https://doi.org/10.1093/oxfordhb/9780199730018.013.0014>
- Gentry, M. (1998). The sexual double standard: The influence of number of relationships and level of sexual activity on judgments of women and men. *Psychology of Women Quarterly, 22*, 505–511. <https://doi.org/10.1111/j.1471-6402.1998.tb00173.x>
- Gill, M. J. (2004). When information does not deter stereotyping: Prescriptive stereotyping can foster bias under conditions that deter descriptive stereotyping. *Journal of Experimental Social Psychology, 40*, 619–632. <https://doi.org/10.1016/j.jesp.2003.12.001>
- *Glass, J. C. (1972). Premarital sexual standards among church youth leaders: An exploratory study. *Journal for the Scientific Study of Religion, 11*, 361–367. <https://doi.org/10.2307/1384676>
- *Greene, K., & Faulkner, S. L. (2005). Gender, belief in the sexual double standard, and sexual talk in heterosexual dating relationships. *Sex Roles, 53*, 239–251. <https://doi.org/10.1007/s11199-005-5682-6>
- Greenwald, A. G., Banaji, M. R., Rudman, L. A., Farnham, S. D., Nosek, B. A., & Mellott, D. S. (2002). A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychological Review, 109*, 3–25. <https://doi.org/10.1037/0033-295X.109.1.3>
- Greenwald, A. G., Poehlman, T. A., Uhlmann, E. L., & Banaji, M. R. (2009). Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. *Journal of Personality and Social Psychology, 97*, 17–41. <https://doi.org/10.1037/a0015575>
- *Guo, Y. (2019). Sexual double standards in White and Asian Americans: Ethnicity, gender, and acculturation. *Sexuality & Culture, 23*, 57–95. <https://doi.org/10.1007/s12119-018-9543-1>
- *Haavio-Mannila, E., & Kontula, O. (2003). Single and double standards in Finland, Estonia, and St. Petersburg. *The Journal of Sex Research, 40*, 36–49. <https://doi.org/10.1080/00224490309552165>
- *Hackathorn, J., & Harvey, R. (2011). Sexual double standards: Bias in perceptions of cyber-infidelity. *Sexuality & Culture, 15*, 100–113. <https://doi.org/10.1007/s12119-010-9082-x>
- Han, J., & Kamber, M. (2006). *Data mining: Concepts and techniques* (2nd ed.). Morgan Kaufmann Publishers.
- *Hanassab, S., & Tidwell, R. (1993). Change in the premarital behavior and sexual attitudes of young Iranian women: From Tehran to Los Angeles. *Counselling Psychology Quarterly, 6*, 281–289. <https://doi.org/10.1080/09515079308254122>

- *Hearn, K. D. (2001). *Predictors of health protective sexual communication among black women and men* [Doctoral dissertation]. ProQuest Dissertations Publishing. (Order No. 3000960).
- Hill, C., & Kearl, M. A. (2011). *Crossing the line: Sexual harassment at school*. American Association of University Women.
- Hill, J. P., & Lynch, M. E. (1983). The intensification of gender-related role expectations during early adolescence. In J. Brooks-Gunn & A. C. Peterson (Eds.), *Girls at puberty* (pp. 201–228). Plenum Press.
- *Hong, J. (2013). *Sexuality professionals' attitudes toward five dimensions of human sexuality* [Doctoral dissertation]. ProQuest Dissertations Publishing. (Order No. 3574230).
- Hort, B. E., Fagot, B. I., & Leinbach, M. D. (1990). Are people's notion of maleness more stereotypically framed than their notions of femaleness? *Sex Roles*, 23, 197–212. <https://doi.org/10.1007/BF00289866>
- *Howell, J. L., Egan, P. M., Giuliano, T. A., & Ackley, B. D. (2011). The reverse double standard in perceptions of student-teacher sexual relationships: The role of gender, initiation, and power. *The Journal of Social Psychology*, 151, 180–200. <https://doi.org/10.1080/00224540903510837>
- Hyde, J. S. (2005). The gender similarities hypothesis. *American Psychologist*, 60, 581–592. <https://doi.org/10.1037/0003-066X.60.6.581>
- Hyde, J. S., & DeLamater, J. (1997). *Understanding human sexuality* (6th ed.). McGraw-Hill.
- Inglehart, R., & Norris, P. (2003). The true class of civilizations. *Foreign Policy*, 135, 62–70. <https://doi.org/10.2307/3183594>
- *Isarabhakdi, P. (1999). Sexual attitudes and behaviors of never-married, rural Thai youth. *Dissertation Abstracts International*, 58(5), 1932A.
- *Javier, K. N. (2018). *Asian American women: The influence of sororities and little sister programs* [Doctoral dissertation]. ProQuest Dissertations Publishing. (Order No. 10822876).
- *Jonason, P. K., & Marks, M. J. (2009). Common vs. uncommon sexual acts: Evidence for the sexual double standard. *Sex Roles*, 60, 357–365. <https://doi.org/10.1007/s11199-008-9542-z>
- *Jozkowski, K. N., & Ekbia, H. R. (2015). “Campus craft”: A game for sexual assault prevention in universities. *Games for Health Journal*, 4, 95–106. <https://doi.org/10.1089/g4h.2014.0056>
- *Kaats, G. R., & Davis, K. E. (1970). The dynamics of sexual behavior of college students. *Journal of Marriage and Family*, 32, 390–399. <https://doi.org/10.2307/350102>
- *Kettrey, H. H. (2016). What's gender got to do with it? Sexual double standards and power in heterosexual college hookups. *The Journal of Sex Research*, 53, 754–765. <https://doi.org/10.1080/00224499.2016.1145181>
- Kiefer, A. K., Sanchez, D. T., Kalinka, C. J., & Ybarra, O. (2006). How women's nonconscious association of sex with submission relates to their subjective sexual arousability and ability to reach orgasm. *Sex Roles*, 55, 83–94. <https://doi.org/10.1007/s11199-006-9060-9>
- *Kistler, M. E. (2011). *Media involvement and maternal communication effects on emerging adults' acceptance of the traditional sexual script* [Doctoral dissertation]. <http://hdl.handle.net/2376/3006>
- *Kleinfelter, K. J. (1999). *An investigation of college students' personal attitudes and subjective norms regarding acceptability of assertive and nontraditional sexual behavior in women: Re-examining the sexual double standard* [Doctoral dissertation]. ProQuest Dissertations Publishing. (Order No. 9939187).
- Knapp, G., & Hartung, J. (2003). Improved tests for a random effects meta-regression with a single covariate. *Statistics in Medicine*, 22, 2693–2710. <https://doi.org/10.1002/sim.1482>
- *Koon-Magnin, S., & Ruback, R. B. (2012). Young adults' perceptions of non-forcible sexual activity: The effects of participant gender, respondent gender, and sexual act. *Sex Roles*, 67, 646–658. <https://doi.org/10.1007/s11199-012-0201-z>
- Krahé, B., Scheinberger-Olwig, R., & Kolpin, S. (2000). Ambiguous communication of sexual intentions as a risk marker of sexual aggression. *Sex Roles*, 42, 313–337. <https://doi.org/10.1023/A:1007080303569>
- Kreager, D. A., Staff, J., Gauthier, R., Lefkowitz, E. S., & Feinberg, M. E. (2016). The double standard at sexual debut: Gender, sexual behavior and adolescent peer acceptance. *Sex Roles*, 75, 377–392. <https://doi.org/10.1007/s11199-016-0618-x>
- Lai, Y., & Hynie, M. (2011). A tale of two standards: An examination of young adults' endorsement of gendered and ageist sexual double standards. *Sex Roles*, 64, 360–371. <https://doi.org/10.1007/s11199-010-9896-x>
- Lakens, D. (2017). *TOSTER: Two one-sided tests (TOST) equivalence testing* (R Package Version 0.2.5). <https://cran.r-project.org/web/packages/TOSTER/index.html>
- Lee, I. C., Pratto, F., & Johnsen, B. T. (2011). Intergroup consensus/disagreement in support of group-based hierarchy: An examination of socio-structural and psycho-cultural factors. *Psychological Bulletin*, 137, 1029–1064. <https://doi.org/10.1037/a0025410>
- *Lefkowitz, E. S., Shearer, C. L., Gillen, M. M., & Espinosa-Hernandez, G. (2014). How gendered attitudes relate to women's and men's sexual behaviors and beliefs. *Sexuality & Culture*, 18, 833–846. <https://doi.org/10.1007/s12119-014-9225-6>
- *Levin, D. S. (2010). *Let's talk about sex . . . education: Exploring youth perspectives, implicit messages, and unexamined implications of sex education in schools* [Doctoral dissertation]. <http://hdl.handle.net/2027.42/75809>
- Lippa, R. A. (2009). Sex differences in sex drive, sociosexuality, and height across 53 nations: Testing evolutionary and social structural theories. *Archives of Sexual Behavior*, 38, 631–651. <https://doi.org/10.1007/s10508-007-9242-8>
- López-Sáez, M., & Lisbona, A. (2009). Descriptive and prescriptive features of gender stereotyping. Relationships among its components. *Revista de Psicología Social*, 24, 363–379. <https://doi.org/10.1174/021347409789050560>
- *Manalastas, E. J., & David, C. C. (2018). Valuation of women's virginity in the Philippines. *Asian Women*, 34, 23–48. <https://doi.org/10.14431/aw.2018.03.34.1.23>
- *Marks, M. J. (2008). Evaluations of sexually active men and women under divided attention: A social cognitive approach to the sexual double standard. *Basic and Applied Social Psychology*, 30, 84–91. <https://doi.org/10.1080/01973530701866664>
- *Marks, M. J., & Fraley, R. C. (2005). The sexual double standard: Fact or fiction? *Sex Roles*, 52, 175–186. <https://doi.org/10.1007/s11199-005-1293-5>
- *Marks, M. J., & Fraley, R. C. (2007). The impact of social interaction on the sexual double standard. *Social Influence*, 2, 29–54. <https://doi.org/10.1080/15534510601154413>

- *Marks, M. J., Young, T. M., & Zaikman, Y. (2019). The sexual double standard in the real world: Evaluations of sexually active friends and acquaintances. *Social Psychology, 50*, 67–79. <https://doi.org/10.1027/1864-9335/a000362>
- Milhausen, R. R., & Herold, E. S. (1999). Does the sexual double standard still exist? Perceptions of university woman. *The Journal of Sex Research, 36*, 361–368. <https://doi.org/10.1080/00224499909552008>
- *Milhausen, R. R., & Herold, E. S. (2001). Reconceptualizing the sexual double standard. *Journal of Psychology & Human Sexuality, 13*, 63–83. https://doi.org/10.1300/J056v13n02_05
- *Miller, K. R. (2012). *Influence of parental communication of sexual messages on late adolescent sexual assertiveness and sexual experience and the influence of adolescent adherence to the sexual double standard: An exploratory study* [Doctoral dissertation]. http://rave.ohiolink.edu/etdc/view?acc_num=osu1337093690
- Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Annals of Internal Medicine, 151*(4), 264–269. <https://doi.org/10.7326/0003-4819-151-4-200908180-00135>
- Mosher, C. E., & Danoff-Burg, S. (2005). Agentic and communal personality traits: Relations to attitudes toward sex and sexual experiences. *Sex Roles, 52*, 121–129. <https://doi.org/10.1007/s11199-005-1199-2>
- Moss-Racusin, C. A., Phelan, J. E., & Rudman, L. A. (2010). When men break the gender rules: Status incongruity and backlash against modest men. *Psychology of Men & Masculinity, 11*, 140–151. <https://doi.org/10.1037/a0018093>
- *Moyano, N., Monge, F. S., & Sierra, J. C. (2017). Predictors of sexual aggression in adolescents: Gender dominance vs. rape supportive attitudes. *European Journal of Psychology Applied to Legal Context, 9*, 25–31. <https://doi.org/10.1016/j.ejpal.2016.06.001>
- Muehlenhard, C. L., & Quackenbush, D. M. (1998). Sexual Double Standard Scale. In C. M. Davis, W. L. Yarber, R. Bauserman, G. Schreer, & S. L. Davis (Eds.), *Handbook of sexuality-related measures* (pp. 199–201). SAGE.
- *Murstein, B. I., & Mercy, T. (1994). Sex, drugs, relationships, contraception, and fears of disease on a college campus over 17 years. *Adolescence, 29*(114), 303–322.
- Oliver, M. B., & Hyde, J. S. (1993). Gender differences in sexuality: A meta-analysis. *Psychological Bulletin, 114*, 29–51. <https://doi.org/10.1037/0033-2909.114.1.29>
- *Ortiz, R. R. (2012). *Subverting the sexual double standard: How exposure to and identification with characters in televised counter sexual scripts affect young women's sexual well-being* [Doctoral dissertation]. ProQuest Dissertations Publishing. (Order No. 3512671).
- *Ortiz, R. R., White, S., & Rassmussen, E. (2016). Do individual perceptions matter in pornography effects? How perceived general acceptance and influence of pornography may impact agreement with sex-role attitudes. *Communication Research Reports, 33*, 88–95. <https://doi.org/10.1080/08824096.2015.1122583>
- O'Sullivan, L. F. (1995). Less is more: The effects of sexual experience on judgments of men's and women's personality characteristics and relationship desirability. *Sex Roles, 33*, 159–181. <https://doi.org/10.1007/BF01544609>
- O'Sullivan, L. F., & Allgeier, E. R. (1998). Feigning sexual desire: Consenting to unwanted sexual activity in heterosexual dating relationships. *The Journal of Sex Research, 35*, 234–243. <https://doi.org/10.1080/00224499809551938>
- *Papp, L. J., Hagerman, C., Gnoleba, M. A., Erchull, M. J., Liss, M., Miles-McLean, H., & Robertson, C. M. (2015). Exploring perceptions of slut-shaming on Facebook: Evidence for a reverse sexual double standard. *Gender Issues, 32*, 57–76. <https://doi.org/10.1007/s12147-014-9133-y>
- Penhollow, T. M., Young, M., & Nnaka, T. (2017). Alcohol use, hooking-up, condom use: Is there a sexual double standard? *American Journal of Health Behavior, 41*, 92–103. <https://doi.org/10.5993/AJHB.41.1.10>
- 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. <https://doi.org/10.1037/a0017504>
- Prentice, D. A., & Carranza, E. (2002). What women and men should be, shouldn't be, are allowed to be, and don't have to be: The contents of prescriptive gender stereotypes. *Psychology of Women Quarterly, 26*, 269–281. <https://doi.org/10.1111/1471-6402.t01-1-00066>
- *Quinn-Nilas, C., & Kennett, D. J. (2018). Reasons why undergraduate women comply with unwanted, non-coercive sexual advances: A serial indirect effect model integrating sexual script theory and sexual self-control perspectives. *The Journal of Social Psychology, 158*, 603–615. <https://doi.org/10.1080/00224545.2018.1427039>
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods (Vol. 1)*. SAGE.
- R Development Core Team. (2016). *R: A language and environment for statistical computing* (Version 3.5). R Foundation for Statistical Computing. <http://www.R-project.org>
- *Reid, J. A., Elliott, S., & Webber, G. R. (2011). Casual hookups to formal dates: Refining the boundaries of the sexual double standard. *Gender & Society, 25*, 545–568. <https://doi.org/10.1177/0891243211418642>
- Reiss, I. L. (1964). The scaling of premarital sexual permissiveness. *Journal of Marriage and Family, 26*, 188–198. <https://doi.org/10.2307/349726>
- *Robinson, I. E., & Jedlicka, D. (1982). Change in sexual attitudes and behavior of college students from 1965 to 1980: A research note. *Journal of Marriage and Family, 44*, 37–240. <https://doi.org/10.2307/351277>
- Rosenthal, R. (1995). Writing meta-analytic reviews. *Psychological Bulletin, 118*, 183–192. <https://doi.org/10.1037/0033-2909.118.2.183>
- Ruble, T. L. (1983). Sex stereotypes: Issues of change in the 1970s. *Sex Roles, 9*, 397–402. <https://doi.org/10.1007/BF00289675>
- Rudman, L. A. (1998). Self-promotion as a risk factor for women: The costs and benefits of counterstereotypical impression management. *Journal of Personality and Social Psychology, 74*, 629–645. <https://doi.org/10.1037/0022-3514.74.3.629>
- Rudman, L. A., & Fairchild, K. (2004). Reactions to counterstereotypic behavior: The role of backlash in cultural stereotype maintenance. *Journal of Personality and Social Psychology, 87*, 157–176. <https://doi.org/10.1037/0022-3514.87.2.157>

- *Rudman, L. A., Fetterolf, J. C., & Sanchez, D. T. (2013). What motivates the sexual double standard? More support for male versus female control theory. *Personality and Social Psychology Bulletin*, *39*, 250–263. <https://doi.org/10.1177/0146167212472375>
- *Sahl, D., & Keene, J. R. (2010). The sexual double standard and gender differences in predictors of perceptions of adult-teen sexual relationships. *Sex Roles*, *62*, 264–277. <https://doi.org/10.1007/s11199-009-9727-0>
- *Sakaluk, J. K., & Milhausen, R. R. (2012). Factors influencing university students' explicit and implicit sexual double standards. *The Journal of Sex Research*, *49*, 464–476. <https://doi.org/10.1080/00224499.2011.569976>
- Sanchez, D. T., Fetterolf, J. C., & Rudman, L. A. (2012). Eroticizing inequality in the United States: The consequences and determinants of traditional gender role adherence in intimate relationships. *The Journal of Sex Research*, *49*, 168–183. <https://doi.org/10.1080/00224499.2011.653699>
- Schmitt, D. P. (2015). The evolution of culturally-variable sex differences: Men and women are not always different, but when they are . . . it appears not to result from patriarchy or sex role socialization. In V. A. Weekes-Shackelford & T. K. Shackelford (Eds.), *The evolution of sexuality* (pp. 221–256). Springer.
- Schwartz, P., & Rutter, V. (2000). *The gender of sexuality* (2nd ed.). Pine Forge Press.
- *Shapurian, R., & Hojat, M. (1985). Sexual and premarital attitudes of Iranian college students. *Psychological Reports*, *57*, 67–74. <https://doi.org/10.2466/pr0.1985.57.1.67>
- *Sheeran, P., Spears, R., Abraham, C. S. S., & Abrams, D. (1996). Religiosity, gender, and the double standard. *Journal of Psychology: Interdisciplinary and Applied*, *130*, 23–33. <https://doi.org/10.1080/00223980.1996.9914985>
- Shen, A. C. T., Chiu, M. Y. L., & Gao, J. (2012). Predictors of dating violence among Chinese adolescents: The role of gender-role beliefs and justification of violence. *Journal of Interpersonal Violence*, *27*, 1066–1089. <https://doi.org/10.1177/0886260511424497>
- *Shin, K. R., Park, H., & Cha, C. (2011). Sex education during the school-aged years influences sexual attitudes and sexual health in college: A comparative study from Korea. *Nursing & Health Sciences*, *13*, 328–334. <https://doi.org/10.1111/j.1442-2018.2011.00622.x>
- *Sierra, J. C., Gutierrez-Quintanilla, R., Bermudez, M. P., & Buela-Casal, G. (2009). Male sexual coercion: Analysis of a few associated factors. *Psychological Reports*, *105*, 69–79. <https://doi.org/10.2466/PR0.105.1.69-79>
- *Sierra, J. C., Moyano, N., Vallejo-Medina, P., & Gómez-Berrocá, C. (2017). An abridged Spanish version of Sexual Double Standard Scale: Factorial structure, reliability and validity evidence. *International Journal of Clinical and Health Psychology*, *18*, 69–80. <https://doi.org/10.1016/j.ijchp.2017.05.003>
- *Sierra, J. C., Rojas, A., Ortega, V., & Ortiz, J. D. M. (2007). Evaluation of sexist attitudes with college students: First psychometric data of the Spanish versions of Double Standard Scale (DSS) and the Rape Supportive Attitude Scale (RSAS). *International Journal of Psychology & Psychological Therapy*, *7*(1), 41–60.
- Signorella, M. L., Bigler, R. S., & Liben, L. S. (1993). Developmental differences in children's gender schemata about others: A meta-analytic review. *Developmental Review*, *13*, 147–183. <https://doi.org/10.1006/drev.1993.1007>
- *Smith, G., Mysak, K., & Michael, S. (2008). Sexual double standards and sexually transmitted illnesses: Social rejection and stigmatization of women. *Sex Roles*, *58*, 391–401. <https://doi.org/10.1007/s11199-007-9339-5>
- *Soller, B., & Haynie, D. L. (2017). Variation in sexual double standards across schools: How do they matter for adolescent sexual behavior? *Sociological Perspectives*, *60*, 702–721. <https://doi.org/10.1177/0731121416668865>
- *Spreadbury, C. L. (1982). The “permissiveness with affection” norm and the labeling of deviants. *The Personnel and Guidance Journal*, *60*, 280–282. <https://doi.org/10.1002/j.2164-4918.1982.tb00663.x>
- *Sprecher, S. (1989). Premarital sexual standards for different categories of individuals. *The Journal of Sex Research*, *26*, 232–248. <https://doi.org/10.1080/00224498909551508>
- *Sprecher, S., McKinney, K., & Orbuch, T. L. (1987). Has the double standard disappeared? An experimental test. *Social Psychology Quarterly*, *50*, 24–31. <https://doi.org/10.2307/2786887>
- Sprecher, S., McKinney, K., & Orbuch, T. L. (1991). The effect of current sexual behavior on friendship, dating, and marriage desirability. *The Journal of Sex Research*, *28*, 387–408. <https://doi.org/10.1080/00224499109551615>
- *Sprecher, S., Treger, S., & Sakaluk, J. K. (2013). Premarital sexual standards and sociosexuality: Gender, ethnicity, and cohort differences. *Archives of Sexual Behavior*, *42*, 1395–1405. <https://doi.org/10.1007/s10508-013-0145-6>
- Stangor, C., & Schaller, M. (1996). Stereotypes as individual and collective representations. In N. C. Macrae, C. Stangor, & M. Hewstone (Eds.), *Stereotypes and stereotyping* (pp. 3–37). Guilford Press.
- Tabachnick, B. G., & Fidell, L. S. (2012). *Using multivariate statistics* (6th ed.). HarperCollins.
- *Thompson, A. E., Hart, J., Stefaniak, S., & Harvey, C. (2018). Exploring heterosexual adults' endorsement of the sexual double standard among initiators of consensually nonmonogamous relationship behaviors. *Sex Roles*, *79*, 228–238. <https://doi.org/10.1007/s11199-017-0866-4>
- Trivers, R. L. (1972). Parental investment and sexual selection. In B. Campbell (Ed.), *Sexual selection and the descent of man 1871-1971* (pp. 136–207). Aldine de Gruyter.
- *Ubillos, S., Goiburu, E., Puente, A., & Pizarro, J. P. (2016). Adaptation and validation of the Double Standard Scale in Basque adolescents [Adaptación y validación de la Escala de Doble Estándar en adolescentes vascas-os]. *Revista de Psicología Social*, *31*, 368–397. <https://doi.org/10.1080/02134748.2016.1152683>
- United Nations Development Program. (2017). *Human development report 2016*. http://hdr.undp.org/sites/default/files/2016_human_development_report.pdf
- Van den Noortgate, W., López-López, J. A., Marín-Martínez, F., & Sánchez-Meca, J. (2013). Three-level meta-analysis of dependent effect sizes. *Behavior Research Methods*, *45*, 576–594. <https://doi.org/10.3758/s13428-012-0261-6>
- van Zanten, B. T., Verburg, P. H., Koetse, M. J., & van Beukering, P. J. (2014). Preferences for European agrarian landscapes:

- A meta-analysis of case studies. *Landscape and Urban Planning*, 132, 89–101. <https://doi.org/10.1016/j.landurbplan.2014.08.012>
- Viechtbauer, W. (2015). *Meta-analysis package for R*. <https://cran.r-project.org/web/packages/metafor/metafor.pdf>
- *Walfield, S. M. (2018). “Men cannot be raped”: Correlates of male rape myth acceptance. *Journal of Interpersonal Violence*. Advance online publication. <https://doi.org/10.1177/0886260518817777>
- *Ward, L. M. (2002). Does television exposure affect emerging adults’ attitudes and assumptions about sexual relationships? Correlational and experimental confirmation. *Journal of Youth and Adolescence*, 31, 1–15. <https://doi.org/10.1023/A:1014068031532>
- *Ward, L. M., & Rivadeneyra, R. (1999). Contributions of entertainment television to adolescents’ sexual attitudes and expectations: The role of viewing amount versus viewer involvement. *The Journal of Sex Research*, 36, 237–249. <https://doi.org/10.1080/00224499909551994>
- Warner, K. (2000). Sentencing in cases of marital rape: Towards changing the male imagination. *Legal Studies*, 20, 592–611. <https://doi.org/10.1111/j.1748-121X.2000.tb00161.x>
- *Weaver, A. D., Claybourn, M., & MacKeigan, K. L. (2013). Evaluations of friends-with-benefits relationship scenarios: Is there evidence of a sexual double standard? *The Canadian Journal of Human Sexuality*, 22, 152–159. <https://doi.org/10.3138/cjhs.2128>
- Weis, K. G. (2009). “Boys will be boys” and other gendered accounts: An exploration of victims’ excuses and justifications for unwanted sexual contact and coercion. *Violence Against Women*, 15, 810–834. <https://doi.org/10.1177/1077801209333611>
- Weis, K. G. (2010). Too ashamed to report: Deconstructing the shame of sexual victimization. *Feminist Criminology*, 5, 286–310. <https://doi.org/10.1177/1557085110376343>
- Wiederman, M. W., & Allgeier, E. R. (1992). Gender differences in mate selection criteria: Sociobiological or socioeconomic explanation? *Ethology and Sociobiology*, 13, 115–124. [https://doi.org/10.1016/0162-3095\(92\)90021-U](https://doi.org/10.1016/0162-3095(92)90021-U)
- Wood, W., & Eagly, A. H. (2002). A cross-cultural analysis of the behavior of women and men: Implications for the origins of sex differences. *Psychological Bulletin*, 128, 699–727. <https://doi.org/10.1037/0033-2909.128.5.699>
- Wood, W., & Eagly, A. H. (2012). Biosocial construction of sex differences and similarities in behavior. In J. M. Olson & M. P. Zanna (Eds.), *Advances in experimental social psychology* (Vol. 46, pp. 55–123). Academic Press.
- World Economic Forum. (2017). *Global gender gap report 2017*. http://www3.weforum.org/docs/WEF_GGGR_2017.pdf
- *Young, M., Cardenas, S., Donnelly, J., & Kittleson, M. J. (2016). Perceptions of peer sexual behavior: Do adolescents believe in a sexual double standard? *Journal of School Health*, 86, 855–863. <https://doi.org/10.1111/josh.12455>
- *Young, M., Penhollow, T. M., & Bailey, W. C. (2010). Hooking-up and condom provision: Is there a double standard? *American Journal of Health Studies*, 25(3), 156–164.
- *Zaikman, Y., & Marks, M. J. (2014). Ambivalent sexism and the sexual double standard. *Sex Roles*, 71, 333–344. <https://doi.org/10.1007/s11199-014-0417-1>
- Zaikman, Y., & Marks, M. J. (2017). Promoting theory-based perspectives in sexual double standard research. *Sex Roles*, 76, 407–420. <https://doi.org/10.1007/s11199-016-0677-z>
- *Zaikman, Y., Marks, M. J., Young, T. M., & Zeiber, J. A. (2016). Gender role violations and the sexual double standard. *Journal of Homosexuality*, 63, 1608–1629. <https://doi.org/10.1080/00918369.2016.1158007>
- Zaikman, Y., Vogel, E. A., Vicary, A. M., & Marks, M. J. (2016). The influence of early experiences and adult attachment on the exhibition of the sexual double standard. *Sexuality & Culture*, 20, 425–455. <https://doi.org/10.1007/s12119-015-9332-z>
- *Zhang, Y., Miller, L. E., & Harrison, K. (2008). The relationship between exposure to sexual music videos and young adults’ sexual attitudes. *Journal of Broadcasting & Electronic Media*, 52, 368–386. <https://doi.org/10.1080/08838150802205462>
- *Zurbriggen, E. L., & Morgan, E. M. (2006). Who wants to marry a millionaire? Reality dating television programs, attitudes toward sex, and sexual behaviors. *Sex Roles*, 54, 1–17. <https://doi.org/10.1007/s11199-005-8865-2>