Differential Evaluation of Straight and Gay Men for Nonverbal Effeminate Behavior

Art D. Marsden
Syracuse University
Abstract

The purpose of this study was to determine how violation of gender-based expectancies might influence attitudes toward men who differ by sexual orientation (i.e., straight or gay). While other studies have examined attitudes toward gay and straight men who differ by gender expression, their designs may have been susceptible to demand effects; this study was specifically designed to avoid such methodological issues. This research was informed by Expectancy-Violation Theory (EVT) and the Black Sheep Effect, which together suggest that an effeminate straight man should be evaluated by other straight men more negatively than an effeminate gay man because the former target negatively violated expectations. Additionally, EVT suggests that a masculine gay man should be evaluated more positively than a masculine straight man because the former positively violated expectations. Self-identified straight men evaluated a male target whose sexual orientation and gender conformity were manipulated through a photo and vignette. Moderated mediation analyses were performed to determine if perceived expectancy violation mediated the relationship between sexual orientation and evaluations for both effeminate and masculine men. When sexual prejudice was used as a covariate, straight masculine targets were evaluated more favorably than gay masculine targets. Perceived expectancy violation did not mediate the relationship between sexual orientation and evaluations regardless of gender expression. More research should be conducted to identify the mechanism through which evaluations of straight and gay targets differ based on gender expression.

Keywords: Gender expression, sexual orientation, sexual prejudice
DIFFERENTIAL EVALUATION OF STRAIGHT AND GAY MEN FOR NONVERBAL EFFEMINATE BEHAVIOR

by

Art D. Marsden

B.S., University of North Texas, 2017

Thesis
Submitted in partial fulfillment of the requirements for the degree of Master of Science in Psychology.

Syracuse University
July 2022
Table of Contents

Introduction..................................................................................................................... 1
Expectancy-Violation Theory ......................................................................................... 4
The Black Sheep Effect ................................................................................................. 6
Integrating EVT and the BSE ......................................................................................... 6
Evidence for EVT and BSE Assumptions in the Sexual Prejudice Literature ............... 8
Gender Conformity Domains ....................................................................................... 12
Study Aims ................................................................................................................... 14
Hypotheses .................................................................................................................. 17
Method ........................................................................................................................ 19
Participants ................................................................................................................... 19
Procedure Overview .................................................................................................... 20
Targets ......................................................................................................................... 21
Measures ..................................................................................................................... 23
Primary Evaluations .................................................................................................... 23
Exploratory Evaluations .............................................................................................. 24
Perceived Expectancy Violation .................................................................................. 25
Sexual Prejudice .......................................................................................................... 26
Socially Desirable Responding .................................................................................... 26
Manipulation and Attention Check .............................................................................. 26
Results ........................................................................................................................ 27
Preliminary Analyses ................................................................................................... 27
Hypothesis Testing ....................................................................................................... 30
Exploratory Analyses .................................................................................................. 33
Exploring Evaluations Without a Covariate ................................................................ 33
Expectancy Violation ................................................................................................. 34
Certainty of John’s Sexual Orientation ....................................................................... 37
Suspicion Check ......................................................................................................... 38
Moral Disgust .............................................................................................................. 38
Warmth and Competence ......................................................................................... 39
Discussion .................................................................................................................... 39
Evidence for the Assumptions Informed by EVT and the BSE ................................. 40
List of Illustrative Materials

Summary of Reviewed Literature ................................................................. 49
Assumptions, Associated Hypotheses, Informing Theories, and Suggested Relationships .... 51
Sample Demographics \( (n = 196) \) ........................................................................ 52
Bivariate Correlations Between and Descriptive Statistics for Primary Variables of Interest .... 53
Estimated Marginal Means and Standard Deviations by Condition for Primary and Exploratory ANCOVA Models .................................................................................. 54
ANCOVA Results and Interaction Descriptives for Liking, Proximity, Global Evaluation, and Affiliation Models........................................................................................................ 55
Moderated Mediation Effects for Liking Model............................................................. 57
Interaction Plot for John’s Sexual Orientation and Gender Expression on the Estimated Marginal Means for Liking........................................................................................................... 58
Interaction Plot for John’s Sexual Orientation and Gender Expression on the Estimated Marginal Means for Proximity ........................................................................................................... 59
Interaction Plot for John’s Sexual Orientation and Gender Expression on the Estimated Marginal Means for Global Evaluation.................................................................................................... 60
Interaction Plot for John’s Sexual Orientation and Gender Expression on the Means for Perceived Expectancy Violation .......................................................................................... 61
Liking Moderated Mediation Model............................................................................. 62
Differential Evaluation of Straight and Gay Men for Nonverbal Effeminate Behavior

Gender conformity refers to adherence to the roles and expectations traditionally assigned to one’s gender (Parent et al., 2012). For males, violations of these expectations include wearing feminine clothing or makeup, speaking with a high or soft voice, swaying the hips while walking, excessively gesturing with the wrist or displaying a “limp wrist,” interests in stereotypically feminine hobbies (e.g., fashion) or occupations (e.g., florist), or a preference for the company of women (Levant et al., 2007; Lippa, 2002; Schatzberg et al., 1975; Taywaditep, 2002). In Western cultures, there is an emphasis on men adhering to masculine norms and the superiority of masculinity over femininity (Connell, 1987; Eguchi, 2009; Pleck et al., 1998). Indeed, the importance of acting masculine is reinforced at an early age; boys are consistently taught that they should “play with trucks, not dolls” (Schope & Eliason, 2004, p. 74) and otherwise conform to male gender expectations (Mahalik et al., 2003; Martin, 1990). Boys who violate these expectations are often discouraged by parents and peers by means of ostracism or verbal or physical abuse (Beard & Bakeman, 2001; Carter & McCloskey, 1983; Harry, 1989; Landolt et al., 2004; Smith & Leaper, 2005). Historically, gender nonconforming boys and transgender girls¹ have even been referred for clinical treatment of effeminacy (i.e., the expression of feminine traits in boys and men; Bartlett et al., 2000; Hooilim & Bottomley, 1983; Stoller, 1978; Wrate & Gulens, 1986), a practice which still occurs today in the forms of gender or sexual identity conversion therapy (Mallory et al., 2019; Turban et al., 2019). In adulthood, effeminate men may continue to experience discrimination and ostracism (Clarke & Smith,

¹Traditionally, there was no clinical distinction between an effeminate boy and a feminine transgender girl such that both were considered effeminate males (e.g., Wrate & Gulens, 1986). The use of “effeminacy” in this work instead matches the contemporary usage of the term, which is restricted to those who identify as male or masculine in gender.
For example, men who value their reputation as masculine may be reluctant to form friendships with effeminate men (Gul & Uskul, 2020).

Historically, effeminacy has been associated with gay men (Taywaditep, 2002); men who are effeminate are assumed to be gay and gay men are perceived to be less masculine (Blashill & Powlishta, 2009a). Similarly, men in stereotypically-feminine occupations (e.g., nursing, teaching) are often assumed to be gay (Dyck et al., 2009; Moss-Racusin & Johnson, 2016), while gay men are perceived as more suitable for these occupations than straight men but less suitable for stereotypically-masculine occupations (e.g., engineering; Rule et al., 2016). There is some truth to these stereotypes, as gay men are more likely to be gender nonconforming (Rieger et al., 2008; Skidmore et al., 2006; Taywaditep, 2002). However, not every man who violates gender norms (e.g., speaks with a high voice, wears make-up) identifies as gay or is sexually attracted to men (Beaver, 2015). While boyhood gender nonconformity may be associated with gay identification in adulthood, not every gender nonconforming boy grows up to identify as a gay, bisexual, or otherwise non-straight man; some grow up to identify as transgender women—who may additionally identify as straight, gay, or any other sexual orientation (Kuper et al., 2012)—while others identify as straight, cisgender men (Bailey & Zucker, 1995; Bartlett et al., 2000; Bryant, 2008). Additionally, not every gay man is effeminate; some gay men “defeminize” before reaching adulthood (Taywaditep, 2002), while others report they never violated gender norms (Bailey & Zucker, 1995). Indeed, some “straight-acting” gay men firmly adhere to male gender role expectations (Eguchi, 2009) and may even disparage other gay men who violate gender role norms (Miller & Behm-Morawitz, 2016; Taywaditep, 2002).

Thus, there is obviously no perfect association between gender conformity and sexual orientation such that all effeminate men are gay and all masculine men are straight, or that all
gay men are effeminate and all straight men are masculine. However, the effeminate gay man remains a prevalent stereotype (Mize & Manago, 2018; Steffens et al., 2019), and this may lead to gay and straight men being perceived differently when they are gender nonconforming. While gay men still face prejudice and discrimination, there has been a steady decrease in sexual prejudice over the past two decades (Poushter & Kent, 2020). Thus, it may be that some increasingly believe that “gay is okay” while still associating gay men with effeminacy. Even though an effeminate gay man may be considered a “double violator” of norms (i.e., violating expectations about both gender and sexual roles; Lehavot & Lambert, 2007), to those lower in sexual prejudice, effeminacy may be acceptable for gay men. In that way, a gay man’s combined gender and sexual identity (“gay man”) may be more salient to others than his gender identity alone (“man”), and this combined identity may be associated with a different set of expectations. On the other hand, an effeminate straight man would be unlikely to benefit from the association between gay men and effeminacy and would still be perceived as a gender norm violator. If so, this would suggest that an effeminate straight man would be evaluated more negatively—and may even face greater ostracism or other discriminatory outcomes—than an effeminate gay man.

Despite the existence of straight men who are effeminate (Beaver, 2015), there are few studies that focus on this population. Specific to this research, I have found only seven studies that sought to determine how evaluations of straight and gay men might differ as a function of gender conformity; however, as discussed later, the designs of these studies may have left them susceptible to demand effects or social desirability bias. The purpose of this study is to add to the scant literature in this area by attempting to clarify if effeminate straight and gay men are differentially evaluated by other men and how these evaluations are influenced by perceived violations of expectations. Specifically, this study will test three assumptions informed by two
theories that address the influence of norm violations on person evaluation: Expectancy-Violation Theory (assumptions 1 and 2) and the Black Sheep Effect (assumption 3).

**Expectancy-Violation Theory**

According to Expectancy-Violation Theory (EVT), when someone violates a stereotype-based expectation, they will be judged more extremely in the same “direction” (i.e., positively or negatively) as the violation than they would have if no violation had occurred (Jussim et al., 1987). In other words, someone who violates stereotype-based expectations by engaging in a behavior that is perceived negatively should be evaluated less favorably, while someone who violates expectations by engaging in a positively-received behavior should be evaluated more favorably. For example, Jackson et al. (1993) found that participants evaluated a White high school student’s abilities more negatively than a Black high school student’s when both had equally weak academic credentials; the authors suggest that, by performing worse than expected, the White student negatively violated expectations. Research has also found that men who are more communal (modest, cooperative, caring) are perceived as less suitable than equally communal women for roles that emphasize leadership (Bosak et al., 2018; Moss-Racusin et al., 2010). Likewise, women who violate expectations by being more agentic (self-reliant, assertive, independent) are rated as less suitable for jobs that benefit more from communality than agentic men (Rudman & Glick, 1999).

Because femininity is a negative trait for men to express (e.g., Pleck et al., 1998), when straight men are expected to be masculine, EVT suggests that they would be evaluated more negatively for effeminacy than a gay man (assumption 1). However, EVT also accounts for directionality such that, when someone positively violates a stereotype-based expectation, they should also be evaluated more positively than someone engaging in the same behavior but who
did not violate expectations. For example, when agentic women are considered for roles where agency is perceived as beneficial (e.g., leadership roles), they are considered to be a better fit than agentic men (Schaumberg & Flynn, 2017). Thus, when gay men are expected to violate gender norms (i.e., by being effeminate), EVT suggests they would be evaluated more positively for gender conformity than a straight man (assumption 2). For that reason, this research will also examine differences in evaluations of masculine straight and masculine gay men.

It is important to note that global evaluation (i.e., general positive or negative affective reactions) of violators in the EVT literature do not always mirror other judgments (e.g., Rudman & Glick, 1999, 2001; Schaumberg & Flynn, 2017). Indeed, it may be easier to detect expectancy violation effects through global evaluation rather than judgments that require trait inferences (Biernat et al., 1999). Thus, it may be that the current study’s groups of interest (masculine or effeminate gay or straight men) are evaluated differently across different judgment domains. Research informed by Fiske et al.’s (2002) stereotype content model suggests that gay men are perceived as warmer (e.g., friendly, helpful) but also as less competent (e.g., skilled, intelligent) than straight men (Mize & Manago, 2018), and effeminate gay men are similarly perceived as warmer but less competent than masculine gay men (Brooks et al., 2019; Sink et al., 2018). People who stereotype social groups as higher in warmth but lower in competence are more likely to socially exclude or avoid these groups but also to help them (Cuddy et al., 2007). However, because judgments of competence (but not warmth) require trait inferences (Biernat et al., 1999), it was unclear how judgments associated with social inclusion and helping would differ as a function of expectancy violation. For exploratory purposes, this research assessed not only liking, but also warmth, competence, social inclusion (specifically, physical proximity and affiliation), and helping.
The Black Sheep Effect

While we all tend to favor and trust those who belong to the same social group as us (i.e., ingroup members; Tajfel, 1982; Turner & Reynolds, 2010), research on the Black Sheep Effect (BSE) finds that when someone violates ingroup expectations, that person will be more negatively evaluated by highly-identified ingroup members than if they belonged to a different social group (i.e., an outgroup; Marques & Paez, 1994; Marques et al., 1988). For example, Belgians rated association football fans who started a riot more negatively when they imagined them as Belgian than when they imagined them as German (Marques et al., 1988, Experiment 3). This strategy serves to protect the status of the ingroup and one’s self-esteem by symbolically distancing the violator and their deviancy from the ingroup.

Thus, the BSE suggests that highly-identified straight men should evaluate effeminate straight men more negatively than effeminate gay men because the straight man violated ingroup norms (assumption 1). However, according to the BSE (or more specifically Social Identity Theory, from which it is derived; Tajfel, 1982), if the straight man does not violate ingroup norms (e.g., by being masculine), then other highly-identified straight men should evaluate him more positively than they would an outgroup member (e.g., a gay man) through inherent ingroup bias. Thus, even though EVT suggests a masculine gay man would receive more positive evaluations than a masculine straight man because the former positively violated expectations while the latter only met expectations (assumption 2), when these targets are evaluated by straight men, ingroup bias effects may lead to the masculine straight man being evaluated more positively than the masculine gay man (assumption 3).

Integrating EVT and the BSE
Despite the potential conflict noted above (assumptions 2 and 3) when comparisons are made between an outgroup member who positively violates a group expectancy (e.g., masculine gay men when evaluated by a straight man) and an ingroup member who does not violate ingroup norms (e.g., a masculine straight man), Biernat et al. (1999) propose that EVT and the BSE can be peacefully integrated. Indeed, the authors suggest that the two theories are actually “intimately connected” (p. 536) because both suggest that evaluations are influenced by violation of expectations—for the BSE, because we have favorable expectations for those who share an ingroup with us.

Using Jackson et al.’s (1993) design as an example, according to EVT, a White participant should evaluate a Black high school student more positively than a White student when both have equally strong academic credentials because the Black student positively violated expectations. However, according to the BSE, a White participant high in racial identification should judge the White student more positively than the Black student because the White student is a member of the participant’s ingroup (i.e., “White”). In this scenario, Jackson et al.’s (1993) results supported the former theory (i.e., EVT). According to Biernat et al. (1999), this is not actually a contradiction of either theory; instead, it may be that results like Jackson et al.’s (1993) are due to a clear violation of expectations having occurred, which may have been more salient to participants than the White target’s ingroup status. However, when expectancy violations are ambiguous (and thus less salient) and targets share an ingroup with the participant, we might observe more positive evaluations for targets through an ingroup bias effect. For these reasons, Biernat et al. (1999) recommend directly measuring perceived expectancy violation in EVT research.
Biernat et al. (1999) included such a measure in their own study looking at differential evaluations of White and Black targets. White participants who were high in racial identification did not rate Black targets more favorably than White targets when both performed equally well on a task, and these targets did not differ in terms of perceived expectancy violation. Instead, White targets were rated marginally more favorably, suggesting a slight ingroup bias effect (although this difference was non-significant). However, White targets who performed poorly on the task did violate expectations more and were evaluated more negatively than poor-performing Black targets. Because ingroup information is more personally-relevant, it is processed more systematically; this extends to information about social norm deviancy and may explain harsher derogation of ingroup deviants as suggested by the BSE (Coull et al., 2001; Reese et al., 2013). Thus, negative violations of stereotype-based expectancies may be more salient when committed by ingroup members. This suggests the strongest EVT effect may be observed for ingroup targets who negatively violate expectations (i.e., assumption 1), such as when straight men evaluate other straight men who defy gender norms. Additionally, ingroup bias effects may not be observed if gay men are perceived as clearly and positively violating expectations by being masculine (Biernat et al., 1999). Alternatively, if a masculine gay man does not clearly and positively violate expectations, the BSE suggests that the straight man would be evaluated more positively by highly-identified straight men (assumption 3).

**Evidence for EVT and BSE Assumptions in the Sexual Prejudice Literature**

Researchers have long sought to determine the role that effeminacy plays in negative attitudes toward gay men (e.g., MacDonald & Games, 1976; Storms, 1978). In attempting to parse these variables, researchers will sometimes present participants with gay male targets whose level of gender conformity is manipulated through vignettes or other stimuli. When these
study designs include a sexual orientation manipulation, it allows for comparisons between evaluations for gay and straight men who engage in gender nonconformity. Seven of such articles have been identified (for a summary of this literature, see Table 1).

Only one of these seven studies was directly informed by EVT. Gowen and Britt (2006) had their participants listen to an audiotape of a man who mentioned having a boyfriend or girlfriend and who spoke with a “gay” (with more dynamic intonation and sibilant /s/ sounds) or “straight” voice. Their participants indicated they would allocate less scholarship funding for straight targets than gay targets when both were effeminate (supporting assumption 1). This may reflect a greater desire to help effeminate gay men over effeminate straight men (perhaps because effeminate gay men were perceived as higher in warmth but lower in competence than effeminate straight men; Cuddy et al., 2007). However, participants did not significantly differ in their decisions to admit, nor did they desire more social distance from, effeminate gay men compared to effeminate straight men (contradicting assumption 1). Additionally, participants desired more social distance from masculine gay over masculine straight men (contradicting assumption 2 and supporting assumption 3). Unfortunately, because Gowen and Britt’s (2006) participants did not complete a measure of perceived expectancy violation, it is unclear where targets violated or met participants’ expectations.

The remaining six articles reviewed in this area also do not prima facie support this research’s assumptions re EVT and the BSE. For example, there is little in this literature to indicate that effeminate straight men would be evaluated more harshly because they violated expectations (assumption 1). In none of these articles were straight men evaluated significantly less favorably than gay men when both were effeminate. Additionally, while EVT might lead us to assume that masculine gay men positively violate expectations and would thus be evaluated
more positively than masculine straight men (assumption 2), this research has largely suggested the opposite (i.e., that participants rate straight men more favorably than gay men when both are masculine; Horn, 2007; Laner & Laner, 1979; Storms, 1978). This may be because these gay men did not clearly violate expectations by being masculine, and thus straight men were liked more due to ingroup bias (assumption 3; Biernat et al., 1999). Indeed, Laner and Laner’s (1979) participants estimated that only 16-25% of gay men in the general population are effeminate and 46-55% are “undetectable” (i.e., normatively masculine). It could be that, despite prevalent stereotypes (e.g., Parmenter et al., 2019), the existence of gay men who are largely gender conforming is widely known. Thus, while an effeminate gay man may be expected, a masculine gay man may not be unexpected.

It may also be that there is little support for either of these assumptions because sexual prejudice toward gay men overpowered any expectancy-violation effects—that is, perhaps gay targets’ sexual orientation played a larger role in their evaluations than whether they violated expectations or not. It should be noted that taken together, assumptions 1 and 2 suggest, perhaps implausibly, that regardless of whether a man is gender conforming or gender nonconforming, he should be evaluated more positively if he is gay than if he is straight. However, in this literature, gay men are consistently rated significantly less favorably than straight men regardless of gender conformity (Blashill & Powlishta, 2009b, 2012; Laner & Laner, 1979; Lehavot & Lambert, 2007; Storms, 1978). Four of the seven reviewed studies included a measure of sexual prejudice (Blashill & Powlishta, 2009b, 2012; Gowen & Britt, 2006; Lehavot & Lambert, 2007), but results were inconsistent between studies. Lehavot and Lambert (2007) found that,  

---

2 While Storms (1978) and Laner and Laner (1979) reported higher mean evaluations for straight masculine over straight gay men, because they were not interested in comparisons across sexual orientation, they did not report whether these differences were significant.
unsurprisingly, it was those high in sexual prejudice who most disliked and rated as immoral gay men over straight men and effeminate men over masculine men. Those who were low in sexual prejudice showed an opposite (although nonsignificant) pattern for liking, and rated masculine gay men as significantly less immoral than masculine straight men (partially supporting assumption 2). In contrast, Blashill and Powlishta (2009b) found that gay and effeminate men were less liked regardless of participants’ level of sexual prejudice. Gowen and Britt (2006) also found that sexual prejudice did not influence target evaluations, but was positively associated with a desire for greater social distance from gay and effeminate men. Finally, Blashill and Powlishta (2012) reported that there were no differences on any study variables by sexual prejudice. Overall, then, past research does not easily lend itself to the interpretation that sexual prejudice toward gay men simply overpowers any expectancy-violation effects.

It is worth noting that these four studies administered to participants the Attitudes toward Lesbians and Gay Men scale (ATLG; Herek, 1988). The ATLG measures overt prejudice (e.g., “I think male homosexuals are disgusting”), but because people have become more knowledgeable about gay men and gay rights issues, people may be less likely to endorse “blatantly homophobic items” such as those included in this measure (Morrison & Morrison, 2003; Raja & Stokes, 1998, p. 115). The lack of consistent results in this literature does not appear to be an issue of inadequate variance, as response means on the ATG were near the scale midpoint for two of the three studies that supplied this information (i.e., Blashill & Powlishta, 2009b; Gowen & Britt, 2006).³ Regardless, because negative attitudes toward gay men have become more subtle (Massey et al., 2013; Nadal, 2013), a measure that better taps into this “modern” prejudice is warranted to clarify how sexual prejudice may overpower EVT effects.

---

³ Lehavot and Lambert (2007) obtained a sample mean toward the low-prejudice end of the scale.
Gender Conformity Domains

The seven studies reviewed above (see Table 1) have targeted a variety of domains to manipulate gender conformity, and often combined traits and various elements of effeminacy within their vignettes (for a discussion on effeminacy domains, see Hennen, 2001). For example, Storms (1978) and Laner and Laner (1979) presented undergraduate students with vignettes in which gender conformity was manipulated through college major (fashion design vs. business), activity (dance vs. intramural sports), and appearance (dress: flowery shirts, tight-fitting European slacks, and stacked-heel shoes vs. sweaters, jeans, and hiking boots). Horn (2007) presented tenth and twelfth-graders with vignettes where targets were either gender conforming, gender nonconforming through appearance (GNC–appearance), or gender nonconforming through activity (GNC–activity). Male targets either played baseball (gender conforming, GNC–appearance) or performed ballet (GNC–activity) and either dressed and acted similarly to (gender conforming, GNC–activity) or differently than the other students at school (GNC–appearance). Blashill and Powlishta (2009b) presented undergraduate students with vignettes in which gender conformity was manipulated through desired occupation (engineer vs. dietician), activities (riding a motorcycle and shooting pool vs. gymnastics and baking cookies), and traits (leader and strong vs. affectionate and well-mannered). Finally, Blashill and Powlishta (2012) manipulated gender conformity through activities (goes fishing, builds with tools, fixes cars vs. bakes cookies, baby-sits, reads romance novels), traits (brave, adventurous, brags a lot vs. emotional, shy, talkative [GNC–trait]), and appearance (deep voice, broad shoulders, rough hands vs. hips sway when walking, plucked eye-brows, delicate).

It is understandable that these researchers would include multiple gender conformity domains in their stimulus materials (e.g., in the same vignette) in an attempt to ensure their
targets are perceived as sufficiently effeminate or masculine. Unfortunately, this makes it difficult to compare results between studies and to determine whether certain domains may elicit more of an expectancy-violation effect than others. However, the results from two of these studies may provide insight into how these domains are differentially evaluated. Horn (2007) found that, for both straight and gay men, GNC–appearance targets (those who dressed and acted differently from other students) were rated as less “acceptable” than GNC–activity targets. Blashill and Powlishta (2012) similarly found that GNC–appearance targets (hips sway when walking, plucked eye-brows, delicate) were rated more negatively than GNC–activity and GNC–trait targets. Thus, it may be that visible gender role violations of this type are more salient and lead to more negative evaluations than activity (e.g., baking cookies) or trait violations (e.g., emotional). Indeed, Hennen (2001) notes that it is the more-visible types of effeminacy that best fit the contemporary stereotype of the “effeminate homosexual” (p. 133).

Sánchez and Vilain (2012) provide further insight into which type of male gender nonconformity may be perceived as most egregious. In their study, gay men were asked to identify how important it was that they look (e.g., “your clothes, hair”) and behave (e.g., “your speech, mannerisms”) masculine while in public (p. 113). Both were considered important, but gay men least wanted to behave effeminately; this may have been the most salient marker of effeminacy for them. Because gay men (in the West) are conditioned to value masculinity through the same culture as straight men (Connell, 1987; Eguchi, 2009; Pleck et al., 1998), it would be unsurprising if behavior was also a more salient marker of effeminacy for straight men. Indeed, nonverbal behavior seems to be a particularly salient marker of sexual orientation.\footnote{Notably, nonverbal effeminacy was included in Blashill and Powlishta’s (2012) GNC–appearance vignette (i.e., “hips sway when walking”; p. 1296) and may have been inferred from part of Horn’s (2007) GNC–appearance vignette (i.e., “acts differently from most of the other guys at school”; p. 367). Thus, their appearance conditions may have been confounded with nonverbal behavior.}
people consistently identify male targets as gay when they walk (Johnson et al., 2007), gesture (Ambady et al., 1999), or speak in a stereotypically-feminine manner (Linville, 1998). For these reasons, this research focused on nonverbal behavior as the gender conformity domain.

**Study Aims**

This study sought to extend the EVT literature by examining if and how gay and straight men are differentially evaluated when they violate gender norm expectations through nonverbal behavior and to clarify how these evaluations are influenced by sexual prejudice. Additionally, it fills a gap in the literature by focusing on a rarely studied population: effeminate straight men. Although I am only aware of one study that was directly informed by EVT and examined differential evaluation of targets who vary by sexual orientation and gender conformity (i.e., Gowen & Britt, 2006), other literature in this area utilized methods that also allow for comparisons between this study’s groups of interest (see Table 1). These studies largely provided evidence that contradicted this study’s assumptions; however, because they sought to answer a different research question (i.e., what role does effeminacy play in sexual prejudice toward gay men?) and made little effort to control for demand effects, their methods were not best suited to testing the questions proposed in this work.

For example, no researchers (including Gowen & Britt, 2006) administered to their participants a measure of perceived expectancy violation, so it is unclear if any targets violated expectations. As recommended by Biernat et al. (1999), my study included such a measure to determine if targets violated expectations and would thus be expected to receive more extreme evaluations as predicted by EVT. Additionally, researchers either utilized a measure of overt (as opposed to subtle) sexual prejudice or none at all (see Table 1), which makes it difficult to determine the role sexual prejudice might play in target evaluations. This study instead included
a measure of modern sexual prejudice, which should more accurately capture contemporary negative attitudes toward gay men (Morrison & Morrison, 2003; Raja & Stokes, 1998). The inclusion of this measure allowed the examination of how sexual prejudice interacts with perceived expectancy violation. Researchers also often included different gender conformity domains in their manipulations (i.e., Blashill & Powlishta, 2009b; Horn, 2007; Laner & Laner, 1979; Storms, 1978); because these domains may differ in salience, this makes comparisons between these studies difficult. This study focused on one domain: nonverbal behavior, which should be more a more salient marker of gender expression than activities, traits (Blashill & Powlishta, 2012; Horn, 2007), and appearance (Sánchez & Vilain, 2012).

Researchers examining how gender conformity influences sexual prejudice have typically used liking or general global evaluation as their outcome measure (Blashill & Powlishta, 2009b, 2012; Laner & Laner, 1979; Lehavot & Lambert, 2007; Storms, 1978). Some have also included measures that may have greater real-world implications, such as willingness to work with the target on a task (Blashill & Powlishta, 2009b, 2012), scholarship funding allocation, or social distance (Gowen & Britt, 2006). It is likely these different domains would be correlated with one another, but they may be differentially influenced by sexual orientation and gender conformity; for example, Gowen and Britt (2006) found that effeminate gay men were not evaluated more positively than effeminate straight men, but participants chose to allocate more scholarship funding to them. Additionally, there is evidence in the EVT literature that evaluations of expectancy violators are not always positive or negative across all domains (e.g., Rudman & Glick, 1999, 2001; Schaumberg & Flynn, 2017). Thus, to examine potential differences by judgment domain, this study included a measure of liking as well as measures with greater real-world implications: those that measure willingness to be physically close to, to affiliate with, and
to help targets. Because gay men are often stereotyped in a way (i.e., higher in warmth, but lower in competence) that suggests people are more likely to socially neglect but also to help them (Brooks et al., 2019; Cuddy et al., 2007; Mize & Manago, 2018; Sink et al., 2018), these first two measures were included to broadly capture forms of ostracism, while the third measure was included to capture helping. Items were also included to directly capture perceived target warmth and competency.

Additionally, this research attempted to avoid demand effects that may have influenced the results of some of the reviewed studies. Demand effects refer to changes in participants’ responses due to cues that indicate the true purpose of the study; participants may respond in a way that they feel will confirm study hypotheses to “help” the researchers (Iyengar, 2011; Orne, 1962). One such cue may be using a within-subject design such that participants rate multiple manipulated targets (e.g., Horn, 2007; Laner & Laner, 1979). By receiving multiple targets that differ only in sexual orientation and/or gender conformity, participants may be able to discern that researchers expect different evaluations between targets based on these characteristics. To avoid this potential issue, this study used a between-subjects design such that each participant only evaluated one manipulated target. There may also be a benefit to including additional “filler” targets that differ from the primary target of interest and are not manipulated. While none of the reviewed studies utilized them, the inclusion of filler targets may also help reduce demand effects. If participants only receive one target to evaluate and its manipulated characteristics (i.e., sexual orientation, gender conformity) are thus particularly salient, it may make the purpose of the study clearer. Thus, this study included two filler targets to further obfuscate the true purpose of the study. Demand effects may also be cued by the stimulus materials themselves—for example, explicitly describing targets as “just like” or “differently from” other males when
manipulating gender conformity (Horn, 2007, p. 367). Alternatively, including a large list of feminine or masculine behaviors, even with the inclusion of some neutral behaviors (Lehavot & Lambert, 2007), may suggest to participants that gender conformity is a variable of interest. In this study, gender conformity was manipulated through nonverbal behaviors included in the target of interest’s vignette. Additional markers of gender conformity were provided through a photo manipulation in which participants are posed in a more masculine or effeminate manner.

Finally, because straight men report more negative attitudes toward gay men (Herek & McLemore, 2013; Ratcliff et al., 2006) and gender nonconformity (Adams et al., 2016; Levant et al., 2007) than straight women do, the attitudes of the former population have potentially greater implications for how gender-nonconforming men are evaluated and treated. Thus, this study’s primary population of interest was straight men. Studying this population allowed for indirect testing of the integration of the BSE with EVT with a smaller sample. If a target’s ingroup status (i.e., as a man or a straight man) is more salient to participants than an outgroup target’s positive expectancy violation (i.e., a gay man being masculine), then we would expect to find that straight men would evaluate masculine straight men more positively than masculine gay men (contrary to assumption 2 but consistent with assumption 3).

**Hypotheses**

This study directly tested two assumptions informed by EVT and the BSE. First, EVT and the BSE suggest that straight men who expect other straight men to be masculine should evaluate effeminate straight men more negatively than effeminate gay men for negatively violating stereotype and ingroup expectations (assumption 1; Jussim et al., 1987; Marques et al., 1988). Thus, I hypothesized that, \( H_{1a} \) when men are effeminate, straight men would be evaluated more negatively than gay men. Despite the existence of gender conforming gay men
and gender nonconforming straight men (e.g., Beaver, 2015; Laner & Laner, 1979), effeminacy is still primarily associated with gay (but not straight) men (Taywaditep, 2002). Thus, I hypothesized that, \((H_{1b})\) when men are effeminate, straight men would be perceived as more unexpected than gay men.

Second, EVT suggests that, when gay men are expected to be effeminate but are instead masculine, they should be evaluated more favorably than masculine straight men for positively violating expectations (assumption 2). Thus, I hypothesized that, \((H_{2a})\) when men are masculine, straight men would be evaluated more negatively than gay men. Because masculinity is associated with straight men, gay men (but not straight men) expressing masculinity should be perceived as violating expectations. Thus, I hypothesized that, \((H_{2b})\) when men are masculine, gay men would be perceived as more unexpected than straight men.

Alternatively, the BSE suggests that, because of ingroup bias, straight men would evaluate another straight man more positively than they would a masculine gay man (assumption 3). Because violations of ingroup norms are more salient for those who more strongly identify with their group (Marques & Paez, 1994), to directly evaluate assumption 3, this study would need to measure the extent to which participants identify as a man or a straight man. Because assumption 2 and 3 predict opposite relationships, and to simplify the research design, this study focused directly on testing assumption 2. However, because all study participants were straight men, more positive evaluations of masculine straight men than masculine gay men might suggest that ingroup bias effects—but not expectancy violation effects—influenced evaluations. Thus, I
also tested the hypothesis that \( H_3 \) when men are masculine, straight men would be evaluated more positively than gay men\(^5\).

According to EVT, because femininity is a negative trait for men to express, violations of expectations through effeminacy should be associated with more negative judgments. Additionally, because masculinity is a positive trait for men to express, violations of expectations through masculinity should be associated with more positive judgments. Thus, I hypothesized that \( H_4 \) perceived expectancy violation would mediate the relationship between sexual orientation and evaluations, and these relationships should be moderated by gender expression such that, for effeminate men, straight sexual orientation should be associated with more perceived expectancy violation, and in turn, these (negative) violations should be associated with less favorable evaluations. For masculine men, gay sexual orientation should be associated with more perceived expectancy violation, and in turn, these (positive) violations should be associated with more favorable evaluations. A summary of all hypotheses and their associated assumptions can be found in Table 2.

**Method**

**Participants**

Participants were initially recruited through the university’s psychology department research participation pool (SONA), where undergraduate students can participate in research in exchange for partial course credit. To be eligible to participate in this study, participants were required to be at least 18 years of age and to identify as a man\(^6\). To supplement my sample, the competing hypotheses \( H_{2a} \) and \( H_3 \) were tested through the same analysis (two-way ANCOVA). Support for \( H_{2a} \) would be indicated by more positive evaluations for masculine straight men over masculine gay men while support for \( H_3 \) would be indicated by more positive evaluations for masculine gay men over masculine straight men.

\(^5\) The competing hypotheses \( H_{2a} \) and \( H_3 \) were tested through the same analysis (two-way ANCOVA). Support for \( H_{2a} \) would be indicated by more positive evaluations for masculine straight men over masculine gay men while support for \( H_3 \) would be indicated by more positive evaluations for masculine gay men over masculine straight men.

\(^6\) While my population of interest was people who identify as straight men, I was unable to prescreen for sexual orientation through SONA. So that both samples would match one another in demographics, identification as straight was also not an inclusion criterion for my MTurk sample.
participants were also recruited through Amazon Mechanical Turk (MTurk) using the CloudResearch MTurk Toolkit. Inclusion criteria for the MTurk sample were the same as the SONA sample, but also included residence in the U.S. and a maximum age of 25 years to match the former sample more closely. I attempted to collect a sample of 300 participants; this would allow sufficient power to detect small-to-medium path coefficients through moderated mediation (Preacher et al., 2007). A total of 304 participants completed the survey ($n_{\text{SONA}} = 47$; $n_{\text{MTurk}} = 257$); the mean completion time was approximately 11 minutes. The final question of the survey asked participants if they consented to have their data used for research. Participants who indicated they wanted their data discarded ($n = 21$) or who did not answer this question ($n = 18$) were removed from analysis. Participants who did not meet inclusion criteria ($n = 30$) and who failed the attention check item (see below; $n = 1$) were also excluded from analysis. Finally, because straight men were my primary population of interest, participants who identified as anything other than straight ($n = 38$) were excluded from primary analysis, leaving a final sample of $n = 196$ (48% White, $M_{\text{age}} = 21.09$; see Table 3 for participant demographics).

**Procedure Overview**

Before data collection began, this study was approved by the university’s Institutional Review Board and the study’s hypotheses were preregistered online through the Open Science Framework\(^7\). All parts of the study were completed remotely through Qualtrics. Prior to their participation, all participants completed informed consent. Participants were then presented with photos and vignettes of three targets. Each vignette included the target’s name, age, race, occupation, and sexual orientation, as well as two behaviors they engage in (e.g., “He maintains eye contact when someone is talking to him”) and two of their preferences (e.g., “He likes to do

---

\(^7\) The pre-registration for this study is located at [https://osf.io/2jrp](https://osf.io/2jrp)
crossword puzzles”). After each vignette, participants answered several questions about the target.

The first two targets participants evaluated were filler targets designed to obfuscate the purpose of the study (“Marcus” and “Susan”; see below). Participants were then randomly assigned to one of four experimental conditions, in which the third and final target (“John”) was either straight and masculine \((n = 50)\), straight and effeminate \((n = 50)\), gay and masculine \((n = 51)\), or gay and effeminate \((n = 45)\). After John was evaluated, participants completed self-report measures including sexual prejudice, demographics, a suspicion check, and a clarifying question about their perceived expectancy violation ratings (see below). Participants were then presented with the study debriefing form through Qualtrics.

**Targets**

Marcus and Susan, the first two targets, were included to mitigate demand effects. If participants were presented solely with, for example, a straight and effeminate John, they might determine the true purpose of the study and respond differently than they otherwise would have (Iyengar, 2011; Orne, 1962). Marcus and Susan’s photos\(^8\) and vignettes were designed to include demographics other than those represented by John (e.g., Latino, older adult) for the purpose of obscuring the study’s actual focus on sexual orientation and gender conformity (for the Marcus and Susan materials, see Appendices A and B. Additionally, because participants evaluated these targets first, the true purpose of certain items (e.g., the sexual orientation manipulation check) should have been less obvious than if they were only presented with John, the final target.

John’s sexual orientation was explicitly manipulated through the vignette with information that he was straight or gay, while his gender conformity was manipulated through a

---

\(^8\) Images of Marcus and Susan were purchased from stock photo websites.
photo and the behaviors listed in his vignette (for the John materials, see Appendices C and D).

Effeminate behaviors for the photos and vignettes were adapted from a scale used to identify effeminacy in men (Schatzberg et al., 1975) and were paired with masculine analogues.

Participants received a photo of one of two different college-aged White men: one who was sitting on a bench or one who was standing. Both photo sets were manipulated such that the man was posed in a more masculine (i.e., sitting back with legs spread apart or standing with feet apart and hands on hips) or effeminate manner (i.e., sitting up straight with legs tightly crossed or standing with arms and legs crossed). Two separate sets of masculine and effeminate nonverbal behaviors were created to manipulate gender conformity and participants were randomly presented with one of these sets. Effeminate John either “gestures flamboyantly when he gets excited” and “speaks with a soft voice,” or he “sways his hips as he walks” and “has a soft handshake.” Likewise, masculine John either “gestures assertively when he gets excited” and “speaks with a deep voice,” or he “swaggers his shoulders as he walks” and “has a firm handshake.” These behaviors were pretested among undergraduate students in a social psychology class (N = 126), and as expected, the masculine behaviors were rated as significantly more masculine than the effeminate behaviors.

The purpose of creating two separate sets of photos and behaviors was to enhance the construct validity of the study (see Wells & Windschitl, 1999); by including multiple exemplars of “masculinity” and “effeminacy,” it is more likely that scores on dependent measures differ as a function of these constructs and not due to a more general dislike of specific behaviors (e.g., not a dislike of effeminate men, but of men with a soft handshake). All other information in the

---

9 “John” models were hired through the freelance marketplace website Fiverr and were instructed how to pose; to wear neutral, moderately casual clothing; and to maintain a neutral expression across photographs. Each model provided me with sets of approximately 20–30 photos; photos were chosen that best represented the requested poses and which remained consistent in framing and background content across manipulations.
John vignettes was chosen to be as gender-neutral as possible (e.g., college student, likes to play cards; the latter adapted from Blashill & Powlishta, 2009b) and remained constant between conditions.

**Measures**

**Primary Evaluations**

To measure liking, two items were adapted from Lehavot and Lambert (2007): “What are your general impressions of John?”⁵⁰ (from 1 = Not at all favorable to 7 = Very favorable) and “How much do you think you would like John if you were to meet him?” (from 1 = Not at all to 7 = Very much). Scores on these two items (r = .86) were averaged to create an overall score of liking.

In addition to liking, this study included three additional measures related to social inclusion: willingness to be near, to affiliate with, and to help John. To measure participants’ willingness to be near John, two items were adapted from Schope and Eliason (2004): “How willing would you be to sit next to John on a bus?” and “How willing would you be to share an office with John at work?” Two items were created to measure willingness to affiliate with John: “How willing would you be to be seen in public by strangers while hanging out with John?” and “How willing would you be to let your friends see you hanging out with John?” To measure willingness to help John, three items were loosely adapted from McGuire (1994): “If John looked like he needed a pen and you had an extra one, how willing would you be to let him borrow it?”; “If you saw John struggling to use a parking meter you're familiar with, how willing would you be to help by giving him advice?”; and “If you saw John drop $5 without realizing it, how willing would you be to tell him or return it to him?” These helping items were

---

⁵⁰ Because the evaluative scores of interest are those toward John and not the other targets, all example target evaluation items include John’s name. In the actual survey, names matched the target being evaluated.
chosen to represent casual, substantial, and emergency helping behaviors, respectively, as identified by McGuire (1994). Participants responded to all seven of these items on a 7-point scale, from 1 = Very unwilling to 7 = Very willing. For each measure, applicable items were averaged to create an overall score: willingness to be near (proximity; \( r = .80 \)), willingness to affiliate with (affiliation; \( r = .90 \)), and willingness to help John (helping; \( \alpha = .87 \)).

All nine primary evaluation items were positively correlated with one-another at \( p < .001 \) with correlations ranging from .52 to .90. Thus, an overall composite score was created from these items to measure global evaluation toward John. Because one more item was used to measure helping than the other three primary evaluations, the averages for each evaluation composite score were used to calculate this variable (\( \alpha = .88 \)).

**Exploratory Evaluations**

Lehavot and Lambert (2007) measured participants’ perceptions of their targets’ immorality; for participants who were low in sexual prejudice, masculine gay men were rated as less immoral than masculine straight men (partially supporting assumption 2). To further explore this relationship, I included a 7-point semantic differential item to measure immorality (moral–immoral). Because moral condemnation is strongly associated with feelings of disgust (e.g., see Schnall et al., 2008), I also included a 7-point semantic differential item to measure disgust (pleasant–disgusting). These two items were correlated at \( r = .64 \) \( p < .001 \); thus, for exploratory purposes, these items were averaged to create an overall measure of moral disgust.

Gowen and Britt’s (2006) participants allocated less scholarship funding for an effeminate straight target than an effeminate gay target, which may reflect a greater desire to help the latter target. Gay men and effeminate gay men are perceived as warmer but less competent than straight and masculine gay men, respectively (Brooks et al., 2019; Mize &
Manago, 2018); when groups are stereotyped this way, they are more likely to be socially excluded but also to be helped (Cuddy et al., 2007). Thus, four 7-point semantic differential items were included to measure target warmth (cold–warm, friendly–unfriendly) and competence (competent–incompetent, unintelligent–intelligent). Items were reverse-scored where applicable and averaged to create overall scores of target warmth ($r = .50$) and competence ($r = .55$).

**Perceived Expectancy Violation**

To measure perceived expectancy violation (as recommended by Biernat et al., 1999), I included three semantic differential scale items created for this study. All items were paired with 7-point scales with the adjectives on polar ends: unique–ordinary, unusual–conventional, and expected–unexpected. The first two of these items were reverse-scored and all items were averaged so that higher scores indicated more perceived violation of expectations (expectancy violation; $\alpha = .68$). To obscure the study’s hypotheses, these items were presented along with all other semantic differential items (see below).

To avoid demand effects, these items did not explicitly target perceived expectations based on gender (non)conforming behavior. Because of this, it may be that participants rated targets as violating expectations due to target characteristics unrelated to gender conformity. To account for this, at the very end of the survey, participants were once again presented with the John target they received earlier in the survey as well as their responses on one perceived expectancy violation item: expected–unexpected. We chose this item because it was the most face-valid item in our measure of perceived expectancy violation. Participants were asked to indicate through an open-ended item why they rated the target as they did on the expected–unexpected item. A second item followed on the next page, again paired with the John target and their response to the expected–unexpected item. This time, participants were asked to indicate
through a multiple-choice item if they were primarily influenced to rate the target as they did due to the target’s gender, sexual orientation, both, or neither.

**Sexual Prejudice**

The 12-item Modern Homonegativity Scale for Gay Men (MHS-G; Morrison & Morrison, 2003) was used to measure modern sexual prejudice toward gay men. The MHS-G captures more modern, subtle attitudes toward gay men (e.g., “Gay men no longer need to protest for their rights”). Participants responded to all items on a 7-point scale from 1 = Strongly Disagree to 7 = Strongly Agree. Items were averaged to create an overall score of sexual prejudice toward gay men (sexual prejudice; α = .94).

**Socially Desirable Responding**

Because participants may have evaluated gay or effeminate targets differently out of a desire to avoid appearing prejudiced, the 16-item Balanced Inventory of Desirable Responding Short Form (BIDR-16; Hart et al., 2015) was included to measure socially desirable responding. Participants responded to all items (e.g., “I am a completely rational person.”) on a 7-point scale from 1 = Strongly Disagree to 7 = Strongly Agree. Items were averaged to create an overall score of socially desirable responding (α = .81).

**Manipulation and Attention Check**

One item was included as a manipulation check for gender conformity: a 7-point feminine–masculine semantic differential. Although the vignettes describing John’s behavior were pretested, the photos also used to manipulate masculinity and effeminacy were not included in pretesting. Thus, it was important to ensure that masculine and effeminate targets were still perceived as masculine and feminine, respectively.
Following completion of the semantic differential items, participants were asked to recall and report the target’s age, race, occupation, and sexual orientation through multiple-choice items. For example, the sexual orientation item asks, “What is John’s sexual orientation?”; response options include, straight, gay, bisexual, pansexual, asexual. This item was intended to serve as a manipulation check for sexual orientation. The primary purpose of the additional recall items were to further obscure the study’s hypotheses and avoid potential demand effects.

Because participants may have interpreted John’s stated sexual orientation differently depending on John’s gender expression, we included a follow-up item after the main survey that asked participants, “Which statement best represents your initial thoughts about this person's listed sexual orientation?” Participants responded on a 5-point scale from 1 = I was certain it was UNTRUE to 5 = I was certain it was TRUE.

Results

Preliminary Analyses

Preliminary analyses were conducted to ensure there were no violations of test assumptions. An examination of histograms for all continuous variables revealed that scores for proximity, affiliation, helping, and global evaluation were negatively skewed. Skewness and kurtosis coefficients were within -1 and 1 for proximity, affiliation, and global evaluation, while kurtosis was within this range for helping; this indicates the distributions for these variables were suitably normal. However, skewness for helping was -1.32. I performed data transformations to determine which (if any) best reduced the absolute value of skewness for helping; a reflected square root transformation reduced this value to .86. This transformed variable was retained for all applicable analyses.
To provide evidence that the gender conformity manipulation was effective, an independent samples t-test was conducted to compare feminine–masculine ratings of effeminate and masculine targets. Masculine John \((M = 4.79, SD = 1.41)\) was rated as significantly more masculine than effeminate John \((M = 3.78, SD = 1.31, t(194) = -5.21, p < .001)\). Because evaluations of masculinity are likely to differ as a function of both gender expression and stated sexual orientation, a two-way ANOVA was also performed where John’s sexual orientation \((1 = \text{gay}, 0 = \text{straight})\), John’s gender expression \((1 = \text{effeminate}, 0 = \text{masculine})\), and the interaction between these two variables were entered as predictors and the feminine–masculine semantic differential item was entered as the outcome. John’s sexual orientation \((F[1,196] = 16.82, p < .001, \text{partial } \eta^2 = .08)\), John’s gender expression \((F[1,196] = 31.73, p < .001, \text{partial } \eta^2 = .14)\), and the interaction between these \((F[1,196] = 13.13, p < .001, \text{partial } \eta^2 = .06)\) all significantly predicted masculinity ratings. Straight/masculine John \((M = 5.50, SD = 1.20)\) was perceived as most masculine, followed by gay/masculine John \((M = 4.10, SD = 1.25)\), straight/effeminate John \((M = 3.82, SD = 1.40)\), and gay/effeminate John \((M = 3.73, SD = 1.21)\). While masculine John was perceived as more masculine regardless of his sexual orientation, there was a greater difference in masculinity ratings between masculine and effeminate John when he was straight. Thus, evaluations of John’s masculinity were influenced by his gender expression as well as his sexual orientation.

Images and vignettes for each target were presented on each target evaluation page except for the final page, on which participants were asked to recall information. Including these recall items for the first two targets (Marcus, Susan) served the purpose of motivating participants to pay attention to target information before they reached the target of interest (John). Any participants who were attending to the photos and vignettes should have been able to
easily identify John’s race (White) and sexual orientation (straight or gay, depending on condition). However, the sexual orientation question (“What is John’s sexual orientation?”) may have been interpreted differently by participants depending on their condition—for example, those who were presented with an effeminate, straight John may have believed the question was asking about John’s “actual,” but not listed, sexual orientation. For this reason, only identification of John’s race was retained as an attention check item; one participant who failed to correctly identify John’s race was removed from analysis. Notably, only two additional participants would have been excluded from analysis if we retained sexual orientation as an attention check item.

Correlations between all primary study variables were examined (see Table 4). Sexual prejudice scores were significantly associated with all judgment domains at \( p < .001 \): negatively with liking \((r = -.38)\), proximity \((- .37)\), affiliation \((- .40)\), helping \((- .29)\), global evaluation \((- .41)\), warmth \((- .20)\), and competence \((- .34)\); and positively with disgust \((.33)\), immorality \((.37)\), and moral disgust \((.39)\). However, as could be expected, correlations between sexual prejudice and judgment domains differed by condition. For example, liking was uncorrelated with sexual prejudice in the straight/masculine condition \((r = -.23, p = .10)\), but negatively correlated in the gay/masculine \((r = -.42, p = .002)\), straight/effeminate \((- .51, p < .001)\), and gay/effeminate conditions \((- .34, p = .023)\). This general pattern, where sexual prejudice is less strongly correlated with variables in the straight/masculine condition than in the other three conditions, is also apparent for proximity, warmth, immorality, disgust, and moral disgust. Because these relationships differ by condition and because past research (e.g., Lehavot & Lambert, 2007) has
found differences in the evaluation of targets based on sexual prejudice, sexual prejudice was entered as a covariate in all following analyses that included judgment domains.\textsuperscript{11}

Socially desirable responding scores correlated only with warmth ($r = .19$, $p = .008$) and sexual prejudice scores ($r = .18$, $p = .011$). These correlations were also separately examined by condition. In the straight/masculine condition, socially desirable responding was positively associated with warmth ($r = .43$, $p = .002$). In the gay/effeminate condition, socially desirable responding scores were positively associated with affiliation ($r = .31$, $p = .04$) and warmth ($r = .32$, $p = .037$). Socially desirable responding was uncorrelated with all judgment domains in the remaining conditions. Thus, socially desirable responding was not entered as a covariate in any of the following models.

\textbf{Hypothesis Testing}

To test the hypotheses $H_{1a}$ (effeminate straight men should be evaluated more negatively than effeminate gay men), $H_{2a}$ (masculine straight men should be evaluated more negatively than masculine gay men), and $H_{3}$ (masculine straight men should be evaluated more positively than masculine gay men), five two-way ANCOVA analyses were conducted where each of the four judgment domains (liking, proximity, affiliation, helping) and the global evaluation scores were entered as the outcome variable and sexual prejudice scores were entered as a covariate. In each model, John’s sexual orientation (1 = gay, 0 = straight), John’s gender expression (1 = feminine, 0 = masculine), and the interaction between John’s sexual orientation and gender expression were included as the predictors.

\textsuperscript{11}The sexual prejudice measure was included for exploratory purposes and was not initially intended to be used as a covariate. To avoid demand effects, sexual prejudice was measured after our manipulation; this means that the target presented to participants may have influenced participants’ scores on this variable, rendering it unsuitable as a covariate. To examine this, a two-way ANOVA was conducted where John’s sexual orientation, John’s gender expression, and the interaction between these two were entered as predictors and sexual prejudices scores as the outcome. John’s sexual orientation ($p = .125$), gender expression ($p = .298$), and the interaction between these ($p = .759$) did not significantly predict sexual prejudice.
For the liking model, sexual prejudice \((F[1,195] = 57.52, p < .001, \text{partial } \eta^2 = .15)\) and the interaction between John’s gender expression and John’s sexual orientation \((F[1,233] = 11.99, p = .008, \text{partial } \eta^2 = .04)\) were significant. Simple main effects were examined to explore this interaction. When targets were effeminate, straight John (estimated marginal mean \([EMM] = 4.44\)) was evaluated more negatively than gay John \((EMM = 4.87)\); however, this comparison was nonsignificant \((p = .113)\). When targets were masculine, straight John \((EMM = 4.97)\) was evaluated more positively than gay John \((EMM = 4.41, p = .031)\), providing support for \(H_3\) over \(H_2a\) (see Figure 1).

For the proximity model, sexual prejudice \((F[1,195] = 66.32, p < .001, \text{partial } \eta^2 = .15)\) and the interaction between Johns’ gender expression and John’s sexual orientation \((F[1,195] = 8.16, p = .045, \text{partial } \eta^2 = .02)\) were significant. Once again, when targets were effeminate, straight men \((EMM = 5.06)\) were evaluated more negatively than gay men \((EMM = 5.24)\), but this comparison was nonsignificant \((p = .525)\). When targets were masculine, straight men \((EMM = 5.40)\) were evaluated significantly more positively than gay men \((EMM = 4.77, p = .027; \text{see Figure 2})\).

For the global evaluation model, sexual prejudice \((F[1,194] = 44.80, p < .001, \text{partial } \eta^2 = .19)\) and the interaction between Johns’ gender expression and John’s sexual orientation \((F[1,194] = 4.48, p = .036, \text{partial } \eta^2 = .02)\) were significant. As with the previous two models, when targets were effeminate, straight men \((EMM = 5.01)\) were evaluated more negatively than gay men \((EMM = 5.27)\), but when targets were masculine, straight men \((EMM = 5.43)\) were evaluated more positively than gay men \((EMM = 5.07; \text{see Figure 3})\). However, neither of these comparisons were significant \((p = .102 \text{ and } p = .178, \text{respectively})\).

\(^{12}\) Using listwise deletion removed one participant with missing data for the helping and global evaluation variables.
For the remaining two models, sexual prejudice was the only significant predictor of affiliation \((F[1,195] = 37.98, p < .001, \text{partial } \eta^2 = .17)\) and helping \((F[1,194] = 18.13, p < .001, \text{partial } \eta^2 = .09)\). However, estimated marginal means for these models follow a similar pattern such that the straight/effeminate John was evaluated more negatively than the gay/effeminate John, while the straight/masculine John was evaluated more positively than the gay/masculine John (see Table 5). For the full results for all ANCOVA models, see Table 6.

To test the remaining hypotheses \((H_{1b}, H_{2b}, H_4)\), five moderated multiple mediation analyses were performed for each of the four primary judgment domains (liking, proximity, affiliation, helping) and the global evaluation scores through the PROCESS macro (model 59, 10,000 bootstrapping resamples; Hayes, 2017) for R. In each model, John’s sexual orientation (1 = gay, 0 = straight) was entered as the predictor, expectancy violation as the mediator, and sexual prejudice as a covariate. John’s gender expression (1 = effeminate, 0 = masculine) was entered as a moderator for all paths.

John’s sexual orientation \((b = 1.05, p < .001, 95\% \text{ CI [0.63, 1.48]})\), John’s gender expression \((b = .69, p = .002, 95\% \text{ CI [0.26, 1.12]})\), and the interaction between these \((b = -.72, p = .023, 95\% \text{ CI [-1.33, -0.10]})\) predicted perceived expectancy violation (a path; overall model, \(F(4,190) = 7.97, p < .001, R^2 = .14\)). Gay John was perceived as more unexpected \((EMM_{\text{masculine}} = 4.15, EMM_{\text{effeminate}} = 4.12)\) than straight John \((EMM_{\text{masculine}} = 3.09, EMM_{\text{effeminate}} = 3.78)\) regardless of gender expression, contradicting \(H_{1b}\). Effeminate John was also perceived as more unexpected than masculine John regardless of sexual orientation. However, the difference in perceived expectancy violation was greater between the two gender expression conditions when John was straight than when John was gay (see Figure 4).
Expectancy violation (b paths) did not predict liking \( (p = .972) \), proximity \( (p = .760) \), affiliation \( (p = .475) \), helping \( (p = .544) \), or global evaluation \( (p = .99) \) regardless of gender expression. The 95% confidence intervals for the indices of moderated mediation for all models contained zero; thus, there is no evidence for moderated mediation for any of the five models. Expectancy violation was not a mediator of the relationship between sexual orientation and any of the five judgment domains, regardless of gender expression, meaning that hypothesis \( H_4 \) was also unsupported. For the moderated mediation results for the liking model, see Table 7 and Figure 5.\(^{13}\)

**Exploratory Analyses**

*Exploring Evaluations Without a Covariate*

For exploratory purposes, the analyses used to test hypotheses \( H_{1a}, H_{2a}, \) and \( H_3 \) were performed again, but this time without sexual prejudice entered as a covariate (i.e., five two-way ANOVAs). For the liking model, the interaction between John’s sexual orientation and John’s gender expression was significant, \( F(1,195) = 4.36, p = .037, \) partial \( \eta^2 = .02 \). Means followed a similar pattern as the liking model that included sexual prejudice as a covariate; when targets were effeminate, straight John \( (M = 4.35, SD = 1.53) \) was evaluated more negatively than gay John \( (M = 4.79, SD = 1.49) \), but when targets were masculine, straight John \( (M = 4.94, SD = 1.36) \) was evaluated more positively than gay John \( (M = 4.52, SD = 1.36) \). However, an examination of simple main effects revealed that no cell comparisons were significant. There were no significant predictors in the remaining four models (proximity, affiliation, helping,

---

\(^{13}\) When participants whose identities include sexual orientations other than straight are included for primary analyses, results follow similar overall patterns as when the sample is limited to only straight participants. However, for the ANCOVA models, the interaction between John’s sexual orientation and John’s gender expression only significantly predicts (at \( p < .05 \)) liking.
global evaluation). Thus, there was no support for any hypotheses when sexual prejudice was not included in the models as a covariate.

**Expectancy Violation**

It is possible that expectancy violation did not predict any of the five judgment domains because the expectancy violation items did not explicitly mention the target’s sexual orientation or gender expression. Thus, it may be that participants were not thinking of these demographics when responding to the items. To assess this, I examined data for the two items where participants indicated why they rated the target as they did on the expected–unexpected item. If participants did not generally indicate that John’s gender expression, sexual orientation, or the combination of these influenced their responses to this item, then that may indicate that the expectancy violation items poorly assessed their intended construct. If expectancy violation influenced more negative evaluations of straight/effeminate John than gay/effeminate John, as predicted, then I would particularly expect participants in the straight/effeminate condition to endorse these characteristics through these two items.

The first of these items was open-ended. Responses were coded by two researchers (the author and a research assistant) to determine whether participants indicated their expected–unexpected rating was influenced by John’s behavior, sexual orientation, or gender. Three dichotomous variables were created for each characteristic, where 1 = characteristic was mentioned in response and 0 = characteristic was unmentioned. Inter-rater reliability was assessed with Cohen’s κ; agreement was high for all characteristics (behavior [κ = .73], sexual orientation [.96], and gender [.89]).

---

14 For behavior, a 1 indicated mention of any of John’s manipulated vignette or photo behaviors (e.g., the way he was standing) or any words such as “behavior” or “act.” For sexual orientation, a 1 indicated mention of John as “gay” or “straight” or words such as “sexual orientation” or “sexuality.” For gender, a 1 indicated mention of the target as a “man” or “male” or words such as “gender” or “sex”).
Overall, as an explanation for their rating on the expected–unexpected item, 25% of participants mentioned John’s behavior, 18.9% mentioned his sexual orientation, and 21.9% mentioned his gender. Percentages were examined by condition for those who rated John above the midpoint (> 4) on this item, indicating an evaluation of John as unexpected. Participants who viewed John as unexpected comprised only 28.1% of the overall sample (n = 55). Of these, 8 belonged to the masculine/straight condition, 17 to the masculine/gay condition, 19 to the effeminate/straight condition, and 20 to the effeminate/gay condition. Forty percent of these 55 participants mentioned John’s behaviors, 36.4% mentioned his sexual orientation, and 29.1% mentioned his gender. These characteristics were mentioned by at least half of participants in a condition only once: four of eight participants in the masculine/straight condition mentioned John’s behavior as an explanation for their rating of John as unexpected.

Overall, it seems that John’s gender, sexual orientation, or the interaction between these did not clearly drive a majority of participants’ evaluations of him as unexpected. Examination of responses to this open-ended item among those in the masculine/gay, effeminate/straight, and effeminate/gay conditions did not provide much insight into which other characteristics may have informed participants evaluations. For example, some participants mentioned components that were constant across conditions like John’s job (“I typically thought that people who are more nerdy tend to work in offices”) or hobbies (“Due to that he likes to play cards, which many people do not often do”), but many provided vague responses (“He seemed unique and not like an everyday person”; “I just thought John seemed unpredictable and unique”). No pattern was evident from these data that would indicate any other single characteristic was consistently influencing evaluations of John as unexpected.
Responses were also examined for the follow-up multiple-choice item which asked if participants rated John as they did on the expected–unexpected item because of his gender, sexual orientation, both, or neither. Cross-tabulations were examined between participant condition (straight/masculine, gay/masculine, straight/effeminate, gay/effeminate) and the expected–unexpected multiple-choice follow-up item, where participants indicated whether their responses to this item were influenced by the target’s gender, sexual orientation, both, or neither. Four cells contained counts less than five, so a chi-squared test was not performed on these variables. A majority of participants chose neither (62.2%); the remaining participants chose sexual orientation (17.3%), both (16.8%), and gender (3.6%). A majority of participants in the straight/masculine (72%), straight/effeminate (72%), and gay/effeminate (55.6%) indicated neither John’s sexual orientation nor gender influenced their response to the expected–unexpected item; in the gay/masculine condition, just under half (49%) chose this option.

A new variable was created where *neither* was coded as 0 and any other response (gender, sexual orientation, both) were coded as 1; when cross-tabulated with condition, these variables now met the assumption for minimum cell count. A chi-squared test for independence was significant, $\chi^2(3, n = 196) = 8.70, p = .034$. Of participants who indicated their responses to the expected–unexpected question were influenced by John’s sexual orientation, gender, or both, 18.9% were assigned to the straight/masculine condition, 35.1% to the gay/masculine condition, 18.9% to the straight/effeminate condition, and 27% to the gay/effeminate condition.

A second variable was created where responses to the expected–unexpected semantic differential item were dichotomized into responses above the midpoint (> 4; “unexpected”) and responses at or below the midpoint. A chi-squared test for independence between this variable and condition was nonsignificant. Of participants who responded above the midpoint on the
expected–unexpected semantic differential item \((n = 55)\), only 14.5% were assigned to the straight/masculine condition, while 27.3% were assigned to the gay/masculine condition, 27.3% to the straight/effeminate condition, and 30.9% to the gay/effeminate condition. Overall, patterns indicate that there were differences by condition—particularly between the straight/masculine and remaining conditions—but participants reported being largely uninfluenced by John’s sexual orientation or gender when responding to the expected–unexpected item.

*Certainty of John’s Sexual Orientation*

The vast majority of participants retained for analysis correctly identified John’s stated sexual orientation\(^{15}\) (99%). However, there is still a risk that participants may have interpreted John’s sexual orientation as incorrect. Participant responses were examined by condition to the follow-up item that assessed participant’s certainty that John’s sexual orientation was true. A one-way ANOVA was conducted with condition as the predictor and the sexual orientation certainty variable as the outcome. The overall model was significant, \(F = 9.56, p < .001\). Post-hoc comparisons were conducted with Tukey’s Honest Significant Difference test. Participants were less confident in John’s stated sexual orientation in the straight/effeminate condition \((M = 2.96, SD = 1.09)\) compared to all other conditions: straight/masculine \((M = 3.88, SD = 0.96, p < .001)\), gay/masculine \((M = 3.67, SD = 0.95, p = .002)\), and gay/effeminate \((M = 3.84, SD = 0.89, p < .001)\). No other comparisons were significant. In other words, participants were significantly less likely to feel certain that John’s sexual orientation was true when John was straight but effeminate than any other combination of sexual orientation and gender expression.

\(^{15}\) Two participants incorrectly identified John’s sexual orientation, one from the straight/effeminate condition and one from the gay/effeminate condition.
Suspicion Check

A single open-ended item was included as a suspicion check after all evaluations of John but before follow-up measures (i.e., sexual prejudice, socially desirable responding) and demographics. Participants were asked to indicate what they thought the purpose of the study was. Responses to this item were examined to determine how well the study avoided demand effects. Only one participant (straight/effeminate condition) mentioned effeminacy, masculinity or femininity, or specific gender expression manipulation behaviors included in the vignette (“[. . . ] for example the last one John being described somewhat effeminately”).

Twenty-three participants mentioned sexual orientation in their responses to this item. However, 19 of these mentioned sexual orientation along with other demographics (e.g., “How we perceive people based on sexuality, race, age, etc.”). Four participants either mentioned sexual orientation and no other demographics (e.g., “views on gays”) or explicitly mentioned gay or straight sexual orientation in their responses (e.g., “I’m assuming this study was to gauge people’s attitudes towards different types of people [for example some were straight and one was gay].”). Three of these participants were assigned to the gay/effeminate condition and one to the gay/masculine condition. Overall, responses to this item do not suggest many participants were aware of the study’s hypotheses.

Moral Disgust

A two-way ANCOVA was conducted with John’s sexual orientation, John’s gender expression, and the interaction between these entered as the predictor variables; moral disgust as the outcome variable; and sexual prejudice as a covariate. Sexual prejudice was the only significant predictor of moral disgust, $F[1,195] = 36.45, p < .001$, partial $\eta^2 = .16$. A second two-way ANCOVA was conducted where the immorality item was instead the outcome variable.
Sexual prejudice ($F[1, 195] = 33.37, p < .001, \text{partial } \eta^2 = .15$) and John’s sexual orientation ($F = 4.156, p = .043, \text{partial } \eta^2 = .02$) both predicted immorality ratings. Gay John ($EMM = 3.52, SD = 1.35$) was rated as more immoral than Straight John ($EMM = 3.18, SD = 1.17$) regardless of gender expression.

**Warmth and Competence**

Measures of warmth and competence were included for exploratory reasons; research has found that gay men and effeminate gay men are perceived as warmer but less competent than straight men and masculine gay men, respectively (Brooks et al., 2019; Mize & Manago, 2018). An additional two 2-way ANCOVAs were conducted with John’s warmth and competence entered as outcome variables. For each model, John’s sexual orientation, gender expression, and the interaction between these were entered as the predictor variables and sexual prejudice was entered as a covariate. For the warmth model, only John’s gender expression ($F[1, 195] = 10.68, p = .003, \text{partial } \eta^2 = .05$) and sexual prejudice ($F[1, 195] = 10.39, p < .001, \text{partial } \eta^2 = .05$) were significant predictors. For the competence model, only sexual prejudice ($F[1, 195] = 30.51, p < .001, \text{partial } \eta^2 = .12$) was a significant predictor. Thus, there is no evidence that gay men were perceived as warmer but less competent than straight men or that effeminate gay men were perceived as warmer and less competent than masculine gay men; only that effeminate men were perceived as warmer than masculine men.

**Discussion**

The purpose of this study was to investigate how straight men evaluate other men who differ in gender expression (masculine or effeminate) and sexual orientation (straight or gay). Specifically, this study was informed by Expectancy Violation Theory (EVT) and the Black Sheep Effect (BSE). While the assumptions informed by these two theories receive little support
from the sexual prejudice literature (see Table 1), studies in this area may have used methodologies that were susceptible to demand effects or were otherwise unsuitable for testing this study’s primary research question. Thus, a study was warranted that looked at these relationships through methodology designed to mitigate demand effects, measure expectancy violation and modern sexual prejudice, and focus on a potentially more-salient marker of gender expression (nonverbal behavior). Also included in this study were multiple judgment domains.

**Evidence for the Assumptions Informed by EVT and the BSE**

Taken together, EVT and the BSE suggest that straight men should evaluate an effeminate straight man more negatively than an effeminate gay man because the former negatively violated expectations (assumption 1). However, participants did not like effeminate/straight John less, nor were they less willing to be near him, affiliate with him, or help him than effeminate/gay John. Thus, $H_{1a}$ (informed by assumption 1) was unsupported.

Additionally, EVT suggests that a masculine gay man should be evaluated more positively than a masculine straight man because the former positively violated expectations (assumption 2), while the BSE suggests that straight men should evaluate a masculine straight man more positively than a masculine gay man because the former shares an ingroup (assumption 3). These opposing assumptions were tested. Participants liked masculine/straight John more and were more willing to be near him than masculine/gay John. Although nonsignificant, this pattern—where masculine/straight John was evaluated more positively than masculine/gay John—was also evident for global evaluation of John and willingness to affiliate with and help him. Thus, there was support for $H_3$ (informed by assumption 3) but not $H_{2a}$ (assumption 2); when he was masculine, straight men evaluated a straight man more positively than a gay man. While this result is contrary to the second assumption informed by EVT, this
may be because participants did not perceive masculine behavior as unexpected for gay men.

Over four decades ago\(^\text{16}\), Laner and Laner’s (1979) participants estimated that only 16-25% of gay men are effeminate and 46-55% are normatively masculine. Thus, masculine/gay John may not have clearly violated expectations, which is necessary to observe EVT effects (Biernat et al., 1999). Instead, it may be that John’s ingroup status (when he was straight) was more salient to participants, resulting in more positive evaluations as predicted by the BSE.

Because effeminacy is associated with gay but not straight men (e.g., Parmenter et al., 2019), then behaving effeminately should be perceived as a violation of expectations for straight men. If evaluations of John were driven by a violation of expectations, then effeminate/straight John should have been perceived as more unexpected than effeminate/gay John, who should not have violated gender norms due to his sexual orientation. However, participants reported that effeminate/gay John was more unexpected than effeminate/straight John, meaning that \(H_{1b}\) was not supported. Additionally, because masculine/gay John positively violated gender norms while masculine/straight John did not violate gender norms in either direction, EVT suggests that masculine/gay John should be perceived as more unexpected. As predicted, masculine John was perceived as more unexpected when he was gay than when he was straight, supporting \(H_{2b}\).

However, perceived expectancy violation was not significantly associated with any of the primary evaluations and did not mediate the relationship between sexual orientation and evaluations for masculine or effeminate men. Thus, \(H_{4}\) was unsupported; while effeminate/straight John was evaluated more negatively than effeminate/gay John and masculine/gay John was evaluated more positively than masculine/straight John, there is no indication that these results were because targets violated participants’ expectations.

\(^{16}\) I have been unable to find a more-recent study examining people’s perceptions of the prevalence of effeminate or gender-conforming gay men.
This may be because the measure of expectancy violation used in this study did not explicitly mention the target’s sexual orientation or gender expression to avoid demand effects. This is counter to Biernat et al. (1999), who specifically asked participants whether their confederate partner’s actual performance during a game violated participants’ expectations about how their partner would perform. Thus, my participants may have evaluated whether targets violating expectations or not due to target characteristics unrelated to gender expression or sexual orientation. In this study, when asked to indicate why they rated John as they did on the unexpected–expected semantic differential item, a majority of participants indicated that they were not influenced by John’s sexual orientation or gender expression across three of the four conditions (straight/masculine, straight/effeminate, gay/effeminate); in the fourth condition (gay/masculine), this was reported by just under half of participants. However, an examination of open-ended data did not suggest there was a dominant alternative explanation for participants’ evaluations of John as unexpected. Overall, it seems that John’s sexual orientation or gender were the characteristics that most influenced participants’ ratings of John as unexpected, but these characteristics did not seem to influence a majority of participants’ responses.

It may also be that the vignette and photo manipulations for this study were too subtle. Effeminate John, while perceived as less masculine than masculine John, may not have been perceived as effeminate enough to be unexpected for participants, regardless of John’s sexual orientation. The photo manipulations may have been particularly subtle; for the effeminate John manipulations, models were instructed to sit with legs crossed or stand with arms and legs crossed (see Appendix C). It may be that participants did not interpret these poses as effeminate.

Alternatively, it may simply be that the BSE is solely influencing this study’s results and that expectancy violation did not influence participant’s evaluations of straight/effeminate John.
Overall, gender and sexual orientation did interact to influence evaluations of John, but not as predicted by EVT when John was masculine. Additionally, while this study’s results supported assumption 1, this assumption was informed by not just EVT, but also the BSE; straight/effeminate John may have been liked less than gay/effeminate John only because the former violated ingroup norms. However, as I did not include a measure of identification as a man or a straight man, I cannot examine how these evaluations differ as a function of the strength of these identities.

**Contribution to the Sexual Prejudice Literature**

While it is still unclear how evaluations of this study’s targets were influenced by expectancy violation, ingroup bias, or ingroup deviancy, this study adds to the literature that examines how evaluations of men differ as a function of their sexual orientation and gender expression. When targets were masculine, this study’s participants liked a gay man less and were less willing to be near him than a straight man. Indeed, mean scores for gay/masculine John were lower across all evaluations than for any other target. It may be that participants liked gay/masculine John less because his masculine behavior, while not unexpected, was perceived as inauthentic given his sexual orientation. The masculine vignette behaviors chosen for this study (e.g., swaggering shoulders, speaking with a deep voice) may have been perceived by participants as exaggeratedly masculine. When masculine John was gay, he may have been perceived as “overcompensating” for his sexual orientation, leading to participants holding more negative attitudes toward him.

The finding that a straight man was evaluated more positively than a gay man when both are masculine is contrary to the reviewed literature; in none of the studies did they report a significant difference between this study’s targets of interest on measures of liking. However,
there were differences on other judgment domains. For example, Gowen and Britt’s (2006) participants allocated more scholarship funding to an effeminate gay man over an effeminate straight man. Allocating more scholarship funding may better represent a greater desire to help these targets than their liking of them; however, in this study, there was no difference in participants’ willingness to help targets regardless of sexual orientation and gender expression. Lehavot and Lambert (2007) had participants evaluate targets on their immorality. When targets were effeminate, participants high (but not low) in sexual prejudice evaluated the gay man as more immoral than the straight man. However, when targets were masculine, low (but not high) prejudice participants rated the straight man as more immoral than gay man. This study included a measure of immorality for exploratory purposes, but gay targets were evaluated as more immoral than straight targets regardless of gender expression. Because moral condemnation is associated with feelings of disgust (e.g., Schnall et al., 2008), this study also included an item to measure disgust toward John. Responses on this item were combined with morality ratings to create an overall measure of moral disgust. Only sexual prejudice predicted this variable; this aligns with research that has found that disgust sensitivity is associated with negative attitudes toward gay men (Hodson et al., 2013; Olatunji, 2008).

Overall, this study’s results do not match that of the reviewed sexual prejudice literature; however, there is also no consensus within this literature on how straight men or gay men are evaluated differently when they are masculine and when they are effeminate (see Table 1). Additionally, any discrepancies between these studies’ results and mine may be explained by differences in design—specifically, that this study was designed to avoid demand effects which may have been present in the reviewed studies (e.g., by using a within-subjects design).
**Exploratory Results**

This study included additional measures of warmth and competence for exploratory purposes. Other researchers have reported that gay men and effeminate gay men were perceived as warmer but less competent than straight and masculine men, respectively (Brooks et al., 2019; Mize & Manago, 2018). Effeminate John was perceived as warmer than masculine John, but gay John was not perceived as warmer than straight John. Additionally, neither gender expression nor sexual orientation predicted competency evaluations of John.

Exploratory results that have implications for this study’s assumptions were that all but two participants correctly identified John’s listed sexual orientation, yet many participants were not confident that John’s stated sexual orientation was true. This suggests that the question assessing recall of the target’s sexual orientation (“What is John’s sexual orientation?”) was interpreted by many participants as asking about John’s stated—but not actual—sexual orientation. Scores on this item differed by condition; participants were less confident that John’s sexual orientation was true when he was straight and effeminate than any other combination of sexual orientation and gender expression. This may suggest that a confound was introduced through the manipulation of these two characteristics; however, such a confound would be difficult to avoid in an experimental design, since effeminacy is strongly associated with gay men. Disbelief about a straight man’s sexual orientation may also be a key influencer of negative evaluations toward effeminate straight men. It may be that someone would perceive a straight/effeminate man as a “closeted” gay man and thus dishonest or inauthentic, and this in turn may result in more negative evaluations toward this target compared to one who is effeminate but also gay. Regardless of the potential for a confound in this study’s design, results
indicated that straight/effeminate John was not evaluated significantly more negatively than gay/effeminate John.

**Limitations and Future Directions**

This study was limited in several ways. First, the items for this study’s measure of perceived expectancy violation did not explicitly mention John’s sexual orientation or gender expression. As a result, participants’ responses to these items may have been influenced by other characteristics from John’s vignette or photo (e.g., his job or clothing). While there may be a concern that the inclusion of questions directly calling participants’ attention to gender expression may introduce demand effects, they would also allow one to assess specifically whether and how sexual orientation and gender expression influence perceptions of a target’s expectancy violation. The inclusion of items that explicitly assess expectancy violation for target demographics other than sexual orientation (e.g., race, age) may allow researchers to collect this data while still mitigating demand effects.

Additionally, the evaluative items used in this study were on unipolar scales—for example, participants indicated how much they liked John from “Not at all” to “Very much.” Using unipolar scales meant that participants were not able to indicate that they, for example, very much disliked (vs. didn’t much like) John or were willing to harm (vs. not help) him. The use of bipolar scales for evaluative items may allow researchers to capture more attitudinal variance. Alternatively, additional items that measure opposing constructs could be added—for example, items assessing how much participants dislike targets in addition to those that assess how much they like them.

Furthermore, while this study was informed in-part by the BSE, I did not include a measure of identification as a man or a straight man to simplify the research design. This study’s
results provide some support for the BSE, but without such a measure, I cannot determine if and how ingroup bias or ingroup deviancy may have influenced evaluations of John. Future studies should include a measure of group identification (e.g., Eidelman & Biernat, 2003) to parse these relationships. As before, to mitigate demand effects, this measure of group identification may include items that assess strength of participants’ identities other than sexual orientation or gender (e.g., race).

Finally, sexual prejudice was entered as a covariate for all of this study’s primary analyses. To avoid demand effects, this measure was administered after participants finished evaluating John, and thus after this study’s manipulations. While there were no significant differences by John’s sexual orientation or gender expression on this variable, future studies that seek to use sexual prejudice as a covariate should, when possible, include this measure before any manipulations—ideally separated in time from the main experiment to avoid signaling to participants the purpose of the study.

Conclusion

Despite this study’s limitations, it also had several strengths. First, it focused on an understudied population: effeminate straight men. While other studies have included effeminate straight male targets, their primary focus was to examine how effeminacy affects evaluations of gay men. Additionally, many of these studies may have been susceptible to demand effects. This study was designed to avoid demand effects by using a between-subjects design, including filler targets who differed on characteristics other than those of interest (e.g., age, race), and included neutral information (e.g., likes to play cards) for all targets in addition to information about their sexual orientation and gender expression. This study also included a measure of modern sexual prejudice to assess more subtle prejudiced attitudes and focused on nonverbal behavior as the
single gender conformity domain, which should be more salient to participants than other domains (e.g., appearance). Finally, it examined not only liking as a judgment domain, but also those with greater real-world implications: willingness to be physically near, affiliate with, or help.

While this study cannot clarify the already discordant results in literature that examined evaluations of masculine or effeminate straight or gay men, these results suggest that straight men may evaluate a masculine gay man more negatively than a masculine straight man. Additionally, they may desire more physical distance from the masculine gay man. However, it is important to note that these results disappeared when sexual prejudice was not included as a covariate in this study’s statistical models. While there was a significant interaction between John’s sexual orientation and John’s gender expression and mean patterns indicated that straight/masculine John was liked more than gay/masculine John, this difference was nonsignificant when sexual prejudice was not entered into the statistical models as a covariate. Thus, this study’s results are only suggestive. More research is necessary to clarify if and how gay and straight men are differentially evaluated when they are masculine or effeminate. Additionally, future research should seek to further determine the means through which differential evaluations might occur—whether through violation of stereotype expectancies, ingroup deviancy, or some other mechanism.
### Table 1

**Summary of Reviewed Literature**

<table>
<thead>
<tr>
<th>Citation</th>
<th>Sample</th>
<th>SP Measure</th>
<th>GC Domains</th>
<th>Evaluation Domain(^a)</th>
<th>Assumption 1</th>
<th>Assumption 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Storms, 1978</td>
<td>258 college students (130 male)(^b)</td>
<td>None</td>
<td>College major, appearance, activity</td>
<td>Liking</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Laner &amp; Laner, 1979</td>
<td>206 male college students(^c)</td>
<td>None</td>
<td>College major, appearance, activity</td>
<td>Liking</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Gowen &amp; Britt, 2006</td>
<td>120 college students (~49 male)(^d)</td>
<td>ATLG</td>
<td>Nonverbal behavior</td>
<td>College admission</td>
<td>ns</td>
<td>ns</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Scholarship funding</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Social distance</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Horn, 2007</td>
<td>103 male 10(^{th}) &amp; 12(^{th}) graders(^c)</td>
<td>None</td>
<td>Appearance, activity</td>
<td>Acceptability</td>
<td>No (activities)</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>No (appearance)</td>
<td></td>
</tr>
<tr>
<td>Lehavot &amp; Lambert, 2007</td>
<td>213 college students (71 male)(^b)</td>
<td>ATLG</td>
<td>Activity</td>
<td>Liking</td>
<td>No (high SP)</td>
<td>No (high SP)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Yes (low SP)</td>
<td>Yes (low SP)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blashill &amp; Powlishtha, 2009b</td>
<td>177 male college students(^c)</td>
<td>ATLG</td>
<td>Occupation, activity, trait</td>
<td>Liking</td>
<td>ns</td>
<td>ns</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Hang out with</td>
<td>ns</td>
<td>ns</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Willing to work with on task</td>
<td>ns</td>
<td>ns</td>
</tr>
<tr>
<td>Citation</td>
<td>Sample</td>
<td>SP Measure</td>
<td>GC Domains</td>
<td>Evaluation Domain&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Assumption 1</td>
<td>Assumption 2</td>
</tr>
<tr>
<td>-------------------</td>
<td>-------------------------------------</td>
<td>------------</td>
<td>-----------------------------</td>
<td>--------------------------------</td>
<td>--------------</td>
<td>--------------</td>
</tr>
<tr>
<td>Blashill &amp; Powlishta, 2012</td>
<td>305 male college students&lt;sup&gt;c&lt;/sup&gt;</td>
<td>Unknown&lt;sup&gt;e&lt;/sup&gt;</td>
<td>Appearance, activity, trait</td>
<td>Liking</td>
<td><em>ns</em></td>
<td><em>ns</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Willing to work with on task</td>
<td></td>
<td><em>ns</em></td>
<td><em>ns</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Psychological adjustment</td>
<td></td>
<td><em>ns</em></td>
<td><em>ns</em></td>
</tr>
</tbody>
</table>

Note. SP = sexual prejudice. GC = gender conformity. ATLG = the Attitudes toward Lesbians and Gay Men scale (Herek, 1988). Assumption 1 = results supported more positive evaluations for gay/effeminate over straight/effeminate targets. Assumption 2 = results supported more positive evaluation for gay/masculine over gay/masculine targets. Assumption 3 is the inverse of assumption 2; thus, support for this assumption is indicated by a “No” in the Assumption 2 column. Bolding indicates a significant difference between groups, while italics indicate a non-significant difference or unreported significance. Where *ns* is present, it means there were no significant differences between groups and group means were unreported.

<sup>a</sup>Not all dependent variables from each study were included in this table—for example, Blashill and Powlishta (2009b) also had participants rate targets’ intelligence and how boring they were.<sup>b</sup>Study sample included male and female participants, but results did not significantly differ by participant gender. <sup>c</sup>Study sample included male and female participants, but I only reviewed and included results for men.<sup>d</sup>Study sample included male and female participants and gender differences are unknown. <sup>e</sup>The authors reported there were no differences in sexual prejudice on study variables, but did not report which measure they used.
Table 2
Assumptions, Associated Hypotheses, Informing Theories, and Suggested Relationships

<table>
<thead>
<tr>
<th>Assumption</th>
<th>Hypothesis</th>
<th>According to…</th>
<th>When men are…</th>
<th>Then…</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$H_{1a}$</td>
<td>EVT &amp; BSE</td>
<td>Effeminate</td>
<td>Straight men should be evaluated more <strong>negatively</strong> than gay men</td>
</tr>
<tr>
<td>2</td>
<td>$H_{2a}$</td>
<td>EVT</td>
<td>Masculine</td>
<td>Straight men should be evaluated more <strong>negatively</strong> than gay men</td>
</tr>
<tr>
<td>3</td>
<td>$H_3$</td>
<td>BSE</td>
<td>Masculine</td>
<td>Straight men should be evaluated more <strong>positively</strong> than gay men</td>
</tr>
</tbody>
</table>
Table 3
Sample Demographics (n = 196)

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White or Caucasian</td>
<td>94</td>
<td>48.0</td>
</tr>
<tr>
<td>Asian or Pacific Islander</td>
<td>43</td>
<td>21.9</td>
</tr>
<tr>
<td>Black or African American</td>
<td>28</td>
<td>14.3</td>
</tr>
<tr>
<td>Hispanic or Latino</td>
<td>17</td>
<td>8.7</td>
</tr>
<tr>
<td>Multiracial</td>
<td>13</td>
<td>6.6</td>
</tr>
<tr>
<td>Not listed</td>
<td>1</td>
<td>0.5</td>
</tr>
<tr>
<td>Religious Affiliation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Christian</td>
<td>77</td>
<td>39.9</td>
</tr>
<tr>
<td>Agnostic</td>
<td>44</td>
<td>22.4</td>
</tr>
<tr>
<td>Atheist</td>
<td>35</td>
<td>17.9</td>
</tr>
<tr>
<td>Buddhist</td>
<td>8</td>
<td>4.1</td>
</tr>
<tr>
<td>Jewish</td>
<td>8</td>
<td>4.1</td>
</tr>
<tr>
<td>Muslim</td>
<td>6</td>
<td>3.1</td>
</tr>
<tr>
<td>Hindu</td>
<td>4</td>
<td>2.0</td>
</tr>
<tr>
<td>Not listed</td>
<td>14</td>
<td>7.1</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>21.09</td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>1.77</td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>18–24</td>
<td></td>
</tr>
<tr>
<td>1. Liking</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>2. Proximity</td>
<td>0.78§</td>
<td></td>
</tr>
<tr>
<td></td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>3. Affiliation</td>
<td>0.79§ 0.81§ –</td>
<td></td>
</tr>
<tr>
<td>4. Helping</td>
<td>0.43§ 0.48§ 0.47§ –</td>
<td></td>
</tr>
<tr>
<td>5. Global Evaluation</td>
<td>0.88§ 0.92§ 0.90§ 0.61§ –</td>
<td></td>
</tr>
<tr>
<td>6. Unexpected</td>
<td>-0.07 -0.12 -0.19‡ -0.05 -0.14† –</td>
<td></td>
</tr>
<tr>
<td>7. Warmth</td>
<td>0.53§ 0.56§ 0.45§ 0.34§ 0.57§ -0.03 –</td>
<td></td>
</tr>
<tr>
<td>8. Competence</td>
<td>0.54§ 0.56§ 0.50§ 0.37§ 0.58§ -0.10 0.60§ –</td>
<td></td>
</tr>
<tr>
<td>9. Disgust</td>
<td>-0.61§ -0.62§ -0.57§ -0.33§ -0.67§ 0.14 -0.69§ -0.65§ –</td>
<td></td>
</tr>
<tr>
<td>10. Immorality</td>
<td>-0.52§ -0.51§ -0.45§ -0.40§ -0.55§ 0.09 -0.56§ -0.63§ 0.64§ –</td>
<td></td>
</tr>
<tr>
<td>11. Moral Disgust</td>
<td>-0.63§ -0.63§ -0.57§ -0.41§ -0.68§ 0.12 -0.69§ -0.71§ 0.91§ 0.90§ –</td>
<td></td>
</tr>
<tr>
<td>12. Sexual Prejudice</td>
<td>-0.38§ -0.37§ -0.40§ -0.29§ -0.41§ 0.09 -0.20‡ -0.34§ 0.33§ 0.37§ 0.39§ –</td>
<td></td>
</tr>
<tr>
<td>13. SDR</td>
<td>0.04 0.04 0.05 0.07 0.04 0.11 0.19‡ 0.11 -0.06 -0.02 -0.04 0.18† –</td>
<td></td>
</tr>
</tbody>
</table>

| M             | 4.65 5.09 4.96 6.00 5.16 3.78 4.63 4.82 3.34 3.35 3.34 3.62 4.03 |
| SD            | 1.44 1.56 1.60 1.17 1.34 1.16 1.13 1.17 1.31 1.26 1.16 1.33 0.82 |

Note. SDR = socially desirable responding. Helping was transformed for analysis; means and standard deviations in parentheses represent the transformed variable’s values.

†p < .05. ‡p < .01. §p < .001.
<table>
<thead>
<tr>
<th>Predictor</th>
<th>Straight/Masculine</th>
<th>Gay/Masculine</th>
<th>Straight/Effeminate</th>
<th>Gay/Effeminate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Liking</td>
<td>4.97</td>
<td>4.41</td>
<td>4.44</td>
<td>4.87</td>
</tr>
<tr>
<td>Proximity</td>
<td>5.40</td>
<td>4.77</td>
<td>5.06</td>
<td>5.24</td>
</tr>
<tr>
<td>Affiliation</td>
<td>5.21</td>
<td>4.76</td>
<td>5.00</td>
<td>5.02</td>
</tr>
<tr>
<td>Helping</td>
<td>6.26</td>
<td>5.89</td>
<td>5.94</td>
<td>5.92</td>
</tr>
<tr>
<td></td>
<td>(1.27)</td>
<td>(1.40)</td>
<td>(1.38)</td>
<td>(1.39)</td>
</tr>
<tr>
<td>Global</td>
<td>5.46</td>
<td>4.96</td>
<td>5.10</td>
<td>5.26</td>
</tr>
<tr>
<td>Immorality</td>
<td>3.19</td>
<td>3.65</td>
<td>3.16</td>
<td>3.39</td>
</tr>
<tr>
<td>Moral</td>
<td>3.23</td>
<td>3.59</td>
<td>3.22</td>
<td>3.31</td>
</tr>
<tr>
<td>Disgust</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Warmth</td>
<td>4.48</td>
<td>4.32</td>
<td>4.86</td>
<td>4.88</td>
</tr>
<tr>
<td>Competence</td>
<td>4.77</td>
<td>4.71</td>
<td>4.87</td>
<td>4.98</td>
</tr>
</tbody>
</table>

*Note. EMM = estimated marginal mean. Global = global evaluation. Bolded predictors are from models where the interaction between John’s sexual orientation and gender expression were significant. Bolded EMMs are the highest mean for positively valenced predictors and the lowest mean for negatively valenced predictors. Italicized EMMs are the lowest mean and highest mean for positively and negatively and positively valenced predictors, respectively. Values in parentheses represent the transformed helping variable.*
### Table 6

**ANCOVA Results and Interaction Descriptives for Liking, Proximity, Global Evaluation, and Affiliation Models**

<table>
<thead>
<tr>
<th></th>
<th>Liking Model</th>
<th>Proximity Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( F )</td>
<td>( p )</td>
</tr>
<tr>
<td>Sexual Prejudice</td>
<td>33.94</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>SO</td>
<td>0.21</td>
<td>.649</td>
</tr>
<tr>
<td>GE</td>
<td>0.07</td>
<td>.792</td>
</tr>
<tr>
<td>SO × GE</td>
<td>7.08</td>
<td>.008</td>
</tr>
<tr>
<td>Masculine/Straight</td>
<td>4.94</td>
<td>.497</td>
</tr>
<tr>
<td>Masculine/Gay</td>
<td>4.52</td>
<td>4.41</td>
</tr>
<tr>
<td>Effeminate/Straight</td>
<td>4.35</td>
<td>4.44</td>
</tr>
<tr>
<td></td>
<td>Global Evaluation Model</td>
<td>Affiliation Model</td>
</tr>
<tr>
<td>Sexual Prejudice</td>
<td>44.80</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>SO</td>
<td>1.31</td>
<td>.255</td>
</tr>
<tr>
<td>GE</td>
<td>0.06</td>
<td>.809</td>
</tr>
<tr>
<td>SO × GE</td>
<td>4.48</td>
<td>.036</td>
</tr>
<tr>
<td>Masculine/Straight</td>
<td>5.43</td>
<td>5.46</td>
</tr>
<tr>
<td>Masculine/Gay</td>
<td>5.07</td>
<td>4.96</td>
</tr>
<tr>
<td>Effeminate/Straight</td>
<td>5.01</td>
<td>5.10</td>
</tr>
<tr>
<td>Effeminate/Gay</td>
<td>5.27</td>
<td>5.26</td>
</tr>
</tbody>
</table>

*Note. SO = John’s sexual orientation. GE = John’s gender expression. EMM = estimated marginal mean. Numbers in brackets represent 95% confidence intervals for EMMs. The helping model was excluded from this table, but its results are similar to those from the affiliation model (i.e., sexual prejudice is the only significant predictor and the means by condition follow a similar pattern).*
Table 7

*Moderated Mediation Effects for Liking Model*

<table>
<thead>
<tr>
<th>Path</th>
<th>B</th>
<th>SE</th>
<th>p</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total effect</td>
<td>0.00</td>
<td>.21</td>
<td>.997</td>
<td>[-0.41, 0.41]</td>
</tr>
<tr>
<td>SO → EV</td>
<td>1.05</td>
<td>.22</td>
<td>&lt; .001</td>
<td>[0.63, 1.48]</td>
</tr>
<tr>
<td>GE × SO → EV</td>
<td>-0.72</td>
<td>.22</td>
<td>.023</td>
<td>[-1.33, -0.10]</td>
</tr>
<tr>
<td>EV → Liking</td>
<td>0.00</td>
<td>.13</td>
<td>.972</td>
<td>[-0.25, 0.26]</td>
</tr>
<tr>
<td>GE × EV → Liking</td>
<td>0.01</td>
<td>.17</td>
<td>.957</td>
<td>[-0.33, 0.35]</td>
</tr>
<tr>
<td>GE × SO → Liking</td>
<td>0.99</td>
<td>.40</td>
<td>.014</td>
<td>[0.21, 1.78]</td>
</tr>
</tbody>
</table>

Conditional indirect effects

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>SE</th>
<th>p</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Masculine</td>
<td>0.01</td>
<td>.14</td>
<td>–</td>
<td>[-0.29, 2.84]</td>
</tr>
<tr>
<td>Effeminate</td>
<td>0.01</td>
<td>.05</td>
<td>–</td>
<td>[-0.10, 0.13]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Index of moderated mediation</td>
<td>0.00</td>
<td>.15</td>
<td>–</td>
<td>[-0.30, 0.31]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Direct effect</th>
<th>B</th>
<th>SE</th>
<th>p</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-0.57</td>
<td>.29</td>
<td>.054</td>
<td>[-1.15, 0.01]</td>
</tr>
</tbody>
</table>

*Note. SO = John’s sexual orientation (where 1 = gay, 0 = straight). GE = John’s gender expression (where 1 = effeminate, 0 = masculine). EV = expectancy violation. B coefficients are unstandardized.*
Figure 1
Interaction Plot for John’s Sexual Orientation and Gender Expression on the Estimated Marginal Means for Liking
Figure 2
*Interaction Plot for John’s Sexual Orientation and Gender Expression on the Estimated Marginal Means for Proximity*
Figure 3
Interaction Plot for John’s Sexual Orientation and Gender Expression on the Estimated Marginal Means for Global Evaluation
Figure 4

Interaction Plot for John’s Sexual Orientation and Gender Expression on the Means for Perceived Expectancy Violation
Figure 5
Liking Moderated Mediation Model

Note. John’s sexual orientation is coded such that 1 = gay, 0 = gay. John’s gender expression is coded such that 1 = effeminate, 0 = masculine.

\( p < .05. \quad ^{5} p < .001. \)
Appendix A

Marcus (Filler Target 1) Photo and Vignette

Marcus is a 32-year-old Latino man.

He is an accountant.

He is straight.

He maintains eye contact when someone is talking to him.

He prefers to have smoothies for breakfast.

He likes to do crossword puzzles.

He walks with a slight limp.
Appendix B

Susan (Filler Target 2) Photo and Vignette

Susan is a 72-year-old Black woman.

She is an office receptionist.

She is straight.

She looks up and to the left when she thinks deeply about something.

She prefers to play dominoes instead of going out.

She likes to read books.

She is naturally agile.
Appendix C

Masculine (left) and Effeminate Photo Manipulations for John
Appendix D

Vignette Manipulations for John

Sexual orientation manipulations are included in brackets and gender conformity manipulations are included in parentheses (masculine/effeminate).

Vignette with behavior set 1:

John is a 21-year-old White man.

He is an office worker.

He is [straight/gay].

He gestures (assertively/flamboyantly) when he gets excited.

He prefers to see new movies in theaters.

He likes to play cards.

He speaks with a (deep/high) voice.

Vignette with behavior set 2:

John is a 21-year-old White man.

He is an office worker.

He is [straight/gay].

He (swaggers his shoulders/sways his hips) as he walks.

He prefers to see new movies in theaters.

He likes to play cards.

He has a (firm/soft) handshake.
References


Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and

https://doi.org/10.1037/0022-3514.82.6.878


https://doi.org/10.1177/0261927X06292769


https://doi.org/10.1177/014616720963665


https://doi.org/10.1177/2158244015621113


https://doi.org/10.1080/00224498809551476


https://doi.org/10.1159/000021447


https://doi.org/10.1300/J082v01n01_02

https://doi.org/10.1037/1524-9220.4.1.3


https://doi.org/10.1080/14792779543000011

https://doi.org/10.1002/ejsp.2420180102

https://doi.org/10.1007/BF00288188


https://doi.org/10.1177/01461672992512005

https://doi.org/10.1016/S0140-1971(86)80004-5
Vita

Art D. Marsden

admarsde@syr.edu | artdmarsden@gmail.com

Education

**Syracuse University, Syracuse, NY**

Ph.D. in Social Psychology  
M.S. in Psychology  
Expected 2022

**University of North Texas (UNT), Denton, TX**

B.S. in Psychology, *Summa Cum Laude*  
Dec 2017

*Thesis:* Does Empathy Mediate the Relationship between Political Ideology and Sexual Prejudice?

Select Awards and Honors

<table>
<thead>
<tr>
<th>Award/Grant</th>
<th>Institution</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>CloudResearch Grant Award ($2500), <em>Innovations in Online Research 2021</em></td>
<td></td>
<td>Fall 2021</td>
</tr>
<tr>
<td>Project: Validating an AI-Generated Database of Social Psychology Face Stimuli</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SPSP Graduate Travel Award ($500), <em>SPSP 2022</em></td>
<td></td>
<td>Fall 2021</td>
</tr>
<tr>
<td>Undergraduate Student Poster Award-First Place ($100), <em>SPSP 2018</em></td>
<td></td>
<td>Spring 2018</td>
</tr>
<tr>
<td>Diversity Fund Undergraduate Registration Award ($160), <em>SPSP 2018</em></td>
<td></td>
<td>Fall 2017</td>
</tr>
<tr>
<td>Brownlee-Moore Scholarship ($1000), <em>UNT</em></td>
<td></td>
<td>Spring 2017</td>
</tr>
<tr>
<td>Louis C. Weber Outstanding Undergraduate Student Award ($750), <em>UNT</em></td>
<td></td>
<td>Spring 2017</td>
</tr>
<tr>
<td><em>One available annually for undergraduate psychology students</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Best Poster Prize ($300), <em>ISSWSH 2017</em></td>
<td></td>
<td>Spring 2017</td>
</tr>
<tr>
<td>Off-Campus Student Services Book Award ($200), <em>UNT</em></td>
<td></td>
<td>Fall 2016</td>
</tr>
<tr>
<td>Transfer Scholarship I ($6000), <em>UNT</em></td>
<td></td>
<td>Fall 2015</td>
</tr>
</tbody>
</table>

Employment

**University of Texas Health Science Center at San Antonio, STRONG STAR and the Consortium to Alleviate PTSD—Fort Hood Research Site**

**Research Area Specialist-Associate**  
Jan 2018 – July 2019

*Project:* STRONG STAR Training Initiative

*Principal Investigators:* Katherine A. Dondanville, Psy.D.; Brooke A. Fina, LCSW
Project coordinator for national grant-funded program that trained over 900 mental health providers in evidence-based therapies (CPT, PE, CBT-I&N, CRP) for treating veterans with PTSD.

Project: Treatment of Comorbid Sleep Disorders and PTSD

Principal Investigator: Daniel J. Taylor, Ph.D.

Co-Principal Investigators: Patricia A. Resick, Ph.D.; Alan L. Peterson, Ph.D.; Doug E. Williamson, Ph.D.

Research assistant for U.S. Departments of Defense and Veterans Affairs-funded randomized controlled trial to examine how to incorporate cognitive-behavioral intervention in treating active duty service members suffering from PTSD and comorbid sleep issues.

Articles Published in Peer-Reviewed Journals


Presentations (* denotes undergraduate mentee)


providers and their patients [Poster presentation]. UT Health San Antonio Department of Psychiatry 6th Annual Research and Quality Improvement Day, San Antonio, TX.


Marsden, A. D., & Barnett, M. D. (2018, March). Dimensions of empathy as mediators between social political ideology and homophobia [Poster presentation]. Society for Personality and Social Psychology, Atlanta, GA. (Winner of Student Poster Award-First Place, SPSP 2018)


Teaching Experience

**Syracuse University**

*Instructor of Record*
**PSY 472: Laboratory in Social Psychology**  
*Spring 2022*

**Supervisor:** Leonard S. Newman, Ph.D.

Instructor of record for laboratory course in advanced research methods in social psychology (25 students). Designed all course lessons, created and graded all assignments.

**PSY 274: Social Psychology**  
*Summer 2020, 2021*

**Supervisor:** Leonard S. Newman, Ph.D.

Instructor of record for online asynchronous (Summer ‘20) and online synchronous (‘21) courses during 6-week summer sessions (43 total students). Designed all course lessons, created and graded all assignments.

**Teaching Assistant**

**PSY 205: Foundations of Human Behavior**  
*Fall 2019 – Spring 2020, Fall 2021*

**Supervisors:** Shannon C. Houck, Ph.D.; Meredith J. Martin, Ph.D.; Jennifer A. Clarke, Ph.D.

Lecturer for weekly recitation sections (9 total sections, 225 total students) for introductory psychology course. Designed all recitation lessons and weekly quizzes and graded all assignments.

**PSY 335: Psychology of Childhood**  
*Fall 2020 – Spring 2021*

**Supervisor:** Shannon M. Sweeney, Ph.D.

Graded student assignments and held office hours (3 sections, 395 total students).

**PSY 274: Social Psychology**  
*Spring 2020*

**Supervisor:** Jessie B. Joyce, Ph.D.

Graded student assignments and held office hours (129 total students).

**University of North Texas**

**Guest Lecturer**

**PSYC 4470: Sexual Behavior**  
*Spring 2017*

**Instructor:** Michael D. Barnett, Ph.D.

Presented lecture on sexual definitions research and considerations for defining sexual behavior for individuals in same-sex sexual relationships.

**Memberships**

Member, *Phi Kappa Phi*  
2017 – Present
Student Member, *Society for Personality and Social Psychology* 2017 – Present
Member, *Psi Chi* 2016 – Present
Student Affiliate, *American Psychological Association* 2016 – Present