A Meta-Analytic Review of Gender Differences in Perceptions of Sexual Harassment

Maria Rotundo
University of Minnesota, Twin Cities Campus

Dung-Hanh Nguyen
California State University, Long Beach

Paul R. Sackett
University of Minnesota, Twin Cities Campus

Research on gender differences in perceptions of sexual harassment informs an ongoing legal debate regarding the use of a reasonable person standard instead of a reasonable woman standard to evaluate sexual harassment claims. The authors report a meta-analysis of 62 studies of gender differences in harassment perceptions. An earlier quantitative review combined all types of social-sexual behaviors for a single meta-analysis; the purpose of this study was to investigate whether the magnitude of the female-male difference varies by type of behavior. An overall standardized mean difference of 0.30 was found, suggesting that women perceive a broader range of social-sexual behaviors as harassing. However, the meta-analysis also found that the female-male difference was larger for behaviors that involve hostile work environment harassment, derogatory attitudes toward women, dating pressure, or physical sexual contact than sexual propositions or sexual coercion.

Since the 1980s, interest in defining sexual harassment has increased. An important issue that arises when trying to define sexual harassment is to identify which behaviors are harassing. However, before the field can reach consensus as to which behaviors constitute sexual harassment, researchers need to obtain a full understanding of the degree to which individual differences affect perceptions of social-sexual behaviors. An individual characteristic that has received a great deal of attention in the literature is gender. Although research reports mixed results, when a gender difference exists, it suggests that women perceive a broader range of behaviors as harassing. In a series of narrative reviews (e.g., American Psychological Association, 1993; Frazier, Cochran, & Olson, 1995; Gutek, 1985) and a recent quantitative review (Blumenthal, 1998), authors have tried to summarize what is known today about the extent to which men and women differ in their judgments about what constitutes sexual harassment. Although these reviews offer insightful observations regarding gender differences, the magnitude of the difference reported in the literature varies widely. Neither the American Psychological Association’s brief nor the qualitative reviews provide any conclusive evidence in this regard. The purpose of this article is to achieve a better understanding of the actual size of the gender difference and conditions under which it varies.

Attention to differences in men’s and women’s judgments about what constitutes sexual harassment was largely motivated by problems faced by the courts in trying to resolve claims of sexual harassment. The Equal Employment Opportunity Commission (EEOC; 1980) defines sexual harassment as verbal or physical conduct of a sexual nature that unreasonably interferes with an individual’s job performance or creates an intimidating or hostile work environment. However, from whose perspective should the courts evaluate whether a set of circumstances creates a hostile work environment? Traditionally, the courts have interpreted the circumstances surrounding a case from the perspective of a reasonable person in similar circumstances (e.g., Ratidue v. Osceola Refining Co., 1986/1987; Radke v. Everett, 1991). Evidence demonstrating that men and women evaluate incidents of social-sexual behaviors differently led to some landmark rulings in which the reasonable woman standard was used in place of the reasonable person standard to evaluate claims of sexual harassment (e.g., Ellison v. Brady, 1991). This standard relies on the assumption that to make an accurate assessment as to whether harassment has occurred or not, decision makers need to take the perspective of the harassee, which in most cases is women (e.g., Abrams, 1989; EEOC, 1993, Ehrenreich, 1990). However, the courts made the decision to implement the reasonable woman standard without any conclusive evidence about the magnitude of the assumed gender difference.

In an attempt to provide a quantitative answer to the question of the degree to which men and women differ in how they perceive social-sexual behaviors, Blumenthal (1998) conducted a meta-analytic review of this body of literature. He meta-analyzed 83 independent effect sizes over 34,350 subjects and 111 studies and
reported an overall standardized mean difference of 0.35, suggesting that women perceive a broader range of behaviors as harassing than do men. Blumenthal examined whether certain characteristics of the studies or the subjects moderated the size of the gender difference. Study characteristics included year of publication, stimulus material (e.g., behavioral scenario vs. checklist), and whether the study was published or unpublished. Characteristics of the sample included national origin and subject group (graduate, undergraduate, worker, faculty, etc.). Blumenthal found that the effect sizes for these moderators were not significantly different from zero, indicating that these characteristics of the study or the sample did not moderate the size of the gender difference. The results of Blumenthal’s quantitative review indicate that across a broad range of behaviors, a gender difference existed and in the hypothesized direction. However, the difference was not large (Cohen, 1988). Furthermore, certain characteristics of the study and the subject sample did not moderate the size of the gender difference.

The reasonable woman standard relies on the assumption that men and women differ in their perceptions of which behaviors constitute sexual harassment. Therefore, this standard requires that decision makers adopt the perspective of the harassee, which often tends to be women, when evaluating the circumstances in a sexual harassment claim. The results of Blumenthal’s (1998) study do not provide strong support for the assumptions underlying the reasonable woman standard. However, research evidence suggests that the magnitude of the gender difference varies according to the specific type of behavior that is evaluated (e.g., Braine, Bless, & Fox, 1995; Fitzgerald & Ormerod, 1991; Fraizer et al., 1995; Garlick, 1994; Gutek, 1985). Therefore, an overall effect size computed across a broad range of behaviors may conceal important gender differences in perceptions of social-sexual behaviors, casting unnecessary doubt on the assumptions underlying the reasonable woman standard. Our quantitative review examined the size of the gender difference within specific behavioral categories in an attempt to identify those behaviors that produce the largest difference and to provide more accurate information about the actual size of the gender difference. Furthermore, Blumenthal’s quantitative review included studies that assessed perceptions other than sexual harassment (e.g., how serious, appropriate, or severe the social–sexual behaviors were; the harasser’s feeling of guilt). Our meta-analysis, however, was restricted to studies in which the dependent variable of interest was perceptions of whether the behavior in question constituted sexual harassment. We believe that measures of the seriousness or appropriateness of social–sexual behaviors tap different perceptions. Finally, research evidence suggests that the harasser’s status moderates the size of the gender difference (U.S. Merit Systems Protection Board, 1988). That is, men’s and women’s perceptions of sexual harassment are influenced by the position or role of the harasser relative to the harassee (e.g., supervisor, peer, subordinate). Therefore, our quantitative review also examined the harasser’s status as a moderator of the gender difference. The notion that gender differences in perceptions of sexual harassment vary according to the type of behavior in question creates the need for a category scheme that groups together similar behaviors. Researchers have offered a variety of such category schemes (Braine et al., 1995; Fitzgerald, Geldaf, & Drasgow, 1995; Fitzgerald, Shullman et al., 1988; Gutek, Nakamura, Grabant, & Handschumacher, 1980; Hippensteele, 1991; Loredo, Reid, & Deaux, 1995; Till, 1980). Our goal was to sort as many as possible of the social–sexual behaviors examined in the harassment literature into specific groups that differentiate between key behaviors. We carefully reviewed existing empirical research, legal cases, and discussions of legal precedent regarding social–sexual behaviors and gender differences in perceptions of sexual harassment (e.g., Adams, Kottke, & Padgitt, 1983; Adler & Peirce, 1993; Bursik, 1992; Delgado v. Lehman, 1987; Fitzgerald et al., 1995; Meritor Savings Bank v. Vinson, 1986; Terpstra & Baker, 1987; Till, 1980). We found some interesting patterns.

In the psychological literature, the most widely used instrument, the Sexual Experiences Questionnaire, was developed to assess the prevalence of sexual harassment (Fitzgerald, Shullman et al., 1988). The items included therein were designed to reflect one of five general categories of sexual harassment identified by sexual coercion, sexual assault, or sexual assault. Although a considerable amount of research on sexual harassment has used the Sexual Experiences
Questionnaire, Fitzgerald et al. (1995) revised the instrument to reflect three broad and conceptually distinct dimensions: gender harassment, unwanted sexual attention, and sexual coercion. The first two categories map onto the legal definition of hostile work environment harassment, whereas the latter category maps onto quid pro quo harassment. We reviewed additional category schemes and concluded that they share some common elements. Various categories include behaviors that are sexual or nonsexual, physical or nonphysical, and personal or nonpersonal in nature. Furthermore, in the legal arena, an important element that needs to be present in a claim of hostile environment harassment is evidence that the target is the object of harassment because of his or her sex. Evidence that certain social–sexual behaviors have occurred satisfies this criterion (Adler & Peirce, 1993). The first type of social–sexual behavior includes behavior in which the target is exposed to explicit sexual advances such as the touching of intimate body parts (Meritor Savings Bank v. Vinson, 1986). The second type includes more subtle behavior directed at the target that is not sexually explicit but that reflects derogatory attitudes toward women (Delgado v. Lehman, 1987; Lipsett v. University of Puerto Rico, 1988). The third type of social–sexual behavior includes subtle behavior that is not directed at a specific woman but that is offensive or derogatory toward women in general (e.g., sexually explicit materials such as photographs; Arnold v. City of Seminole, 1985; Holl v. Gas Construction Co., 1988).

This review of the psychological research and legal precedent revealed that some of the existing category schemes are too broad and do not allow for the differentiation between some key social–sexual behaviors (e.g., gender harassment, unwanted sexual attention, seductive behavior, verbal comments). Using these broad categories to classify behaviors could potentially conceal important gender differences. Furthermore, it is not always clear into which category (e.g., sexual bribery or sexual coercion) certain social–sexual behaviors fall (e.g., refusal to go out with someone who would hurt job situation). Therefore, we combined some of the categories presented in prior research and subdivided others in an attempt to classify as many as possible of the social–sexual behaviors examined in the harassment literature.

The resulting classification scheme included seven groups of behaviors (see Table 1). The first two categories (derogatory attitudes—impersonal and derogatory attitudes—personal, respectively) deal with behavior that the legal arena describes as subtle social–sexual behaviors that reflect derogatory attitudes toward women and that can be directed toward women in general or the target in particular (Adler & Peirce, 1993). The difference between the two is that the first category deals with behaviors that we labeled impersonal (i.e., not aimed directly at the target person but at women or men in general), whereas the second category deals with behaviors that we labeled personal (i.e., aimed specifically at the target person). We labeled derogatory comments about women in general derogatory attitudes—impersonal and gender-based derogatory comments about the target woman derogatory attitudes—personal. The third and fourth categories included behaviors in which the harasser makes some type of proposition. Category 3, labeled unwanted dating pressure, was restricted to unwanted, repeated requests for a date or propositions of a nonsexual nature (i.e., personal invitation to lunch, persistent requests for a date after refusal). In contrast, Category 4, labeled sexual propositions, included propositions that were explicitly sexual in nature (i.e., unwarned pressure or request for sexual involvement). Therefore, an invitation to a drink after work fell under the third category, whereas an invitation to have a sexual encounter was included in Category 4. Whereas the previous four categories included nonphysical behaviors, the next two included phisical social–sexual behaviors. Category 5, labeled physical sexual contact, included social–sexual behaviors in which the harasser actually made physical sexual contact with the target (i.e., fondle, kiss, press or caress leg). This category included behaviors in which the target was exposed to explicit sexual advances (Adler & Peirce, 1993). Category 6, labeled physical nonsexual contact, also included behaviors in which the harasser made physical contact with the target. However, the contact was not of a sexual nature (e.g., congratulatory hug). The harasser did not coerce the target in any of the social–sexual behaviors included in Categories 1 through 6. The key feature that distinguished Category 7, labeled sexual coercion, from the others was the presence of force or coercion (i.e., rape, request for a sexual encounter as a condition of employment or promotion). The first six categories reflect what the courts define as hostile work environment harassment, and the seventh category reflects quid pro quo harassment. We classified behaviors into these categories using the definitions provided in Table 1 and report a separate meta-analysis for each category. We believe that

<table>
<thead>
<tr>
<th>Category</th>
<th>Description</th>
<th>Behavioral examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Derogatory attitudes—impersonal</td>
<td>Behaviors that reflect derogatory attitudes about men or women in general</td>
<td>Obscene gestures not directed at target.</td>
</tr>
<tr>
<td></td>
<td>Behaviors that are directed at the target that reflect</td>
<td>Sex-stereotyped jokes.</td>
</tr>
<tr>
<td></td>
<td>derogatory attitudes about the target’s gender</td>
<td>Obscene phone calls.</td>
</tr>
<tr>
<td></td>
<td>Persistent requests for dates after the target has refused</td>
<td>Belittling the target’s competence.</td>
</tr>
<tr>
<td></td>
<td>Explicit requests for sexual encounters</td>
<td>Repeated requests to go out after work or school.</td>
</tr>
<tr>
<td></td>
<td>Behaviors in which the harasser makes physical sexual contact with the target</td>
<td>Proposition for an affair.</td>
</tr>
<tr>
<td></td>
<td>Behaviors in which the harasser makes physical nonsexual contact with the</td>
<td>Embracing the target.</td>
</tr>
<tr>
<td></td>
<td>target</td>
<td>Kissing the target.</td>
</tr>
<tr>
<td></td>
<td>Requests for sexual encounters or forced encounters that</td>
<td>Congratulatory hug.</td>
</tr>
<tr>
<td></td>
<td>are made a condition of employment or promotion</td>
<td>Threatening punishment unless sexual favors are given.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sexual bribery.</td>
</tr>
</tbody>
</table>
examining the size of the gender difference within specific behavior categories provides more definitive information about whether men and women differ in their perceptions as to which social-sexual behaviors constitute sexual harassment. Furthermore, it may provide insight into conditions in which the reasonable woman standard is most applicable. If the classification of behaviors into different categories does not explain a sizable proportion of the variation in the effect sizes, other characteristics of the study can be explored as moderators.

Evidence suggests that behaviors are more likely to be perceived as harassing by both sexes if they are engaged in by someone who has higher status or formal authority over the harasser (e.g., Bursik, 1992; Cleveland & Kerst, 1993; Lester et al., 1986; Popovich, Gehlauf, Jolton, Somers, & Godinho, 1992; Pryor & Day, 1988). However, men’s and women’s judgments about social-sexual behaviors differ when there is no formal status differential between the harasser and the target (e.g., peer or coworker; U.S. Merit Systems Protection Board, 1988). It could be argued that the implications for employment opportunities are clearer when the harasser has formal authority over the harassee. However, when the harasser is a peer, the immediate threat is not apparent, which may elicit actual gender differences in how events are interpreted or in how social-sexual behaviors are perceived. Whereas men may perceive the behavior as harmless social interaction, women may perceive an element of threat. This meta-analysis tested the hypothesis that the size of the gender difference varies according to the status of the harasser. We posited that the female–male difference would be greater when the harasser is a peer, coworker, or fellow student of the harassee than when the harasser has formal authority over the harassee.

**Method**

**Description of Database**

**Location of studies.** In an attempt to create a comprehensive database of effect sizes that represent gender differences in perceptions of sexual harassment, three sources were consulted. First, an effort was made to obtain studies that were included in Blumenthal’s (1998) quantitative review. Of the 70 articles and reports cited in his meta-analysis, we were able to obtain 62; the remainder were unpublished, and we were unable to locate them. Second, 83 additional studies were identified from a computerized search of the PsycINFO database for the years spanning 1969–1999. Third, we reviewed the reference lists from articles on the topic.

**Inclusion criteria.** We reviewed each of the 145 articles and included it in the meta-analysis if it satisfied the following three criteria. First, the dependent variable in the study had to measure explicitly perceptions of sexual harassment (e.g., use the following 5-point scale to rate the degree to which each behavior reflects sexual harassment). Therefore, this meta-analysis excluded studies in which the dependent variable of interest measured (a) the degree to which the behavior was serious, (b) whether the target or harasser should be blamed or experience feelings of guilt for his or her actions, and (c) whether the subject had previously experienced the given social-sexual behavior. Second, enough data had to be reported to allow us to compute a gender difference. Third, the study had to report statistics from which a standardized mean difference could be computed (e.g., $F$ statistic, $t$ statistic, correlation coefficient, $p$ value, chi-square, or percentage). Various formulas were used to convert the reported statistic to the standardized mean difference (Croceter & Algina, 1986; Hunter & Schmidt, 1990; Rozenthal, 1984). The final database consisted of 62 studies that contained 66 independent samples. The total sample size (number of participants) across the 66 samples was 33,164. Studies that were included in our meta-analyses are marked with an asterisk in the reference section.

**Behavioral categories.** We made a list of all of the 122 social-sexual behaviors that were examined in the 62 studies. Two of the authors independently sorted each behavior into one of the following seven categories: derogatory attitudes—impersonal, derogatory attitudes—personal, unwanted dating pressure, sexual propositions, physical sexual contact, physical nonsexual contact, or sexual coercion. They reached agreement on 98% of the behaviors. For the three behaviors on which the authors disagreed, the third author was consulted, and a decision was made that the items either were too vague or did not clearly fall into any category. These three behaviors were dropped from the analyses. For the purposes of the broader distinction made by the courts, the first six categories represented hostile work environment harassment, and the seventh category represented quid pro quo harassment. We conducted two sets of meta-analyses. The first included a separate meta-analysis for hostile work environment and quid pro quo harassment. The second included a separate meta-analysis for each of the seven behavioral categories. The proportion of men and women within the categories was approximately equal.

In a meta-analysis, theoretically predicted moderators can be detected using significance tests. After the moderator is coded and the original group of studies is divided into subgroups (e.g., behavioral categories), a separate meta-analysis is conducted within each subgroup. Statistical formulas that test for the significance of the difference between the subgroup means assume that the data points that contribute to each subgroup come from different study samples (Hunter & Schmidt, 1990, pp. 452–454). However, in the present meta-analysis, it was possible for a given study sample to contribute a data point to more than one behavioral category (e.g., a multiple-item survey including different types of social-sexual behaviors), thus violating the independence assumption. However, we conducted significance tests (Hunter & Schmidt, 1990, pp. 436–438) and acknowledge that violating this assumption increases the probability of Type II errors.

**Status.** Evidence suggests that men and women evaluate social-sexual behaviors differently when the harasser has formal authority over the harassee. We created a dichotomous variable to differentiate between studies in which there was a power differential between the harasser and the harasser. For example, we coded the study with a 1 when the harasser was a supervisor or a faculty member and the harassee was a worker or a student. The study received a 0 when the harasser and the target had equal status (fellow student, peer, or coworker).

Although a majority of the research on sexual harassment deals with situations in which the offender has formal authority over the harassee, evidence suggests that the power differential can operate in the opposite direction when someone of lower status is the harasser (Fitzgerald, Weitzman, Gold, & Ormerod, 1988). Only one of the studies included in our sample examined contrapower harassment; therefore, it was not possible to test this as a moderator.

**Other moderators.** Evidence suggests that the nature of the stimulus materials and year of publication influence perceptions of sexual harassment. Although these moderators were examined by Blumenthal (1998), we also coded stimulus material and year of publication in an attempt to replicate his findings (i.e., we used different decision rules in our meta-analysis; therefore, our database included a different set of studies). We created a dichotomous variable to differentiate between studies in which a participant was presented with a list of behaviors and asked to check whether each one constituted sexual harassment (study was coded with a 1) and studies in which individuals were presented with scenarios and asked to indicate whether the behavior depicted in the scenarios constituted sexual harassment (study was coded with a 0). Finally, we also coded year of publication.
Meta-Analytic Procedure

We used Hunter and Schmidt’s (1990, p. 185) psychometric meta-analytic procedure to test the hypotheses in this study. Meta-analysis is a statistical technique that can be used to estimate the amount of observed variance in the findings across studies that can be attributed to statistical and methodological artifacts. Statistical artifacts include sampling error, unreliability in the criterion, range restriction, and dichotomization of variables, to name a few. Using this method, the variance that is attributable to these artifacts is subtracted from the observed variance. If the residual variance is small, one can conclude that there is true variance across studies. However, if a large amount of variance remains after these corrections, this raises the possibility that other uncorrected artifacts or methodological differences (e.g., moderators) between studies play a role.

Ideally, it is the goal of the researcher to correct the observed effect sizes for biases caused by those artifacts that are relevant to the set of variables under investigation. For the purposes of this meta-analysis, the pertinent artifacts included sampling error and unreliability in the dependent variable. We coded the 62 studies for information related to the reliability of perceptions of sexual harassment. Although 13 samples reported this information, 12 of them included measures of internal consistency. However, test-retest reliability was the relevant form of reliability for this meta-analysis because we were interested in the reliability of ratings of whether a single behavior was perceived as harassment. Therefore, we were unable to correct effect sizes for unreliability in the dependent variable because of insufficient data. However, we were able to estimate the amount of observed variance in the findings across studies due to sampling error. This is one of the features that distinguishes our meta-analysis from Blumenthal’s (1998), in which no attempts were made to correct for any statistical artifacts.

We used $d$, the standardized mean difference between men’s and women’s perceptions of sexual harassment, as the effect size measure. Positive values of $d$ indicate that women are more likely than men to perceive behaviors as harassment. A majority of the studies that we included in the meta-analysis were conducted on independent samples. In two studies, authors used multiple independent samples to measure the same dependent variable (Frazier et al., 1995; Moyer & Nath, 1998). We coded each sample as a separate study. However, in seven studies, more than one effect size could be computed on the same sample (Hemmasi, Graf, & Russ, 1994; Kenig & Ryan, 1986; McKinney, 1999; Moore, 1995; Popovich, Licata, Nokovich, Martelli, & Zoloty, 1986; Tang, Yik, Cheung, & Choi, 1995; Thacker & Gohmann, 1993). For these studies, we computed the average effect size in the overall meta-analyses to avoid violations of the independence assumption.

Table 2
Meta-Analyses of Gender Differences in Perceptions of Sexual Harassment

<table>
<thead>
<tr>
<th>Category of analysis</th>
<th>$N$</th>
<th>$K$</th>
<th>Mean $d$</th>
<th>90% credibility interval</th>
<th>$SD_d$</th>
<th>$SD_{res}$</th>
<th>% variance $SE$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall effect size</td>
<td>33,164</td>
<td>66</td>
<td>0.30</td>
<td>0.04, 0.57</td>
<td>0.18</td>
<td>0.16</td>
<td>24</td>
</tr>
<tr>
<td>Moderator analysis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Legal definition of harassment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HWE</td>
<td>27,354</td>
<td>45</td>
<td>0.33</td>
<td>0.10, 0.56</td>
<td>0.17</td>
<td>0.14</td>
<td>24</td>
</tr>
<tr>
<td>QPQ</td>
<td>9,646</td>
<td>16</td>
<td>0.18</td>
<td>0.12, 0.25</td>
<td>0.09</td>
<td>0.04</td>
<td>82</td>
</tr>
<tr>
<td>Behavioral category</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HWE: derogatory attitudes—impersonal</td>
<td>9,604</td>
<td>18</td>
<td>0.34</td>
<td>0.09, 0.59</td>
<td>0.18</td>
<td>0.15</td>
<td>25</td>
</tr>
<tr>
<td>HWE: derogatory attitudes—personal</td>
<td>23,675</td>
<td>27</td>
<td>0.33</td>
<td>0.09, 0.56</td>
<td>0.16</td>
<td>0.14</td>
<td>19</td>
</tr>
<tr>
<td>HWE: pressure for dates</td>
<td>17,075</td>
<td>16</td>
<td>0.28</td>
<td>0.08, 0.48</td>
<td>0.14</td>
<td>0.12</td>
<td>21</td>
</tr>
<tr>
<td>HWE: sexual propositions</td>
<td>7,257</td>
<td>13</td>
<td>0.18</td>
<td>0.01, 0.34</td>
<td>0.13</td>
<td>0.10</td>
<td>41</td>
</tr>
<tr>
<td>HWE: physical sexual contact</td>
<td>9,244</td>
<td>16</td>
<td>0.36</td>
<td>0.10, 0.62</td>
<td>0.18</td>
<td>0.16</td>
<td>21</td>
</tr>
<tr>
<td>HWE: physical nonsexual contact</td>
<td>8,419</td>
<td>11</td>
<td>0.14</td>
<td>0.00, 0.28</td>
<td>0.12</td>
<td>0.09</td>
<td>38</td>
</tr>
<tr>
<td>QPQ: sexual coercion</td>
<td>9,646</td>
<td>16</td>
<td>0.18</td>
<td>0.12, 0.25</td>
<td>0.09</td>
<td>0.04</td>
<td>82</td>
</tr>
<tr>
<td>Harasser status</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Formal authority</td>
<td>4,616</td>
<td>14</td>
<td>0.26</td>
<td>-0.02, 0.53</td>
<td>0.20</td>
<td>0.17</td>
<td>31</td>
</tr>
<tr>
<td>No formal authority</td>
<td>1,259</td>
<td>6</td>
<td>0.42</td>
<td>-0.24, 1.08</td>
<td>0.43</td>
<td>0.40</td>
<td>11</td>
</tr>
<tr>
<td>Stimulus material</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Checklist</td>
<td>19,375</td>
<td>18</td>
<td>0.31</td>
<td>0.19, 0.44</td>
<td>0.10</td>
<td>0.08</td>
<td>39</td>
</tr>
<tr>
<td>Scenario</td>
<td>13,789</td>
<td>48</td>
<td>0.29</td>
<td>-0.08, 0.67</td>
<td>0.26</td>
<td>0.23</td>
<td>21</td>
</tr>
</tbody>
</table>

Note: $K =$ number of effect sizes; Mean $d =$ mean observed effect size; $SD_d =$ observed standard deviation; $SD_{res} =$ residual standard deviation; % variance $SE =$ percentage of variance due to sampling error; HWE = hostile work environment; QPQ = quid pro quo.
harassment ($Z = 4.48, p < .05$). Sampling error explained a large amount of the observed variance for quid pro quo harassment.

The second set of meta-analyses focused on the seven behavioral categories. The effect sizes for sexual propositions ($Z = 2.87$ for derogatory attitudes—impersonal, $Z = 3.16$ for derogatory attitudes—personal, $Z = 1.99$ for pressure for dates, and $Z = 3.08$ for physical sexual contact), physical nonsexual contact ($Z = 3.61$ for derogatory attitudes—impersonal, $Z = 4.01$ for derogatory attitudes—personal, $Z = 2.79$ for pressure for dates, and $Z = 3.77$ for physical sexual contact), and sexual coercion ($Z = 3.33$ for derogatory attitudes—impersonal, $Z = 3.94$ for derogatory attitudes—personal, $Z = 2.40$ for pressure for dates, and $Z = 3.53$ for physical sexual contact) were significantly smaller than the effect sizes for the other four categories ($p < .05$). Sampling error did not explain a sizable proportion of the variance in the six categories that represented hostile work environment harassment.

**Harasser Status**

Next, we examined status of the harasser. It was not possible to analyze this moderator within the categories of social—sexual behaviors because of an insufficient number of effect sizes (i.e., in some cases fewer than five) that would give rise to second-order sampling error. Therefore, we meta-analyzed the overall effect sizes separately for studies in which the harasser had formal authority over the harasssee (e.g., target was a student or worker and harasser was a faculty member or supervisor) and in which there was no formal status differential (e.g., harasser was a peer, coworker, or fellow student). The results are reported in Table 2.

Although the standardized mean difference was larger when there was no obvious status differential between harasser and target, significance tests revealed that the difference was not significant ($Z = 0.87, p < .05$). Considering that the separate meta-analyses conducted within behavioral categories revealed that the gender difference was larger for more ambiguous examples of social—sexual behaviors, the larger $d$ for the studies in which there was no formal authority could have been produced if a majority of the studies included in this analysis used more ambiguous social—sexual behaviors (e.g., derogatory attitudes), whereas the studies included in the formal authority condition used more extreme examples of social—sexual behaviors (e.g., sexual coercion). However, behavior category was not confounded with the status of the harasser.

**Stimulus Material**

We also examined stimulus material as a moderator. It was not possible to analyze this moderator within each of the categories of behavior because of an insufficient number of effect sizes. Therefore, we meta-analyzed the overall effect sizes separately for studies in which the stimulus material was a checklist and for studies in which the stimulus material was presented as a scenario. The results are reported in the last two rows of Table 2. These results indicated that stimulus material did not moderate the gender difference in perceptions of sexual harassment. Although the standardized mean difference was larger when participants were presented with a list of behaviors, as opposed to scenarios, the difference was not significant ($Z = 0.05, p < .05$). Behavior category was not confounded with stimulus material.

Another potential moderator was the year in which the study was published. This represented a continuous variable. Therefore, we correlated year of study with the overall effect size, which was computed across the wide range of social—sexual behaviors. The correlation was not significantly different from zero ($r = .07, p > .10$).

**Discussion**

"Sexual harassment is a problem with a long past but a short history" (American Psychological Association, 1993, p. 3). Although some progress has been made in defining this workplace phenomenon, questions remain as to which behaviors constitute sexual harassment. The task is complicated by the fact that individual differences influence perceptions of social—sexual behaviors. A substantial body of research has examined the extent to which men and women differ in their judgments of sexual harassment. The purpose of this meta-analysis was to answer quantitatively questions about the size of this difference.

Our meta-analysis produced an overall standardized mean difference of 0.30, suggesting that women are more likely than men to define a broader range of behaviors as harassing. These findings support general assertions that although the gender difference exists and in the hypothesized direction, it is not large. The overall effect size provides minimal support for the assumptions underlying the reasonable woman standard. However, research suggests that the gender difference varies according to the type of social—sexual behavior studied. Therefore, an overall effect size computed across behaviors may conceal important differences and provide inconclusive evidence about the size of the gender difference. Furthermore, our meta-analysis showed that only 24% of the variance in the observed gender difference was accounted for by sampling error, suggesting that other statistical or methodological factors play a role. A finding that other factors moderate the gender difference indicates that whether men and women perceive social—sexual behaviors as harassing depends on other features of the situation.

Researchers have called for the need to identify moderators of the gender difference (Frazier et al., 1995). We used two different classification schemes to test whether the gender difference varied according to different types of behaviors. First, we conducted meta-analyses on the two broad groups of behaviors identified by the EEOC (1990) as hostile work environment and quid pro quo harassment. We found the mean effect size to be significantly larger for hostile work environment harassment. Second, in an effort to differentiate between the social—sexual behaviors included in hostile work environment harassment, we proposed a category scheme that consisted of seven groups of behaviors. These categories included derogatory attitudes (personal and impersonal), dating pressure, sexual propositions, physical contact (sexual and nonsexual), and sexual coercion. Separate meta-analyses within each category showed that the gender difference was larger for the less extreme and more ambiguous behaviors like derogatory attitudes and dating pressure than for sexual propositions and sexual coercion. However, behaviors included in the category labeled physical sexual contact were extreme but produced a larger gender difference relative to other categories, which also included extreme social—sexual behaviors (sexual propositions and sexual coercion). This finding, however, is consistent
with prior research that demonstrated a similar effect size for this type of behavior (Kenig & Ryan, 1986). Gutek (1985) also found that women were more likely than men to perceive sexual touching as sexual harassment (e.g., 59% for men and 89% for women). Furthermore, a reviewer raised the possibility that men are more likely to view physical sexual contact as a compliment whereas women are more likely to view it as a threat. In an effort to explain this anomaly, we examined various study characteristics to see whether any of these features were confounded with the behavior categories. However, study characteristics were not confounded with behavior category. Behaviors included in the category labeled physical nonsexual contact were less extreme but demonstrated a smaller gender difference relative to derogatory attitudes and dating pressure. However, although these behaviors may have been less extreme, they were also less ambiguous (a congratulatory hug). Furthermore, researchers found that when social–sexual behaviors are so benign that they are not harassment, the gender gap closes (Gutek & O’Connor, 1995). This may explain the small gender difference for physical nonsexual contact. Although behaviors in this category involved physical contact, they were not sexual, nor did they imply sexist attitudes or propositions.

In sexual harassment claims, the courts often struggle with trying to determine whether an intimidating, hostile, or offensive working environment has been created (Gutek & O’Connor, 1995). Our finding that the size of the gender difference depends on the type of social–sexual behavior in question brings us one step closer to understanding which behaviors may contribute to perceptions that such an environment has been created. For example, men and women agree that sexual coercion and sexual propositions constitute sexual harassment. However, they do not necessarily agree that sex-stereotyped jokes or repeated requests for dates after refusal do. Therefore, a woman may perceive that sexual harassment has occurred after a number of the latter types of social–sexual behaviors have taken place, whereas a man may be less inclined to do so. Furthermore, a finding that physical sexual contact elicits a standardized mean difference of 0.36 indicates that men and women tend to perceive these types of behaviors in different ways. Men may interpret the behavior as flattery, whereas women may perceive it as something that may escalate to harassment. Although a gender difference exists, it is larger for behaviors that are consistent with the courts’ definition of hostile work environment harassment. Furthermore, significance tests rely on the assumption that the effect sizes included in the two categories come from different study samples. The nature of the data did not permit this assumption to be satisfied, making it harder to detect differences. Therefore, finding that some of the mean effect sizes between the categories were significantly different from one another is evidence that the corresponding female–male differences are real.

Our meta-analysis also tested whether status of the harasser moderated the gender difference. Although the mean effect sizes suggest that women were more likely than men to evaluate behaviors as harassing when the harasser was a peer, coworker, or fellow student, the difference was not significant. Sexual harassment has been defined to include social–sexual behaviors that are made a condition of employment or that create a hostile work environment. Considering that people of higher authority have greater control over conditions of employment, it is not surprising to find that men and women demonstrate greater agreement when the harasser is in a position of higher authority. When the harasser is of the same status as the harassee, a different type of working relationship exists. Therefore, it could be argued that men interpret social–sexual behaviors in these circumstances as casual or informal interaction as opposed to harassment per se.

We also tested whether the format by which behaviors are evaluated moderates the gender difference. Our results supported those of Blumenthal (1998). Both formats produced effect sizes around 0.30. Apparently, providing more detail about the incident and the type of interaction between harasser and target does not affect whether men and women interpret social–sexual behaviors as constituting sexual harassment. The finding that presentation format did not moderate the gender difference has important implications for how sexual harassment is studied. It suggests that a lot of weight is placed on the behavior itself as opposed to circumstances surrounding the event.

Although progress has been made at defining sexual harassment, it is still unclear as to whose perspective should be taken when the circumstances surrounding a case of harassment are evaluated in the courts. This issue, paired with research evidence that women define a broader range of behaviors as harassing, has given rise to the implementation of a legal standard that relies on the point of view of a reasonable woman as opposed to a reasonable person. However, the decision to implement this standard was made without any conclusive evidence as to the size of the gender difference. Blumenthal’s (1998) review provided the first attempt to quantify this difference. He found that across a broad range of behaviors, women were more likely than men to label behaviors as sexual harassment. However, the magnitude of the difference was small. Our study extended the research of Blumenthal by investigating whether the variation in the gender difference could be explained by different types of behaviors. We found that the female–male difference was greater for social–sexual behaviors that reflect derogatory attitudes and involve dating pressure or physical sexual contact. The gender difference was significantly smaller for behaviors that involve sexual propositions and sexual coercion. Although these two quantitative reviews show that a gender difference does exist and that it is larger for behaviors that the courts define as hostile work environment harassment, the magnitude of the difference is not large. Overall, these findings do not provide strong support for a legal standard that calls for the courts to use the reasonable woman standard when evaluating claims of sexual harassment. It has been argued that the adoption of the reasonable woman standard in the presence of a gender difference may be unfair to men because it does not take their viewpoint into consideration when deciding if sexual harassment has occurred (Meads, 1993). Furthermore, the present study shows that behavior category does not explain all of the variation in the gender difference. Therefore, future research needs to identify and examine other moderators before strong assertions can be made about the appropriateness of the reasonable woman standard.

The finding that a gender difference is larger for certain social–sexual behaviors calls for the need to determine the origin of this difference (Gutek & O’Connor, 1995). At present, it is not clear whether gender differences are innate or a product of socialization and a person’s value system. Men and women may be socialized to perceive different social–sexual behaviors as appropriate or inappropriate. Therefore, it is conceivable that a series of behaviors may be perceived as flattery by one group and as harassment...
by another solely on the basis of one’s value system or how one is socialized. Furthermore, prior experiences with social–sexual behaviors that eventually led to harassment may cause people to interpret isolated occurrences of social–sexual behaviors more seriously, fearing that they will escalate to sexual harassment. Although the two quantitative reviews examined a number of moderators of the gender difference, there may be other factors that explain additional variance.

References

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


Ellison v. Brady, 924 F.2d 872, 879 (9th Cir. 1991).


*Lipsitt, v. University of Puerto Rico, 864 F.2d 881 (lst Cir. 1988).*


*Revision received October 16, 2000

*Accepted October 27, 2000*