The Gender Similarities Hypothesis

Janet Shibley Hyde
University of Wisconsin—Madison

The differences model, which argues that males and females are vastly different psychologically, dominates the popular media. Here, the author advances a very different view, the gender similarities hypothesis, which holds that males and females are similar on most, but not all, psychological variables. Results from a review of 46 meta-analyses support the gender similarities hypothesis. Gender differences can vary substantially in magnitude at different ages and depend on the context in which measurement occurs. Overinflated claims of gender differences carry substantial costs in areas such as the workplace and relationships.

Keywords: gender differences, gender similarities, meta-analysis, aggression

The mass media and the general public are captivated by findings of gender differences. John Gray’s (1992) Men Are From Mars, Women Are From Venus, which argued for enormous psychological differences between women and men, has sold over 30 million copies and been translated into 40 languages (Gray, 2005). Deborah Tannen’s (1991) You Just Don’t Understand: Women and Men in Conversation argued for the different cultures hypothesis: that men’s and women’s patterns of speaking are so fundamentally different that men and women essentially belong to different linguistic communities or cultures. That book was on the New York Times bestseller list for nearly four years and has been translated into 24 languages (AnnOnline, 2005). Both of these works, and dozens of others like them, have argued for the differences hypothesis: that males and females are, psychologically, vastly different. Here, I advance a very different view—the gender similarities hypothesis (for related statements, see Epstein, 1988; Hyde, 1985; Hyde & Plant, 1995; Kimball, 1995).

The Hypothesis

The gender similarities hypothesis holds that males and females are similar on most, but not all, psychological variables. That is, men and women, as well as boys and girls, are more alike than they are different. In terms of effect sizes, the gender similarities hypothesis states that most psychological gender differences are in the close-to-zero (d ≤ 0.10) or small (0.11 < d < 0.35) range, a few are in the moderate range (0.36 < d < 0.65), and very few are large (d = 0.66–1.00) or very large (d > 1.00).

Although the fascination with psychological gender differences has been present from the dawn of formalized psychology around 1879 (Shields, 1975), a few early researchers highlighted gender similarities. Thorndike (1914), for example, believed that psychological gender differences were too small, compared with within-gender variation, to be important. Leta Stetter Hollingworth (1918) reviewed available research on gender differences in mental traits and found little evidence of gender differences. Another important reviewer of gender research in the early 1900s, Helen Thompson Woolley (1914), lamented the gap between the data and scientists’ views on the question:

The general discussions of the psychology of sex, whether by psychologists or by sociologists show such a wide diversity of points of view that one feels that the truest thing to be said at present is that scientific evidence plays very little part in producing convictions. (p. 372)

The Role of Meta-Analysis in Assessing Psychological Gender Differences

Reviews of research on psychological gender differences began with Woolley’s (1914) and Hollingworth’s (1918) and extended through Maccoby and Jacklin’s (1974) watershed book The Psychology of Sex Differences, in which they reviewed more than 2,000 studies of gender differences in a wide variety of domains, including abilities, personality, social behavior, and memory. Maccoby and Jacklin dismissed as unfounded many popular beliefs in psychological gender differences, including beliefs that girls are more “social” than boys; that girls are more suggestible; that girls have lower self-esteem; that girls are better at rote learning and simple tasks, whereas boys are better at higher level cognitive processing; and that girls lack achievement motivation. Maccoby and Jacklin concluded that gender differences were well established in only four areas: verbal ability, visual-spatial ability, mathematical ability, and aggression. Overall, then, they found much evidence for gender similarities. Secondary reports of their findings in textbooks and other sources, however, focused almost exclusively on their conclusions about gender differences (e.g., Gleitman, 1981; LeFrancois, 1990).

Preparation of this article was supported in part by National Science Foundation Grant REC 0207109. I thank Nicole Else-Quest, Sara Lindberg, Shelly Grabe, and Jenni Petersen for reviewing and commenting on a draft of this article.

Correspondence concerning this article should be addressed to Janet Shibley Hyde, Department of Psychology, University of Wisconsin—Madison, 1202 West Johnson Street, Madison, WI 53706. E-mail: jshyde@wisc.edu
Gender meta-analyses generally proceed in four steps: (a) The researcher locates all studies on the topic being reviewed, typically using databases such as PsycINFO and carefully chosen search terms. (b) Statistics are extracted from each report, and an effect size is computed for each study. (c) A weighted average of the effect sizes is computed (weighting by sample size) to obtain an overall assessment of the direction and magnitude of the gender difference when all studies are combined. (d) Homogeneity analyses are conducted to determine whether the group of effect sizes is relatively homogeneous. If it is not, then the studies can be partitioned into theoretically meaningful groups to determine whether the effect size is larger for some types of studies and smaller for other types. The researcher could ask, for example, whether gender differences are larger for measures of physical aggression compared with measures of verbal aggression.

**The Evidence**

To evaluate the gender similarities hypothesis, I collected the major meta-analyses that have been conducted on psychological gender differences. They are listed in Table 1, grouped roughly into six categories: those that assessed cognitive variables, such as abilities; those that assessed verbal or nonverbal communication; those that assessed social or personality variables, such as aggression or leadership; those that assessed measures of psychological well-being, such as self-esteem; those that assessed motor behaviors, such as throwing distance; and those that assessed miscellaneous constructs, such as moral reasoning. I began with meta-analyses reviewed previously by Hyde and Plant (1995), Hyde and Frost (1993), and Ashmore (1990). I updated these lists with more recent meta-analyses and, where possible, replaced older meta-analyses with more up-to-date meta-analyses that used larger samples and better statistical methods.

Hedges and Nowell (1995; see also Feingold, 1988) have argued that the canonical method of meta-analysis—which often aggregates data from many small convenience samples—should be augmented or replaced by data from large probability samples, at least when that is possible (e.g., in areas such as ability testing). Test-norming data as well as data from major national surveys such as the National Longitudinal Study of Youth provide important information. Findings from samples such as these are included in the summary shown in Table 1, where the number of reports is marked with an asterisk.

Inspection of the effect sizes shown in the rightmost column of Table 1 reveals strong evidence for the gender similarities hypothesis. These effect sizes are summarized in Table 2. Of the 128 effect sizes shown in Table 1, 4 were unclassifiable because the meta-analysis provided such a wide range for the estimate. The remaining 124 effect sizes were classified into the categories noted earlier: close-to-zero \((d \leq 0.10)\), small \((0.11 < d < 0.35)\), moderate \((0.36 < d < 0.65)\), large \((d = 0.66–1.00)\), or very large \((>1.00)\). The striking result is that 30% of the effect sizes are in the close-to-zero range, and an additional 48% are in the small range. That is, 78% of gender differences are

\[
d = \frac{M_M - M_F}{s_w},
\]

where \(M_M\) is the mean score for males, \(M_F\) is the mean score for females, and \(s_w\) is the average within-sex standard deviation. That is, \(d\) measures how far apart the male and female means are in standardized units. In gender meta-analysis, the effect sizes computed from all individual studies are averaged to obtain an overall effect size reflecting the magnitude of gender differences across all studies. In the present article, I follow the convention that negative values of \(d\) mean that females scored higher on a dimension, and positive values of \(d\) indicate that males scored higher.

**Table 1**

<table>
<thead>
<tr>
<th>Category</th>
<th>Effect Sizes</th>
<th>Studies</th>
<th>Mean Effect Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cognitive variables</td>
<td>32</td>
<td>42</td>
<td>0.10</td>
</tr>
<tr>
<td>Verbal communication</td>
<td>20</td>
<td>33</td>
<td>0.20</td>
</tr>
<tr>
<td>Social personality</td>
<td>12</td>
<td>15</td>
<td>0.25</td>
</tr>
<tr>
<td>Psychological well-being</td>
<td>18</td>
<td>26</td>
<td>0.30</td>
</tr>
<tr>
<td>Motor behaviors</td>
<td>18</td>
<td>22</td>
<td>0.40</td>
</tr>
<tr>
<td>Miscellaneous constructs</td>
<td>12</td>
<td>18</td>
<td>0.50</td>
</tr>
</tbody>
</table>

*Note: The number of studies is marked with an asterisk.*
Previous research suggested that sex differences in personality traits are larger in prosperous, healthy, and egalitarian cultures in which women have more opportunities equal with those of men. In this article, the authors report cross-cultural findings in which this unintuitive result was replicated across samples from 55 nations (N = 17,637). On responses to the Big Five Inventory, women reported higher levels of neuroticism, extraversion, agreeableness, and conscientiousness than did men across most nations. These findings converge with previous studies in which different Big Five measures and more limited samples of nations were used. Overall, higher levels of human development—including long and healthy life, equal access to knowledge and education, and economic wealth—were the main nation-level predictors of larger sex differences in personality. Changes in men’s personality traits appeared to be the primary cause of sex difference variation across cultures. It is proposed that heightened levels of sexual dimorphism result from personality traits of men and women being less constrained and more able to naturally diverge in developed nations. In less fortunate social and economic conditions, innate personality differences between men and women may be attenuated.

Keywords: sex differences, personality traits, culture

Why can’t a woman be more like a man?
—Alan Jay Lerner, My Fair Lady

In many studies, including several meta-analytic investigations, it has been found that men tend to be more assertive and risk taking than women, whereas women are generally higher than men in anxiety and tender-mindedness (Brody & Hall, 2000; Byrnes, Miller, & Schafer, 1999; Feingold, 1994; Kring & Gordon, 1998; Lynn & Martin, 1997; Maccoby & Jacklin, 1974). These sex differences in personality traits can be detected in early childhood (Else-Quest, Hyde, Goldsmith, & Van Hulle, 2006; Wilgenbusch & Merrell, 1999) and remain fairly constant across adulthood (Feingold, 1994; McCrae & Costa, 1984). The effects of these sex differences lead to predictable differences in men’s and women’s leisure behaviors, occupational preferences, and health-related outcomes (Browne, 1998; Collaer & Hines, 1995; Lippa, 2005). Although sex differences in personality traits are not as large as sex differences in mate preferences, permissive sexual behaviors, or physical strength (Feingold, 1992; Schmitt, 2005b; Thomas & French, 1985), sex differences in personality traits do appear to be larger and more robust than sex differences in other domains such as cognitive ability, attributional style, and self-esteem (Else-Quest et al., 2006; Hyde, 2005).

Observed sex differences in personality traits such as assertiveness and anxiety also appear to be culturally pervasive (Costa, Terracciano, & McCrae, 2001; Lynn & Martin, 1997). Feingold (1994) found that women in Canada, China, Finland, Germany, Poland, and Russia tended to score higher than men on scales related to the personality traits of neuroticism, agreeableness, and conscientiousness. Men, in contrast, scored higher in the extraversion-related trait of assertiveness across cultures. In a much larger study, self-report responses to the Revised NEO Personality Inventory (NEO-PI-R) across 36 cultures revealed that women in most countries are higher in several traits related to neuroticism, agreeableness, warmth, and openness to feelings, whereas men score higher on scales measuring assertiveness and openness to ideas (Costa et al., 2001; McCrae, 2002).
Many of these sex differences in personality traits appear to transcend data sources (e.g., Williams & Best, 1990). In a large study of 50 cultures, college students were asked to identify an adult or a college-aged man or woman whom they knew well and to rate that person’s personality traits, again using the NEO-PI-R (McCrae, Terracciano, & 78 Members of the Personality Profiles of Cultures Project, 2005). Men were rated by observers as being higher than women in assertiveness, excitement seeking, and openness to ideas. Women were rated by observers as being higher on many traits, especially in anxiety, vulnerability, aesthetics, feelings, and tender-mindedness (McCrae et al., 2005). Thus, sex differences in personality traits seem to be rather robust, persisting across a diverse array of measures, data sources, ages, and cultures.

Cultural Variability in the Size of Sex Differences in Personality

Sex differences in most personality traits, however, are not uniform in magnitude across all samples. At times, sex differences can be much larger in some cultures than in others (Fischer & Manstead, 2000; Guimond et al., 2007; Schwartz & Rubel, 2005). One unexpected finding has been that sex differences in personality traits are often larger in prosperous, healthy, and egalitarian cultures in which women have more opportunities equal with men (Costa et al., 2001; McCrae, 2002). Both in self-report and in other-report data, Asian and African cultures generally show the smallest sex differences, whereas European and American cultures—in which living standard and gender equity indexes are generally higher—show the largest differences (McCrae et al., 2005). With improved national wealth and equality of the sexes, it seems differences between men and women in personality traits do not diminish. On the contrary, the differences become conspicuously larger.

It might seem intuitive to think that the more prosperous and egalitarian a society, the more free men and women are to be similar in terms of their personality profiles. This logic appears useful for explaining certain value priorities and sexual strategies pursued by men and women. For instance, Schwartz and Rubel (2005) found that sex differences in the value of self-direction are smaller in richer countries with more individualist and autonomous values than in poorer countries with more collectivist and embedded cultures. Similarly, Schmitt (2005b) found that sex differences in sociossexual orientation are smaller (though still moderate in magnitude) in countries with higher levels of prosperity and sexual equality. The finding that sex differences in personality traits are larger in rich and egalitarian cultures may therefore be somewhat counterintuitive and is certainly contrary to other established patterns of sexual differentiation across cultures. It is important to note that understanding this intriguing cross-cultural pattern might be particularly informative for discerning the ultimate origins of personality traits.

Explaining Cultural Variability in the Size of Sex Differences in Personality

Several theoretical approaches would appear useful in explaining cultural variability in the size of sex differences in personality. In general, these approaches are founded on the same group of theories used to explain the basic origins of psychological sex differences.

1. Social Role Explanations

A leading candidate for explaining variations in the size of sex differences across cultures is the social role model approach. According to this approach, most sex differences are assumed to result from exposure to sex role socialization, a process whereby culture defines the appropriate ways of thinking, feeling, and behaving for men and women (Eagly, 1987; Ruble & Martin, 1998; though see Maccoby, 2000).

Because specified male and female roles are thought to contribute directly to all observed psychological differences between men and women, including personality traits, it is expected that when men and women occupy social roles that are more similar, sex differences will tend to erode (Eagly & Wood, 1999; Wood & Eagly, 2002). Thus, the social role model approach predicts that sex differences in personality traits will be attenuated in more progressive and gender egalitarian cultures and will be accentuated in more traditional cultures.

2. Evolutionary Explanations

Evolutionary approaches consider sex-related differences as arising, in part, from innate dispositional differences between the sexes (Baron-Cohen, 2003; Buss, 1997; Geary, 1998). In this view, the sexes are thought to psychologically differ only in domains in which they have faced different adaptive problems throughout evolutionary history. As a consequence, much of the sex-related differences that appear in modern societies may be due to sexual selection pressures that shaped psychological sex differences in the evolutionary past (Buss & Kenrick, 1998; Mealey, 2000).

Sex differences in levels of obligatory parental investment (Symons, 1979; Trivers, 1972) are thought to have led to sexual selection pressures causing men to be more prone to take risks and to seek social dominance (which benefits the lesser-investing parental sex in a species), whereas women are thought to have been selected to be more cautious and nurturing (which benefits the heavier-investing parental sex; Buss, 1997; Campbell, 2002; MacDonald, 1995). Although evolutionary explanations can readily account for the existence of culturally pervasive differences between men and women, such explanations may seem less adept at explaining the variability in the size of sex differences across cultures (though see Buss, 2001; Gangestad, Haselton, & Buss, 2006; Gangestad & Simpson, 2000).

One evolutionary approach that directly addresses cultural variability is the mismatch perspective (Crawford, 1998; Nesse & Williams, 1994). Evolutionary mismatch theories explain psychological variations across cultures by the degree of mismatch between contemporary environmental conditions and those in which early humans evolved—namely, hunter-gatherer environments (Brown, 1991; Tooby & DeVore, 1987). When contemporary environments are different from hunter-gatherer environments, the adaptive development of innate psychological sex differences can be impeded.

Mismatches between ancestral conditions in which sex differences in personality evolved and contemporary environments might at first glance appear largest in the most modern and
industrialized nation-states. However, this may not be the case (Pasternak, Ember, & Ember, 1997). Schmitt (2005a) has argued that the psychological mismatch between contemporary environments and those in which early humans evolved is not always a linear function of sociocultural time. For example, according to the curvilinear hypothesis of cultural variation (Schmitt, 2005a), modern nation-states may be psychologically closer to hunter-gatherer cultures than are less-developed agricultural or pastoral cultures (Lee & Daly, 1999). Agricultural and pastoral cultures, with extremely large disparities in resource distribution, familial isolation, and relative gender inequality, may represent the largest psychological deviations from our hunter-gatherer past (Korotayev & Kazankov, 2003; Lamb & Hewlett, 2005). Over sociocultural time, therefore, our most modern postagricultural environments may be gradually becoming more similar to, not more different from, the hunter-gatherer psychological conditions in which sex differences in personality traits evolved.

3. Artifact Explanations

A third type of explanation assumes that observed sex differences in personality are caused by forms of measurement error. For example, it is possible that social desirability biases lead men and women to endorse particular gender-relevant traits at different levels. In some cultures, certain traits (such as fearfulness) may be less undesirable for women to endorse than for men. Observed sex differences, in this case, would not reflect actual personality trait differences and would instead reflect each sex’s comfort in revealing undesirable personality characteristics. However, the hypothesis that men and women have different social desirability biases in some cultures is unlikely. For instance, a study of 10 countries from around the world found strong correlations between men’s and women’s favorability ratings (Williams, Satterwhite, & Saiz, 1998).

Another possibility is that different frames of reference for self-description are used in different cultures (Guimond et al., 2007). Costa and his colleagues (Costa, Terracciano, & McCrae, 2001) proposed the following scenario: Self-descriptions in some cultures (but not others) are based on comparisons of the self with others of the same gender. For example, when asked whether she was kind, a traditional woman might rank herself relative to women she knows, but not relative to men. In that case, sex differences would be eliminated, just as they are eliminated by the use of within-sex norms. By contrast, in modern and more egalitarian cultures men and women may compare themselves with others from both sexes and thus reveal true sex differences in personality. If respondents in traditional cultures were explicitly used as a part of the International Sexuality Description Project (ISDP; Schmitt & 121 Members of the ISDP, 2003, 2004). The 55 diverse nations of the ISDP allowed us to explore a wide range of culture-level factors that might influence variability in personality sex differences across cultures.

Basic Aims of the Current Study

The first goal in this study was to replicate the previously observed sex differences in personality traits. Thus far, there have been only two studies in which the widening gap between the personalities of men and women in more modern cultures has been reported. In both cases, the NEO-PI-R was used, either in self-report (Costa et al., 2001; McCrae, 2002) or in observer-report (McCrue et al., 2005) formats. Therefore, it is unknown whether this observed regularity across cultures is produced by the NEO-PI-R instrument itself and is not replicated by other personality measures. Although 36 (McCrae, 2002) and 50 (McCrae et al., 2005) cultures are large datasets, the selection of cultures in these studies has been biased toward European nations. Improved attempts to generalize these findings would include more diverse cultures, especially those from Africa. In the current study, several new African, Asian, and Middle Eastern samples were included.

The second aim in the current study was to provide evidence that could constrain the range of possible explanations for the widening gap between men’s and women’s personality traits in developed and more egalitarian countries. Obviously, this unresolved issue is not due to a lack of theoretical explanations but instead lies in the absence of decisive evidence that could eliminate less plausible theories. The current study was based on one of the largest cross-cultural studies of personality ever conducted, carried out as a part of the International Sexuality Description Project (ISDP; Schmitt & 121 Members of the ISDP, 2003, 2004). The 55 diverse nations of the ISDP allowed us to explore a wide range of culture-level factors that might influence variability in personality sex differences across cultures.

Method

The research reported in this article is a result of the ISDP, a collaborative effort of over 100 social, behavioral, and biological scientists from 56 nations (Schmitt & 121 Members of the ISDP, 2003, 2004). A detailed description of the methodology and sampling techniques used in the ISDP is given elsewhere (Schmitt &
higher. In the present article, I follow the convention that negative
values of \( d \) are positive, and positive values of \( d \) are negative, which measures the magnitude of an effect—in this case, gender differences. Meta-analyses quickly appeared on issues such as aggression (Eagly & Steffen, 1986; Hyde, 1984, 1985), and aggression (Eagly & Carli, 1981), gender differences in influenceability (Eagly & Carli, 1981), gender differences in personality variables (Hyde, 1981; Hyde & Linn, 1988; Linn & Petersen, 1985), gender differences in psychosocial or personality variables, such as aggression or leadership; and those that assessed cognitive variables, such as abilities; and those that assessed motor behaviors, such as throwing distance; and those that assessed verbal or nonverbal communication; and those that assessed miscellaneous constructs, such as moral reasoning. I began with meta-analyses reviewed previously by Hyde and Plant (1995), Hyde and Frost (1993), and Ashmore (1990). I have argued that the canonical method of meta-analysis—ter statistical methods.

Gender meta-analyses generally proceed in four steps: (a) The researcher locates all studies on the topic being reviewed, typically using databases such as PsycINFO and sometimes other databases as well. Findings from samples such as these are important to meta-analyses, as are data from major national surveys such as the National Longitudinal Study of Youth, as well as data from large probability samples, at least when that is possible. (b) Statistics are extracted from each report, and an effect size is computed for each study. (c) A weighted average of the effect sizes is computed (weighting by sample size) to obtain an overall effect size reflecting the magnitude of the gender difference. (d) Homogeneity of effect sizes is assessed. If it is not, then the differences when all studies are combined. (e) The Evidence Crucial to meta-analysis is the concept of effect size, which measures how far apart the male and female means are in standardized units. In gender meta-analyses, the measure of effect size typically is the mean score for males, \( M_M \), divided by the average within-sex standard deviation. That is, \( \frac{M_M - M_F}{SD_w} \). In gender meta-analyses, \( d \) is the mean score for males, \( M_M \), divided by the average within-sex standard deviation. That is, \( \frac{M_M - M_F}{SD_w} \).

Differences are larger for measures of physical aggression compared with measures of verbal aggression. Some types of studies yield large effect sizes, such as those that assessed motor behaviors, such as throwing distance; and those that assessed psychosocial or personality variables, such as aggression or leadership; and those that assessed cognitive variables, such as abilities; and those that assessed verbal or nonverbal communication; and those that assessed miscellaneous constructs, such as moral reasoning. I began with meta-analyses reviewed previously by Hyde and Plant (1995), Hyde and Frost (1993), and Ashmore (1990). I have argued that the canonical method of meta-analysis—ter statistical methods.

Gender meta-analyses generally proceed in four steps: (a) The researcher locates all studies on the topic being reviewed, typically using databases such as PsycINFO and sometimes other databases as well. Findings from samples such as these are important to meta-analyses, as are data from major national surveys such as the National Longitudinal Study of Youth, as well as data from large probability samples, at least when that is possible. (b) Statistics are extracted from each report, and an effect size is computed for each study. (c) A weighted average of the effect sizes is computed (weighting by sample size) to obtain an overall effect size reflecting the magnitude of the gender difference. (d) Homogeneity of effect sizes is assessed. If it is not, then the differences when all studies are combined. (e) The Evidence Crucial to meta-analysis is the concept of effect size, which measures how far apart the male and female means are in standardized units. In gender meta-analyses, the measure of effect size typically is the mean score for males, \( M_M \), divided by the average within-sex standard deviation. That is, \( \frac{M_M - M_F}{SD_w} \). In gender meta-analyses, \( d \) is the mean score for males, \( M_M \), divided by the average within-sex standard deviation. That is, \( \frac{M_M - M_F}{SD_w} \).
Separate and Unequal: Occupation-Establishment Sex Segregation and the Gender Wage Gap

Trond Petersen
University of California, Berkeley

Laurie A. Morgan
University of Illinois at Urbana-Champaign

The authors report the first large-scale empirical investigation of within-job wage differences between men and women in the same occupation and establishment, using data first on blue-collar and clerical employees from 16 U.S. industries in 1974–83 and second on employees in 10 professional and administrative occupations. The authors report three findings. First, wage differences at the occupation-establishment level were small even without controls for individual-level characteristics. Hence, within-job wage discrimination was much less important than occupation-establishment segregation for observed wage differences. Second, establishment segregation was an important cause, although not as important as occupational segregation, of wage differences. Third, establishment segregation was extensive, as was occupational segregation.

I. INTRODUCTION

Wage differences between men and women caused by discrimination can result from several processes. The first is where women are differentially allocated to occupations and establishments that pay lower wages. This

1 We are grateful to Marta Elvira, Kenneth Koput, Katrine Teigen, and Vemund Snartland for research assistance. We also thank Erling Barth, Robert Erikson, Jan Hoem, Geir Hågsnes, Carl LeGrand, David Levine, Karin Martin, Arne Mastekaasa, Eva Meyerson, Charles O'Reilly, Seymour Spilerman, Donald Tomaskovic-Devey, and the AJS reviewers for useful comments and discussions. We also thank Erica Groshen who provided data from several of the 16 Industry Wage Surveys analyzed here. The research was supported by the Institute of Industrial Relations at the University of California, Berkeley, and by National Science Foundation grant SES-8912502. Direct correspondence to Trond Petersen, Walter A. Haas School of Business, 350 Barrows Hall, University of California, Berkeley, California 94720.

© 1995 by The University of Chicago. All rights reserved.
0002-9602/96/10102-0003$01.50

AJS Volume 101 Number 2 (September 1995): 329–65 329
process may involve discrimination partly through differential access to occupations and establishments, that is, the matching process at the point of hire, and partly through subsequent promotions.\(^2\) We call this process “allocative discrimination.” The second is where occupations held primarily by women are paid lower wages than those held primarily by men, although skill requirements and other wage-relevant factors are the same. This is the issue addressed by comparable worth. We call this process “valuative discrimination.” The third is where women receive lower wages than men within a given occupation within a given establishment. We call this process “within-job wage discrimination.”\(^3\) Allocative and valuative discrimination involve the segregation of men and women into different occupations, establishments, or both and may occur without within-job wage discrimination. Thus, it may be the case that where men and women share the same jobs they receive the same pay but that in most cases they simply do not share the same jobs.

One conjecture currently accepted by many researchers is that wage differences are less an issue of within-job wage discrimination and more a matter of allocative and valuative processes. That is, the segregation of women into lower-paying occupations, establishments, or both and lower pay in occupations held primarily by women are more important than pay differences within the same job in explaining the gender wage gap. Treiman and Hartmann (1981, pp. 92–93) write, “Although the committee recognizes that instances of unequal pay for the same work have not been entirely eliminated, we believe that they are probably not now the major source of differences in earnings.”

This conjecture is drawn primarily from a large literature that focuses on pay differences across and within occupations. One pattern of findings is that the wage gap between men and women becomes smaller as occupational controls become finer (Treiman and Hartmann 1981; Marini 1989),

\(^2\) We follow the convention that discrimination occurs when wage differences between men and women are not accounted for by average differences in productive attributes, which is not to say that the attainment of productive attributes, e.g., education, is itself not related to discrimination (e.g., American Association of University Women 1992).

\(^3\) Treiman and Hartmann (1981, pp. 8–9) refer to allocative discrimination as employment discrimination, and valuative and within-job wage discrimination as wage discrimination. Both allocative and within-job wage discrimination are illegal. The former is covered by Title VII of the 1964 Civil Rights Act, while the latter is covered by the Equal Pay Act of 1963 (see England 1992, chap. 5). Valuative discrimination is discrimination against classes of jobs occupied primarily by women but not discrimination against any specific individual. Its legal status is unclear, but the current legal situation can be summarized as one in which the courts do not interpret Title VII to require comparable worth unless the plaintiff can show that a job was intentionally paid less because the incumbents are women (see England 1992, chap. 5).
suggesting that a large proportion of the wage gap is explained by occupational distribution. For example, Treiman and Hartmann (1981, pp. 33–39) explained 10%–20% of the raw wage gap using 222 occupational categories and 35%–40% using 479 categories. These studies usually draw on data from the census or national probability samples that allow no analysis of practices in specific establishments. Additional evidence suggests that, within occupations, the distribution of women across firms or establishments also accounts for some portion of the wage gap. For example, Blau (1977) found that in 11 clerical occupations, differences in men’s and women’s wages were larger between than within establishments.

Yet the prevailing conjecture remains a conjecture. It has not been shown that men and women receive equal pay within given occupations in given establishments. What has been shown is that sex segregation is extensive and pervasive (Bielby and Baron 1984), but not the extent to which sex segregation accounts for the wage gap or that, when sex segregation is absent, the sexes receive equal treatment. To confirm such a claim, one needs data on wages of men and women in the same detailed occupational group or position within the same establishment. Such data are not widely available except on isolated establishments.

This article reports the first relatively large-scale empirical investigation of wage differences between men and women within the same detailed occupational position within the same establishment. We use establishment-level data from a wide variety of industries. In each establishment, individual-level wage data for a large array of detailed occupational groups were collected, providing more accurate wage as well as occupational data than probably any other surveys available (except in some case studies of single establishments, e.g., Hartmann [1987]). We focus first on production and clerical employees in 16 U.S. industries in the 1974–83 period, primarily 1974–78, analyzing data on about 870,000 employees, 700 industry-specific occupations, 6,000 establishments, and 71,000 occupation-establishment pairs, where each occupation within an establishment is an occupation-establishment pair. Second, we focus on seven professional and three administrative occupations across a broad range of industries in 1981, analyzing data on about

---

4 A recent and extensive set of occupational case studies can be found in Reskin and Roos (1990), mostly using data from the 1980 U.S. census (see, e.g., U.S. Bureau of Census 1984). When three-digit occupational groups are studied and there are few or no controls for other characteristics, women’s annual earnings are about 30%–40% below men’s in most of the 11 occupations studied.

5 Treiman and Hartmann (1981, p. 33) write on the standard analyses of wage differences: “This exercise illustrates that further analysis of occupational segregation requires much more detailed data than are currently available from the census or from national sample surveys.”
740,000 employees distributed across 2,162 establishments and 16,433 occupation-establishment pairs.

Apart from Blau's (1977) study of 11 clerical occupations, the studies that most resemble our design are Groshen (1991) and Tomaskovic-Devey (1993). Unlike the present study, neither computes the amount of within-job wage differences between men and women (i.e., in the same occupation and establishment). Tomaskovic-Devey uses a random sample of employees in North Carolina, and so is unable to make this calculation, because he has no data on men and women working in the same jobs. However, he does have information on the sex composition of each respondent's job, which he includes as an independent variable in a regression analysis. Groshen (1991) uses data on six of the 16 industries analyzed here, but she does not compute within-job wage differentials, performing instead the same type of analysis as Tomaskovic-Devey. Both authors report that the sex composition of jobs (i.e., occupation-establishment) accounts for a large portion of the wage gap.

We make no attempt to settle the important conceptual issues that go along with the empirical patterns we address, namely the sources of observed patterns, neither from the demand side, that is, discriminatory behavior by employers, or the supply side, that is, behaviors by employees and prospective employees (see England 1992, chap. 2). Nevertheless, our results have implications for the kinds of theoretical issues that are most in need of being addressed and for the type of data that need to be collected and analyzed.

II. DATA

We use two large-scale data sets. The first data set comes from 16 Industry Wage Surveys (IWSs) conducted by the U.S. Bureau of Labor Statistics (BLS) in the period 1974–83 (see, e.g., U.S. Department of Labor 1976a), corresponding to industry codes at three and more digits as defined in the Standard Industrial Classification Manual (see U.S. Executive Office of the President 1987). Eleven industries were surveyed in 1974–78, while five were surveyed in 1980–83. The populations for the surveys and the sampling from the populations are described in the U.S. Department of Labor publications listed in the note to table 1 (e.g., 1976a, p. 48). Of the 16, 11 are manufacturing industries, while 5 are service industries. The selection of industries was to a large extent determined by availability from the BLS.6 Table 1 lists the industries analyzed.

6 These surveys ceased to be collected in September 1990, after about 100 years in operation. Information on gender of employees was no longer collected by the mid-1980s. The data analyzed below may be the last Industry Wage Survey to be made available to researchers, due to changes in data-processing procedures at the BLS.
New Evidence on Sex Segregation and Sex Differences in Wages from Matched Employee-Employer Data

Kimberly Bayard, Board of Governors of the Federal Reserve System

Judith Hellerstein, University of Maryland and National Bureau of Economic Research

David Neumark, Public Policy Institute of California, Michigan State University, and National Bureau of Economic Research

Kenneth Troske, University of Missouri—Columbia and the Institute for the Study of Labor

We use new matched employer-employee data to estimate the contributions of sex segregation and wage differences by sex within occupation, industry, establishment, and occupation-establishment cells to the overall sex gap in wages. In contrast to earlier data used to study this question, our data cover all industries and occupations across all regions of the United States. We find that segregation of women into lower-paying occupations, industries, establishments, and occupations within establishments accounts for a sizable fraction of the sex gap in wages. Nonetheless, approximately one-half of the sex gap in wages remains attributable to the individual’s sex.

We thank Fran Blau, Bill Evans, Claudia Goldin, Erica Groshen, Larry Katz, Robert LaLonde, Jonathan Leonard, Brooks Pierce, Peter Schmidt, and seminar/
I. Introduction

Women have consistently earned lower wages than men in U.S. labor markets, although this gap has narrowed in recent decades (Blau 1998). Understanding the sources of sex differences in wages is vital to determining why the wage gap between men and women persists. Previous research has focused on the impact of the occupational segregation of men and women on the wage gap (e.g., Macpherson and Hirsch 1995), the effect of industry segregation (e.g., Fields and Wolff 1995), and, to a lesser extent, on the segregation of men and women into different employers (Blau 1977; Bielby and Baron 1984; Carrington and Troske 1998). These studies all find evidence that the wage gap falls considerably after accounting for segregation.

Evidence on the contribution to the wage gap of within-establishment, within-occupation segregation is far harder to find. Indeed, we are not aware of any empirical work on this issue that uses large data sets representative of a wide array of industries. The reason for this is the paucity of data sets containing detailed demographic information for multiple workers in the same establishment. As a result, studies of the effects of establishment and occupation-establishment segregation have used unusual, quite narrow data sets. For example, the best-known study is by Groshen (1991); it uses surveys of wages for a subset of occupations in five specific industries included as part of the Bureau of Labor Statistics (BLS) Industry Wage Surveys (IWS). In earlier work, Blau (1977) used the BLS Area Wage Surveys to provide a decomposition of the sex gap in wages, including evidence on the importance of an individual’s sex within occupation, establishment, and job cell. Her data covered subsets of three broad occupations in three large northeastern cities.

The focus in these studies on a handful of industries or occupations provides something closer to a set of case studies, with the lack of representativeness limiting their usefulness in assessing the forces at work in
generating the sex wage gap in the United States. Our goal in this article is to use a much broader and more nationally representative data set to estimate the contributions of sex segregation by industry, occupation, and occupation-establishment cell (job cell) to the sex wage gap. For our analysis, we construct and use an extended version of the Worker-Establishment Characteristics Database (WECD) to decompose the source of male-female wage differentials. Like the WECD, this data set uses the U.S. Census Bureau’s Standard Statistical Establishment List (SSEL) to identify the employers of individuals who responded to the long form of the 1990 Decennial Census. However, whereas the WECD is limited to manufacturing plants, this new data set (the New Worker-Establishment Characteristics Database, or NWECD) includes workers and establishments from all sectors of the economy and all regions. Nonetheless, because of the constraints imposed by matching employees to employers, some nonrepresentative characteristics of the data set are unavoidable.

Using the NWECD, we provide new estimates of the role of various dimensions of sex segregation in generating sex differences in wages. Although in some respects our evidence may be viewed as complementary to that in the earlier studies, in our view, the NWECD, while having some shortcomings, is clearly better suited to characterizing the effects of sex segregation in U.S. labor markets. Our results indicate that a sizable fraction of the sex gap in wages is accounted for by the segregation of women into lower-paying occupations, industries, establishments, and occupations within establishments. We also find, however, that a very substantial part of the sex gap in wages remains attributable to the individual’s sex.

II. The Data

The data used in this study come from a match between worker records from the 1990 Sample Edited Detail File (SEDF) to establishment records in the 1990 Standard Statistical Establishment List (SSEL). The 1990 SEDF consists of all household responses to the 1990 Decennial Census long form. As part of the Decennial Census, one-sixth of all households receive a “long-form” survey, which asks a number of questions about each member of the household (“person questions”) as well as about the housing unit (“housing questions”). Those receiving the long form are asked to identify each employed household member’s (1) occupation, (2) employer location, and (3) employer industry in the previous week. The Census Bureau then assigns occupational, industrial, and geographic codes to long-form responses. Thus, the SEDF contains the standard demographic information for workers collected on the long form of the Decennial

---

1 See Troske (1998) and Bayard et al. (1999) for descriptions of these data sets.