

## The Structure of Sexual Harassment: A Confirmatory Analysis across Cultures and Settings

MICHELE J. GELFAND, LOUISE F. FITZGERALD, AND FRITZ DRASGOW

*University of Illinois at Urbana-Champaign*

Sexual harassment is increasingly recognized as a serious social problem with important implications for individuals, organizations, and society as a whole. Despite increased research attention, fundamental conceptual issues remain unresolved. The present paper proposes that sexual harassment is a stable behavioral construct distinct from but related to evolving legal formulations. Based on previous research and theory, we propose a tripartite model of this construct (i.e., gender harassment, unwanted sexual attention, and sexual coercion) and test it through confirmatory factor analysis conducted simultaneously in three populations. Results confirm the generalizability of the construct across settings (workplace and higher education) and cultures (United States and Brazil) and are discussed in terms of their implications for theory, research, and social policy. © 1995 Academic Press, Inc.

Sexual harassment has recently emerged as a critical social issue with important implications for individuals, organizations, and society as a whole. Numerous studies document its extent (Fitzgerald et al., 1988; Gutek, 1985; USMSPB, 1981; 1987; Tinsley & Stockdale, 1993), judicial decisions have begun to clarify its legal framework (*Ellison v Brady*, 1991; *Franklin v. Gwinnett County School District*, 1992; *Meritor Savings Bank, FSP v. Vinson*, 1986), and both practical guides (Bravo & Cassedy, 1992) and social policy recommendations (Fitzgerald, 1993; 1994) have begun to appear. Psychologists have made important research and policy contributions in this area (American Psychological Association, 1993; APA Taskforce on Male Violence Against Women, 1994; Fiske & Borgida, in press; Fitzgerald, 1993; Gutek, 1985; Pryor & McKinney, in preparation; Tinsley & Stockdale, 1993) and continue to play important roles in shaping public awareness of this issue.

In the intense spotlight generated by public events such as the Thomas confirmation hearings and the Navy "Tailhook" scandal, it is sometimes

The authors acknowledge the considerable conceptual and practical contributions of Charles Hulin and the Sexharass Research Group to the development of this paper. Correspondence and reprint requests should be addressed to Michele Gelfand, Department of Psychology, New York University, 6 Washington Place, New York, NY 10003-6634.

difficult to remember that sexual harassment emerged as a legal concept only slightly over a decade ago (EEOC, 1980; MacKinnon, 1979) and has been the subject of serious scientific scrutiny for somewhat less than that (Gutek, 1985; USMSPB, 1981), although its existence has been documented for nearly a century (Bularzik, 1978). Reflecting on this situation, Fitzgerald and Hesson-McInnis (1989) observed "As with many topics that are both socially important and somewhat controversial, data collection has proceeded rapidly, without benefit of theory, or even the careful formulation of definitions . . . researchers themselves tend to mean different things by the term, leading to data being collected on a variety of behaviors often with no rationale beyond 'these have been used in previous research' " (p. 310). Similar comments were recently provided by Gruber (1992).

Over the past several years, the legal framework of sexual harassment has become increasingly clear. Landmark Supreme Court decisions have confirmed that Title VII of the 1964 Civil Rights Act prohibits both coerced sexual exchange (*quid pro quo*) and the more general *hostile environment* behaviors (*Meritor Savings Bank, FSB v Vinson*, 1986) and that such prohibitions hold for educational as well as workplace settings (*Franklin v Gwinnett County*, 1992). These decisions essentially elevate the definitions and guidelines issued by the Equal Employment Opportunity Commission (EEOC, 1980) to the status of law and parallel the *quid pro quo vs conditions of work* distinction first suggested by MacKinnon (1979). The courts have thus distinguished two general types of sexual harassment; however, the severity needed to trigger the legal standard in any particular instance and whose judgment of severity is to be accepted remain important questions (*Ellison v Brady*, 1991; *Harris v Forklift Systems Inc.*, 1993; *Robinson v Jacksonville Shipyards*, 1991). The decision in *Harris* was widely expected to clarify the standard of severity in hostile environment cases as well as to settle the issue of whose perspective (i.e., reasonable person, reasonable woman) is to be dispositive; however, the Court chose not to address this latter controversy directly, leaving lower courts and the appellate bench without clear guidance on this issue.

In contrast to the broad-brush analysis of contemporary legal theory, the great majority of psychological research has taken place at the level of specific acts, with little attempt to aggregate behaviors at a higher level of generality. Virtually no theoretical attention has been given to defining the domain of this construct, nor to specifying its structure or dimensions, which are presumably logical preconditions to valid assessment. One consequence of this situation has been lack of agreement concerning exactly what behaviors should be examined and at what level of specificity.

The major incidence studies (USMSPB, 1981; 1987) employed a check list approach, examining seven behavioral clusters (sexual teasing, jokes, remarks, or questions; pressure for dates; letters, phone calls, or materials of a sexual nature; sexually suggestive looks or gestures; deliberate touching,

leaning over, cornering or pinching; pressure for sexual favors; actual or attempted rape or sexual assault). Gutek (1985) examined participants' experience of sexual comments, sexual looks or gestures, sexual touching, nonsexual touching, dating as part of the job, and sexual relations expected as part of the job. She further asked them to rate the behaviors as to whether they were intended to be insulting or complementary. Smaller scale studies have generally adopted a similar methodology, although the actual behaviors examined have varied from study to study.

In addition to employing nonstandard stimulus lists, some researchers have asked respondents to report not only their own experiences, but also whether they have ever heard of similar situations happening to others (Lott, Reilly, & Howard, 1982)—thus likely inflating prevalence rates; others have inquired only about a narrow range of extremely severe situations (Maihoff & Forrest, 1983), correspondingly depressing them. The most recent major prevalence study—a telephone survey of a nationally representative sample of working women (Saunders, 1992)—utilized a uniquely generated list of six behaviors examining exclusively situations involving supervisors, although harassment by co-workers has been found to be by far the most common occurrence (Gutek, 1985; USMSPB, 1981; 1987).

The effects of such variation may best be understood in terms of Brogden and Taylor's (1950) classic discussion of criterion bias. These writers defined a biasing factor as "any variable, except errors of measurement and sampling error, producing a deviation of obtained criterion scores from a hypothetical 'true' criterion score" (p. 161). They then classified the sources of bias as (1) deficiency (the omission of pertinent elements), (2) contamination (the introduction of extraneous elements), (3) scale unit bias (inequality of scale units in the criterion), and (4) distortion (improper weighting in combining score elements).

Although specifically focused on criterion construction and measurement, their discussion is enlightening here, as an individual's profile or score on a harassment survey can be considered the observed criterion score, and the actual experiences of harassment the hypothetical true score. From this perspective, Maihoff and Forrest's (1983) brief list of severe harassing behaviors, as well as the absence of co-worker harassment in the Saunders (1992) study, provide examples of criterion deficiency, whereas reports of harassment experienced by others (Lott et al., 1982) and judgments of perpetrator intentions (Gutek, 1985) are instances of contamination.

Perhaps the most common problem has been scale unit bias, which in the present case may be viewed as oversampling from certain aspects of the construct (e.g., unwanted sexual attention) while simultaneously undersampling from others (most notably, gender harassment, i.e., hostile and misogynistic behavior that is not overtly sexual in the traditional sense of a sexual overture). The combination of these factors results in inappropriate score distortion, which includes each of the other types of bias as well: that is, the

assignment of zero weights to factors that should have nonzero weights; the converse assignment of nonzero weights to factors that should have zero weights; and differentially weighting (i.e., sampling) aspects of the construct that should be equally weighted (cf. Crites, 1969, for a more extended discussion).

The first systematic attempt to map the conceptual domain of sexual harassment and develop a comprehensive classification system was initiated by Till (1980). Writing before any legal framework for harassment existed, he classified the responses of college women to an open-ended call for information about their experiences of sexual harassment, a procedure that yielded five distinct types of behaviors: *gender harassment*, *seductive behavior*, *sexual bribery*, *threat of punishment for noncompliance*, and *sexual imposition or assault*. Till (1980) recommended that these categories be thought of as levels of harassment, as they appeared to form a rough continuum of severity, and argued that they were exhaustive as they could encompass any particular example (act) of harassment.

Based on Till's categories, Fitzgerald and her colleagues (Fitzgerald et al., 1988) developed the *Sexual Experiences Questionnaire* (SEQ)—a self-report inventory designed to assess the prevalence of sexual harassment in both work and educational settings. Exploratory factor analysis of the SEQ (Fitzgerald & Shullman, 1985) suggested that the five categories could be combined to yield a more parsimonious classification as a three-factor solution adequately accounted for the data: *gender harassment* (Level one), *sexual harassment* (Levels two and five), and *sexual coercion* (Levels three and four). They argued that these categories were most appropriately thought of as types rather than levels, noting the existence of both mild and severe instances of each. This distinction between type and severity, which is also reflected in current legal theory, was later borne out in a scaling analysis of ratings of SEQ items that revealed two nearly orthogonal dimensions: *quid pro quo vs conditions of work* (type) and *severity* (Fitzgerald & Hesson-McInnis, 1989).

Based on this body of exploratory research and preliminary theory, we propose that sexual harassment is a behavioral construct composed of three related, but conceptually distinct and nonoverlapping dimensions: *gender harassment*, *unwanted sexual attention*, and *sexual coercion*; that these categories are necessary and sufficient to classify any particular incident of harassment; and that they constitute the irreducible minimum of the construct as it is currently understood, both legally and psychologically. We further propose that the structure of the construct is isomorphic across settings (i.e., work vs education) and, within settings, across cultures. These propositions are tested in the present paper via a simultaneous confirmatory factor analysis of the SEQ in samples of women drawn from three separate populations across two cultures. The following description of the dimensions of the framework draws heavily from Fitzgerald et al.'s (1988) reformulation of Till's original system, with the exception that we reserve the term *sexual harassment* for the overall

construct and substitute the term *unwanted sexual attention* to label their second dimension.

### *Gender Harassment*

This category encompasses a range of verbal and nonverbal behaviors generally not aimed at sexual cooperation; rather, they convey insulting, hostile, and degrading attitudes about women. The EEOC (1993) has recently supplemented its original Sex Discrimination Guidelines (EEOC, 1980) with a statement explicitly prohibiting such gender-based harassment in the workplace. According to the EEOC, gender harassment is "verbal or physical conduct that denigrates or shows hostility or aversion" (p. 51269); examples include epithets, slurs, taunts, and gestures; the display or distribution of obscene or pornographic materials; gender-based hazing; and threatening, intimidating, or hostile acts.

### *Unwanted Sexual Attention*

More easily recognized as harassing by most individuals, *unwanted sexual attention* is just that. Including both verbal and nonverbal behavior, it ranges from repeated, nonreciprocal requests for dates; intrusive letters and phone calls; touching, grabbing, and cornering; and gross sexual imposition or assault. Although frequently experienced as intimidating or coercive, it can be distinguished from the third category (*sexual coercion*) by its lack of job-related losses or benefits, either explicit or implied.

### *Sexual Coercion*

The classic instance of *quid pro quo* sexual harassment, behavior of this type refers to bribes and threats, whether explicit or subtle, that condition some job-related benefit on sexual cooperation. Although it is almost universally recognized and labeled as harassment, it is, perhaps not paradoxically, also the least common. Research studies routinely provide prevalence estimates in the 5 to 10% range, compared to 20 to 25% for unwanted sexual attention and 50 percent and higher for gender harassment. Studies of perceptions of harassing phenomena (e.g., Fitzgerald & Ormerod, 1991) reliably yield the highest harassment ratings (and smallest standard deviations) for instances of job benefits conditioned on sexual cooperation and negative job-related consequences for noncompliance.

In order to examine this proposed dimensionality, we utilize the SEQ, an instrument with known psychometric properties, established reliability and validity, and designed specifically to sample systematically from the entire universe of sexually harassing behaviors. Taking this instrument as the most comprehensive measure of harassment available, we then attempt to demonstrate its measurement equivalence through application of Jöreskog's (1971) procedure for *Simultaneous Factor Analysis in Several Populations* (SIFASP). Equivalence of measurement is obtained if the relationships between the

observed variables (i.e., items) and the hypothesized latent constructs (i.e., gender harassment, unwanted sexual attention, sexual coercion) are identical across samples from the different cultures and settings. Failure to obtain equivalence of measurement across different populations would imply that the construct of sexual harassment differs from group to group; SEQ scores could not be compared across groups as they would not mean the same thing. We restrict our analyses to women, not only because they constitute the overwhelming majority of victims of sexual harassment but also because recent research suggests that it may not be meaningful to apply present conceptual paradigms to men (Berdahl, Magley, & Waldo, 1994).

## METHOD

### *Participants*

*Sample 1.* Sample 1 was drawn from Fitzgerald et al.'s (1988) data set consisting of 1746 female university students from the United States. These women, both graduate and undergraduate students, were enrolled in one of two medium-sized universities located respectively in the Midwest or on the West Coast. Participants from University One consisted of 903 women ( $N = 349$  graduate women and 554 undergraduates) from over 70 academic disciplines, with an average age of 23 for the undergraduates and 32 for the graduate women. Ninety-three percent of the undergraduates were advanced students (i.e., juniors and seniors). The 843 University Two participants were similar in age and class status and consisted of 535 undergraduate women and 309 graduate women.

Data from undergraduate participants were collected in monitored sessions from a stratified random sample of intact classroom groups, with >99% participation rate; graduate women received and returned the research materials by mail, yielding a participation rate of slightly greater than 40% with no follow-up. Despite the differing procedures and particularly the significantly different participation rates, comparisons between the groups indicated no systematic differences, and they were thus combined.

*Sample 2.* Sample 2 consisted of 389 female university students from one of four universities located in two geographically separated areas of Brazil (Fitzgerald & Serra, 1990). The participants' average age was 23.89, and the great majority were in their second or third year of university. Data were collected in monitored sessions from intact classroom groups with >99% participation rate. Although the classrooms themselves were selected mainly on the basis of availability during the period of time when the researcher was in Brazil, examination of the participant demographics indicated no systematic bias in the pattern of participation.

*Sample 3.* Sample 3 consisted of the 307 female university employees described by Fitzgerald et al. (1988) and included academic, professional, clerical, and blue-collar women. The women—who constituted slightly less

than half (48%) of the university's female workforce—received and returned the research materials through campus mail. Of the 307 respondents, 61 were faculty, 61 administrators or academic professionals, and 170 were staff (e.g., clerical, maintenance, security, etc.). Fourteen reported their employment status as "other" and one did not identify her job category. The great majority (94%) were Caucasian, with a mean age of 38 and a median age of 39.

### *Instrumentation*

All participants completed one of two forms of the Sexual Experiences Questionnaire (SEQ), either the student (SEQ-E) or employee (SEQ-W) form (Fitzgerald et al., 1988). Items were developed based on literature searches, focus groups, and consultation with subject matter experts; all are written in behavioral terms, and the words sexual harassment do not appear until the end of the inventory. Although some aspects of harassment are represented more comprehensively than others, all are tapped by multiple items and—importantly—no aspect of the construct goes unassessed. Rather than grouping behaviors together in an a priori fashion (e.g., "letters, phone calls or materials of a sexual nature") the SEQ asks separately about a comprehensive sampling of relevant acts, with each item written in sufficient detail to ensure that respondents interpret it in a similar manner.

The two forms of the SEQ (student and employee) have 25 items in common; a typical item reads "Have you ever been in a situation where a supervisor or coworker (professor or instructor) made repeated attempts to engage you in a romantic sexual relationship despite your attempts to discourage him?" All participants responded on a 3-point scale, labeled *Never*, *Once*, or *More Than Once*; the latter two options were collapsed for scoring purposes due to low frequencies of *More Than Once* for some items. Each participant thus received a dichotomous score on each item.

Of the 25 behavioral items common to both the student and employee forms of the SEQ, 18 were included in this study. Seven of the original items assessing either very severe instances of harassment (e.g., rape) or those that were relatively rare in the samples examined (e.g., "flashing") were deleted due to low base rates of endorsement. Correlating items with low base rates and skewed distributions may yield a biased estimate of the underlying latent structure (Hulin, 1991) as well as increased standard errors of parameter estimates.

Of the remaining items, seven had been originally classified by the SEQ's authors as gender harassment, seven as unwanted sexual attention, and four as sexual coercion. Examination of Fitzgerald and Shullman's (1985) original factor solution as well as our own exploratory work suggested that two had been misclassified; the resulting revision yielded a matrix containing five examples of gender harassment, nine examples of unwanted sexual attention, and four of sexual coercion.

The wording for each of these 18 items was slightly different for the

SEQ-E (student) and the SEQ-W (employee) to fit the particular context being investigated (e.g., the target being a supervisor or a professor). Additionally, for the SEQ-E form, there was both an English version for the U.S. students, and a Portuguese version for the Brazilians.

Regarding the reliability and validity of the SEQ, Fitzgerald et al. (1988) reported internal consistency estimates of .92 for the original student sample and .86 for the employee sample; test-retest stability analyses computed on a small subsample of graduate students ( $n = 46$ ) yielded a coefficient of .86 over a 2-week period. Alpha in the Brazilian sample was .75. Evidence of content and criterion-related validity were reported by Fitzgerald et al. (1988). As noted above, exploratory factor analysis using Kaiser and Caffrey's (1965) alpha method (Fitzgerald & Shullman, 1985) yielded a three-factor solution (*gender harassment*, *quid pro quo*, and *unwanted sexual attention*) rather than the five factor structure suggested by Till (1980).

## RESULTS

The data analysis proceeded in two steps. First, we conducted separate confirmatory factor analyses for each group. This required that we begin by specifying the measurement model defining the relations between the measured variables (i.e., items) and the underlying theoretical constructs (i.e., latent variables such as gender harassment) that were defined previously. For each of the separate confirmatory factor analyses, the free elements of the factor pattern matrix were then estimated by the LISREL program and the fixed elements held constant at zero. Several estimation approaches are available as options in LISREL, and the choice of method depends on the characteristics of one's data. In our case, the data were dichotomous, posing a particularly vexing problem for factor analysis. We employed the PRELIS (Jöreskog & Sörbom, 1988) polychoric option to compute tetrachoric correlations and their asymptotic sampling variance-covariance matrix.

For each of the analyses, the tetrachoric correlation matrix and its asymptotic variance-covariance matrix were then input to LISREL, and the method of weighted least squares (WLS) used to estimate free elements in the factor loading matrix. This estimation method, which is due to Browne (1982, 1984), makes less restrictive assumptions about the distribution of the observed variables than many alternative estimation methods. LISREL's maximum likelihood estimation, for example, assumes that the observed variables are multivariate normal, which is certainly not the case for our dichotomous variables. Moreover, LISREL's maximum likelihood estimation method should not be applied to a matrix of tetrachoric or polychoric correlations because this estimation approach expects Pearson product-moment correlations (or the ordinary variance-covariance matrix) as input.

The degree to which the specified model fit each sample data set was next examined. Because different fit statistics are informative about alternative aspects of the data analysis, investigators typically examine a variety of mea-

TABLE 1  
 Goodness-of-Fit Measures for the Three-Factor Model of Sexual Harassment  
 Fit Separately in Three Samples and Simultaneously

	U.S. Student sample	U.S. Working sample	Brazilian sample	SIFASP analysis
$\chi^2$	350.07	105.23	166.68	424.30
df	132	132	132	426
$\chi^2/df$	2.70	.800	1.10	.996
GFI	.991	.988	.973	.982
Adjusted GFI	.988	.985	.965	—
RMSR	.070	.184	.223	.110

asures for converging evidence of the quality of a solution. One common measure is the overall Chi-Square statistic, which examines the hypothesis that the fitted model adequately describes the associations of the observed variables. The closer the obtained Chi-square value is to its degrees of freedom, the better the fit of the hypothesized model.

Because the Chi-square statistic is dependent on sample size, however (and hence will usually lead to the rejection of any model when the sample is large), it is important to examine other indicators of fit as well. Typically, this includes the Root Mean Square Residual (RMSR), a measure of the magnitude of the differences between the fitted correlation matrix and the sample correlation matrix. LISREL VII provides two additional fit statistics: the Goodness of Fit index (GFI) and the adjusted goodness of the fit (AGFI), which range between zero and one, although their exact statistical distributions are still unknown (Jöreskog & Sörbom, 1986).

The results of all confirmatory factor analyses are shown in Table 1. As can be seen, the various fit statistics depict a clear three factor solution for the SEQ in all three samples, when fit separately. For the two smaller samples, the Chi square statistic approaches the magnitude of its degrees of freedom; the ratio is reasonable for the U.S. student sample given its very large size ( $N = 1746$ ). For all the samples, the GFI and AGFI indices are greater than .95, and the RMSRs are reasonably low for tetrachoric correlations, indicating a very good fit of the model to the data. Although not depicted in Table 1, all factor loadings hypothesized to be nonzero were large and significant at the .001 level, indicating that the measured behaviors (i.e., items) were appropriately assigned to their respective theoretical constructs. In short, the separate analyses provide support for the model in the samples examined.

We next examined the three factor model fit simultaneously to all three data sets using simultaneous factor analysis for several populations (SIFASP). This procedure allows us to test the hypothesis that the factor loading matrices are invariant across the three groups. Accordingly, the factor loading matrix was constrained to be identical across the groups, and weighted least squares

TABLE 2  
SIFASP Estimates (Weighted Least Squares) of the Invariant Factor Loading Matrix

	Gender harassment	Unwanted sexual attention	Sexual coercion
told suggestive stories	1.000	—	—
made crude sexual remarks	1.061	—	—
displays sexist materials	0.788	—	—
treated differently because of gender	0.913	—	—
made sexist remarks	0.852	—	—
made sexual remarks about you	—	1.000	—
was staring, leering, ogling at you	—	0.984	—
attempts discuss sexual matters	—	0.941	—
unwelcome seductive behavior	—	1.088	—
unwelcome sexual attention	—	1.101	—
unwanted attempts to establish relationship	—	1.046	—
propositioned you	—	1.034	—
deliberately touched you	—	0.932	—
attempts to stroke/fondle you	—	0.917	—
subtly bribed with reward for sexual cooperation	—	—	0.915
actually rewarded prior to sexual cooperation	—	—	0.702
subtly threatened for lack of sexual cooperation	—	—	0.944
experienced negative consequences for refusing	—	—	1.000

was used for estimation. As the diagonal elements of the factor variance-covariance matrices were not fixed, a large loading in each column of the factor matrix was fixed at one. Because the Fitzgerald et al. (1988) data set was much larger than our two other samples, a 25% random sample was taken from the total data set of 1746 U.S. student women ( $n = 434$ ). By reducing it to a size comparable to Samples 2 and 3, Sample 1 should not exert an overwhelming influence on the results of our analyses.

The results of this analysis appear in Table 2.<sup>1</sup> As can be seen, each of the factor loadings hypothesized to be nonzero were large and significant ( $p < .001$ ); the SIFASP Chi-Square, which appears in Table 1, was 424.30 with 426 degrees of freedom, yielding a ratio of almost exactly one. Moreover, the other fit statistics (GFI, RMSR) suggest that the invariant three-factor solution fits the data satisfactorily.

In addition, we calculated the difference between the sum of the separate Chi-squares (371.76, with 396  $df$ ) and the SIFASP Chi-square (424.30, with 426

<sup>1</sup> SEQ items appear in abbreviated form.

*df*). When a more restrictive model is a nested submodel of a more general model, the difference in their Chi-square statistics can be interpreted in reference to a Chi-Square distribution, with degrees of freedom equal to the difference between the degrees of freedom of the models. Here, the SIFASP analysis can be viewed as a nested submodel of the separate analyses and therefore its Chi-square can be compared to the sum of the Chi-squares from the separate analyses (Jöreskog, 1971). The difference was 52.54 with 30 degrees of freedom, producing a Chi-Square to degrees of freedom ratio of about 1.75. Together, these statistics provide strong evidence that the SEQ provides equivalent measurement across the different settings and cultures examined, and that the construct can thus be considered equivalent across the groups.

## DISCUSSION

What are the theoretical implications and practical applications of this tripartite model of sexual harassment? At the level of theory, we propose that our results provide a comprehensive description of the nature of sexual harassment as a behavioral construct. Based on previous theory and research, we proposed (and confirmed) that sexual harassment is a multidimensional construct consisting of three distinct but related dimensions: *gender harassment*, *unwanted sexual attention*, and *sexual coercion*. Our model, operationalized via the *Sexual Experiences Questionnaire* (Fitzgerald & Shullman, 1985; Fitzgerald et al., 1988) was confirmed across target populations (employed women and students) and two cultures (the United States and Brazil). Although it is possible to argue that the university provides a unique work environment, thus weakening the generalizability of the present results, related research demonstrates that this is not the case. In a recent paper (Fitzgerald, Gelfand, & Drasgow, 1995, in press), we demonstrated that the structure holds in a heterogeneous non-university-based sample of working women as well. Further, it is striking evidence for the stability and generalizability of the model that the proposed behavioral relationships were confirmed even in a culture whose language (Portuguese) possessed no words for the concept of sexual harassment, at least at the time the study was conducted. It is our belief—and our data support it—that the model provides an answer to the question “What is sexual harassment?” at least from a behavioral perspective.

With respect to research, one obvious practical implication of our data is that the SEQ provides a psychometrically sound measurement strategy that can be used instead of the checklist methodology typically employed in most prevalence studies. A more general implication, which is perhaps less obvious, is that researchers should begin to conceptualize sexual harassment as a construct, with multivariate responses that are related. This represents a marked departure from previous research treating behaviors assessing sexual harassment as separate and isolated from one another. It is our hope that our results will influence researchers to direct their energies to studying patterns of behaviors that represent more general underlying constructs, a crucial shift that has the potential to expand

considerably our theoretical and empirical understanding of sexual harassment. Historically, many researchers have argued that science will advance only when we study manifest behaviors as indicators of underlying constructs (Hulin, 1991; Hulin & Rosseau, 1980; MacCorquodale & Meehl, 1948). A basic assumption underlying this belief is that individuals do not enact behaviors in isolation; rather, behaviors are patterned, and represent manifestations of more general constructs (Hulin, 1991). In the domain of sexual harassment, for example, it is likely that a person who makes sexist remarks, will also use sexist or suggestive materials, and/or tell offensive stories or jokes. Put in this light, it is considerably more theoretically powerful to conceptualize such behaviors as illustrations of *gender harassment* than to study (relatively arbitrarily chosen) exemplars in isolation.

Further, research that does examine behaviors in isolation is often plagued by low base rates and skewed distributions. Correlating infrequent behaviors with other behaviors inevitably yields biased estimates of the relations among underlying constructs (Hulin, 1991). Phrased in statistical jargon, there is a ceiling on the values of Pearson product-moment correlations computed on variables with low base rates. This problem is particularly acute in sexual harassment research, in which many critical behaviors (e.g., those constituting sexual coercion) have very low base rates (see Fitzgerald & Hesson-McInnis, 1989). Thus, continuing to examine "checklists" of isolated behaviors will not only limit our understanding of sexual harassment itself, it will also limit our knowledge of its antecedents, correlates, and consequences.

It is our belief that future attempts to study harassment must begin to treat it as a multidimensional construct, with a theoretically infinite number of indicators, rather than continuing to focus on an essentially arbitrary subset of those indicators. It is likely that the most appropriate indicators may sometimes vary across settings (e.g., the behavioral manifestations of gender harassment in factories, mines and shipyards appear to differ from those in offices, classrooms and clinics) whereas the latent constructs do not. Thus although the SEQ appears widely generalizable, its power lies not in its specific items, but rather in its systematic sampling of important, nonunique indicators of the three dimensions of an important psychological construct. We urge researchers to analyze their target settings with an eye towards identifying indicators with more or less power for assessing those dimensions in each particular context.

Finally, we suggest that this research has considerable potential for legal and social policy implications. For example, if it is theoretically important to disentangle the psychological construct of sexual harassment from the legal standard, it is equally critical to articulate the relationship between the two. For psychological research to be useful for informing legal theory, such connections are essential. Although a complete discussion of this issue is beyond the scope of this paper, we note that there is an isomorphism of the behavioral construct of sexual coercion to the legal concept of *quid pro quo*,

and observe that gender harassment and unwanted sexual attention constitute the two ways in which a hostile environment can be manifested. Not only do the three dimensions parallel the relevant legal concepts, but they also provide a relevant conceptual language via which research data can be related to those concepts. Similarly, they can be used to frame policy language at a variety of levels (e.g., organizational, institutional, legislative) that can then be linked to specific examples, recommendations, and procedures relevant to the particular setting.

### CONCLUSION

Like psychology itself (Boring, 1951), sexual harassment has a long past but a short history. Over the relatively short span of a decade and a half, our knowledge base has developed from a scattering of isolated studies to a relatively sophisticated, integrated body of knowledge. As with many other socially important topics, data have far outstripped theory, in both quantity and quality. Drawing on traditional concepts in the philosophy of science as well as the powerful methods of modern quantitative psychology, the present paper represents an attempt to address that lack and to provide a more sophisticated framework within which to conceptualize this important social problem.

### REFERENCES

- American Psychological Association (1993). *Amicus brief. Harris v. Forklift Systems, Inc.* Washington, DC: Author.
- American Psychological Association Taskforce on Male Violence Against Women. (1994). *No safe haven: Violence against women at home, at work, and in the community.* Washington, DC: APA.
- Boring, E. G. (1951). *The history of experimental psychology.* New York: Appleton-Century-Crofts.
- Bravo, E., & Cassidy, E. (1992). *The 9 to 5 guide to combating sexual harassment.* New York: Wiley & Sons, Inc.
- Berdahl, J., Magley, V., & Waldo, C. (1994). *The sexual harassment of men? A construct in search of definition.* Paper presented at the 14th annual meeting of the Society for Industrial and Organizational Psychology, Nashville, Tennessee, April 1994.
- Brogden, H. E., & Taylor, E. K. (1950). The theory and classification of criterion bias. *Educational and Psychological Measurement*, **10**, 159-186.
- Brown, M. W. (1984). Asymptotically distribution-free methods for the analysis of covariance structures. *British Journal of Mathematical and Statistical Psychology*, **37**, 62-83.
- Brown, M. W. (1982). Covariance structures. In D. M. Hawkins (Ed.), *Topics in applied multivariate analysis* (pp. 72-141). Cambridge: Cambridge Univ. Press.
- Bularzik, M. (1978). Sexual harassment at the workplace: Historical notes. *Radical America*, **12**, 25-43.
- Crites, J. O. (1969). *Vocational psychology.* New York: McGraw-Hill.
- Fiske, S., & Borgida, E. (in press). Sexual harassment. *Journal of Social Issues*, Whole No.
- Fitzgerald, L. F. (1993). Sexual harassment: Violence against women in the workplace. *American Psychologist*, **48**, 1070-1076.
- Fitzgerald, L. F., Gelfand, M. J., & Drasgow, F. (in press). Measuring sexual harassment: Theoretical and psychometric advances. *Basic and Applied Social Psychology*.
- Fitzgerald, L. F., Gelfand, M. J., & Drasgow, F. (1994). *The measurement of sexual harassment.*

- Paper presented at the Fourth Annual Meeting of the Society for Industrial-Organizational Psychology, Nashville, Tennessee, April 1994.
- Fitzgerald, L. F., & Ormerod, A. J. (1991). Perceptions of sexual harassment: The influence of gender and context. *Psychology of Women Quarterly*, *15*, 281–294.
- Fitzgerald, L. F., & Serra, A. (1990). Unpublished raw data. Department of Psychology, University of Illinois at Urbana-Champaign.
- Fitzgerald, L. F., & Hesson-McInnis, M. (1989). The dimensions of sexual harassment: A structural analysis. *Journal of Vocational Behavior*, *35*, 309–326.
- Fitzgerald, L. F., & Shullman, S. L. (1985). *The development and validation of an objective measure of sexual harassment*. Paper presented at the annual meeting of the American Psychological Association, Los Angeles, August.
- Franklin v. Gwinnett County School District, 112 S. Ct. 1028 (1992).
- Gruber, J. E. (1992). A typology of personal and environmental sexual harassment: Research and policy implications for the 1990's. *Sex Roles*, *26*, 447–464.
- Gutek, B. (1985). *Sex and the workplace*. San Francisco: Jossey-Bass.
- Harris v. Forklift Systems, Inc., 976 F.2d 733 (6th Cir. 1992), cert. granted, 61 U.S.L.W. 3600 (March 1, 1993).
- Hulin, C. L. (1991). Adaptation, persistence and commitment in organizations. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial/organizational psychology (2nd ed.)* (pp. 445–505). Palo Alto, CA: Consulting Psychologists Press, Inc.
- Hulin, C. L., & Rousseau, D. M. (1988). Analyzing infrequent events: Once you find them your troubles begin. In K. H. Roberts & L. Burstein (Eds.), *Issues in aggregation: New directions for methodology of social and behavioral sciences* (Vol. 6). San Francisco: Jossey-Bass.
- Jöreskog, K. G. (1971). Simultaneous factor analysis in several populations. *Psychometrika*, *36*, 409–426.
- Jöreskog, K. G., & Sörbom, D. (1988). *PRELIS: A preprocessor for Lisrel, 2nd ed.* Mooresville, IN: Scientific Software, Inc.
- Jöreskog, K. G., & Sörbom, D. (1986). *LISREL VI: Analysis of linear structural relationships by maximum likelihood (4th ed.)*. Morrisville, IN: Scientific Software, Inc.
- Kaiser, H. F., & Caffrey, J. (1965). Alpha factor analysis. *Psychometrika*, *30*, 1–14.
- Lott, B., Reilly, M. E., & Howard, D. (1982). Sexual assault and harassment: A campus community case study. *Signs*, *8*, 296–319.
- MacCorquodale, K., & Meehl, P. (1948). On a distinction between hypothetical constructs and intervening variables. *Psychological Review*, *55*, 95–107.
- MacKinnon, C. A. (1979). *Sexual harassment of working women*. New Haven: Yale Univ. Press.
- Maihoff, N., & Forrest, L. (1983). Sexual harassment in higher education: An assessment study. *Journal of the NAWDAC*, *46*, 3–8.
- Meritor Savings Bank v. Vinson, 477 U.S. 57 (1986).
- Pryor, J., & McKinney, K. Sexual harassment. In preparation.
- Robinson v. Jacksonville Shipyards, Inc. (1991). 59 L. WI 2470 (DC M Fla).
- Saunders, B. E. (1992). *Sexual harassment of women in the workplace: Results from the national women's study*. Presentation at the Eighth Annual North Carolina/South Carolina Labor Law Seminar. Asheville, NC: October 23.
- Till, F. J. (1980). *Sexual harassment: A report on the sexual harassment of students*. Washington, DC: National Advisory Council on Women's Educational Programs.
- Tinsley, H. E. A., & Stockdale, M. S. (1993). Special issue on sexual harassment in the workplace. *Journal of Vocational Behavior*, Whole. No. 3.
- United States Merit Systems Protection Board (1981). *Sexual harassment of federal workers: Is it a problem?* Washington, DC: United States Government Printing Office.
- United States Merit Systems Protection Board (1987). *Sexual harassment of federal workers: An update*. Washington, DC: United States Government Printing Office.

Received: March 8, 1994