Measuring Sexual Harassment: Theoretical and Psychometric Advances

Article *in* Basic and Applied Social Psychology · December 1995

DOI: 10.1207/s15324834basp1704_2

3 authors:

Louise Fitzgerald
University of Illinois, Urbana-Champaign
118 PUBLICATIONS **14,631** CITATIONS
SEE PROFILE

Michele Gelfand
University of Maryland, College Park
299 PUBLICATIONS **37,914** CITATIONS
SEE PROFILE

Fritz Drasgow
University of Illinois, Urbana-Champaign
220 PUBLICATIONS **20,625** CITATIONS
SEE PROFILE

BASIC AND APPLIED SOCIAL PSYCHOLOGY, 1995, *17*(4), 425-445

Measuring Sexual Harassment: Theoretical and Psychometric Advances

Louise F. Fitzgerald, Michele J. Gelfand, and Fritz Drasgow
Department of Psychology
University of Illinois at Urbana-Champaign

This article describes a program of research designed to yield a conceptually grounded, psychometrically sound instrument for assessing the incidence and prevalence of sexual harassment in the workplace. Following the specification of a theoretical framework that is consistent with both legal guidelines and psychological research, we review the development and evaluation of a three-dimensional model of sexual harassment (gender harassment, unwanted sexual attention, and sexual coercion). Based on this model, we describe the development of a revised version of the Sexual Experiences Questionnaire (SEQ; Fitzgerald et al., 1988). Following extensive pilot work, the instrument was field tested in a large regulated utility. Data from 448 employed women (professional, technical, clerical and blue collar workers) support the reliability of the scales, and confirmatory factor analysis in this new sample confirms the stability and generalizability of the theoretical model. Following a brief review of validity data recently reported in the literature, implications for further measurement improvements are discussed.

Whatever exists, exists in quantity and can be measured. —E. L. Thorndike

It seems safe to say that no concept has come as close to defying Thorndike's famous aphorism as sexual harassment. Although its existence has been acknowledged in the psychological literature for well over a decade, and interest in the topic has virtually exploded in the past few years, the legendary passion of psychologists for measuring anything that varies has been notable, in this instance, mainly by its absence. As a result, the state of the art (or, in this case, the science) has remained at a fairly rudimentary level, with the great majority of studies relying on simple

Requests for reprints should be sent to Louise F. Fitzgerald, Department of Psychology, University of Illinois at Urbana-Champaign, 603 East Daniel Street, Champaign, IL 61820.

checklists of unknown reliability and validity, without benefit of any theoretical framework; furthermore, such checklists generally vary from study to study. Even very basic questions remain unanswered; for example, what is the appropriate level of measurement? What are the dimensions that need to be assessed? What is the appropriate relation between the legal and psychological conceptions of harassment? These are only a few of the questions that remain not only unanswered but, for the most part, unasked.

In this article, we describe a program of research designed to yield a conceptually grounded, psychometrically sound instrument for assessing the incidence of sexual harassment in organizations. We begin with a brief review of earlier work on which our present efforts are based. We then lay out our conceptual framework, including its assumptions and the development and testing of our theoretical model of sexual harassment. Following a brief description of the development and piloting of our instrument, we present findings from a large-scale organizational study and close with a few observations about what we have learned so far, as well as the questions that remain unanswered.

BACKGROUND AND EARLY WORK

It is perhaps an overstatement to suggest that no theoretical work has been done in this area, as the first systematic attempt to map the conceptual domain of sexual harassment and develop a comprehensive classification system was initiated in 1980. Writing before any legal framework for harassment had been articulated, Till (1980) classified the self-described experiences of a national sample of college women into five behavioral categories: gender harassment, seductive behavior, sexual bribery, sexual coercion, and sexual imposition or assault. He recommended that these categories be thought of as levels of harassment because they appeared to form a rough continuum of severity, and he argued that they were exhaustive as they could encompass any particular example (act) of harassment. More recently, Gruber (1992) proposed a system composed of 11 specific types of harassment, organized into three higher order categories: verbal requests, verbal remarks, and nonverbal displays. Each of these categories contains a subset of the 11 types of harassment (e.g., verbal requests is composed of sexual bribery, sexual advances, relational advances, and subtle pressure/advances) which Gruber arranged in decreasing order of severity.

In contrast to such systems, which are organized at the level of classes or categories of behavior, the most widely used framework has been that of the U.S. Merit System Protection Board (USMSPB, 1981, 1987), which developed a simple enumeration of seven harassing behaviors that were classified into three levels of severity: less severe (unwelcome sexual remarks,

suggestive looks and gestures, and deliberate touching), moderately severe (pressure for dates, pressure for sexual favors, and unwelcome letters and telephone calls), and most severe (actual or attempted rape or sexual assault). Until recently, most data-collection efforts have used some version of this checklist, asking respondents to indicate whether they have experienced any of the behaviors described. Absent from this literature has been any attempt to link such data-collection efforts to a conceptual framework or to ascertain the reliability or validity of the measures.

Sexual Experiences Questionnaire (SEQ)

To address these issues, Fitzgerald et al. (1988) developed the Sexual Experiences Questionnaire (SEQ), a self-report inventory representing the first attempt to assess the prevalence of sexual harassment in a manner that met traditional psychometric standards. Based on a content validity strategy, the original SEQ was designed to tap each of Till's (1980) five dimensions via multiple items developed through literature searches, focus groups, and consultation with subject matter experts. All items employed a standard stem and were written in strictly behavioral terms; the actual term *sexual harassment* did not appear until the end of the inventory. This last was considered particularly important given the ambiguity surrounding the concept, as well as individual differences in willingness to apply the term among women who reported similar experiences; we thus sought to avoid confounding the experience of offensive sex-related behavior at work with the labeling of that experience as sexual harassment. Respondents were instructed to circle the response most closely describing their own experiences on a scale with three options—never, once, and more than once (the latter two being subsequently collapsed for purposes of scoring)—yielding frequencies or percentages of individuals who indicated that they have experienced the situations included in the inventory. Sample items appear in Table 1.

This original version of the SEQ yielded an internal consistency coefficient of .92 based on a sample of approximately 1,700 college students. Corrected split-half reliability coefficients for the five subscales ranged from .62 to .86 and averaged .75, whereas test-retest stability estimates computed on a considerably smaller sample ($N = 46$) yielded a coefficient of .86 over a 2-week interval. We attempted to build in content validity by writing multiple items to assess each of the five categories, which constituted the facets of our conceptual domain. In addition, we correlated each item with what was labeled the *criterion* item (i.e., the final item on the scale that asked "Have you ever been sexually harassed?") and examined the average item-criterion correlations for each subscale, reasoning that if Till (1980) were correct that the dimensions formed a rough continuum of

TABLE 1
Sample Items From the SEQ

Category	*Sample Items*
Gender harassment	Have you ever been in a situation where a supervisor or coworker habitually told suggestive stories or offensive jokes?
Seductive behavior	Have you ever been in a situation where a supervisor or coworker attempted to establish a romantic sexual relationship with you despite your attempts to discourage him?
Sexual bribery	Have you ever been in a situation where you felt you were being subtly bribed with some sort of reward (e.g., preferential treatment) to engage in sexual behavior with a coworker?
Sexual coercion	Have you ever been in a situation where you actually experienced negative consequences for refusing to engage in sexual activity with a coworker?
Sexual imposition	Have you ever been in a situation where a coworker made unwanted attempts to stroke or fondle you (e.g., stroking your leg or neck, touching your breast, etc.).

severity, the correlation should increase in a linear fashion across the five subscales. With one exception, the correlations conformed to expectation, ranging from .15 for gender harassment to .37 for sexual threats. The coefficient for sexual imposition/assault was lower than expected, most likely because several items showed very little variance. In general, the SEQ appeared to possess sufficient reliability and validity for research purposes.

Over the past several years, the SEQ has been used in a number of prevalence studies, both large and small, in a large variety of educational, occupational, and organizational settings; it has also been translated into other languages and used in cross-cultural settings. In this research, it has been found to be reliable, and a small but growing number of studies have documented its validity. Reviews and commentary have been positive (e.g., Arvey & Cavanaugh, 1995; Beere, 1990), and it is generally acknowledged as the most theoretically and psychometrically sophisticated instrument available. Research with the SEQ has also, however, demonstrated a number of shortcomings; for example, several critical items typically demonstrate markedly skewed base rates (specifically, the bribery and threat items that constitute *quid pro quo* harassment, as well as the more serious forms of sexual imposition), complicating the application of standard statistical techniques and suggesting the desirability of developing more sensitive wording. In addition, standard procedures are lacking for computing and assigning continuous scale scores to individuals because the original scoring methods utilized dichotomous procedures designed to produce only frequency distributions (i.e., the percentage of individuals in a given sample who had experienced a particular type of behavior). Finally, some experimental work (e.g., Fitzgerald & Hesson-McInnis, 1989) sug-

gested the desirability of distinguishing between type and severity of harassment, and at least one study (Fitzgerald & Shullman, 1985) cast doubt on the five-dimensional structure, suggesting that only three dimensions were necessary to explain variation in SEQ data. Such considerations led us to rethink our conceptual framework and revise the instrument accordingly, with hopes of addressing a number of these issues in the process.

SEXUAL HARASSMENT: A CONCEPTUAL FRAMEWORK

Assumptions

We began by distinguishing between sexual harassment as a legal concept and a psychological construct, noting that the two are not completely isomorphic. For example, women have been shown to confront a wide range of psychologically noxious workplace experiences, most of which constitute patterns or processes rather than events. Any particular exemplar may embody behavioral instances of varying types, frequency, intensity, and duration and may or may not meet current legal criteria for sexual harassment. This is particularly so given that legal criteria evolve and change based on regulatory definitions, case law, appellate decisions, and the like.

For example, until the Supreme Court's decision in *Meritor Savings Bank v. Vinson* (1986), it was unclear whether the concept of hostile environment (as opposed to *quid pro quo*) would survive as a viable cause of action under Title VII. Psychological constructs, on the other hand, are defined not by fiat but rather by the relations among the variables, that is, patterned behaviors and their nomological net. In this case, it is clear from even a cursory examination of item correlations that both types of behavior are dimensions of the same construct—an empirical relation that presumably existed before the enactment of civil rights legislation and will continue to do so independent of subsequent legislation and judicial interpretation.

This observation has a number of implications for instrument design. First, each legal determination is always contingent on a variety of factors not assessable by psychometric measures; it is thus important not to equate the two. However, it is desirable from a validity perspective (and imperative from an applied one) that the relation between them be articulated. For practical purposes, we interpret this to mean assessing the full range of experiences that have the potential to meet legal criteria, given an appropriate set of facts and circumstances, while at the same time acknowledging the distinction.

Second, the notion of a construct implies that measures should contain

multiple observable indicators of each aspect or dimension of the construct. This represents a marked departure from the current practice of treating each behavior as a dichotomous variable isolated from other behaviors, and assessed via a brief, global, and frequently ambiguous phrase (e.g., pressure for dates)—in essence, as a separate independent construct, tapped by a lone indicator (see, e.g., USMSPB, 1981, 1987; and many others).

Given these assumptions, our first task was to specify a theoretical model of harassment that is consistent with both the legal framework and psychological theory; we then subjected this model to a series of rigorous tests and, finally, developed an instrument based on this framework to assess sexual harassment in an efficient, reliable, and valid manner. It is this series of studies that is the subject of this article.

Theoretical Model

Over the last several years, the legal framework of sexual harassment has become increasingly clear. The original U.S. Equal Employment Opportunity Commission (EEOC, 1980) guidelines have been found by the courts to prohibit coerced sexual exchange (*quid pro quo*) and other types of generally offensive sex-related behavior that have the effect of creating a noxious workplace environment (hostile environment). In contrast to these broad categories, the psychological research on this topic has examined only specific acts, with little attempt to aggregate them at any higher level of generality. As noted elsewhere, "Virtually no theoretical attention has been given to defining the domain of this construct, nor to specifying its structure or dimensions . . ." (Gelfand, Fitzgerald, & Drasgow, 1995). Given these observations, our first task was to specify the theoretical dimensions (i.e., the domain) of the sexual harassment construct from which suitable, observable indicators (i.e., behaviors) could be systematically sampled. Our conceptual model appears in Figure 1.

Specifically, we proposed that the behavioral construct of sexual harassment is composed of three related, but conceptually distinct, dimensions: sexual coercion, unwanted sexual attention, and gender harassment. Gender harassment refers to a broad range of verbal and nonverbal behaviors not aimed at sexual cooperation but that convey insulting, hostile, and degrading attitudes about women. Some examples include sexual epithets, slurs, taunts, and gestures; the display or distribution of obscene or pornographic materials; gender-based hazing; and threatening, intimidating, or hostile acts. The EEOC (1993) recently supplemented its original Sex Discrimination Guidelines with a statement explicitly prohibiting such gender-based harassment in the workplace.[1] Although this

[1]As this article was completed, those particular guidelines (which also precluded harassment

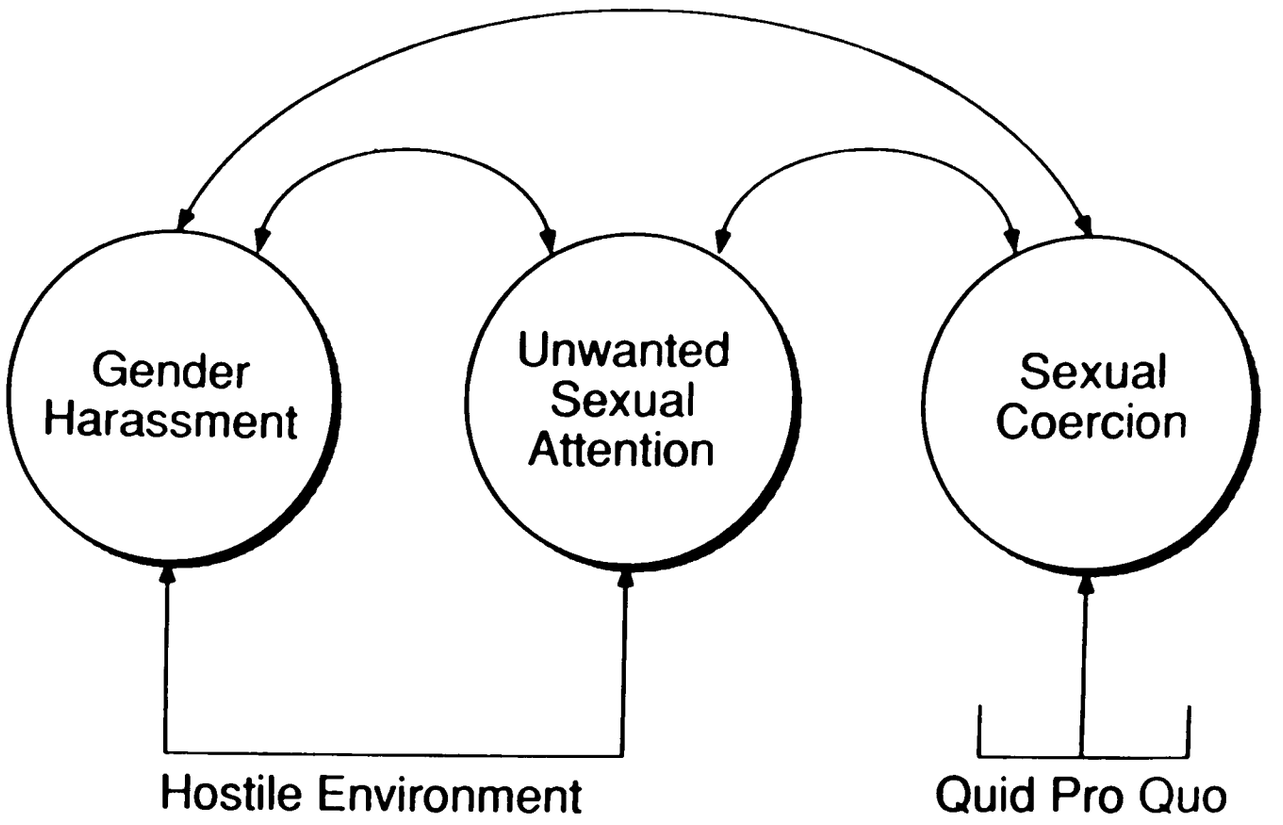

FIGURE 1 The relation between behavioral categories of sexual harassment and parallel legal concepts.

appears to be the most widespread form of harassing behavior (particularly in the nontraditional, blue-collar workplace), researchers have typically ignored it, choosing to focus on more explicitly sexual (as opposed to sex-related) situations.

Unwanted sexual attention is, of course, just that and includes a wide range of verbal and nonverbal behavior that is offensive, unwanted, and unreciprocated, whereas sexual coercion constitutes the canonical example of sexual harassment, that is, the extortion of sexual cooperation in return for job-related considerations. The consistency of these behavioral categories with their legal counterparts is apparent from Figure 1, which depicts the isomorphism of sexual coercion to the legal concept of quid pro quo and suggests that gender harassment and unwanted sexual attention constitute the two aspects of a hostile work environment. Like all theoretical models of human behavior, Figure 1 represents something of an oversimplification; for example, unwanted sexual attention from a supervisor with the power to hire, fire, and promote carries an implicitly coercive message. Similarly, stripped of its context, it is sometimes difficult to determine whether a sexualized conversation is a come on or a put down (the essential distinction

---

of other types) have been withdrawn; the EEOC is planning to reissue a separate set of guidelines specific to gender harassment in the near future (B. Henderson, personal communication, February 5, 1995).

between unwanted sexual attention and gender harassment). Despite such real-world ambiguities, we submit that the model captures the basic dimensions of sexual harassment as defined by law and described by psychological research.

In a previous paper (Gelfand et al., 1995), we described a test of this model in three samples representing different settings (employed women and college students) and cultures (United States and Brazil) via an application of Jöreskog's (1971) procedure for simultaneous factor analysis in several populations. A stringent test, this procedure necessitates a priori model specification, the assignment of each item to its theoretically appropriate factor, and the constraint that each item load in identical fashion across all samples. Taking the SEQ as the most comprehensive operational definition of harassment available, we examined a subset[2] of its original validation sample of university women (n = 434); a second sample (n = 389) enrolled in one of four Brazilian universities (these participants completed a Portuguese translation); and a third sample consisting of 307 female employees including academic, professional, clerical, and blue-collar workers. Results of the separate analyses—as well as the simultaneous three-sample procedure—revealed a clear and compelling three-factor solution closely corresponding to the hypothesized model and confirming the three-dimensional structure first suggested by Fitzgerald and Shullman's (1985) exploratory analysis. Not only did the items load as hypothesized in each sample, but when specifically constrained to the idential structure across samples, the combined data continued to fit the model closely. Based on these results, we argued that our proposed model of sexual harassment is (a) theoretically reasonable; (b) structurally identical across both workplace and educational settings; and (c) cross-culturally generalizable, at least across the cultures examined. A complete account of this research can be found in Gelfand et al. (1995).

SEQ: A REVISION

Following the establishment of this conceptual framework, our next step involved the development of a revised instrument. In addition to conceptual and psychometric desiderata, we sought to (a) develop an instrument short enough for practical use in organizations, (b) provide balanced item coverage for each dimension, and (c) address base rate and associated distributional problems through the development of more sensitive item and scaling procedures. The initial step in this process involved the generation of a revised item pool; in addition to the best SEQ items (e.g., those that are

[2]The original data set contained more than 1,700 individuals; we randomly selected 25% of these for analysis so that this sample would not receive disproportionate weight in the analysis.

unidimensional, most reliable, and least skewed), new items were developed, edited, and revised. As with the original form, all items were behaviorally based and sufficiently detailed to ensure that respondents interpret them in a similar manner.

These procedures resulted in a total of 54 behavioral items, balanced to the degree possible, across the three dimensions of the model. To address distributional problems, a 5-point Likert scale was substituted for the original 3-point scale; items were arranged so that milder forms of gender harassment and unwanted sexual attention appeared at the beginning of the instrument, followed by sexual coercion and the more intrusive physical forms of unwanted sexual attention. The resulting scale was then piloted as part of a larger survey examining a variety of organizational and institutional issues related to sexual harassment. The pilot sample included 150 female graduate students from the Colleges of Law, Commerce, Veterinary Medicine and the Institute for Labor and Industrial Relations at the University of Illinois; it is these data that were used for preliminary item and reliability analysis. Selection criteria included basic item characteristics such as means and standard deviations, item–subscale correlations, and theoretical concerns (e.g., balancing verbal with nonverbal behaviors and mild behaviors with more severe behaviors, and generally ensuring adequate construct coverage via the fewest number of items).

These procedures resulted in a revised scale containing 20 items; in addition to the three subscales and the criterion item, the final unwanted sexual attention item (which meets the legal definition of attempted rape) was retained as a separate category and not scored as part of the subscale. Alpha reliabilities in the pilot sample were .86 for gender harassment, .75 for unwanted sexual attention, and .87 for sexual coercion; alpha reliability for the total scale was .89. Following minor editing, this revised scale (referred to as SEQ–W) was then examined in an extensive validation study conducted in a large, regulated utility company.

METHOD

Sample and Procedure

The sample consisted of 1,188 employees of a west coast public utility company who participated as part of that organization's efforts to develop a harassment-free workplace. In addition to the SEQ–W,[3] participants also

[3]A parallel form, designed for students, is designated the SEQ–E; these instruments, along with a packet of information describing their characteristics and use, are available from Louise F. Fitzgerald (the first author).

completed a number of other scales assessing job attitudes, job stress, organizational tolerance for sexual harassment, organizational withdrawal, and other attitudinal and behavioral responses to organizational and job characteristics. An eight-person research team administered the questionnaires on-site during the course of 5 consecutive working days; data were collected in groups ranging from 1 to 78, under conditions of guaranteed confidentiality and anonymity. The sampling frame was the work group (as opposed to the individual employee); work groups at various sites across the state were randomly selected to participate.

These procedures resulted in a total of 1,156 completed or partially completed questionnaires (a 97% participation rate), of which 448 were from female employees, including professional, technical, clerical, and blue-collar workers. If a participant omitted one or two responses from the SEQ, truncated item means were inserted in place of the missing responses. No scale scores were computed for questionnaires that were missing more than two responses. In addition to the SEQ-W, respondents completed a variety of additional measures. Those results are reported elsewhere (Gelfand & Drasgow, 1994; Schneider & Swan, 1994; Zickar, 1994); however, the portions of that research that bear on the validity of the SEQ are reviewed briefly in the following section.

ANALYSIS

Basic Psychometric Analyses

Alpha reliability coefficients in this sample were comparable to those obtained in the development sample with the exception of sexual coercion, whose value reflects the low base rate of this experience in this sample. These correlations appear in Table 2, along with item-total, item-scale, and scale-total correlations.

Structural Analyses

Based on both theory and our previous work, we specified a measurement model containing five gender harassment items, seven unwanted sexual attention items, and five gender coercion items. Table 3 displays an abbreviated version of the items, arranged according to their theoretical assignment. Despite improved items and scaling procedures, the persistent base rate problems with sexual coercion necessitated that the items be dichotomized. Thus, we employed the PRELIS (Jöreskog & Sörbom, 1988) polychoric option to compute tetrachoric correlations and their asymptotic variance-covariance matrix; the method of weighted least squares (WLS) was then used to estimate the free elements in the factor-loading matrix.

TABLE 2
Item Statistics (SEQ-W)

	Item/Total r	*Item/Scale r*
Gender harassment (α = .82)		
1) . . . told suggestive stories	.62	.64
3) . . . made crude sexual remarks	.74	.73
4) . . . made offensive remarks	.70	.61
8) . . . displayed offensive materials	.51	.51
9) . . . sexist comments	.59	.60
Unwanted sexual attention (α = .85)		
2) . . . attempted to discuss sex	.73	.59
5) . . . unwanted sexual attention	.71	.72
6) . . . staring, leering at you	.66	.66
7) . . . attempts to establish a sexual relationship	.57	.63
10) . . . repeated requests for drinks, dinner, despite rejection	.50	.59
13) . . . touching in a way that made you feel uncomfortable	.49	.49
14) . . . attempts to stroke or fondle	.39	.48
Sexual coercion (α = .42)		
11) . . . subtly bribed you	.38	.12
12) . . . subtly threatened you	.14	.10
17) . . . made it necessary to cooperate to be well treated	.27	.18
18) . . . made you afraid of poor treatment if you didn't cooperate	.25	.45
19) . . . experienced consequences for refusing	.18	.19

Note. One sexual coercion item was dropped from the analysis because of extremely restricted variance.

Because the diagonal elements of the factor variance–covariance matrix were not fixed, a large loading in each column of the factor matrix was fixed at 1.0. LISREL estimates of this matrix appear in Table 4; the associated goodness-of-fit index (GFI) measures are displayed in Table 5, which reveals a clear and compelling three-factor solution. As can be seen, the chi-square statistic approaches the value of its degrees of freedom, the GFI and Adjusted Goodness of Fit Index are greater than .95, and the root mean square residual (RMSR) is reasonably low for tetrachoric correlations, thus providing additional support for the robustness of the theoretical model and the construct validity of the revised instrument.

DISCUSSION

The goal of our research was to develop a conceptually grounded, psychometrically sound instrument for assessing the incidence of sexual harass-

TABLE 3
Measurement Model for the SEQ-W

	Gender Harassment	*Unwanted Sexual Attention*	*Sexual Coercion*
1) . . . told suggestive stories	X	O	O
3) . . . made crude sexual remarks	X	O	O
4) . . . made offensive remarks	X	O	O
8) . . . displayed offensive materials	X	O	O
9) . . . sexist comments	X	O	O
2) . . . attempted to discuss sex	O	X	O
5) . . . unwanted sexual attention	O	X	O
6) . . . staring, leering at you	O	X	O
7) . . . attempts to establish a sexual relationship	O	X	O
10) . . . repeated requests for drinks, dinner, despite rejection	O	X	O
13) . . . touching in a way that made you feel uncomfortable	O	X	O
14) . . . attempts to stroke or fondle	O	X	O
11) . . . subtly bribed you	O	O	X
12) . . . subtly threatened you	O	O	X
17) . . . made it necessary to cooperate to be well treated	O	O	X
18) . . . made you afraid of poor treatment if you didn't cooperate	O	O	X
19) . . . experienced consequences for refusing	O	O	X

ment in organizations. The results suggest that we have made considerable progress in meeting that goal. In particular, the three-factor structural model appears remarkably stable. With respect to the SEQ itself, it appears to tap the construct in a reliable and efficient manner, and the correlations of its scales with other organizationally relevant variables support its predictive validity and practical value (Schneider & Swan, 1994; Zickar, 1994). In the following section, we present a more thorough analysis of our structural and psychometric results, as well as their implications and applications; briefly review the validity data on the SEQ; and conclude with a discussion of the problems and issues that remain to be resolved when assessing sexual harassment.

The Structure of Sexual Harassment

We begin our remarks at the point where all discussions of applied measurement must begin—with an analysis of the construct to be assessed—and submit that the neglect of this analysis is arguably the single greatest

TABLE 4
LISREL Estimates (WLS) of the Invariant Factor Loading Matrix of the SEQ-W

	Gender Harassment	*Unwanted Sexual Attention*	*Sexual Coercion*
GH1	0.815		
USA1		0.851	
GH2	0.875		
GH3	0.861		
USA2		0.904	
USA3		0.831	
USA4		0.958	
GH4	0.789		
GH5	0.743		
USA5		0.979	
SC1			0.948
SC2			0.661
USA6		0.719	
USA7		0.779	
SC4			0.828
SC5			0.906
SC6			0.749

Notes. Items are arranged in the matrix in the order that they appear in the inventory. GH = Gender harassment. USA = Unwanted sexual attention. SC = Sexual coercion.

weakness plaguing the body of sexual harassment prevalence literature as it exists today. Viewed from a certain perspective, such neglect is understandable; harassment research arose not as a matter of academic or theoretical interest, but in response to the need to solve a pressing social problem. Early prevalence studies tended to make a brief examination of the meager legal and regulatory guidance that existed at that time and then got on with the more pressing business of documenting the scope of the problem, often in the face of considerable skepticism that a problem existed at all.

This leap into data collection, however, lacking a careful specification of exactly what was being measured, has led to a number of unfortunate problems that, as the field matures, we can no longer afford to ignore. Of these problems, the failure to articulate the linkage between the legal and operational definitions of harassment and the failure to define the dimensions of its domain appear to be the most significant. We submit that the three-factor model described here represents a promising solution to these problems. Not only is the relation between harassment as a psychological construct and the parallel statutory dimensions clearly articulated, but those dimensions appear to be both necessary and sufficient to capture the variance in sexual harassment in both organizational and educational environments. Similarly, the dimensions appear to be cross-culturally

TABLE 5
GFI Measures of the Three-Factor Model

Measure	*Value*
Chi-square	133.67
Degrees of freedom	116.00
Ration of chi-square/degrees of freedom	1.15
GFI index	.983
RMSR	.207

generalizable (at least to the degree examined so far) and stable across time, as two of the original SEQ samples we (Gelfand et al., 1995) analyzed were collected nearly 10 years before those on which the present analysis is based.

Examining the model from a legal perspective, it is important to reiterate that it does not speak to the conditions under which the three categories of behavior (gender harassment, unwanted sexual attention, sexual coercion) become harassment in the statutory meaning of that term. Such determinations can only be made in light of the totality of the circumstances involved in any particular situation (EEOC, 1990), circumstances that tend to turn on more subjective issues of seriousness, welcomeness, and the like. As a practical matter, a hostile environment claim (gender harassment and unwanted sexual attention) typically requires a showing of a pattern of offensive behavior, whereas in *quid pro quo* harassment (i.e., sexual coercion) a single instance is generally sufficient to trigger Title VII and parallel state statutes. Even the requirement that behavior be repeated and patterned is not invariable, however, depending on other circumstances; for example, the EEOC noted "a single, unusually severe incident of harassment may be sufficient to constitute a Title VII violation . . . [in particular] a single unwelcome physical advance can seriously poison the victim's working environment" (pp. 16–17). Similarly, not all instances of gender harassment are sexual in nature; although still potentially actionable under Title VII as impermissible sex discrimination (EEOC, 1993), such instances may not always qualify as sexual harassment in the legal meaning of the phrase. Psychologically, however, such distinctions are less meaningful, as sexual and nonsexual gender harassment tend to co-occur, as do the three categories more generally, and to produce the same psychological effects (Schneider & Swan, 1994). The advantage of our model is that it articulates the relation between the legal and psychological constructs without in any sense equating them.

From a behavioral perspective, the three categories appear to be both parsimonious and comprehensive, that is, necessary and sufficient. Although it is possible to make finer logical distinctions (e.g., Till, 1980), it is unclear that there is an advantage in doing so. For example, Till (1980)

distinguished between bribery and threat, but it is clear that a statement such as, "If you sleep with me, I'll promote you," logically implies and subsumes its reverse. On the other hand, Gruber's (1992) categories conflate the important legal distinction between *quid pro quo* and hostile environment (e.g., his verbal request category contains instances of both sexual coercion and unwanted sexual attention). Thus, from both a logical and empirical perspective, we propose that our model provides a parsimonious yet comprehensive answer to the question, "What is sexual harassment?"

Harassment as a Construct: Locating the Psychological Process

In addition to its purely structural and operational definitions (i.e., multiple items that correlate with one another in a particular fashion, as well as with other variables in theoretically expected ways), the notion of a construct presupposes that the indicators covary by virtue of their relation to some psychological process. To take a simple example, we would expect items and facets of an arithmetic test (i.e., addition, subtraction, multiplication, and division) to "hang together," that is, covary, because answers to items assessing those facets are a function of the underlying cognitive process of arithmetic reasoning.

Of course, our analogy is not exact; the process of arithmetic reasoning is located within the individual who produces the responses on the correlated items of the arithmetic test, whereas sexual harassment resides in the extra-individual experiences of the victim. Thus, what analogous theoretical process can be proposed to account for observable interitem correlations in this case? We suggest two complementary, contemporaneous processes—one organizational and the other personological.[4] First, it is relatively clear by now that sexual harassment arises largely from organizational conditions that facilitate its existence, not from individual deviance (Culbertson, Rodgers, & Rosenfeld, 1994; Fitzgerald, Hulin, & Drasgow, 1995; Gelfand & Drasgow, 1994; Pryor, LaVite, & Stoller, 1993; Zickar, 1994); thus, it seems reasonable to argue that behaviors covary within and between our conceptual facets because the organizational norms that govern them are similar. For example, Hulin (1993) and Hulin, Fitzgerald, and Drasgow (in press) identified three aspects of organizational climate that have been shown to give rise to a greater incidence of sexual harassment: risk to victims for complaining, the likelihood that their complaints will not be taken seriously, and the probability that offenders will not be sanctioned in any meaningful way. Similarly, Pryor et al. (1993)

[4]We are indebted to John B. Pryor, who first suggested this conceptualization to us in a thoughtful review of an earlier version of this article.

showed in both field and laboratory studies that lenient management norms are related to higher levels of sexual harassment.

Thus, we argue that the organizational conditions and social norms that permit or encourage unwelcome sexual attention apply to each of the various aspects of such attention (e.g., leering, requests for dates, and attempts to stroke or fondle) as well as to other aspects of harassment (i.e., gender harassment and sexual coercion). It follows that when norms are weak and organizational tolerance is high, such behaviors will be not only prevalent but will co-occur.

Second, although we argue that organizational (i.e., environmental) conditions are paramount, we also emphasize that they are not everything. Individual differences in the propensity to harass do exist (Pryor, 1987; Pryor & Stoller, 1994) and can be facilitated or inhibited by environmental norms (Pryor et al., 1993). When men harass, it appears to be the case that their behavior is multidimensional; that is, men who tell suggestive stories are also likely to make crude sexual remarks and sexist, misogynist putdowns, obscene nonverbal gestures, body language, and the like. Of course, the norms and culture of the particular workplace (e.g., hospital, university, shop floor, and coal mine) will have a great deal to do with the form in which such harassment is expressed; this is yet another example of the ways that organizational environments shape individual behavior. We suggest that the combination of these organizational and individual factors are the most promising explanations for the pattern of behaviors assessed by the SEQ.

The Measurement of Sexual Harassment

Judged by the traditional standards of applied measurement, the SEQ appears to be a reliable and valid measure of sexual harassment. In addition, criterion-related validity data appear to be quite promising. For example, higher levels of organizational tolerance (i.e., perceived climate) are reflected in higher scores on the SEQ (Zickar, 1994); similarly, higher SEQ scores are associated with less satisfaction with both coworkers and supervisors (but not with work itself), as well as with organizational withdrawal and commitment (Schneider & Swan, 1994). Such data provide evidence of the high cost that harassment inflicts on organizations.

With respect to the price paid by individual victims, Schneider and Swan (1994) provided one of the first empirical and objectively assessed accounts of the relation between SEQ scores and psychological distress. Although this relation has been discussed in the literature from its inception, virtually all data has been provided by self-identified victims or clinical accounts (see Gutek & Koss, 1993; Koss et al., 1994, for reviews; and Dansky & Kilpatrick, in press, for an exception). Schneider and Swan, however,

related scores on the SEQ to objective psychometric measures of psychological distress (e.g., anxiety, depression, and symptoms of posttraumatic stress disorder) in an unselected sample, providing strong support for the contention that sexual harassment is a threat to women's psychological well-being. In addition, Gelfand and Drasgow's (1994) modeling analysis demonstrated that SEQ scores are strongly related to physical health, both directly and indirectly, via their impact on psychological well-being. Taken together, the data provide convincing evidence of the validity of the SEQ. Problems remain, however, and it is to these that we now turn.

Base rates. Probably the most obvious of these problems is the persistent issue of markedly skewed base rates on items assessing sexual coercion and some of the more blatant forms of unwanted sexual attention. Despite our attempts to develop more sensitive wording and scaling procedures, base rates on these items never rose beyond 5% and, in some cases, were considerably lower. We argue that this simply reflects reality; all prevalence surveys to date have indicated that *quid pro quo* situations are relatively rare, despite the popular stereotype that equates sexual harassment with sex for promotion, retention, and the like. Although it is probably the rare woman who has never had to "be nice" to the boss, smile, or acquiesce gracefully to small indignities, it is clear by now that sexual coercion in the workplace is a relatively infrequent event compared to other forms of sexual harassment. To the degree that this is true, our base rates simply reflect the validity of the items assessing this aspect of the construct. Although complicating the application of various statistical techniques (and thus the lives of researchers interested in modeling the antecedents, process, and outcomes of this problem), this situation presents no practical problem for conducting incidence studies, the application in which organizations are most likely to be interested.

Scoring. The SEQ can be used to assess either individuals or organizations, and the issues having to do with scoring are different for each. With respect to organizational assessment, the process is relatively straightforward, involving the simple computation of frequency distributions at the desired level of analysis (e.g., separate frequency distributions for each of the three types, combined categories, and the like). Although the matter of what constitutes a "case" does arise and must be addressed (e.g., Should an individual who endorses a single gender harassment item and indicates via the scale that this was an isolated happening be classified as sexually harassed?), this is a practical issue that can be decided on the basis of the organization's goals in conducting the assessment. And, the data tell us that such isolated happenings appear to be relatively rare.

Thornier issues arise, however, when determining how to assign scores to

individuals, the situation in which researchers are more likely to be interested. Although it is obviously possible to treat harassment as a dichotomous variable, this seems to be a serious oversimplification. At the very least, it would seem necessary to allow for the effects of both type and frequency, and it is in this manner that we assigned scores for our analysis. Each participant was assigned a score for each of the three subscales reflecting the simple sum of within-scale responses, responses that are weighted by the frequency with which they were experienced. Given that harassment is a psychological stressor, however, it also seems important to capture the respondent's appraisal of the situation as benign, irrelevant, or threatening (Lazarus & Folkman, 1984; see also Fitzgerald, Swan, & Fisher, 1995, for an in-depth discussion of harassment as an instance of psychological stress). Such appraisals invoke the issue of how to score or scale the seriousness of a behavior or situation, a complex issue that we address in the following section.

Seriousness. The entire issue of sexual harassment seriousness, particularly the relative importance of objective versus subjective criteria, is a complex one that has yet to receive the attention it deserves (cf. Gruber, 1995). Although some behaviors are, in an objective sense, clearly more serious than others (e.g., sexual assault is obviously a greater stressor than offensive remarks), the great majority of stimulus-based severity distinctions appear to be made on a more or less arbitrary basis. For example, it is unclear on what basis the USMSPB (1981, 1987) concluded that being deliberately touched was less serious than receiving unwanted letters and telephone calls; similarly, although Gruber (1992) articulated a logical rationale for his seriousness ratings, there is, at the least, room for reasonable people to disagree. For example, it is unclear whether, as he asserted, personal remarks made to a woman are always more offensive than remarks made to others about her. If nothing else, it would appear to depend on the remarks being made.

It is our belief that much of the fuzziness concerning this concept has to do with a failure to specify exactly what is meant by seriousness, under what circumstances, and for what purpose. From a psychological perspective, it is clearly not possible to achieve a purely stimulus-based classification, if only because individual differences in target (i.e., victim) history, perceptions, and evaluations inevitably account for so much of the variance in reactions. However, it is also clear that certain stimulus characteristics (e.g., frequency) clearly contribute to the seriousness of outcomes (Schneider & Swan, 1994). Similarly, although the law has determined that *quid pro quo* harassment is generally more serious than a hostile environment, this is not necessarily true from the perspective of victim outcomes and is not even always true in the courtroom. Space precludes a more extensive discussion

of this issue here; rather, we content ourselves with noting that, for practical purposes, the relative power and sensitivity of weighted versus unweighted scores is an empirical question that can only be solved via appropriate data and analysis. In this case, we included only frequency as an empirical indicator of seriousness; however, theory and some data suggest that other, more subjective, evaluations may, in the end, prove superior. For example, Swan (1994) demonstrated that simple ratings of perceived offensiveness can double the amount of variance accounted for in outcomes. This appears to be an important direction for future research.

Relationship among SEQ subscales. The final issue we raise has to do with the substantial intercorrelation among the factors of the SEQ (average $r = .74$), suggesting less differentiation than is desirable. To some degree, this reflects a fact of life; in the real world of work, experiences of sexual harassment are often multidimensional phenomena, with various aspects of the problem predominating in the experience of any particular victim. Sexual coercion, by definition, implies unwanted sexual attention (although the reverse is not necessarily or even usually so), and gender harassment is often expressed in sexual terms. It is also the case that the SEQ collapses, so to speak, across different behavioral sequences and perpetrators, thus artificially inflating the relation among the three dimensions. Although these factors account to some degree for the intercorrelation, it is also possible that the SEQ–W in its present form provides somewhat limited coverage of the domain of gender harassment; in particular, the blatantly hostile and often physically dangerous experiences reported by women in the traditionally most male-dominated occupations (e.g., coal miners; railroad workers; and some aspects of the police, fire, and military services) are not well represented. The inclusion of such experiences may serve to lower the correlation among the factors, although, given that misogyny is so often expressed in sexual terms, this is an empirical question.

CONCLUSION

This brings us full circle to the original, central argument of this article—that sexual harassment is most appropriately conceptualized as a psychological construct. Although its dimensions are now known, its population of observable indicators (i.e., items) is potentially infinite. As with other constructs, some indicators may be more salient (i.e., valid) depending on the population being assessed. This suggests the desirability of using our theoretical framework to guide the development of a comprehensive item bank containing large numbers of indicators from which organizationally

(and occupationally) specific versions of the SEQ can be tailored and linked to the validity data available on the base version of the measure. This process, common in intellective measurement but so far untried in the present arena, promises to keep us occupied for some time to come.

ACKNOWLEDGMENTS

An earlier version of this article was presented at the Ninth Annual Conference of the Society for Industrial/Organizational Psychology, April, 1994, Nashville, TN.

We thank Chuck Hulin and the other members of the SexHarass Research Group for their collaboration and contribution to the research project, on which this article is based, and Larry Hubert and John B. Pryor for their assistance in clarifying the concept of harassment as a psychological construct.

REFERENCES

Arvey, R. D., & Cavanaugh, M. A. (1995). Using surveys to assess the prevalence of sexual harassment: Some methodological problems. *Journal of Social Issues.*

Beere, C. A. (1990). *Gender roles: A handbook of tests and measures* (Vol. 1). Westport, CT: Greenwood.

Culbertson, A. L., Rodgers, W., & Rosenfeld, P. (1994, August). *Organizational change: The case of sexual harassment.* Paper presented at the annual meeting of the American Psychological Association, Los Angeles.

Dansky, B., & Kilpatrick, D. (in press). Effects of sexual harassment. In W. O'Donohue (Ed.), *Sexual harassment: Theory, research, and treatment.* Needham Heights, MA: Allyn & Bacon.

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., Hulin, C. L., & Drasgow, F. (1994). The antecedents and consequences of sexual harassment in organizations: An integrated model. In G. P. Keita & J. J. Hurrell, Jr. (Eds.), *Job stress in a changing workforce: Investigating gender, diversity, and family issues* (pp. 55–74). Washington, DC: American Psychological Association.

Fitzgerald, L. F., & Shullman, S. L. (1985, August). *The development and validation of an objectively scored measure of sexual harassment.* Paper presented at the annual meeting of the American Psychological Association, Los Angeles.

Fitzgerald, L. F., Shullman, S. L., Bailey, N., Richards, M., Swecker, J., Gold, Y., Ormerod, A. J., & Weitzman, L. (1988). The dimensions and extent of sexual harassment in higher education and the workplace. *Journal of Vocational Behavior, 32,* 152–175.

Fitzgerald, L. F., Swan, S., & Fischer, K. (1995). Why didn't she just report him? The psychological and legal implications of women's responses to sexual harassment. *Journal of Social Issues, 51,* 117–138.

Gelfand, M., & Drasgow, F. (1994, April). *Antecedents and consequences of sexual harassment in organizations.* Paper presented to the Society for Industrial-Organizational

Psychology, Nashville, TN.

Gelfand, M. J., Fitzgerald, L. F., & Drasgow, F. (in press). The structure of sexual harassment: A confirmatory analysis across cultures and settings. *Journal of Vocational Behavior.*

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., & Koss, M. P. (1993). Changed women and changed organizations: Consequences of and coping with sexual harassment. *Journal of Vocational Behavior, 42,* 28-48.

Hulin, C. L. (1993, May). *A framework for the study of sexual harassment in organizations: Climate, stressors, and patterned responses.* Paper presented at the annual meetings of the Society for Industrial/Organizational Psychology, San Francisco.

Hulin, C. L., Fitzgerald, L. F., & Drasgow, F. (in press). Organizational influences on sexual harassment. In J. Stockdale & B. A. Gutek (Eds.), *Women and work.* Newbury Press, CA: Sage.

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.

Koss, M. P., Goodman, L., Browne, A., Fitzgerald, L. F., Keita, G. P., & Russo, N. F. (1994). *No safe haven: Male violence against women at home, at work, and in the community.* Washington, DC: American Psychological Association.

Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal, and coping.* New York: Springer.

Pryor, J. B. (1987). Sexual harassment proclivities in men. *Sex Roles, 17,* 269-290.

Pryor, J. B., LaVite, C. M., & Stoller, L. (1993). A social-psychological analysis of sexual harassment: The person/situation interaction. *Journal of Vocational Behavior, 42,* 68-83.

Pryor, J. B., & Stoller, L. M. (1994). Social cognition processes in men who are high in the likelihood to sexually harass.

Schneider, K., & Swan, S. (1994, April). *Job-related, psychological and health-related outcomes of sexual harassment.* Paper presented at the annual conference of the Society for Industrial/Organizational Psychology, Nashville, TN.

Swan, S. (1994). [Simple ratings of perceived offensiveness can double the amount of variance accounted for in outcomes]. Unpublished raw data.

Till, F. (1980). *Sexual harassment: A report on the sexual harassment of students.* Washington, DC: National Advisory Council on Women's Educational Programs.

U.S. Equal Employment Opportunity Commission. (1980). Guidelines on discrimination because of sex. *Federal Register, 45,* 74676-74677.

U.S. Equal Employment Opportunity Commission. (1993). Proposed guidelines on harassment based on race, color, religion, gender, national origin, age or disability. *Federal Register, 58,* 5126-51269.

U.S. Merit Systems Protection Board. (1981). *Sexual harassment in the federal government: Is it a problem?* Washington, DC: U.S. Government Printing Office.

U.S. Merit Systems Protection Board. (1987). *Sexual harassment in the federal government: An update.* Washington, DC: U.S. Government Printing Office.

Zickar, M. (1994, April). *Antecedents of sexual harassment.* Paper presented at the annual conference of the Society for Industrial/Organizational Psychology, Nashville, TN.