Conceptualizing the Multidimensional Nature of Self-Efficacy: Assessment of Situational Context and Level of Behavioral Challenge to Maintain Safer Sex

Debra A. Murphy and Judith A. Stein
University of California, Los Angeles

William Schlenger
Research Triangle Institute and Duke University Medical Center

Edward Maibach
Porter-Novelli

National Institute of Mental Health Multisite HIV
Prevention Trial Group
National Institutes of Health

A. Bandura (1991) argued that self-efficacy measurement should be specific both to the situation in which the behavior occurs and level of challenge in that situation. Measures consistent with the 2 dimensions were developed with graded challenge levels and differing gender-appropriate situations. Participants were 1,496 controls in the National Institute of Mental Health Multisite HIV Prevention Trial recruited from STD clinics and health service centers (925 women and 571 men). The authors tested 4 separate-sex confirmatory factor analysis models as follows: (a) Condom negotiation efficacy as a unitary construct across situations and gradation of difficulty; (b) situation as preeminent, which transfers across skills whatever the gradation of difficulty; (c) skill as predominant, irrespective of situation; and (d) a multidimensional design that simultaneously accounts for both situation and graded difficulty. Consistent with Bandura's theory, the multidimensional model provided the best fit for both samples.

Key words: self-efficacy, HIV/AIDS, assessment, structural equation modeling (SEM)

To reduce HIV risk behavior successfully, people must exercise control over their behavior and their social environment. This requires certain risk reduction skills as well as self-regulative skills—but there is a difference between possessing skills and being able to use them under difficult circumstances (Bandura, 1994). Reducing HIV risk behavior requires a strong sense of efficacy—that is, the sense that one can exercise personal control. Even if people are knowledgeable about HIV/AIDS transmission

and have the requisite skills, when they lack self-efficacy for HIV risk reduction they do not manage situations effectively.

Perceived self-efficacy has been defined as a judgment of one's ability to organize and execute given types of performances (Bandura, 1997). Bandura (1990) proposed a self-efficacy model of safer sex behavior that examines HIV risk reduction within the context of social learning theory. In the model, for sexual risk reduction, knowledge and skills to effect safer sex behavior are necessary but not sufficient for a successful outcome. Unless people believe that they can produce a desired effect by their own actions, they have little motivation to act or persevere in the face of obstacles. Efficacy beliefs influence the course of action individuals choose, how much effort they put into the course of action, how long they persevere in the face of barriers, and the level of accomplishment they realize (Bandura, 1999).

The role of perceived self-efficacy in HIV risk reduction has been investigated in a number of studies. For example, perceived self-efficacy to negotiate condom use has been found to predict safer sex practices among adolescents (Basen-Engquist & Parcel, 1992; Jemmott, Jemmott, & Fong, 1992) as well as among adults (Brafford & Beck, 1991; O'Leary, Goodhart, Jemmott, & Boccher-Lattimore, 1992). Wulfert and Wan (1993) used a structural model with self-efficacy as the central mediator of safer sex among college students to test and confirm Bandura's model. Finally, in a longitudinal study of gay men's sexual behavior (McKusick, Wiley, Coates, & Morin, 1986), self-efficacy to control one's sexual behavior emerged as the best predictor of sexual risk-taking.

However, self-efficacy is not a global, cross-situational construct. Self-efficacy judgments are specific to the behaviors that must be enacted in the situations in which they occur (Bandura, 1977, 1986; Hofstetter, Sallis, & Hovell, 1990; Murphy, Multhauf, & Kalichman, 1995). Accordingly, Bandura (1991) has argued that self-efficacy measurement should be specific to both the situation in which the behavior will occur and the level of challenge (or impediments to successful performance) in that situation. All of the studies just cited either used a global measure of self-efficacy (e.g., Wulfert & Wan, 1993) or focused on behaviors that were not within any situational context. It is an indication of just how robust the construct is that self-efficacy measures that are less than optimal are still predictive of safe-sex outcomes.

However, as conceptualized by Bandura (1997), self-efficacy theory would reject depleted or overly simplistic conceptions of perceived capability. For example, perceived efficacy to protect against sexually transmitted infections should not be reduced to a single act of putting on a condom (Bandura, 1997). Self-efficacy in which safe sex is the desired outcome involves several challenges. First, in the context of the performance situation, for example, perceived ability to negotiate condom use with a sexual partner may vary depending on the person's level of sexual arousal, whether this is a new partner or a former partner, whether the person is intoxicated or influenced by drugs, or whether the partner is perceived as a low- or high-risk partner. Each of those situations or scenarios differs, and thus perceived self-efficacy to perform in each of them may vary. Second, the level—or gradation of difficulty—of behaviors to be performed to have a safe sex outcome will vary. For example, the behavior required may range from initiating discussion of condom use with the sexual partner to resisting pressure for unsafe sex to resisting one's own desire to

not use a condom to leaving the situation if the partner refuses to agree to condom use.

In summary, Bandura's (1997) conceptualization identified two factors that strongly influence self-efficacy for safer sex: (a) The situation or scenario in which the person finds himself or herself (e.g., is intoxicated; is with someone with whom he or she has a history of unsafe sex) and (b) the graded difficulty level of the behavior the person must perform to have safe sex (e.g., only needs to bring up the topic of safe sex or condoms; must convince a partner reluctant to use a condom into using one). In the present study, self-efficacy measures consistent with Bandura's conceptualization were developed separately for men and for women, with the intent of determining through structural equation modeling (SEM) the extent to which situation versus level of difficulty drives self-efficacy. The SEM approach enabled us to examine the relative fit of alternative conceptualizations to the data. We hypothesized that perceived self-efficacy for HIV risk reduction would not fit a model in which self-efficacy was viewed as a unitary construct across different situations or across different gradations of difficulty of skill required. Rather, according to social learning theory the best model fit would be one in which both situation and graded difficulty level needed to be accounted for in the model and would impact participants' perceived self-efficacy for maintaining safer sex practices.

Method

The National Institute of Mental Health (NIMH) Multisite HIV Prevention Trial (Multisite Trial) methodology has been described in detail in several papers (NIMH HIV Prevention Trial Group, 1997a, 1997b, 1997c, 1998). However, a brief description of the participants, procedures, and assessment is provided here.

Participants

The primary purpose of the Multisite Trial was to develop and test the efficacy of an HIV preventive intervention with multiple and high-risk populations: men recruited in STD clinics, women recruited in STD clinics, and women recruited in community-based health service organizations. The trial was implemented in community-based clinics located in five metropolitan areas at seven different funded sites: the Bronx and Harlem, New York City; Manhattan and Brooklyn, New York City and northern New Jersey; Baltimore, Maryland; Atlanta, Georgia; Milwaukee, Wisconsin; Los Angeles, California; and Orange and San Bernardino Counties, California. Participants were recruited from January 1994 to February 1996.

Procedures

Participants were screened for HIV risk (Murphy, Rotheram-Borus, Srinivasan, & Hunt, 1997) in the waiting rooms of all recruitment clinics. Eligibility criteria included age (20 years or older for STD participants, 18 years or older for women attending health service organizations); engaging in unprotected vaginal or anal sex within the past 90 days; and having either sex with a new sexual partner, more than one sexual partner, a sexual partner with an STD, a sexual partner they knew also had other sexual partners, or sex with an injection drug user or a person infected with HIV within the past 90 days (NIMH Multisite HIV Prevention Trial Group, 1997c). Eligible participants were asked to return after screening for a 35-min baseline interview, were randomized to either the intervention or control condition, and were interviewed at 3, 6, and 12 months after the end of the seven-session intervention.

Assessment

Participants were interviewed in person using computer-assisted personal interviewing methods. The 3-month interviews assessed most domains covered during the baseline interview as well as hypothesized variables predicted to be mechanisms through which behavior change is mediated, including self-efficacy. Therefore, the data for this study are from the Multisite Trial's 3-month interview. Because the primary focus of this study is the structure of the self-efficacy instrument itself, only the untreated control group was used in this validation study.

The self-efficacy assessment, developed by the collaborators for the Multisite Trial, used a self-efficacy item format previously developed by Murphy et al. (1995). First, individual qualitative interviews were conducted with 50 men and women from the target populations at two of the Multisite Trial sites to identify the situational factors associated with the last time they engaged in unsafe sex, using formative interview questions that have been presented in detail elsewhere (Maibach & Murphy, 1995; Murphy et al., 1995). Second, themes of the relevant situations and internal and external barriers or challenges were identified from the formative interview information. These situations differed by gender: Four primary themes were identified for men and four themes were identified for women. The themes for men were (a) wanting to have sex and having a former partner available; (b) wanting to have sex after drinking alcohol; (c) being with a new partner who appears "clean" or low risk; and (d) wanting to switch from having unsafe sex to safe sex with a main partner. The themes for women were (a) having their sense of worth tied to physically being with a man; (b) wanting to use sex to mend a relationship with a main partner; (c) wanting to have sex after drinking alcohol with a former partner; and (d) being with someone who is a good prospect for a long-term relationship. Third, a specific scene was developed that instantiated each theme (four scenes for men and four scenes for women). The scenes were brief descriptions of sexual situations, and Multisite Trial participants were asked to imagine themselves in those situations. Finally, four specific self-efficacy questions were developed to provide a graded difficulty level approach consistent with Bandura's conceptual framework that level of perceived self-efficacy should reflect perceived capability as measured against task demands at various levels of challenge to successful performance (Bandura, 1986, 1997; Maibach & Murphy, 1995). The four self-efficacy questions, which were asked about each of the scenes, assessed that one could (a) bring up the issue of condoms or safe sex in conversation; (b) convince one's partner to be safe even if the partner says he or she hates condoms; (c) be safe even if both parties prefer sex with the feel of bare skin; and (d) refuse to have sex or leave the situation if the partner will not be safe. Participants utilized an 11-point scale (0–10) to indicate how confident they were about engaging in each of the four specific behaviors for each of the four scenes. Thus, 16 items (4 scenes × 4 efficacy items) reflected self-efficacy each for men and women, although the scenes varied by gender. A sample of one of the scenes for men and one of the scenes for women and the self-efficacy questions are provided in Table 1.

Analysis

Because the self-efficacy assessment incorporated both situation-specific and graded difficulty components, we were able to address a series of research questions concerning the dimensionality of condom negotiation skills using confirmatory factor analysis (CFA). The EQS structural equations program was used in the analysis (Bentler, 1995). To test the hypotheses, the following CFA models were contrasted. First, it could be the case that condom negotiation efficacy is a unitary construct across gradation of difficulty and situations that would be indicated adequately with one overarching latent variable (Model 1). If Model 1 were most viable, the 16 separate efficacy items would load significantly on one latent variable or factor representing efficacy, and further, the fit indexes would indicate that one latent variable is sufficient to capture the relationships among the items regardless of difficulty or situation. In other words, one efficacy latent

Table 1
Examples of Situations and Scenes

Gender	Example
Women	You've been having some problems with your main partner. Because of the problems, he hasn't been spending as much time with you and you're worried. He comes around and wants to make up with you. You don't want him to be mad anymore, and you know if you have sex with him, things won't be so tense. You want to have sex to mend things between the two of you.
Men	You're with a woman you've known casually for a while. You feel like she's been giving you signals that she's maybe interested in you, so you get her alone with you. After laughing and talking for awhile, you start kissing and fondling her. She seems like a nice person, and appears pretty clean.

variable would capture all of the shared variance among the variables to such an extent that partialing out the unitary latent variable would reduce correlations among the observed variables essentially to zero (Bentler & Stein, 1992).

Second, it is possible that the situation is preeminent and determines one's degree of efficacy, which transfers across skills whatever the gradation of difficulty. In this case, four latent variables representing the four situations or scenarios would be adequate to explain the relationships among the 16 measured items and again would account for all or nearly all of the variation in the individual items (Model 2). A hypothetical model depicting Model 2 can be visualized by covering the column of latent variables on the right side of Figure 1.

Third, it is also possible that condom negotiation skills are consistent and predominate irrespective of the situation or scenario. If this were the case, then four latent variables representing the four levels of skills based on their graded difficulty would be sufficient to explain relationships among the 16 items (Model 3). This model can be visualized by covering the left side of Figure 1.

Last, it may be the case that the best fit would be provided by a model based on a multidimensional design that simultaneously accounts for both situation and graded difficulty by systematically partialing out the errors in variables (Model 4). Figure 1 depicts the multidimensional model, Model 4, in which the 16 measured variables, by loading on a situation factor and a difficulty factor, can be represented by eight latent variables. The model allows both the situational factors and the difficulty factors to correlate among themselves, but the factors do not correlate across dimensions.[1]

Goodness of fit of the series of models described above was assessed using the adjusted robust comparative fit index (RCFI) based on the Satorra–Bentler chi-square statistic (Bentler & Dudgeon, 1996) and chi-square/degrees of freedom ratios. The Satorra–Bentler chi-square is preferable when the data are multivariately kurtose (Bentler & Dudgeon, 1996), and the multivariate kurtosis estimate was high in this study (normalized estimate = 109.47 [women] and 91.56 [men]). The RCFI ranges from 0–1 and compares the improvement of fit of the hypothesized model to a model of independence among the measured variables while adjusting for sample size. Values greater than .95 are desirable (Hu & Bentler, 1999). We also report the root mean square errors of approximation (RMSEAs; Browne & Cudeck, 1993). This statistic is helpful in evaluating fit because it indicates the size of the residuals. Values less than .06 indicate a relatively good fit between the hypothesized model and the observed data (Hu & Bentler, 1999).

---

[1] A detailed table of the measured variables correlations is available on request from Debra A. Murphy.
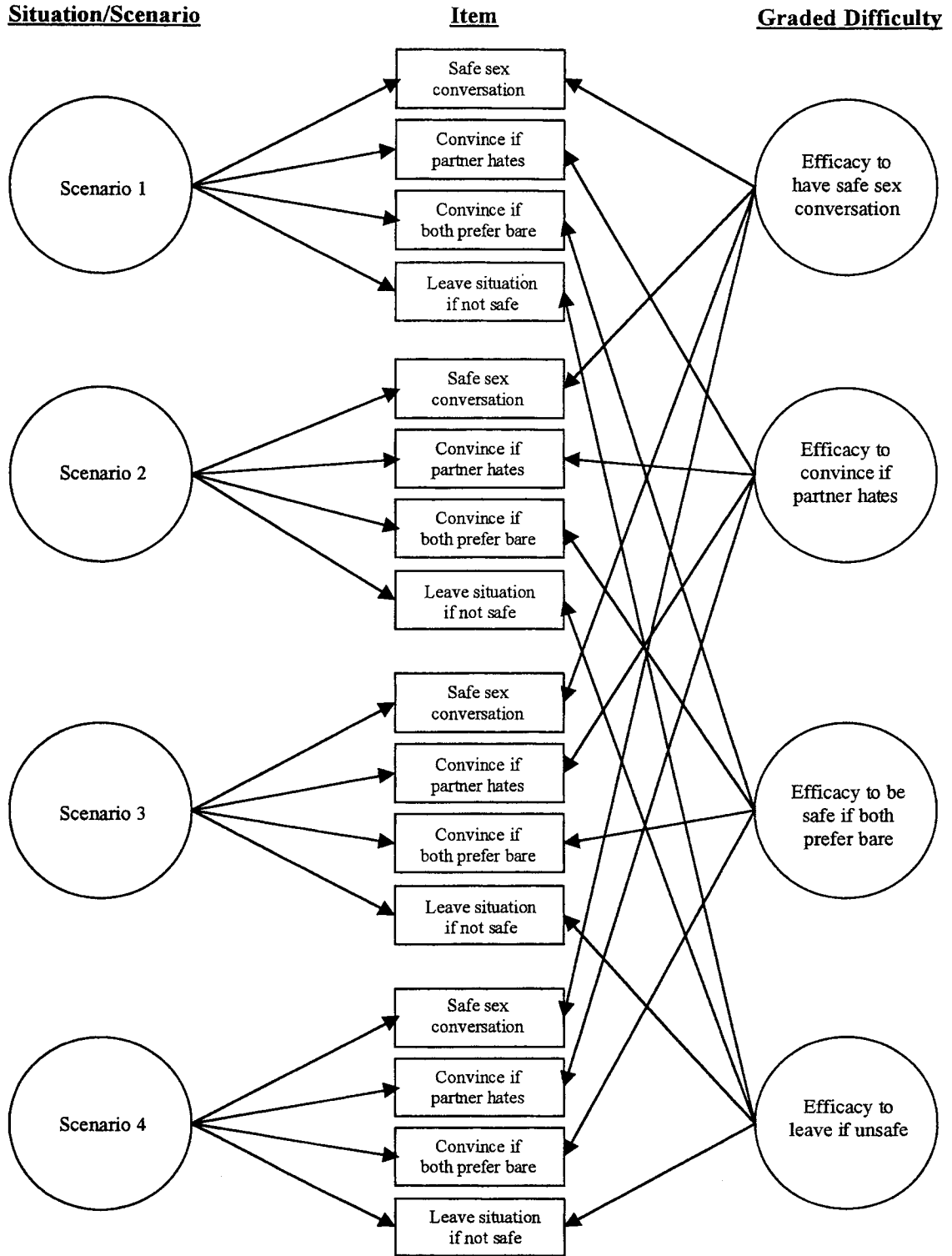
Figure 1. Hypothetical multidimensional confirmatory factor model assessing condom efficacy.

Multiple group analyses. After testing the fits of Models 1–4 for the men and women separately, we used multiple group procedures to compare the men and women (Bentler, 1995; Byrne, 1994). After establishing a baseline multidimensional two-group model, we used multiple-group latent variable models to test the equivalence of the factor structures (measurement models) of the four graded difficulty factors for the men and women. The factor-loading of each measured variable on its latent factor was constrained to equality across the gender groups. As a next more stringent

step, we then tested the equivalence of the covariances and variances among the four graded difficulty latent variables through further equality constraints. The same series of analyses would not have made sense among the situation or scenario factors because they were not equivalent for the men and women.

The plausibility of the equality constraints was determined with the goodness-of-fit indexes described above, chi-square difference tests, and results of the LaGrange multiplier test (LM test; Chou & Bentler, 1990). In the context of a multiple group analysis, the LM test provides information concerning which equality constraints are not reasonable and should be released to improve the fit of the model. Sequential chi-square difference tests are performed in which nonequivalent parameters, reported by univariate and multivariate LM tests, are dropped in succession. The resulting chi-square differences indicate whether differences in the loadings are statistically significant. A chi-square difference value of 3.84 is the criterion for significance for a difference of 1 df, .05 level. Because adjusted chi-squares were used as indicators of fit and a difference between two scaled chi-squares is not distributed as chi-square, the chi-square differences were adjusted using the method of Satorra (2000).

Latent means models. Once the measurement models were contrasted and found to be reasonably similar, differences between the latent means of the groups could be tested. This technique yielded a statistic analogous to a z-score and indicated if one group's mean was significantly higher than the other group's at the latent variable level. Again, we only tested differences in the latent means representing graded difficulty because the situations were not the same in both groups. Also, because the responses arose from different situations or scenarios, gender differences had to be interpreted with some degree of caution.

Results

The sample sizes for the three populations were: 571 men and 366 women from STD clinics and 559 women from health service organizations. Within the combined women's sample, the mean age was 31.6 (SD = 8.5; range = 18–61), and the ethnic/racial breakdown was as follows: 63% African American, 29% Hispanic, 6% White, and 2% other ethnicities. The average education level for the women was 11.1 years (median and mode = 12 years). For the men's sample, mean age was 34.9 (SD = 8.6; range = 18–71), and the ethnic/racial percentages were 78% African American, 17% Hispanic, 2% White, and 3% other ethnicities. The average education level for the men was 11.6 years (median and mode = 12 years).

Single-Factor Model (Model 1)

The first confirmatory model—Model 1 for a single Efficacy variable indicated by the 16 items in the instrument—provided an exceptionally poor fit to the data despite large and significant factor loadings of the individual items on the one Efficacy factor. Chi-square values were extremely large. For the women, the fit indexes were as follows: Robust χ^2(104, N = 925) = 2,194.44, RCFI = .66, RMSEA = 0.147, $p \leq$ .001. For the men, the fit indexes were robust χ^2(104, N = 571) = 1,598.56, RCFI = .66, RMSEA = 0.159, $p \leq$ .001. The LM tests for both the men and women indicated that adding numerous correlated error residuals between items within situations and between similarly worded difficulty items would have provided enormous improvements in goodness of fit.

Situation or Scenario Model Only (Model 2)

Model 2, in which four factors represented the four situations or scenarios, showed improved fit over the single-factor model re-

ported above. However, the fit indexes were not optimal, especially for the men, and chi-square values were quite high. For the women, the fit indexes were robust χ^2(98, N = 925) = 715.43, RCFI = .90, RMSEA = .083, $p <$ .001. For the men, the fit indexes were robust χ^2(98, N = 571) = 682.94, RCFI = .87, RMSEA = 0.102, $p <$ .001. The LM test indicated that great improvements could be made if additional correlated error residuals were included between similarly worded difficulty items. Therefore, scenario alone did not capture the response tendencies of the participants, although again, the factor loadings were large and significant.

Graded Difficulty Model Only (Model 3)

The model in which four factors represent the four sets of graded difficulty items showed only a slightly improved fit over the single-factor model and the fit indexes were quite poor. For the women, the fit indexes were robust χ^2(98, N = 925) = 2,093.23, RCFI = .68, RMSEA = .148, $p <$ .001. For the men, the fit indexes were robust χ^2(98, N = 1,175) = 1,300.33, RCFI = .73, RMSEA = 0.147, $p <$ .001. Again, the factor loadings were large and significant in these models. The LM test indicated that improvements could be made if additional correlated error residuals were included between items from the same scenario. This analysis indicated that difficulty alone does not determine efficacy; specificity about the scenario is important as well. The superiority of the situational models in terms of relatively better model fit was highlighted in this analysis as well as the relative weakness of a graded difficulty model.

Multidimensional Model (Model 4)

The hypothesized model depicted in Figure 1, in which each item was used as an indicator of a situation factor and a graded difficulty factor simultaneously, provided an excellent fit for both the women and the men. No correlated error residuals were necessary to improve model fit, and considering the sizes of the samples, the chi-square values are quite small. For the women, the fit indexes were as follows: robust χ^2(76, N = 925) = 123.20, RCFI = .99, RMSEA = .026, p = .001. For the men, the fit indexes were robust χ^2(76, N = 571) = 93.95, RCFI = .99, RMSEA = 0.020, p = .08. Figures 2 and 3 report the results for the women and men respectively, including correlations among the factors and the standardized factor loadings. All factor loadings are significant. The women (Figure 2) had reasonably large and significant correlations among both the graded difficulty factors and the situation-specific factors. Thus, the women's efficacy was consistent across various degrees of difficulty and across the four situations.

Results differed for the men (Figure 3) in that there was less consistency across the situations for the men than the women. Most covariances between the situational factors were nonsignificant. (However, it should be noted that the scenarios presented to the men and women were different, and qualitative differences in the scenarios may be influencing the outcome. Direct comparisons of the outcomes between the men and women on situational aspects of the model should be avoided.) Correlations among the graded difficulty factors were quite high.

Figure 2. Results of multidimensional confirmatory factor analysis for women. Parameter estimates are standardized. All factor loadings are significant. *** $p \leq .001$.

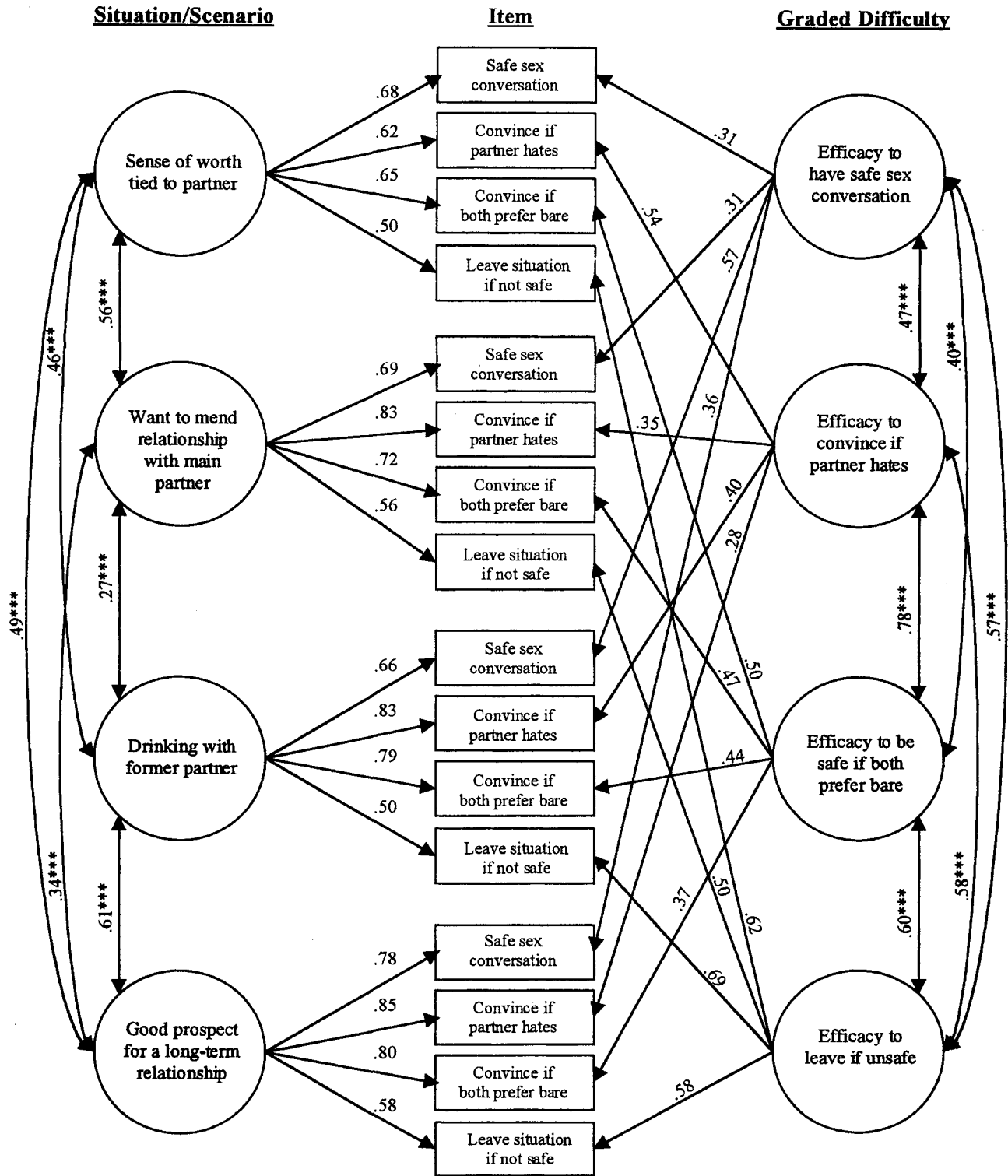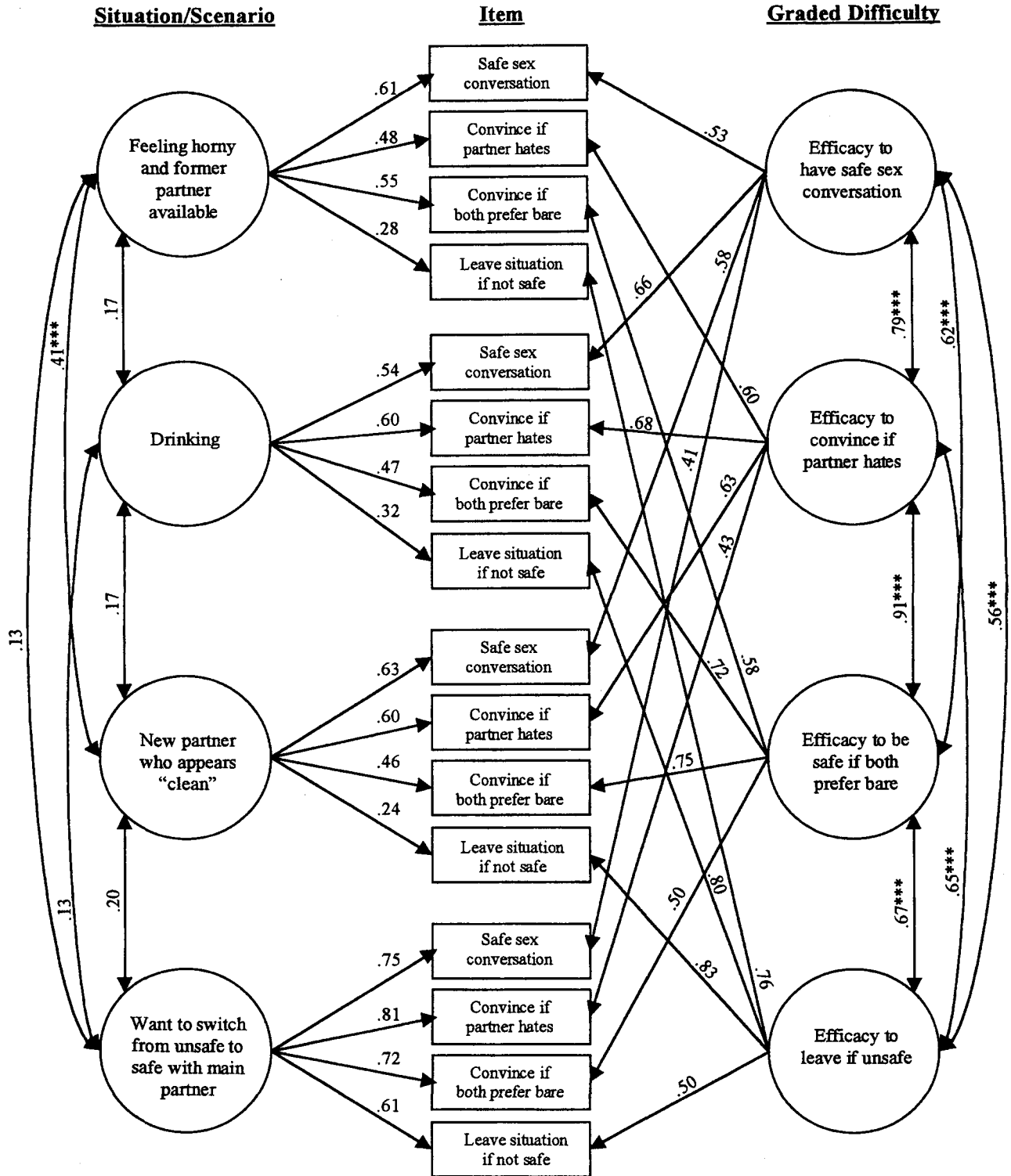*Figure 3.* Results of multidimensional confirmatory factor analysis for men. Parameter estimates are standardized. All factor loadings are significant. *** $p \leq .001$.

Multisample Comparisons

Although the situations were not similar for men and women, by using multiple group techniques we were able to compare response tendencies to graded difficulty level of both groups. This was done by focusing on the "right side" of the hypothesized multidimensional model (see Figure 1), in which the items based on graded difficulty were exactly the same across the genders and the method variance from the dissimilar situations was partitioned and controlled separately. As mentioned above, multiple group comparisons by situations were not possible because the situations differed by gender. An initial model in which the two groups were contrasted but no equality constraints were imposed served as the baseline. That combined model, on the basis of the separate results reported above, had an excellent fit: Robust $\chi^2(152, N = 1,496) = 222.96$, RCFI = .99, RMSEA = 0.018. The factor structures based on graded difficulty were then constrained as equivalent between the men and women. The chi-square difference between the constrained and baseline model was 14.45/12 df after adjusting the scaled chi-squares appropriately (Satorra, 2000). This difference is nonsignificant and remarkably modest considering that the items arise from separate situations for the men and women. Other indicators of fit remained strong: RCFI = .99, RMSEA = .017.

At that point, we constrained the six covariances of the graded difficulty latent variables as equivalent. This provided a significant but relatively modest decrement in fit that is not surprising given the differences in the correlations among the four factors observable in Figures 2 and 3 (adjusted chi-square difference = 31.01, 18 df between baseline and the constrained model). The LM test indicated that the equality constraint on the covariance between efficacy to have a safe sex conversation and efficacy to convince partner if he or she hated to use condoms was particularly untenable (correlation for men = .79, for women = .47, $\chi^2 = 7.12$). The next step of stringency was constraints on the four variances. This analysis provided an adjusted chi-square difference of 52.66/22 df, again a significant difference. The variability was especially marked for the men and women on the difficulty item assessing their efficacy to leave the situation if it was unsafe. These results indicate that the relationships within the factor structure were quite similar for men and women; the questionnaire held the same meaning for them and response tendencies were similar. However, their variability on the latent variables and the relationships among the latent variables tended to diverge by gender.

In the latent means analyses, significant gender differences were observed between the means of the latent variables representing graded difficulty. Table 2 reports the measured means and standard deviations for each item for the men and women. Significant latent mean differences are also reported in Table 2. The women reported greater efficacy in three out of four latent variables representing difficulty.

Discussion

Consistent with Bandura's theory, we found that a global model of a single Efficacy construct made up of the 16 items fit the data poorly among this sample of over 90% minority participants. The situations/scenarios model showed improved fit over the single-factor model, but the fit indexes were not optimal, indicating that situation alone could not account for the participant responses

Table 2
Means and Standard Deviations for Men (N = 571) and Women (N = 925)

Scenario	Men		Women	
	M	SD	M	SD
Safe sex conversations*				
1	7.3	2.9	7.5	3.0
2	6.6	3.3	7.0	3.3
3	7.7	2.8	7.6	3.1
4	7.0	3.4	7.9	2.9
Convince if partner hates				
1	7.2	2.9	7.0	3.1
2	6.6	3.2	6.8	3.3
3	7.3	2.9	7.4	3.0
4	6.7	3.4	7.5	2.9
Safe if both prefer bare**				
1	6.9	3.1	6.9	3.1
2	6.4	3.4	6.6	3.4
3	7.0	3.0	7.3	3.1
4	6.4	3.5	7.5	3.0
Can leave if unsafe***				
1	6.3	3.6	7.0	3.5
2	6.0	3.6	6.5	3.7
3	6.5	3.5	7.7	3.2
4	5.7	3.8	8.0	3.0

Note. Latent means significantly higher for women, two-tailed test. Range for all items = 0–10; higher scores mean greater confidence.
$* p \leq .05.$ $** p \leq .01.$ $*** p \leq .001.$

despite large and significant factor loadings. The graded difficulty model also showed improved fit over the single-factor model, although fit indexes were poor, indicating again that difficulty alone does not determine efficacy despite large and significant factor loadings. Finally, utilizing a multidimensional model, for which items were used as indicators of situation and graded difficulty concurrently, provided an excellent fit for both the men's and the women's samples.

Among the women, efficacy was consistent across degrees of difficulty and situations. Among the men, efficacy was very consistent across the graded difficulty levels but much less consistent across the situations—that is, the men's efficacy relied more on the level of skill required. This may be partially related to the fact that men put condoms on themselves and can control the final skill needed in a safe sexual interaction. However, they cannot necessarily control the situation in which a sexual interaction may occur. This situation specificity may indicate they are quite cognizant of the fact that there are situations in which they themselves may be influenced by not wanting to use a condom (e.g., when they have been drinking or when they are very aroused). The multisample comparisons show the graded difficulty assessment scale has the same meaning for both men and women and that their response tendencies are similar. The relationships among the latent variables do diverge by gender both in the magnitude of the correlations and in their means, with women reporting greater self-efficacy in three of the four latent variables for graded difficulty.

A multidimensional design allowed us to account for the problem of correlated errors arising from method and trait variance in the design of the instrument (Bentler & Stein, 1992). As reported above, improvements may have been found in Models 1–3 if various correlated errors based on the other dimensions of the questionnaire had been added to these models. This trial-and-error nonsystematic procedure would not have had the elegance or efficiency of the theoretically based multidimensional design in which all such relationships were fully accounted for, planned a priori, and explicitly included in the design. We uncovered in a systematic way important gender differences in basic components of safe-sex negotiation efficacy for men and women, and also learned more about safe-sex negotiation efficacy within genders. The viability and strength of the eight separate latent variables demonstrate that both situations and gradation of difficulty are important and salient in safe-sex negotiation efficacy. Thus, realistically, efficacy in safe-sex negotiation cannot be conveniently subsumed under one simplistic unitary construct that adequately reflects behavioral tendencies or that is independent of situation and level of skills. Both of these latter areas should be taken into account by those designing programs to improve safe-sex efficacy skills.

There are limitations to this study. The situations as presented in the questionnaire are different for men and women. This is due to the fact that the situations were based on data from the initial formative interviews that were conducted prior to item development, which indicated that men and women were influenced by somewhat different barriers to having safe sex. However, the effect of that difference in the questionnaires may have driven some differences that we found, even though, depending on the degree of difficulty of the behavior, behaviors in one particular situation were partitioned from tendencies across situations by CFA techniques. We controlled the specific measurement error based on situation for the men and the women and were able to contrast men and women in the substantive trait (graded difficulty) portion of the model, even though their responses arose from differing situations, by partitioning out from each measured variable the part based on the situation with method factors uncorrelated with the difficulty factors (Stein, Newcomb, & Bentler, 1988). It is notable that the factor structures for the men and women were so similar given that they arose from different scenes. In addition, this sample was made up of mainly minority participants. African Americans were well represented (78%), but the study included fewer Hispanics (17%), very few Whites (2%), and very few other ethnicities (3%). Findings may differ for a primarily White sample.

Self-efficacy scales should be tailored to specific domains of functioning; no standard sets of domain-specific items apply to all populations in all situations (Maibach & Murphy, 1995). Although developing such scales may require more formative work and elicitation research, the end results would be an improvement in predictive power. Despite less than optimal measurement, previous studies attest to the fact that self-efficacy has a strong impact on health behavior change. These findings confirm the need to assess self-efficacy in a situation- and level-specific manner. Theoretically precise assessment of self-efficacy will provide an even better picture of the role perceptions of efficacy play in safe-sex behavior change and maintenance of change. Assessing self-efficacy in this manner not only improves understanding of the dynamics of human behavior but, perhaps more important, by

indicating the most likely causes of performance failure, can help focus behavior modification efforts.

References

Bandura, A. (1977). Social learning theory. Englewood Cliffs, NJ: Prentice Hall.

Bandura, A. (1986). Social foundations of thought and action: A social cognitive theory. Englewood Cliffs, NJ: Prentice Hall.

Bandura, A. (1990). Perceived self-efficacy in the exercise of control over AIDS infection. Evaluation and Program Planning, 13, 9–17.

Bandura, A. (1991). Self-efficacy mechanism in physiological activation and health-promoting behavior. In J. Madden (Ed.), Neurobiology of learning, emotion, and affect (pp. 229–269). New York: Raven Press.

Bandura, A. (1994). Social cognitive theory and exercise of control over HIV infection. In R. J. DiClemente & J. L. Peterson (Eds.), Preventing AIDS: Theories and methods of behavioral interventions (pp. 25–59). New York: Plenum Press.

Bandura, A. (1997). Self-efficacy: The exercise of control. New York: Freeman.

Bandura, A. (1999). A sociocognitive analysis of substance abuse: An agentic perspective. Psychological Science, 10, 214–217.

Basen-Engquist, K., & Parcel, G. S. (1992). Attitudes, norms, and self-efficacy: A model of adolescents' HIV-related sexual risk behavior. Health Education Quarterly, 19, 263–277.

Bentler, P. M. (1995). EQS Structural Equations Program Manual. Encino, CA: Multivariate Software.

Bentler, P. M., & Dudgeon, P. (1996). Covariance structure analysis: Statistical practice, theory, and directions. Annual Review of Psychology, 47, 563–592.

Bentler, P. M., & Stein, J. A. (1992). Structural equation models in medical research. Statistical Methods in Medical Research, 1, 159–181.

Brafford, L. J., & Beck, K. H. (1991). Development and validation of a condom self-efficacy scale for college students. Journal of American College Health, 39, 70–80.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), Testing structural equation models. Newbury Park, CA: Sage.

Byrne, B. M. (1994). Structural equation modeling with EQS and EQS/Windows. Thousand Oaks, CA: Sage.

Chou, C.-P., & Bentler, P. M. (1990). Model modification in covariance structure modeling: A comparison among likelihood ratio, Lagrange multiplier, and Wald tests. Multivariate Behavioral Research, 25, 115–136.

Hofstetter, C. R., Sallis, J. F., & Hovell, M. F. (1990). Some health dimensions of self-efficacy: Analysis of theoretical specificity. Social Science and Medicine, 31, 1051–1056.

Hu, L.-T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling, 6, 1–55.

Jemmott, J. B., Jemmott, L. S., & Fong, G. T. (1992). Reductions in HIV risk-associated sexual behaviors among black male adolescents: Effects of an AIDS prevention intervention. American Journal of Public Health, 82, 372–377.

Maibach, E., & Murphy, D. A. (1995). Self-efficacy in health promotion research and practice: Conceptualization and measurement. Health Education Research, 10, 37–50.

McKusick, L., Wiley, J., Coates, T. J., & Morin, S. F. (1986, November). Predictors of AIDS behavioral risk reduction: The AIDS Behavioral Research Project. Paper presented at the New Zealand AIDS Foundation Prevention Education Workshop, Auckland, New Zealand.

Murphy, D. A., Multhauf, K. E., & Kalichman, S. C. (1995). Development and validation of a graded, safe-sex self-efficacy scale. The Behavior Therapist, January, 8–10.

Murphy, D. A., Rotheram-Borus, M. J., Srinivasan, S., & Hunt, W. K. (1997). Recruiting a cohort for the HIV vaccine trial: Sensitivity and specificity of a screening for sexual and substance use acts. *AIDS and Behavior, 1*, 75–80.

National Institute of Mental Health Multisite HIV Prevention Trial Group. (1997a). Endpoints and other measures in a multisite HIV prevention trial: Rationale and psychometric properties. *AIDS, 11*, S37–S48.

National Institute of Mental Health Multisite HIV Prevention Trial Group. (1997b). Methodological overview of a multisite HIV prevention trial for populations at risk for HIV. *AIDS, 11*, S1–S12.

National Institute of Mental Health Multisite HIV Prevention Trial Group. (1997c). Screening, recruiting and predicting retention of participants in a multisite HIV prevention trial. *AIDS, 11*, S13–S20.

National Institute of Mental Health Multisite HIV Prevention Trial Group. (1998, June 19). The NIMH Multisite HIV Prevention Trial: Reducing HIV sexual risk behavior. *Science, 280*, 1889–1894.

O'Leary, A., Goodhart, F., Jemmott, L. S., & Boccher-Lattimore, D. (1992). Predictors of safer sex on the college campus: A social cognitive theory analysis. *Journal of American College Health, 40*, 254–263.

Satorra, A. (2000). Scaled and adjusted restricted tests in multi-sample analysis of moment structures. In R. D. H. Heijmans, D. S. G. Pollock, & A. Satorra (Eds.), *Innovations in multivariate statistical analysis* (pp. 233–247). London: Kluwer.

Stein, J. A., Newcomb, M. D., & Bentler, P. M. (1988). Structure of drug use behaviors and consequences among young adults: Multitrait-multimethod assessment of frequency, quantity, work site, and problem substance use. *Journal of Applied Psychology, 73*, 595–603.

Wulfert, E., & Wan, C. K. (1993). Condom use: A self-efficacy model. *Health Psychology, 12*, 346–353.

AMERICAN PSYCHOLOGICAL ASSOCIATION
SUBSCRIPTION CLAIMS INFORMATION

Today's Date:_____

We provide this form to assist members, institutions, and nonmember individuals with any subscription problems. With the appropriate information we can begin a resolution. If you use the services of an agent, please do NOT duplicate claims through them and directly to us. **PLEASE PRINT CLEARLY AND IN INK IF POSSIBLE.**

PRINT FULL NAME OR KEY NAME OF INSTITUTION

ADDRESS

CITY   STATE/COUNTRY   ZIP

YOUR NAME AND PHONE NUMBER

TITLE

MEMBER OR CUSTOMER NUMBER (MAY BE FOUND ON ANY PAST ISSUE LABEL)

DATE YOUR ORDER WAS MAILED (OR PHONED)

____PREPAID ____CHECK ____CHARGE
CHECK/CARD CLEARED DATE:_____

(If possible, send a copy, front and back, of your cancelled check to help us in our research of your claim.)

ISSUES: ____ MISSING ____ DAMAGED

VOLUME OR YEAR   NUMBER OR MONTH

Thank you. Once a claim is received and resolved, delivery of replacement issues routinely takes 4–6 weeks.

━━━ (TO BE FILLED OUT BY APA STAFF) ━━━

DATE RECEIVED: _____   DATE OF ACTION: _____
ACTION TAKEN: _____   INV. NO. & DATE: _____
STAFF NAME: _____   LABEL NO. & DATE:_____

Send this form to APA Subscription Claims, 750 First Street, NE, Washington, DC 20002-4242

PLEASE DO NOT REMOVE. A PHOTOCOPY MAY BE USED.