A Study of the Validity of a Modified Ordinal Scale of HIV Transmission Risk Among Seropositive Men who Have Sex with Men

DISSERTATION

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Sandra J. Reed, M.B.A.

Graduate Program in Education

The Ohio State University

2012

Dissertation Committee:

Ann A. O’Connell, Ed.D., Advisor

Julianne M. Serovich, Ph.D.

Dorinda J. Gallant, Ph.D.

ABSTRACT

The use of a measure of HIV transmission risk which incorporates seroadaptive behaviors is needed in order to assess the complex effects of disclosure on HIV transmission. The purpose of this study was to explore the use of a modified version of the ordinal scale of HIV risk behavior proposed by Osmond, Pollack, Paul, and Catania (2007) to operationalize the risk of HIV transmission among HIV-positive MSM and to establish the strength of the empirical evidence supporting the use of these scores for inference about HIV transmission. The original measure was modified to include the full potential for seropositioning (i.e., with and without 100% condom use) as a potential preventive strategy used by MSM to reduce the risk of HIV transmission. Additionally, the measure was applied to individual sexual encounters rather than globally so that the frequency of risk behaviors can be accurately modeled. The appropriateness of these methods was explored using data from a study involving the disclosure of serostatus to sexual partners in a sample of HIV-positive MSM. Results of the study are promising for the refinement of measurements of HIV transmission risk, and for the understanding of seroadaptive behavior in MSM. For researchers who are seeking to demonstrate the effectiveness of interventions designed to reduce HIV transmission risk, the ordinal measure provides a means for detecting qualitative shifts in sexual activity which can be critical to the question of effectiveness.

DEDICATION

For Jack & Soots …

"For there we loved, and where we love is home,

home that our feet may leave, but not our hearts…"

--Oliver Wendell Holmes (1871)

ACKNOWLEDGEMENTS

"And he that strives to touch the stars, Oft stumbles at a straw."

-Edmund Spenser

There are few times and places in history where a life such as mine would be possible. I am grateful to those who have gone before and made my journey possible.

I would like to honor the fierce women of The Ohio State University who have mentored, guided, and at times cajoled me as was appropriate. First among them is my advisor Dr. Ann A. O'Connell. I thank you for sharing with me some small portion of your intellectual world. You extended a hand to me and, in so doing, changed everything. Second is Dr. Julianne M. Serovich. I thank you for opening all the doors and for teaching me how the big dogs roll. Third is Dr. Karen Zuga. I thank you for personifying the beauty, honor, and value of scholarly work. Fourth is Julie Carpenter-Hubin. I thank you for showing me how grown-ups use data. Fifth is Dr. Dorinda J. Gallant. I thank you for sharing your passion, energy, and joy in learning and teaching. Finally, I would like to recognize Deb Zabloudil, an extraordinary woman, mentor, sister, and friend who has shown so many of us the way.

I also offer thanks to the men of Technology Education who welcomed me into the fold. First is Peter Lund. You are the quintessential teacher and advocate. Thank you for showing me what it means to stand for something - steadfast, unwavering, and with love. Second is Dr. David Meyer, a mensch who paved the way for us all. My admiration and respect for you glows undiminished. Third is Dr. Mark Mahoney, a young man of remarkable passion and imagination, with an equally remarkable heart. It makes me smile to think of each of you, and I will always remember you.

As to some very special friends, thanks must be offered to Dr. Xueli Wang, my adopted mother from Harbin. You are my wise and funny friend who literally plucked me from the street and set me right. Few would have tried, and fewer could have succeeded. You seriously impress me. Thanks also to my dissertation partner Dr. Katye Miller. Together we did battle and emerged victorious. I can't wait to see how it all turns out. To Sally Jackson – a true friend and woman of letters. Thank you for fumbling along with me through the messy business of living in this world. You bring joy to argument and vigor to reflection. Having plumbed the depths of your vast stores of wisdom, advice, and humor across myriad topics, I consider you to be my most accomplished and learned of friends.

To my family – Jack & Soots, Jay, Sue, Janie, and Nick. Despite our years of deep acquaintance, you continue to teach and inspire me. I count myself among the luckiest of all beings - for I am able to call you *mi familia*. As I embark on yet another journey, I am grateful in the knowledge that all paths will lead me home to you.

Finally, to my son Billy J Allen Reed. To know you is my greatest pleasure, and to be counted among your friends is my greatest achievement. Thank you for your love, support, and understanding despite my weaknesses and imperfections. You make it possible for me to carry on.

VITA

1977……………..…………………………Sylvania Northview High School

1996...B.A. Business, University of Toledo

2001...M.B.A., University of Findlay

2007 to present ....Senior Research Associate, Department of Human Development and Family Science, The Ohio State University

Publications

O'Connell, A. A. & Reed, S. J. (in press). Hierarchical data structures, institutional research, and multilevel modeling. In *New Directions in Institutional Research* – Multilevel Modeling: Techniques and Applications in Institutional Research, J. Lott & J. Antony (Eds.). Jossey-Bass.

Serovich, J. M., Craft, S. M., & Reed, S. (2012). Women's HIV disclosure to immediate family and friends. *AIDS Patient Care and STDs, 26*(4), 241-249.

Serovich, J. M., Reed, S. J., Grafsky, E. L., Hartwell, E. E., & Andrist, D. W. (2011). An intervention to assist men who have sex with men disclose their serostatus to family members: Results from a pilot study. *AIDS and Behavior, 15*(8), 1647-1653.

Serovich, J. M., Grafsky, E. L., & Reed, S. (2010). Comparing regret of disclosing HIV versus sexual orientation information by MSM. *AIDS Care, 22*(9), 1052-1059.

Serovich, J. M., Reed, S., Grafsky, E., & Andrist, D. (2009). An intervention to assist men who have sex with other men disclose their serostatus to casual sex partners: Results from a pilot study. *AIDS Education and Prevention, 21*(3), 207-219.

Fields of Study

Major Field: Educational Policy and Leadership

Quantitative Research, Evaluation and Measurement

TABLE OF CONTENTS

ABSTRACT ii

ACKNOWLEDGEMENTS iv

VITA vi

LIST OF TABLES xii

LIST OF FIGURES xv

Background of the Study 1

Purpose of the Study 8

Significance of the Study 10

Potential Implications of the Study 11

Limitations to Generalizability 11

Limitations to Theory 12

CHAPTER 2: REVIEW OF THE LITERATURE 13

History of HIV and HIV Prevention in the United States 14

HIV in Men Who Have Sex with Men (MSM) 18

Risk Reduction in MSM 20

The Importance of Serostatus Disclosure 22

The Need for Valid Measures of Sexual Risk Behavior 26

The Current Measurement and Analysis of Sexual Risk Behavior 35

Validity 47

Current APA Standards 61

Summary 61

Conclusion 62

CHAPTER 3: METHODOLOGY 64

Introduction 64

Context of the Larger Study 64

Data Set Preparation 78

Descriptive Analysis 78

Review of Goals and Statistical Analysis 79

Summary 96

CHAPTER 4: RESULTS 98

Participant Characteristics 99

Participant Global Sexual Behavior 103

Participant Scale Scores 110

Encounter-Specific Measures 113

The Ordinal Measure of HIV Transmission Risk 138

Goal 1: Evaluation of Content Representativeness 143

Goal 2: Evaluation of Content Relevance 147

Goal 3: Evaluation of technical quality 156

Goal 4: Evaluation of Substantive Validity 162

Goal 5: Evaluation of Structural Validity 172

Goal 6: Evaluation of Validity Considering External Factors 189

Goal 7: Evaluation of the Potential for Generalizability 213

Goal 8: Evaluation of Consequential Validity 215

Summary 224

CHAPTER 5: DISCUSSION 226

Purpose and Goals 226

Discussion of Content and Technical Quality 227

Discussion of Substantive and Structural Validity 229

Discussion of Validity Considering External Factors 232

Discussion of Generalizability and Consequential Validity 233

Implications 235

Limitations and Further Research 235

REFERENCES 238

APPENDIX A: Global Measures of Sexual Behavior.. 256

APPENDIX B: Global Measures of Disclosure 257

APPENDIX C: Sexual Encounter Form... 258

APPENDIX D: Disclosure Form.. 266

APPENDIX E: Instructions to Participants 270

APPENDIX F: Original Items for Use in Scale Construction.. 275

APPENDIX G: Table of Methodologies 282

LIST OF TABLES

Table 1. Items Used to Construct the Ordinal Measure of HIV Transmission Risk 75

Table 2. Participant Demographic Characteristics (n_p =145) 100

Table 3. Participant Gender and Sexuality (n_p =145) 101

Table 4. Participant Serostatus Information (n_p =145) 103

Table 5. Participant Global Sexual Activity, Last 30 Days (n_p =145) 105

Table 6. Participant Global Reports of Anal Intercourse, Last 30 Days (n_p =145)........ 106

Table 7. Participant Global Reports of Oral Intercourse, Last 30 Days (n_p =145) 107

Table 8. Participant Global Reports of Non-Disclosure, Last 30 Days (n_p =145)......... 108

Table 9. Participant Global Reports of Disclosure, Last 30 Days (n_p =145) 109

Table 10. Participant Scale Scores.. 111

Table 11. Item Scores and Intercorrelations for Disclosure Outcome Expectancy 112

Table 12. Number of Participants Reporting Encounters, by Reporting Order............. 114

Table 13. Partner characteristics, By the Number of the Encounter............................. 117

Table 14. Encounter Context, By the Number of the Encounter 118

Table 15. Substance Use Behaviors, By the Number of the Encounter 122

Table 16. Partner Serostatus, By the Number of the Encounter 123

Table 17. Partner Disclosure Behaviors (Positive Partners), By Encounter.................. 124

Table 18. Partner Disclosure Behaviors (Negative Partners), By Encounter 126

Table 19. Participant Disclosure, Summing the Last 5 Encounters (n_e =469).............. 127

Table 20. Participant Disclosure Behaviors, By the Number of the Encounter 128

Table 21. Participant Non-Disclosure Behaviors, By the Number of the Encounter 130

Table 22. Anal Intercourse, By the Encounter Number.. 132

Table 23. Reasons for Considering the Encounter Safe, By Encounter 135

Table 24. Reasons for Considering the Encounter Unsafe, By Encounter 137

Table 25. Items Used to Construct the Ordinal Scale 138

Table 26. Levels of Ordinal HIV Transmission Risk 139

Table 27. Level of Ordinal HIV Transmission Risk, By Encounter............................. 141

Table 28. N/A Responses, by Item and Encounter 149

Table 29. Skipped Items, By Item and Encounter 152

Table 30. N/A Responses, Predicted by Participant Characteristics 154

Table 31. Item Readability... 157

Table 32. Frequency of Skipped Responses, By Encounter 160

Table 33. Correlations With Number of Skipped Responses, By Encounter 161

Table 34. Number of Encounters[d] , Cross-Tab By Section 165

Table 35. Multinomial Regression of Reporting Consistency (n_p = 133) 169

Table 36. Frequency of Dichotomous Risk Behaviors, By Encounter 174

Table 37. Summary of Item Intercorrelations, Encounter 1 Only 177

Table 38. Correlations Between Ordinal Scores and Items, By Encounter 179

Table 39. Corrected Item-Total Correlations, By Encounter.. 181

Table 40. Modified Dichotomous Risk Behaviors 183

Table 41. Dichotomous Risk Behaviors by Partner Serostatus[a] 187

Table 42. Dichotomous Environmental Risk Factors 191

Table 43. Level of Ordinal Environmental Risk... 192

Table 44. Tied Rank Scores, HIV Risk Transmission Total.. 193

Table 45. Tied Rank Scores, Environmental Scale Total 194

Table 46. Multi-Trait Multi-Method Matrix, Encounter 1 Only (n_p = 128) 195

Table 47. Correlations Between Scale Scores Risk, By Encounter 197

Table 48. Correlations[a] Between Participant Characteristics and Risk, By Encounter . 199

Table 49. Participant-Level Predictors of HIV Transmission Risk 202

Table 50. Encounter-Level Predictors of HIV Transmission Risk 207

Table 51. Cumulative Analyses for an Ordinal Model of Risk Category...................... 210

Table 52. Risk Scores Using Global Measures... 216

Table 53. Risk Scores Using Encounter-Specific Measures... 217

Table 54. Correlations Between Risk Measures 218

Table 55. Relative Performance of Alternative Measures[a].. 221

Table 56. Original Items for Use in Scale Construction 276

Table 57. Table of Methodologies 283

LIST OF FIGURES

Figure 1. CONSORT Diagram of the Larger Study 67

Figure 2. Hierarchy of Transmission Risk, Ordinal Scale 74

Figure 3. Output from Rasch Analysis of Scale Items... 184

Figure 4. Sexual Encounter Form 259

Figure 5. Disclosure Form 267

Figure 6. Instructions to Participants 271

CHAPTER 1: INTRODUCTION

In 2009, the Centers for Disease Control (CDC), through the Division of HIV/AIDS Prevention (DHAP), awarded over $500,000,000 funding for HIV prevention activities throughout the United States (CDC, 2011a). In association with these funds, the CDC has increasingly stressed the importance of evidence-based interventions focused on the promotion of behaviors which eliminate the potential for transmission of the HIV virus. Efforts to determine which of these HIV prevention activities hold promise for the reduction of behavior risk and HIV transmission rely on the ability to adequately measure risk behavior as well as the ability to apply appropriate statistical techniques to estimate intervention effectiveness.

Background of the Study

The problem of transmission is particularly important among men who have sex with men (MSM) who, regardless of race, represent the group which is most impacted by HIV infection in the United States. The designation MSM is based on sexual behavior rather than sexual identity. It includes men who may identify their sexuality as gay, bisexual, or heterosexual. Though MSM represent only about 4% of the male population aged 13 or older, almost 50% of those persons currently living with HIV/AIDS in the US are men who report engaging in same-sex sexual behaviors, and more than half of all new infections reported in the US are among MSM. Some estimates place the rate of diagnosis among MSM as high as 44 times that of other men (CDC, 2011b).

While HIV can also be transmitted through blood transfusion or mother-to-child transmission, the primary means of transmission among MSM is through sexual risk behavior. Many aspects of sexual behavior must be considered in the estimation of transmission risk among MSM including type of sex, condom use, and partner type. Unprotected anal intercourse (UAI) has been associated with the highest risk of disease transmission (CDC, 2011c). However, an accurate estimate of risk must also take into consideration the serostatus of both partners and their positioning during sexual intercourse. Vittinghoff, Douglas, Judson, McKirnan, MacQueen, and Buchbinder (1999) reported that, in a study of 2,189 homosexual and bisexual men, the per-contact risk to seronegative men of contracting HIV in receptive unprotected anal sex with a partner of unknown status (.27%) was significantly higher than insertive anal sex with the same type of partner (.06%). Use of a condom reduced those risks to .18% and .04% respectively. In a study of HIV transmission risk in MSM, Varghese, Maher, Peterman, Branson, and Steketee (2002) reported that the relative risk to a seronegative individual of having UAI with a partner of unknown serostatus was 47 times higher than with a partner who had tested negative for HIV. If the partner was known to be HIV positive, the risk was 4,706 times higher than UAI with a negative partner.

Traditional HIV prevention activities have focused on abstinence or delay of sexual activity, decreasing the number of sexual partners, and condom use as the key strategies (Global Working Group, 2010) for transmission reduction. However, current research suggests that MSM individuals' personal strategies for risk reduction may be more complex. Men may practice a variety of methods of risk reduction which are not based on condom use. Non-condom strategies reported by MSM include withdrawal prior to ejaculation, negotiated safety, and seroadaptation (Highlaymen, 2009; Van Griensven, 2008, Parsons et al., 2005). In negotiated safety, MSM who are seronegative agree to have unprotected sex only within concordant primary partnerships. These agreements may involve a commitment to monogamy, mutual HIV testing, or a restriction of the types of sex permitted outside the relationship (Guzman et al., 2005). Seroadaptation refers to a collection of strategies employed by MSM to impact the risk of transmission including serosorting and seropositioning. Serosorting MSM, whether they are HIV-positive or not, select seroconcordant partners, particularly when they wish to engage in risky sexual behaviors. For example, seropositive men may choose sexual partners who are also positive in order to reduce the risk of UAI, while seronegative MSM may choose to avoid UAI with any partner who is serodiscordant. Seropositioning is a technique used by serodiscordant partners to reduce transmission risk based on the lower risk of transmission when the seropositive partner bottoms (i.e., is the receptive partner in UAI).

The use of seroadapative strategies among MSM has been well-documented (Fendrich, Mackesy-Amiti, Johnson, & Pollock, 2010; Halkitis, Moeller, & Pollock, 2008; McFarland et al., 2011; Snowden, Raymond, & McFarland, 2009; Truong et al., 2006), but their appropriateness has been controversial. For people living with HIV (PLWH), serosorting might be a way to reduce anxiety about transmitting HIV while increasing sexual enjoyment and intimacy (Sincolfi & Moeller, 2007). Additionally, research has shown that serosorting can reduce the risk of HIV transmission (Eaton, Kalichman, & Cherry, 2010; Morin et al., 2008; Philip, Yu, Donnell, Vittinghoff, & Buchbinder., 2010). In a multisite study of seronegative MSM, Philip and colleagues (2010) reported a significant reduction in the rate of HIV acquisition due to serosorting. Eaton and colleagues (2010) reported that lower rates of seroadaptive behavior in African American MSM might be a contributor to the disproportionately high infections rates in that population. Others have found less promising results (Golden, Stekler, Hughes, & Wood, 2008; Truong et al., 2006). While Truong et al. (2006) suggested that observed stabilization in the HIV infection rate among MSM in San Francisco from 1998-2004 might be due to increased serosorting, they also reported increases in the transmission of other sexually transmitted infections (e.g., rectal gonorrhea, syphilis) during the same period. Among seronegative men, the risk of serosorting is greater because HIV positive partners may be unaware of their positive serostatus (Wilson, Regan, Heymer, Prestage, & Grulich, 2010) or because disclosure of serostatus between the partners may be inadequate (Eaton, Cherry, Cain, & Pope, 2011; Philip et al., 2010; Zablotska et al., 2009).

The importance of disclosure in HIV prevention is increasingly a subject of interest among researchers. Several recent studies highlight the potential importance of disclosure as a prevention strategy (Chiasson, Shaw, Humberstone, Hirshfield, & Hartel, 2009; Klitzman et al., 2007; Rosser et al., 2008; Serovich, Reed, Grafsky, & Andrist, 2009; Tieu et al., 2011). For example, in a multisite study among 675 MSM in six U.S.

cities, Rosser and colleagues (2008) reported a significant association between disclosure to secondary partners and a number of positive outcomes including unprotected anal intercourse among discordant partners. Others have suggested that disclosure is not an effective HIV prevention strategy. Citing skepticism among their participants about the truthfulness of their partners' disclosures and the lack of evidence of safer sex following disclosure, Sheon and Lee (2009) suggest that efforts to increase disclosure among MSM may not result in decreased HIV transmission.

Calls for increased research and intervention to assist PLWH in disclosing their serostatus have been made (Grossman et al., 2011; Tieu et al., 2011). However, efforts to evaluate the effectiveness of new disclosure-based interventions may be hampered by the lack of adequate measures of HIV transmission risk. In a recent review of articles reporting sexual risk behaviors in *AIDS Prevention and Education* during 2006, of the 15 empirical studies reporting some measure of HIV risk behavior, all but two were limited to the measurement of individual continuous count (e.g., number of sexual partners, number of sexual acts) or dichotomous status (i.e., unprotected sex) variables and only one (Coyle et al., 2006) utilized multilevel analysis across multiple partners or encounters. In one example of this approach to analysis, Hong and colleagues (2006) studied the effect of perceived partner serostatus on sexual behavior. In the study, they asked participants to report sexual behavior for the last five partners including the type of partner, number of vaginal, insertive anal, and receptive anal sex acts, and condom use in each act. Data on sexual acts were then reclassified. Four individual dichotomous variables were created: 1) 100% condom use, 2) any unprotected vaginal sex act, 3) unprotective insertive sex, and 4) unprotected receptive sex. Analysis was limited to contingency table methods and chi-squared statistics for categorical responses (i.e., yes or no on each behavior). The effects of potential covariates were assessed separately for each response variable using binary logistic regression, and odds ratios were reported. While this approach to the analysis of risk behavior provides useful information, it fails to capture the complex interrelationships between risk behaviors and ignores the underlying order of risk implied by the characterization of sexual risk. Additionally, it ignores the effect of the frequency of risky behavior on transmission risk as each participant is coded only once based on the highest reported risk level.

Ordinal scales of HIV transmission have recently been proposed which could provide a viable alternative to the one-at-a-time analysis of individual risk factors; however, despite the fact that many variables in the social sciences are ordinal in nature, the use of ordinal response measures in behavioral research presents unique challenges (Cliff & Keats, 2003; Hedeker, 2007; Hedeker & Gibbons, 1994; O'Connell, 2006). Methods for modeling ordinal outcomes are now available which are appropriate for use with ordinal scales of behavioral risk. For example, the *cumulative logit model for ordinal responses* (Agresti, 2007) has been applied in a variety of behavioral science contexts. In this generalization of the commonly used logistic regression model, the cumulative logit is defined as:

(1)

where the logit of the cumulative probability that response *Y* falls at or below a particular category is equal to the natural log of the cumulative odds for that category. The use of a model for ordinal response, such as the *cumulative logit model*, provides enhanced analytical power in that it utilizes the ordering of the underlying ordinal categories.

More recently, generalizations of these models to the hierarchical analysis of structured data sets provides even greater analytical power in the assessment of HIV prevention effectiveness. In the area of HIV prevention, the use of randomized and longitudinal designs, regarded as the strongest approach for the evaluation of effectiveness, often requires that data are collected across multiple sites or observations. In these situations, homogeneity within the various sites or occasions introduces non-independence of individual observations. Use of a model that ignores this homogeneity results in inflation of Type I error. Multilevel models inherently address this issue. Fortunately, models for ordinal responses, such as the *cumulative logit model*, have been extended to the multilevel context (O'Connell, Goldstein, Rogers, & Peng, 2008; Raudenbush & Bryk, 2002). Such models can easily accommodate the multilevel structure of the data, and the estimation of the effects of contextual variables on the ordinal response. The multilevel contextual model for ordinal response which accounts for explanatory variables at both the individual level (*X)* and the group level (*W*) is as follows:

(2a)

(2b)

where the cumulative logit is a function of a random intercept which is permitted to vary across groups, the effects of individual level predictors (i.e., slopes) which are also permitted to vary across groups, and a category specific factor which incorporates a different value of the intercept (referred to a “threshold”) for each category. The addition of the Level 2 equation permits modeling with explanatory variables at the group level which affect either the intercept or the slopes in Level 1.

The application of ordinal response scales and multilevel ordinal logistic regression to the detection of intervention effects on HIV risk behavior holds promise. Through ordinal scaling, risk measures can more adequately reflect the underlying continuum of risk behavior and strengthen the validity of the scores which arise from measurement. By incorporating ordinal responses into modeling strategies, researchers can exploit this underlying ordering to more accurately reflect the impact of external variables on sexual risk behavior. By extending these models to the hierarchical framework, ordinal scales of risk can be utilized to assess intervention in multi-site or repeated measure trials which can be used to strengthen the case for intervention effectiveness. However, to fully understand the implications of the use of ordinal response in the study of HIV transmission, a thorough exploration of the approach within the complex context of sexual behavior is needed.

Purpose of the Study

The use of a measure of HIV transmission risk which incorporates seroadaptive behaviors is needed in order to assess the complex effects of disclosure on HIV transmission. The purpose of this study was to explore the use of a modified version of the ordinal scale of HIV risk behavior proposed by Osmond, Pollack, Paul, and Catania (2007) to operationalize the risk of HIV transmission among HIV-positive MSM and to establish the strength of the empirical evidence supporting the use of these scores for inference about HIV transmission. The original measure was modified to include the full potential for seropositioning (i.e., with and without 100% condom use) as a potential preventive strategy used by MSM to reduce the risk of HIV transmission. Additionally, the measure was applied to individual sexual encounters rather than globally so that the frequency of risk behaviors can be accurately modeled. The appropriateness of these methods was explored using data from a study involving the disclosure of serostatus to sexual partners in a sample of HIV-positive MSM. Application of selected models focused on the following goals:

1. To evaluate the validity of a new ordinal scale of HIV transmission risk modified from a scale proposed by Osmond, et al. (2007) by assessing its:
    a. Content validity, including the representativeness and relevance of the items used to generate the scales, and their technical quality
    b. Substantive validity involving conformance of both items and response patterns to theory-based hypotheses
    c. Structural validity supported by evidence of the unidimensionality and local independence of the observed scores

d. Validity considering select external factors, including seroconcordance and serostatus disclosure
e. Potential for generalizability, based on estimates of both person and item reliability
f. Consequential validity based on sensitivity and specificity comparisons with typically used nominal measures of HIV transmission risk

Significance of the Study

An initial review of the literature reveals that persons interested in studying the effect of interventions to reduce HIV transmission have a limited number of risk scales available and that those scales which are available do not adequately reflect the complexity of the sexual decision-making behavior of HIV-positive MSM. This is particularly true in the evaluation of interventions which seek to reduce transmission risk through increased serostatus disclosure. Recent passage of disclosure laws in many states has intensified the focus on the interaction between disclosure and sexual behavior. Such laws make it a crime, subject to prosecution, for a person living with HIV to engage in sexual intercourse without disclosure (The Center for HIV Law and Policy, 2012). While the research on serostatus disclosure has generally established the rate of disclosure in various groups including HIV-positive MSM, the study of the effect of disclosure on the rate of HIV transmission is relatively new. This study focuses on the modification of an ordinal scale of sexual risk behavior which could ultimately lead to a better understanding of the role of serostatus disclosure and of the effectiveness of interventions designed to reduce the risk of HIV transmission in the population of MSM. The use of multilevel ordinal models to assess the effect of disclosure, and any number of other behaviors, on the risk of HIV transmission within individual sexual encounters may help to overcome the limitations imposed by more commonly used multivariate techniques, permitting researchers to better understand the complexity of the context within which disclosure occurs.

Potential Implications of the Study

It is hoped that the results of this study will assist researchers in the development of ordinal scales appropriate for the study of HIV risk behaviors at the encounter level, provide a guide for the application of logistic regression models for ordinal responses within the context of HIV disclosure, introduce new and effective tools in the effort understand the role of disclosure in sexual risk behavior, and identify problematic areas in the measurement of sexual behavior risk in MSM which might still need to be addressed. It is also hoped that information regarding the application of the hierarchical logistic models employed in this study will provide relevant support for the application of these techniques in other disciplines where hierarchical data are encountered. Most importantly, it is hoped that the use of ordinal scales of HIV transmission risk will make the measurement of the effectiveness of evidence-based interventions more precise, leading to an increased ability to discern the effects of these interventions on the sexual risk behavior of MSM living with HIV.

Limitations to Generalizability

A convenience sample was used to collect information on sexual behavior in MSM and serostatus disclosure. The nature and variety of methods employed in recruitment made it impossible to estimate the response rate. Additionally, those who volunteered for the study may have been more comfortable with discussing sexual activity and disclosure. As such, it is difficult to assure the representativeness of the sample or the collected data. As of 2009, 2754 HIV positive males were reported living in Franklin County, Ohio. Of these men, 1883 (56%) were white, non-Hispanic, 1289 (38%) were Black, non-Hispanic, and 111 (3%) were Hispanic (ODH, 2010). The sample recruited for this study was comprised of 145 males, of which 73 (50%) identified their race as white and 65 (45%) as Black or African American. Additionally, 7 (5%) of the participants identified their ethnicity as Hispanic. While problems resulting from the fact that participants were volunteers remain, the argument for generalizability was strengthened by the diversity observed in the sample.

Limitations to Theory

The study of the effect of serostatus disclosure on HIV transmission risk in MSM is relatively new. The availability of measures to accurately investigate this phenomenon is limited. While the contribution of the study is strengthened by this fact, it also represents a challenge to the investigation of construct validity. Because relationships between serostatus disclosure, HIV transmission risk, and other variables are not yet well-established, tests of the hypothesized relationships between these variables may have unexpected but important results.

CHAPTER 2: REVIEW OF THE LITERATURE

The impact of HIV and its transmission is demonstrated by the large amount of related research. The Centers for Disease Control (CDC) has approved a variety of interventions shown to curb the spread of HIV (Centers for Disease Control, 2012a). However, the ability to gage the relative effectiveness of these interventions relies, in part, on the ability to successfully measure reduced risks of transmission. A review of the relevant literature was conducted to discover the nature of the measures currently being used to assess HIV transmission risk and the adequacy of those measures to evaluate intervention effectiveness. The goal of the first section of this review was to provide a brief description of the history of HIV and HIV prevention efforts in the United States. Second, a brief review of the importance of serostatus disclosure to HIV prevention efforts is undertaken. In this section, the primary focus is on the potential impact of disclosure on HIV transmission risk. In the third section, the review seeks to briefly describe and review examples of the basic types of sexual risk behavior measures which are currently being used by professionals in the study of sexual behavior and the challenges associated with the use of the measures, particularly in their use as measures of HIV prevention efforts. This section also describes briefly the types of statistical analyses commonly used to analyze data on sexual risk behavior. In the final section, the topic of validity is discussed. Emphasis in this section is placed on the meaning of validity and the analytical techniques used to explore it.

History of HIV and HIV Prevention in the United States

It was during the 1980's that the effects of HIV were first observed in the United States. Early in that decade, unusual cases of cancer (Kaposi's Sarcoma), and pneumonia (Pneumocystis carinii pneumonia; PCP) were being reported in populations of gay men in New York and California. The Centers for Disease Control (CDC) reported in July of 1981 that there was apparently no danger of infection to persons who were not homosexual, but by the end of that year, the spread of HIV to other populations had been observed. During the early years of the epidemic, the gradual spread of the disease to intravenous drug users and persons receiving blood transfusions suggested that the disease was related to an infectious agent which could be transmitted through non-sexual means. It was not until 1983, the year that the HIV virus was first identified in France, that the CDC began formal prevention efforts (AVERT, 2011). It was not until a few years later that the CDC issued the first prevention guidelines to US health care workers, requiring the initiation of precautions similar to those used to prevent the transmission of Hepatitis B (CDC, 1988).

The first blood test to detect the HIV virus was licensed by the U.S. Food and Drug Administration in 1985. At the time, rumors about how the virus was spread were widespread. Suspecting sexual transmission among gay men, the mayor of San Francisco had closed all of the bathhouses and private sex clubs in the city. Concurrently, efforts to safeguard the U.S. blood supply through testing of donated blood were also identified as a priority. During this year, the first small-scale HIV prevention efforts based on needle-exchange were reported in Amsterdam, Netherlands. In 1986, the Surgeon General of the United States, C. Everett Koop, made the first major announcement about HIV prevention. In the statement, Koop suggested that education was the only effective way to prevent transmission. He called for the initiation of AIDS education at the earliest possible age in elementary school and called out to parents to communicate frankly with their children about the disease and about risks for transmission. Koop also attempted to dispel the notion that HIV could be spread through casual contact and formally announced that the disease was primarily transmitted through sexual contact, both homosexual and heterosexual (AVERT, 2011). Specific strategies for risk prevention at the time were behavioral: abstinence, HIV testing, condom use, and avoidance of male-to-male rectal intercourse, sexual intercourse with sex workers, and shared needle use (Boffey, 1986).

Early on in the epidemic, strategies for the prevention of HIV transmission through sexual behavior focused on changing the sexual behaviors of those at risk of contracting the virus. Beginning in 1982, programs to increase awareness of HIV, reduce misinformation and stigma, and initiate risk reduction were undertaken. These early efforts were primarily targeted for delivery among communities of homosexual men. However, by late in the decade, prevention efforts had been extended to other populations. Additional populations initially identified as being at high risk of contracting HIV included high-school and college-aged youth, racial and ethnic minorities, and health-care workers (CDC, 2011d).

During the 1980s and 1990s, a variety of behavior interventions operating at the individual, small group, community, structural, and medical/technical levels were found to be effective in the fight against the spread of HIV. While initial efforts at targeted education were moderately successful, it was not long before more focused, program-based efforts toward reduction were underway. To date, the evaluation of the effectiveness of these interventions has been accomplished largely through quasi-experimental study designs. For example, the AIDS Community Demonstration Projects (ACDP) was the first effort at coordinated community-level intervention. The ACDP sought to reduce sexual risk behavior by utilizing peer networks and small media to spread risk-reduction messages in high risk communities. Implementation of ACDP was part of a larger CDC-coordinated study of the intervention's effectiveness. The parent study was conducted by the CDC in five US cities between 1989 and 1994 and focused on increasing condom use among intravenous drug users, sex workers, at-risk youth, and men who have sex with men (MSM). Key outcomes for the ACDP included intention to use condoms with main and non-main sex partners, condom carrying, and bleach use; the last-named was used as a measure to prevent transmission among intravenous drug users. Given that ACDP was based on the stages-of-change model (Prochaska & DiClemente, 1992), significant increases in stage-of-change scores (i.e., precontemplation, contemplation, preparation, action, and maintenance) were also employed as a measure of intervention effectiveness (CDC, 2011e).

The ACDP is only one of myriad studies which have evaluated intervention effectiveness. Meta-analytic reviews by Herbst and colleagues (2005), Mullen, Ramirez, Strouse, Hedges, and Sogolow (2002), Neumann and colleagues (2002), Noar (2008), and Semaan and colleagues (2002), have demonstrated that behavioral intervention is an effective way of reducing HIV risk in a variety of high-risk populations including MSM, sexually active adolescents, drug users, and sexually active heterosexual adults. A review of these meta-analyses provides valuable insight into the evaluation of sexual risk behaviors associated with the transmission of HIV. Historically, researchers have chosen a variety of intervention outcomes measures including number of partners (i.e., total, casual), condom use (i.e., % of times), percent of encounters involving unprotected sex (i.e., oral, receptive anal, insertive anal), mean number of unprotected sexual acts, unprotected sex outside of a monogamous relationship, unprotected sex outside of a seroconcordant relationship, incidence of sexually transmitted diseases, and condom use at last intercourse.

With the introduction of protease inhibitors in the late 1990's, health outcomes of people living with HIV (PLWH) improved dramatically. As a result improved treatments, the mortality rates of PLWH decreased. With larger numbers of people living with HIV infection, efforts to prevent HIV transmission turned toward PLWH, or *prevention for positives* (Janssen & Validserri, 2004). In 2003, the CDC launched the Advancing HIV Prevention (AHP) which emphasized early diagnosis and access to medical care and HIV prevention for PLWH. While some research has indicated that PLWH engage in fewer risky behaviors immediately after diagnosis, others have found that the tendency toward risky behavior may re-emerge. As a part of the AHP strategy, efforts were made to develop interventions which were suitable to assist PLWH in maintaining greater physical, psychosocial, and sexual health. Along with prevention case management delivered by AIDS service organizations (ASO), a variety of evidence-based interventions focusing on reducing risk behaviors in PLWH were developed and disseminated. One of these interventions focused on disclosure of serostatus to family, friends, and sexual partners. Healthy Relationships (Kalichman et al., 2001) is a group-based intervention which focuses on skill development, self-efficacy, and positive expectation. Decision-making about disclosure of serostatus is one of the skills targeted by the intervention (CDC, 2011f; Kalichman et al., 2001).

The most recent efforts in HIV prevention involve the combination of both biomedical and behavioral approaches (Merson, O'Malley, Serwadda, & Apisuk, 2008). New and more effective treatments for PLWH not only dramatically increase their quality of life, but result in significantly lower viral loads, and in turn, significantly lower risks of transmitting the virus. Additionally, developments in pre-exposure prophylaxix (PrEP) with antiretroviral drugs represent a promising additional avenue for prevention. In a recent study of seronegative MSM, more than 40% expressed an interest in the use of PrEP. However, it is important to note that interest in using PrEP was not related to sexual risk behavior (Barash & Golden, 2010).

HIV in Men Who Have Sex with Men (MSM)

The first cases of HIV reported in the United States were among men who have sex with men. Since those first cases were reported, more than 1.7 million people in the United States have been diagnosed with HIV, and more than 500,000 have died from AIDS-related causes (Kaiser Family Foundation, 2009). More than one million people are currently living with HIV, and it estimated that over 20% of those persons are undiagnosed. Among those affected by the disease, the largest share is comprised of gay and bisexual men who have sex with men. While MSM account for approximately 2% of the population, they account for over 50% of PLWH. In the United States, MSM have consistently represented the largest portion of new infections, the greatest percentage of those living with HIV, and the greatest number of those dying after being diagnosed with AIDS (CDC, 2011g).

The primary means of transmitting and contracting HIV among MSM is through unprotected anal intercourse. In a recent study of sexual behavior in PLWH, 32% of men and 39% of women reported having unprotected sexual intercourse with a partner who was either seronegative, or whose serostatus was unknown (Kalichman et al., 2011). The greater risks of contracting HIV faced by MSM are due in part to the larger number of men living with the disease. With each sexual encounter, seronegative MSM face a greater risk of having sex with a person who is already HIV positive. This is complicated by the fact that many MSM are not aware that they are infected (CDC, 2011h). The risk of transmission of HIV in MSM is also higher because MSM are more likely to engage in anal intercourse; unprotected anal intercourse is the sexual activity with the highest risk of transmission.

Increased risk of infection among MSM is also associated with a variety of other factors including drug and alcohol use, lack of knowledge of HIV risks, multiple sexual partners, and complacency (CDC, 2011h; Kalichman et al., 2011). While these risk factors are not unique to MSM, when combined with a higher incidence of anal intercourse, they may contribute to the greater transmission risk in this population. Sexual positioning also has an effect on the risk of transmission. Among MSM, when a negative partner *bottoms* (receives anal sex) with a positive insertive partner, the risk of transmission is the highest (Fox & Fidler, 2010). Though positional preferences have been reported among MSM, a recent study by Wei and Raymond (2011) suggests that men's positional preferences are mutable.

Risk Reduction in MSM

Efforts to reduce the risk of HIV transmission among MSM have typically focused on the use of condoms during intercourse. In a recent study of 1199 MSM, consistent condom use was found to be the most frequently used risk reduction strategy (Wei et al., 2011). Despite the effectiveness of condoms in preventing HIV transmission, some MSM demonstrate a resistance to their use. This resistance can persist despite a positive HIV diagnosis. In a study of 28 individuals with acute or early stage HIV infection, participants showed a significant post-diagnosis drop in the number of sexual partners and in the frequency of sex with seronegative partners. However, the use of condoms did not significantly increase (Steward et al., 2009).

While condom use is possibly the most widely recognized risk reduction strategy, non-condom based strategies are also evident. Seroadaptive strategies based on the deliberate selection of partners, positions, or practices as a means of reducing the risk of HIV transmission are widely used (McConnell, Bragg, Shiboski, & Grant, 2007). More specifically, serosorting among MSM refers to the restriction of unprotected anal sex to partners who are seroconcordant (Zablotska et al., 2009). Serosorting can be practiced as a risk reduction strategy by any person regardless of serostatus. For those who are HIV negative, unprotected sex is restricted to others who are not infected; seropositive persons may agree to untprotected sex only with other positive partners. Significant reduction in the risk of transmission is possible with serosorting (Marks et al., 2010; Philip et al., 2010). However, other research has found that flaws in seroadaptive strategies may actually lead to increased risk of transmission (Eaton, Kalichman, & Cherry, 2011). For example, Wilson et al. (2010) found that serosorting, while effective in reducing the absolute risk of transmission, is likely to increase the relative risk of acquiring HIV among serosorting MSM in populations with a relatively high prevalence rate of undiagnosed HIV infection.

Along with serosorting, a variety of other seroadaptive strategies have been reported in MSM. Seropositioning refers to the practice of strategically positioning discordant partners in the sexual act based on serostatus. The risk of HIV transmission is reduced if the insertive partner is negative. A reduction in the risk of transmission can also be achieved through early withdrawal. Negotiated safety refers to the practice among seronegative men of having unprotected sex with a main partner who is known to be seroconcordant and using condoms in 100% of all sexual encounters outside of that relationship. Men may also choose to reduce risk by engaging in sexual activities other than anal intercourse. Examples of these activities include rimming, mutual masturbation, oral sex, digital penetration, or the use of sex toys (Reisner, Mimiaga, Skeer, & Mayer, 2009).

The use of serosorting and other seroadaptive risk reduction strategies in populations of MSM has been widely documented (Eaton, Kalichman, O'Connell, & Karchner, 2009; Eaton, West, Kenny, & Kalichman, 2009; Wei & Raymond, 2011, Wei et al., 2011). However, great concern about the effectiveness of serosorting in preventing HIV transmission has also been noted (Eaton et al., 2009a; Sheon & Lee, 2009; Zablotska et al., 2009). In order to limit unprotected sex to seroconcordant partners, seronegative MSM must believe that their potential partners are disclosing truthfully and that they have been tested recently enough to ensure that they are seronegative. This is complicated by the fact that during the acute stage of infection when the seropositive individual is most infectious, the standard enzyme immunoassay HIV tests cannot detect the HIV virus (Eaton et al., 2009a). Seropositive men also face risks associated with the strategy. Engaging in unprotected intercourse with another positive person does increase the risk of reinfection with a second, and additional, variant of the HIV virus (known as *superinfection*; Marcus, McConnell, & Grant, 2011; Smith, Wong, Daar, Richman, & Little, 2004). Additionally, unprotected sex with any partner can lead to the transmission of other sexually transmitted infections which can accelerate the progression of HIV (Eaton et al., 2009a).

The Importance of Serostatus Disclosure

The disclosure of serostatus to sex partners is an important issue for PLWH. The passage of the Ryan White Act in 1990 tied the receipt of federal funds to the ability to prosecute individuals for criminal nondisclosure. To date, twenty-four states have passed laws which specifically criminalize nondisclosure of serostatus in particular situations.

The nature of these laws varies from state to state as do the penalties for nondisclosure. Despite this variability, awareness of the existence of HIV-specific laws and understanding of their provisions has been reported at fairly high levels among PLWH. This may be due to the work done my AIDS service organizations, HIV/AIDS support groups, and reading materials targeted to seropositive individuals (Galletly, DiFranceisco, & Pinkerton, 2008). It is important to note that despite the serious potential consequences of nondisclosure to sexual partners, current research shows a significant proportion of PLWH report who did not disclose their serostatus to sexual partners (Allen et al., 2008; Ciccarone et al., 2003; Gaskins, Foster, Sowell, Lewis, & Parton, 2011; Holt et al., 2011).

Despite the potential importance of disclosure to sexual health, many MSM do not disclose their serostatus to their partners. In a study of 1,828 HIV-positive MSM, Klitzman et al. (2007) found that only 41.8% of participants reported disclosing their serostatus to all of their casual sexual partners and 21.5% reported disclosing to none of these partners. More than 36.5% of these men reported engaging in unprotected anal intercourse with a partner whose serostatus was either negative or unknown. Two recent studies suggest nondisclosure may be related to the disproportionate increase in the rate of HIV transmission among ethnic minorities. In a study of African American MSM, Tieu and colleagues (2011) observed a high rate of same race/ethnicity partnerships (91.6%) among participants, a high prevalence of alcohol (47.3%) and drug use (38.7%) during sex, and high rates of nondisclosure (31.2-42.8%). A large percentage of participants (27.2%) also reported unprotected anal intercourse with a serodiscordant partner. Bird, Fingerhut, and McKirnan (2011) found that African American MSM were less likely than White men to disclose their serostatus to their partners; however, those who did disclose to seronegative or sero-unknown partners were less likely than White men to engage in unprotected anal intercourse with their serodiscordant partners. Other factors found to be associated with disclosure to sexual partners include HIV stigma, partner type, type of sexual activity, substance use, perceived responsibility to disclose, depression, social support, and Internet use (Gore-Felton et al., 2008; Holt et al., 2011; McCready & Halkitis, 2008; Niccolai, D'Entremont, & Pritchett, 2006; Poppen, Reisen, Zea, Bianchi, & Echverry, 2005; Serovich et al., 2009; Tieu et al., 2011)

Studies of serostatus disclosure demonstrate that the reasons for disclosure or nondisclosure of serostatus to sexual partners vary widely among PLWH (Derlega, Winstead, Greene, Serovich, & Elwood, 2004; Gaskins et al., 2011; Serovich & Mosack, 2003). Among the reasons given for disclosure, PLWH may include a sense of responsibility to the partner, a desire to instruct or educate, a desire to test the effect of disclosure on the relationship, a desire for a relationship, or a desire for emotional release (Driskell, Salomon, Mayer, Capistrant, & Safren, 2008; Gaskins et al, 2011; Serovich & Mosack, 2003). Reasons given for nondisclosure may include internalized shame/stigma, fear of a negative reaction, fear that the person will tell others, rejection/possible missed opportunity for sex, assumed partner positive serostatus, or a desire for privacy (Driskell et al., 2008; Gaskins et al., 2011; Serovich and Mosack, 2003). Disclosure decisions also vary based on the type of partner and the nature of the relationship (McCready & Halkitis, 2008). In a study of HIV positive MSM, Serovich & Mosack (2003) found that the highest scoring reasons for nondisclosure were, “we didn’t know each other well,” “our relationship was pretty casual,” and, “we weren’t very close to one another.” Gaskin and colleagues (2011) reported that MSM may identify the need to disclose to sexual partners but still not report disclosing to anonymous casual partners. In a qualitative study of disclosure among MSM in Seattle, Washington and Los Angeles, California, Gorbach et al. (2004) found that some men regard their seropositive status as private and personal information and do not feel the need to disclose their serostatus. Other reasons for nondisclosure given by the participants in the study included denial of serostatus, type of sex, and low viral load.

The effectiveness of serostatus disclosure as a means of reducing the risk of HIV transmission is still under study. In a study of methamphetamine-using MSM, McCready and Halkitis (2008) reported that disclosers were less likely to exchange semen in a sexual encounter than nondisclosers. In a test of the effectiveness of an intervention to assist MSM in disclosing serostatus to their casual partners, Serovich and colleagues (2009) showed a reduction in risk behavior among HIV positive MSM in the intervention group. In a study of the risk of HIV transmission, Pinkerton and Galletly (2007) found that serostatus disclosure reduced the risk of HIV transmission by 18% or more across a range of sexual risk conditions. Other studies have found associations between disclosure and unprotected sex (Holt et al., 2011). However, in these studies it is often unclear as to whether unprotected sex took place after seroadaptation. Clearly, serostatus disclosure makes possible a variety of risk-reduction strategies that might otherwise be difficult. If serosorting and other seroadaptive strategies are to be effectively used, then the serostatus of both partners must be openly and accurately disclosed (Eaton et al., 2009a; Eaton et al., 2009b; Wei et al., 2011). The importance of full and accurate disclosure does present a challenge to persons choosing to use these strategies. Those who are living with HIV must face the challenges associated with disclosure of their positive status. For those who are negative, seroadaptation might not be particularly effective. While a seropositive status lasts a lifetime, a seronegative status must be continually evaluated. This situation is further complicated by the fact that, during its acute phase, the virus may not be detectable by standard tests (Pilcher et al., 2005). Thus, for MSM who are sexually active, with multiple partners, and who engage in high risk behavior, it might be impossible to know with certainty that they and their partners are seronegative prior to any specific sexual encounter. Consequently, a high degree of skepticism about partners' disclosures of a negative serostatus has been observed in MSM (Sheon & Lee, 2009). Other researchers have underscored the importance of measurement in the study of disclosure. Niccolai et al. (2006) found that measurements which permit examination of the timing of disclosure (i.e., before, during or after sexual intercourse) are required if the effect of disclosure on HIV transmission is to be accurately estimated.

The Need for Valid Measures of Sexual Risk Behavior

A review of several articles published between 1990 and 2003 which focus on discussion of the current "state of the art" in behavioral risk measurement is the starting point for this brief review of the measurement of sexual risk behavior. These articles review a variety of issues in measurement, from validity and reliability to scaling and analysis. For each article, those issues identified will be summarized and discussed in connection to the proposed study. Additionally, the recommendations for improvements to measurement made by the authors will be summarized. At the conclusion of this section, a summary of the evolution of sexual risk behavior measurement, as well as a discussion of the remaining challenges relevant to the proposed study, will be undertaken.

In 1990, Catania, Gibson, Chitwood, and Coates described a series of methodological problems which affected the study of AIDS-related sexual behaviors. At the time, the extant literature on the topics of AIDS, HIV, and the associated risk behaviors was rapidly expanding. The authors of the article expressed a deep concern for the observed weaknesses in the reliability and validity of the measures being used to assess sexual risk behaviors and the role of sexual behavior in the spread of HIV, beginning with the lack of consensus concerning what activities constitute risky sexual behavior. While the focus of the literature at the time was placed on penetrative vaginal and anal intercourse, this was at least in part due to the fact that the risks associated with oral sex, oral-anal sex, and all forms of protected intercourse where condoms were used were not well-known. Thus, the authors called for additional study of these behaviors in order to improve understanding of the connections between sexual behavior and HIV transmission. The authors included an extensive discussion of potential measurement errors arising from respondent, instrument, interviewer, and mode of delivery effects. After a review of the literature from that time period, the authors provided an overview of the information currently available on these types of measurement errors and called for the replication of studies across groups based on gender, age, and ethnic risk group, to pave the way toward the adoption of a battery of sexual risk behavior measures with strong psychometric support. While many have answered the call for improvements in the measurement of risk behavior, in many ways the challenges observed in the 1990s have not yet been met.

In the same year, Leigh and Stall (1993) published an article which discussed the importance of measurement to the formation of conclusions as a result of HIV-related research, specifically with respect to the connections being studied between alcohol and drug use and sexual risk behavior. At the time, substance use was viewed as an important, modifiable risk factor which could be targeted to reduce the risk of HIV transmission. The authors explored the evidence available concerning the connections between substance use and high risk sexual behavior, and the limitations in the research designs, measurement methods, and interpretation of results which they had observed during a review of the literature. The discussion of study design described three basic approaches to data collection at the global, situational, and event levels. Studies employing global measures sought to establish general tendencies toward higher risk sexual behavior based on increased substance use. Data in such studies described the frequency of high risk sexual behaviors and the frequency of substance use. The inability to connect specific instances of risky behavior and substance use limit the ability of such designs to provide evidence of causality. In the more specific situational studies, participants indicated the frequency with of high risk sexual behavior while using substances or responded to dichotomous measures indicating the use of substances during sex. While these studies permit the study of the co-occurrence of substance use and high risk behavior, they often did not reflect the absolute risk associated with the overall level of sexual activity of the participant. Finally, studies associated with event-level data were also discussed. In such studies, data are collected on specific sexual encounters or episodes and reflect the co-occurrence of substance use and sexual risk behavior as well the frequency of co-occurrence for each participant relative to overall sexual activity. The strength of this approach to data collection and analysis is that it permits the evaluation of within-person differences which hold constant other person-related characteristics. This approach was considered optimal to the study of the relationship between potential causal risk factors such as substance use and HIV sexual risk behavior. While the importance of within-subjects analysis of behavior is underscored by the authors, current research on HIV transmission risk still includes many studies which employ designs which make this analysis impossible. With respect to limitations concerning measurement, Leigh and Stall (1993) revealed that the predominance of the literature at the time involved a simplistic, single item-based approach to sexual risk behavior measurement. For example, the authors cited a study by McCusker, Stoddard, Zapka, Zorn, and Mayer (1989) which sought to evaluate the prediction of preventive-behavior in homosexual men. In the study, the authors utilized two different dichotomous behavioral outcome measures to quantify risk behavior. The first measure indicated whether or not the participant was involved in a single monogamous relationship and the second indicated whether or not the participant avoided unprotected anal intercourse (either insertive or receptive). Each of the dichotomous outcomes was used as the dependent variable in a multivariate logistic regression. While the limitations of this approach to the study of complex risk behaviors are significant, Leigh and Stall (1993) actually decried the use of this type of strategy for measurement based on the inconsistent definition of the dichotomous measures.

One such alternative approach mentioned in the article by Leigh and Stall involved the establishment of a risk index (Biglan et al., 1990). In this study, a single scale of risk behavior was derived from a series of individual risk activity-based items involving the frequency of unprotected anal intercourse, the number of sexual partners, and condom use. In this approach, participant risk was operationalized as the mean of the participant's standard score on 6 different items. The use of the single risk index permitted analysis of the predictors of sexual risk using hierarchical multiple regression. The strengths of this approach were the ability to incorporate a variety of risk behaviors in a single analysis and the ability to include partner characteristics (e.g., intravenous drug use, familiarity) as part of the estimation or risk. However, the study provided no assessment of the validity or reliability of the scaling approach.

More than 10 years later, Schroder, Cary, and Vanable (2003) described a variety of methodological issues in the current research on HIV-related sexual risk behavior. Despite years of research on the sexual transmission of the HIV virus, many of the issues raised in the articles echoed those raised much earlier in the epidemic. In the first of three related articles, the authors focused attention on the observed inconsistencies in the research with respect to item content, risk scaling, and analysis. The authors identified two general approaches to the measurement of risk which were evident in the research at the time. The first approach to measurement involves estimates of absolute risk arising from the number of occurrences of various types of sexual risk behaviors. The second measurement approach employs relative frequency measures of sexual risk behaviors. The article by Schroder et al. (2003) included an extensive comparison of the relative strengths and weaknesses of the two approaches. In the first approach, data are collected as event frequencies on a ratio scale. Such an approach provides the greatest potential measurement precision and the ability to measure at the event-level. However, the analysis of such data can be challenging due to non-normality in the variable distribution. In the latter approach, data are collected in a variety of ways including percentages, categorical Likert-type responses ranging from "none of the time" to "all of the time," and dichotomous responses scales such as "always" or "never" in connection with discrete sexual risk behaviors. The relative frequency approach results in measures which are less precise than those associated with absolute frequencies and also result in distributions which are non-normal. In a review of 116 contemporary articles, the authors noted that more than half (n = 74) involved the use of relative risk measures. The authors also observed a trend toward the use of counts (i.e., absolute risks). Reasons given for the trend included an increase in the use of mathematical models of HIV transmission risk reduction to evaluate the effectiveness of HIV prevention interventions (Pinkerton & Abramson, 1993; Pinkerton & Abramson, 1994; Susser, Desvarieux, & Wittkowski, 1998). Such models necessitate the use of absolute measures of risk frequency. However, several other approaches to the analysis of non-normal, absolute risk measures are discussed as alternatives to the mathematical approach.

In the discussion of available analytical methods, Schroder et al. (2003) first focus on the use of linear, parametric models of sexual risk behavior. Such models, including variations of the ANOVA and ANCOVA models, require that participant scores on measures of sexual risk behavior be normally distributed. This assumption is rarely met with absolute risk scores. As a result, use of these models often requires that outcome variables be transformed with one of a variety of functions to achieve normality, and special consideration of outliers. Transformations recommended by the authors range from relatively straight-forward logarithmic transformations to iterative methods such as the Box-Cox transformation. In any case, the authors recommend that such transformations be followed up with normality tests to ensure their effectiveness. In the case of outliers, the authors prescribe a variety of methods for dealing with both univariate and multivariate outliers which may also lead to violations of model assumptions. In behavioral counts, it is quite common to observe large percentages of zero counts and a few extremely high scores. While the authors do not recommend exclusion of outliers as a remedy for non-normality, their potential effect on the results of parametric analyses is acknowledged.

Though Schroder et al. (2003) strongly recommend the use of absolute risk measures obtained on the ratio scale as measures of HIV risk behavior, a discussion of the use of categorical measures of risk, including ordinal measures, is also provided. With respect to the use of proportions or percentage ratings as outcome measures of relative behavioral risk, the authors caution that these measures often result in non-normal distributions and should typically be regarded as ordinal in nature. Categorical measures such as Likert-type rating scales and dichotomous measures are acknowledged as viable alternatives to behavioral counts but are discounted by Schroder and colleagues because of their lack of precision. For example, they argue that dichotomous measures of sexual behavior such as condom use must be regarded as relative measures because they, "provide information only about condom use relative to the total number of intercourse occasions." (Schroder et al., 2003, p. 77) However, this criticism does not take into account that condom use, even within the context of only one sexual encounter, is a relative measure. When participants respond to a behavior count item, such as the number of times condoms were used during a given period, it cannot be clearly determined whether or not they are referring to specific sexual encounters or specific acts of intercourse. Additionally, depending on the nature of their behavior and their ability to remember, participants may be providing information which is more reasonability regarded as ordinal. Unfortunately, the authors do not address the issue of comparative validity or reliability of the two approaches nor do they address the use of multiple-item scales comprised of either absolute risk or relative risk measures. While it is clear that increased precision will result from the use of a single absolute risk item when compared to a single item measured as a relative risk, no comparison is made to the reliability or precision resulting from the use of multiple item scales.

In a response to the debate regarding measurement, Patterson and Strathdee (2005) introduced an issue of the journal *Annals of Behavioral Medicine* with a call for methodological vigor and inquiry. In their article, the authors stressed the complexity of measuring the actual risk of HIV transmission. Myriad factors, including condom use, relative frequencies of both high and low risk sexual activity, stage of illness and viral load, and the presence of ulcerative sexually transmitted infections, affect the absolute risk of transmission. This complexity presents measurement challenges beyond the distinction between absolute and relative risk. In a response to Schroder et al., (2003, 2005), Catania et al., (2005) raised the issue of precision in a rather different light. The authors argued that the degree of precision necessary for any analysis is based on the research question at hand and that the costs associated with the collection of unnecessarily precise data cannot be justified. These authors suggest that currently the specific level of measurement precision necessary to assess the levels of HIV transmission which are acceptable from a public health perspective or to assess the effectiveness of interventions designed to achieve those levels have not been empirically determined. For a variety of reasons, the actual increase in precision achieved by adopting absolute risk measures remains in doubt. In light of the increased burden to the participant, the increased difficulty in estimating behavioral frequencies, and the increased challenges of accurately defining terms like *sexual encounter,* data are not available to suggest at what point the measurement of behavioral frequencies might actually reduce measurement precision. Alternatively, the authors argue that categorical measures of risk may provide adequate precision for the measurement of risk, and they provide the opportunity to characterize valid groupings of individuals based on broader risk categories. These groupings could be empirically based and their boundary values validated based on clinical and statistical meaningfulness. The authors further argue that non-linear models, such as generalized linear models (GLMs) and generalized linear mixed models (GLMMs) are available which can be effectively used to assess changes in sexual risk behavior and intervention effectiveness. Included with GLMs are the family of logistic regression models, both binary and ordinal, and zero-inflated models such as the zero-inflated Poisson model and the zero-inflated negative binomial model.

The Current Measurement and Analysis of Sexual Risk Behavior

The prior section of the literature review presented a brief overview of the expressed needs of the research community with respect to the measurement of sexual risk behavior. Though a consensus about the “gold standard” of measurement has not yet been reached, the need for additional research focusing on the usefulness and validity of a variety of approaches is certainly evident. In preparation for this study, a review of the current research on sexual risk behavior was undertaken to evaluate contemporary approaches to the measurement of HIV transmission risk. Articles published in selected peer-reviewed journals from January of 2006 through the June of 2011 were included in the review. Articles that described empirical studies involving risk behavior and that included direct behavioral measurements were selected from among the published issues. For the purposes of this review, articles involving measurement of intentions or attitudes but not actual sexual behaviors, those involving health promoting behaviors, and those measuring participants’ perceptions of risk rather than risky behaviors were excluded. Correlational and methodological articles dealing specifically with HIV transmission risk, or in the measurement of risk or disclosure, were also included in the review.

One of the most commonly used strategies to evaluate sexual risk involves the use of absolute frequencies. The absolute frequency measures typically employed in the study of sexual risk include the numbers of sexual partners, sexual acts, unprotected sexual acts, and unprotected anal intercourse. While these articles all yielded the more precise measures of risk behavior as defined by Schroder et al. (2003), the use of the data varied widely. In many cases, the non-normal distributions of these variables were accommodated through a process of dichotomous recoding. In one example, Huang, Jacobs, and Derevensky (2010) evaluated the prevalence of sexual risk-taking among college athletes. Sexual risk was conceptually defined as either engaging in unprotected sex or as engaging in sex with multiple partners in the prior 12 months. Operationally, the definition of sexual risk combined measures of both of these aspects of sexual behavior. Condom use was measured as a relative frequency, including the options, "I did not have sexual relations," "Used all the time," "Used most of the time," "Used some of the time," and "Never used a condom." The number of sexual partners was measured as an absolute frequency. Both outcome measures were recoded to dichotomous response for analysis. Unprotected sex was operationalized as the self-reporting of any sexual intercourse without a condom, and multiple sex partners was operationalized as self-reporting of two or more partners in the prior 12 months. Multivariate logistic regression models were then used to determine the predictors of unprotected sex and multiple sex partners. Though the data on sexual risk were gathered on a ratio scale based on absolute frequencies, the analysis conducted by the researchers did not utilize the more precise measure. Additionally, no evidence of the validity of the observed scores on the measure was provided.

In studies of HIV risk, particularly those involving persons with a positive serostatus, the use of absolute frequency data can be expanded to include a variety of other risk characteristics. The chosen risk characteristics appeared to vary significantly across studies. For example, partner serostatus is often incorporated into the definition of *risky sex* based on the assumption that only persons who are seronegative are at risk of contracting HIV. In such cases, a risky sexual encounter is defined as one in which the participant engages in unprotected intercourse with a serodiscordant partner, or with a partner of unknown serostatus (Morin et al., 2007; Myers, et al., 2010; Poppen et al., 2005). In studies involving MSM only, the definition of risky sex may be further restricted to unprotected intercourse (Carpenter, Stoner, Mikko, Dhanak, & Parsons, 2010; Wilton et al., 2009) or to sex with casual sex partners (Halkitis, Mukherjee, & Palamar, 2009).

Other authors chose to utilize non-parametric approaches to analysis. In the report of an evaluation of an HIV prevention intervention in bisexual African American men, Operario, Smith, Arnold, and Kegeles (2010) asked participants to report the number of female, male, and transsexual sex partners over the prior 3 months. Participants also reported the number of sex acts based on the type of sex (e.g., oral insertive, anal receptive). No evidence of validity related to the participants' responses was provided. In the analysis, the authors reported the significance of pre- and post-intervention changes in the number of sex partners and the number of acts using the Wilcoxon signed-rank test. The Wilcoxon test is appropriate for dependent group comparisons on variables measured on an interval or ratio scale. However, unlike the student's *t* test, the assumption of a normally distributed dependent variable does not apply. Each value of the dependent is assumed to be ordered on a continuum, but only the ability to make "greater than" or "less than" comparisons are assumed. Dichotomous measures of unprotected insertive or receptive anal intercourse by partner type were also analyzed for pre/post changes using the McNemar test. The McNemar test is employed when assessing two nominal variables in matched pairs of data.

In other articles, the authors chose to utilize absolute frequencies as a means of calculating the proportion of unprotected sex (NIMH, 2010; Amirkhanian et al., 2010). As an example, in an article from 2010, the results of the National Institutes of Mental Health (NIMH) Multisite HIV/STD Prevention Trial for African American Couples Group describe the primary outcome variable as the proportion of the participants' vaginal and anal intercourse which was protected by condom use. These data were collected using audio-computer assisted self-interviewing (ACASI), and the authors state the items used to gather information on sexual behavior had been used effectively in prior studies. However, no specific information on the validity of observed responses was provided. In order to determine the percentage, the researchers collected information from the participants regarding the total number of sexual encounters by type (e.g., anal, vaginal) and the total number of times that condoms were used during the same period. They then divided the total number of encounters into the total number of times condoms were used to compute and index of protected sex. A similar index was also computed for oral intercourse. In all cases, the study utilized a 90 day time period. The resulting proportions were used as outcome variables in group contrasts based on a variety of relationships characteristics including the gender of the partners and the relationship type (i.e., steady, casual). Analysis using the proportion of protected sex was limited to the use of the independent *t*-test. No mention was made of any transformations used to correct for non-normality.

A variety of other approaches to dealing with absolute frequency data were also observed. In another NIMH sponsored study authored by Steward and colleagues (2009), participants detailed their recent sexual histories through retrospective reporting of the counts of partners by serostatus and counts of sexual behaviors by type. The number of times condoms were used in conjunction with oral, vaginal, and anal sex acts was also collected. The authors reported the use of visual stimuli to aid participant recall, but no specific evidence of validity was provided. The resulting counts were standardized based on the number of days in the reporting period but not with respect to the variation in the behavior scores. Paired-sample *t* tests and were then used to assess the differences in sexual behavior before and after diagnosis with acute/early HIV infection.

In a study of the impact of group therapy on HIV sexual risk behaviors, Hien and colleagues (2009) collected data on the number of sexual encounters over the prior 30 day period using an ordinal measure of relative sexual activity. Data on sexual behavior were collected through one-on-one interviews. No specific evidence of the validity of participant responses was provided. Participants' sexual activity was rated on a scale of 1 to 60, with 1 representing 1 sexual encounter in the period and 60 representing 2 or more sexual encounters per day. Another relative measure, the proportion of times condoms were used, was ranked on a fractional scale of zero (never) to 1 (always). The two ordinal measures were then multiplied together, using one minus the fraction of condom use, to create a continuous variable describing the number of unprotected sexual encounters during the period. The resulting variable was highly skewed and zero-inflated. Therefore, the authors chose a zero-inflated negative-binomial mixed effects model to accommodate the deviation from normality.

Rosser et al. (2009) chose to employ a number of different techniques in the analysis as a way of dealing with the non-normality of the relative frequency data collected on sexual behavior. In a study of HIV sexual risk behavior of men who use the internet, participants were asked to report the number of partners met online, the number with whom they engaged in sexual activity, and the number with whom they engaged in unprotected sex. Data were collected using an online survey format. Though items used to gather information, and possible responses, were explained, no specific evidence of the validity of the observed responses was provided. The resulting distributions were reported to be severely skewed and zero-inflated. Three different techniques were used in the analysis. First, the outcomes were dichotomized to represent a risk or no risk condition. The dichotomous variable was then used as the dependent variable in a logistic regression procedure. The authors then constructed an ordinal risk variable and attempted to use ordinal regression procedures in the analysis. However, the authors reported that the proportional hazards assumption associated with the procedure was not met, and the results of the ordinal logistic regression were not reported. Details of the construction of the ordinal scale were also not provided. The third and final approach involved the use of a negative-binomial regression procedure.

Several approaches to the scaling of sexual risk were observed in the review of the literature. In a study by Mattson and colleagues (2010) a sexual propensity scale was constructed from a series of dichotomous and ordinal behavioral items. The resulting 18-item scale sought to measure the effect of circumcision on sexual risk behavior in Kenyan men aged 18-24. Sexual history data were gathered as absolute frequencies (e.g., number of encounters, condom use) provided for each sexual partner over the last 6 months up to a maximum of 12 partners. Absolute frequencies were then rescaled to dichotomous or polytomous categories depending the on the amount of observed variability in the data set. Item-response theory (IRT) was then utilized to combine item-level data into a latent trait scale measuring sexual risk propensity. Scale data were then subjected to a series of procedures to evaluate its psychometric properties including reliability, monotonicity, differential item functioning, and criterion validity. Results of these procedures confirmed the usefulness of the approach in the development of composite scales of sexual risk propensity. However, the scale was not designed to evaluate specific sexual behaviors within encounters, thus it included many items not relevant to specific sexual behaviors. Additionally, it was created for use with men who have sex with women. Fendrich and colleagues (2009) also used a Rasch-based scaling approach to develop a scale of sexual risk for use in adult MSM. The resulting multi-item scale was comprised of an array of sexual behaviors ranging oral sex with withdrawal prior to ejaculation to unprotected insertive anal intercourse during group sex. Rasch scaling was used to evaluate the validity of the resulting scale scores with favorable results. The validity of the resulting scale scores was explored through the examination of the item fit statistics, the meaningfulness of the resulting hierarchy of items, and a variety of other methods. This technique for scale development shows promise as the feasibility of the approach was adequately demonstrated. However, some limitations in the construction of items were noted. Primarily, the constructed scale did not incorporate the serostatus of the partner. As a result, scale scores did not reflect the lower risk associated with serosorting or seroadapative behaviors.

The use of ordinal measures of risk and multi-item based risk scales was rare, but several examples were found among those articles reviewed. Cases in which these models were employed varied based on the nature of the data structure. For example, a mixed model was employed by Zea, Reisen, Poppen, and Bianchi (2009) to assess the relationship between person and encounter characteristics and the likelihood of unprotected sex in immigrant Latino MSM. An ACASI instrument was used to gather the data on participant sexual behavior over the last 3 months, and during specific sexual encounter. An ordinal variable describing the relative frequency of unprotected sex over the last three months was constructed and used for as the outcome in an ordinal logistic regression of the frequency of unprotected sex on selected characteristics the partner and encounter. This variable operationalized the frequency of behavior on a scale from 0 to 4, with zero representing no unprotected sex acts over the prior three months and four representing more than 15 unprotected sexual encounters over the past 3 month period. While the authors used an ordinal outcome measure for a portion of the analysis, the reason given for the creation of the ordinal measure was to accommodate the skewness inherent in the count measure. No specific evidence of the validity of the count items used to create the ordinal score was provided. At the encounter level, hierarchical logistic regression was also conducted on the dichotomized outcome of unprotected sex (i.e., unprotected anal intercourse, no unprotected anal intercourse). In another example of the use of an ordinal behavioral measure, Tate, Singh, Ndubani, Kamwanga, and Buckner (2010) operationalized risk by creating an ordinal scale of the number of sexual partners in a 12 month period, with "low" describing a participant with no new partners and no multiple partners, "moderate" describing a participant with new or multiple partners in the past 12 months, and "high" describing a person with a new partner in the past 4 weeks. This ordinal measure was used in conjunction with several other dichotomous risk-related items including condom use at last sex and receiving money for sex to contrast the sexual behavior of the groups of interest in the study. Analyses in this study were limited to chi-square comparisons. Data were collected by personal interview. No specific evidence of the validity of participant responses was provided. Nappi et al.(2009) created a composite risk score to describe adolescent sexual risk behavior based on a scale of 0 to 5. On this scale, 0 was used to describe risk in participants who had never had sex and a 5 was used to describe risk in participants who had never engaged in protected sex. Data in this study were collected from adolescents using a structured interview delivered with computerized voice assistance. No specific evidence of the validity of participants' responses to the interview questions was supplied. In this study, no mention was made concerning the distribution of the ordinal variable, although it was used as the dependent variable in a linear regression analysis. In a study reporting the results of a cluster randomized controlled trial of an adolescent HIV prevention program, Chen and colleagues (2009) created a five-step scale of sexual behavior. On the scale, a score of "1" was assigned to participants who had never engaged in sex and reported no intentions to have sex, "2" was assigned to participants who had never had sex but reported the intention to engage in sexual behavior, "3" was assigned to those who reported a single sexual encounter in 6 months, "4" was assigned to sexually experienced youth who reported always using a condom, and "5" was assigned to youth who were sexually experienced and reported unprotected sex. The composite score was used as the repeated measure a study of risk prevalence at baseline, 6 month post-intervention, and 12 month post-intervention. The resulting score was used to detect subgroups among the participants with respect to the progression of sexual behavior across the study period. Employing a method proposed by Nagin (1999), the authors explored the resulting data to identify subgroups of participants based on their trajectory of sexual risk. Data for this study were collected using a paper-and-pencil questionnaire. No specific information regarding the validity of responses to the instrument was provided.

Two studies cited by Schroder and colleagues (2003, 2005) which employed mathematical modeling approaches to the quantification of transmission risk were also reviewed. In 1998, Susser, Desvarieux, and Wittkowski and colleagues published the description of the vaginal episode equivalent (VEE). This risk index was obtained from summing the episode counts of unprotected vaginal, anal, and oral episodes, with each type of sexual act multiplied by a weighting factor (i.e., vaginal weight = 1, anal weight = 2, oral weight = .1) associated with the relative risk of that behavior. In this paper, the authors also suggested an extension to the VEE scale called the Multivariate Ordinal Risk method (MOR). Using this method, the VEE can be extended to include any number of risk behaviors, each weighted by the relative degree of transmission risk. To compute the ordinal risk score, each individual in the study is ranked based on risk behavior and according to outcome (e.g., seroconversion; yes or no). The total risk is based on the average of the risk behaviors, with each risk behavior weighted by the average of its correlation with the seroconversion outcome. In this way, behaviors which distinguished persons in the sample who seroconverted are given more weight than those which did not result in seroconversion. While this approach to risk measurement has advantages, including its empirical development of risk based on actual data on seroconversion, such a measure it not useful in studies involving samples of persons who are all seropositive or where data on seroconversion are not available. The question of the validity of such an approach would be similar to the question faced by other researchers measuring HIV transmission risk or risky sexual behavior. Before applying a scaling method, the researcher would have to deal with demonstrating the validity of the participant responses used to construct the scale.

The work of Pinkerton and Abramson (1993, 1994) on the mathematical modeling of HIV transmission risk was also mentioned by Schroder (2005). In the 1993 article, entitled, "A Bernoulli Process Model of HIV Infection and Risk Reduction," the authors introduced a model of HIV risk transmission in order to estimate the infection risks associated with various sexual behaviors. The resulting model estimated the probability of contracting HIV, given *n* sexual acts, based on the prevalence of the infection in the population of potential sexual partners and the probability of infection for a specific sexual act. In the subsequent article, the authors extended the model to the computation of the reproductive rate of infection (R_0). This rate estimates the number of expected secondary infections for each primary infection. In both cases, the proposed models demonstrated the greater effectiveness of condom use in reducing HIV transmission rates than of a reduction of the number of sexual partners (i.e., monogamy; Pinkerton & Abramson, 1994). Perhaps the work most relevant to the proposed study is Pinkerton and Galletly's 2007 article on the reduction of HIV transmission risk based on serostatus disclosure. In this article, the authors proposed a model of transmission risk which is based on the probability of engaging in intercourse after disclosure, and the probability of condom use after disclosure relative to the probability that a condom is used in the absence of disclosure. The model predicts the greatest improvement in HIV transmission rates in cases where intercourse is refused after disclosure. Reductions in HIV transmission are also observed where disclosure results in increased condom use. Other work by Pearson and colleagues (2007) employed a modified version of the mathematical approaches proposed by Pinkerton and Abramson (1993, 1994) and Weinhardt, Forsyth, Carey, Jaworski, and Durant (1998) to study the risk of transmission among HIV-positive persons in Mozambique prior to initiating HIV treatment. The mathematical model was modified to incorporate several different risk factors including male circumcision, comorbid sexually transmitted diseases, and the stage of HIV infection. One advantage of the mathematical modeling approach to HIV transmission is that risk estimation equation can be modified to incorporate a variety of risk factors that are specific to the population under study. However, the ability to tailor the equation to specific contexts is limited by data availability, and on known estimates of risks associated with a variety of sexual behaviors.

Validity

The question of validity in measurement is a contentious one. The meaning of the word itself foreshadows the debate. Merriam-Webster (2012) defines the word valid as, “well-grounded or justifiable: being at once relevant and meaningful.” However, this definition is only one of those provided. The word valid is also defined as, “appropriate to the end in view.” Examination of the definitions of validity offered by prominent scholars in the field also suggests underlying tension among those in the field of psychometry. In 1949, Cronbach defined validity as, “the extent to which a test measures what it purports to measure,” and 5 years later, Anastasi (1954) offered, “Validity is what the test measures and how well it does so” (p. 120). Modern definitions of the term have moved beyond the simpler conception of validity as a characteristic of the instrument. In his seminal chapter on validity, Messick (1995) wrote, “Validity is an integrated evaluative judgment of the degree to which empirical evidence and theoretical rationales support the adequacy and appropriateness of inferences and actions based on test scores or other models of assessment” (p. 13). In a discussion of the played by values in the study of measurement, Messick suggested that, “because validity and values go hand in hand, the value implications of score interpretation should be explicitly addressed as part of the validation itself. It is the tension between an evidentiary approach to validation based in method, and a values-based approach to validation which considers the larger social consequences of measurement which define at least one important aspect of the ongoing controversy surrounding validation in the social sciences.

The ongoing argument about the meaning of validity and the appropriate approach to validation of psychological and educational tests is long-standing. In their foundational article of 1955, Cronbach and Meehl describe four newly improved categories of validity which had recently been recommended by the American Psychological Associate (APA). These categories included two types of criterion-related validity (i.e., predictive and concurrent), content validity, and construct validity. Prior to the publication of the revised APA Technical Recommendations for Psychological Tests in 1955, the discussion of validity had been generally limited to criterion validity (Gray, 1997). By definition, criterion validity refers to four subtypes described as predictive, concurrent, convergent, and discriminant validity (Trochim, 2006). Predictive validity refers to measure's ability to predict something that should be related to the construct being measured. For example, a measure of test anxiety should be predictive of test performance. Concurrent validity refers to the ability of the measure to distinguish between groups which exhibit different levels of the construct of interest. In the case our hypothetical measure of test anxiety, those who report experiencing test anxiety should in fact score higher on the measure than those who do not. Convergent validity refers to the relationship of the measure to other measures which, based on theory, should be similar. Scores on a measure of test anxiety should be similar to other measures of anxiety, particularly those which also purport to measure anxiety specific to examinations.

Conversely, divergent validity refers to the degree of dissimilarity between the measure and other measures which, based on theory, should be different or distinct. The degree of test anxiety observed should not be related to poor study habits or poor preparation.

In addition to criterion validity, Cronbach and Meehl (1955) described two new types of validity; content and construct validity were added to the list of requirements for educational and psychological tests. The term “content validity” has been defined as a process for comparison of the measure to the content domain for the construct being measured (Trochim, 2006). This definition first requires a complete description of the construct, a comprehensive identification of the boundaries of the content domain, and a representative sampling of items from that domain. In the hypothetical case of a measure of test anxiety, the question of content validity would first require the adoption of one or more theoretical perspectives on test anxiety (e.g., drive model, skills deficit model). Once the theoretical basis of the construct was defined, the domain of the construct would require definition. The definition of the construct domain is typically the result of some type of expert judgment. Allen and Yen (2002) describe the simple process of subjective judgment of item relevance, either by experts or the examinees themselves, as “face validity”, and reserve the name “logical validity” for a more rigorous approach based on the development of comprehensive test specifications to guide item development.

With the introduction of the concept of construct validity, Cronbach and Meehl (1955) made a significant contribution to the evolution of the practice of psychometrics. Much of what was introduced in 1955 remains part of the discussion of validity today.

The authors defined a construct as, " some attribute or quality which is not operationally defined" (p. 282). These qualities or attributes are thought to explain responses to the items on the instrument. The authors further suggested that the demonstration of construct validity was related to the difficulties of establishing criterion and content validity. Where the tester who is seeking to validate the instrument determines that the criteria available for comparison and the defined domain of the construct are not adequate or acceptable, the validation of the construct through indirect means must be undertaken. While the authors were careful to point out that construct validity is not based on the type of method employed, they did describes a series of methodologies which could be used to investigate it. First among these, examination of differences in scores for persons in groups which are known to differ on the construct was described. Basically, the degree to which group membership and test score are correlated reflects the degree of construct validity, with stronger the observed correlation providing stronger evidence of construct validity. Interestingly, the authors also point out that relationships which are too strong are problematic as they suggest a degree of insensitivity to the expected overlap in scores between the groups (Cronbach & Meehl, 1955). Correlation matrices were also seen as useful tools in connection with the investigation of construct validity. In an example of how this type of validity could be explored, the authors described a procedure whereby the relationship between scores were correlated with select variables, both external to the construct of interest and observable, which comprised a nomological net in the form of predicted positive and negative relationships. The nomological net was suggested as a way of connecting the theoretical definition of the construct to a scheme for observing it empirically (Trochim, 2006). Where variability in the scores on the construct of interest coincided with theoretically relevant observable variability in the nomological net, evidence of construct validity was obtained. In their original article, correlation matrices and factor analyses were suggested as the means for evaluating the nomological net. Shortly after, the multi-trait multi-method matrix (MTMM; Campbell & Fiske, 1959) was suggested as a method for construct validation of the nomological net. With the introduction of the MTMM, Campbell and Fiske introduced two new types of validity; "convergent validity" was described as the degree to which constructs which should be related according to theory were observed to be correlated, and "discriminate validity" was defined as the degree to which constructs which should not be related according to theory were observed to be uncorrelated. Investigators of construct validitywere instructed to return to theory to determine which of the relationships should be positive and which negative. An instrument was then to be considered valid to the extent that the expected relationships were observed.

Cronbach and Meehl (1955) also suggested that construct validity be demonstrated through the exploration of the internal structure of the instrument, and the change in scores over time. Typically described as "internal consistency reliability," the first recommendation involves the examination of item intercorrelations, and item-test correlations. The authors suggest that even low correlations among items can support construct validity. As was the case with known group differences, the authors also caution that very high item intercorrelations and item-test correlations might actually suggest poor construct validity. Score stability over time, usually evaluated as test-retest reliability either with or without intervention was also recommended as a potential source of evidence for construct validity. The authors recognized that the presence of stability in scores could argue both for and against validity, depending on the construct under examination. If the construct is thought to be volatile, stability could suggest insensitivity to changes in the construct.

Use of the classic approach to validation has endured despite arguments that several of its basic premises were flawed. In his discussion of the controversies surrounding validity, Gray (1997) focused on three of these premises. Classical descriptions of different types of validity, characterized by the terms “criterion validity,” “content validity”, and “construct validity” were the first to be questioned. Leading the attack on the so-called “Trinitarian” concept of validity, Guion (1980) argues that criterion-related validity and content validity are each special cases of the more broadly defined construct validity. He states, “both the kinds of evidence know as content validity and as criterion-related validity may contribute to evaluations of how well the operations represent the underlying concept, but they do so only insofar as they are special cases of construct validity” (p. 393). To Guion, criterion-related validity, as demonstrated by relationships between the scores obtained from the measurement of construct of interest and other standards for judging that construct, is a demonstration of the alignment between the observed scores and the underlying construct. Content validity was also seen as a special case of construct validity in that inferences based on observed scores require that connection between test performance and performance in non-testing conditions (i.e., the domain of the test and the performance domain) be demonstrated. Stated in another way, it must be demonstrated that variability due to measurement processes does not mask variability in the construct the scores purport to describe.

The second premise of the classical approach to validity that has been challenged is the role of values and consequences in the validation of measurement (Gray, 1997). Critics of the classical approach to validity argue that values play a critical role in the interpretation of test scores, and as such, should be included in the definition and demonstration of validity. In his influential chapter entitled Validity, included in the 3rd edition of Educational Measurement, Messick (1995) argued that the construct labels, theories underlying construct meaning, and broad social ideologies which are involved in the interpretation of scores are each value-laden, and as such, validation must include an exploration of the role played by values in score interpretation. Kane (2009) states, “The evolution of validity theory has involved the development of a range of statistical and analytic models, but more fundamentally, it has involved changes in the kinds of interpretations and uses to be considered in validation” (p. 43). The consideration of interpretations, uses, and consequences of scoring means that each validation exercise is different, and that the evidence that is required to support validity also changes based on the intended use of the scores. This shift in the focus of validation has resulted in a significant broadening of the types of evidence considered, as well as a proliferation of methodologies for demonstrating validity (Kane, 2009).

The third area of controversy described by Gray (1997) involves the subject of validation. Under the classical definition, validity was viewed as a property of the instrument. Since then, a shift has taken place, and the subject of validation has changed from the instrument to the scores obtained from its administration. This idea began as a rather concrete appreciation of the fact that the same instrument has different properties when administered in different populations. However, the notion that it is actually the scores resulting from measurement which are valid to some degree, arises from more than potential differences among populations of examinees. The same instrument administered to the same population but for different uses, or with differing underlying theoretical and philosophical approaches to interpretations, is subject to different validation processes. It has therefore been argued that it is the scores obtained from measurement, and their interpreted for a particular purpose, that should be the subject of validation (Wainer & Braun, 1988).

Among the modern approaches to validity which sought to address some of the shortcomings of classical approaches, the unified theory (Guion, 1980; Messick, 1989) is perhaps the best recognized. The popularity of this approach to validity is underscored by its adoption in the 1999 edition of the AERA/APA Standards for Educational and Psychological Testing and by the inclusion of Messick's explanation of unified theory in the 3rd edition of Educational Measurement. In the latter, Messick argues that the classical concept of distinct types of validity cannot be supported. He posits three important distinctions which characterize modern validity theory. First, Messick proffers that, though a broad variety of methods and evidences exist which are relevant to the question of validity, the concept of validity itself is a unitary one. Second, he states that validity is a property of the scores obtained from measurement, and the inferences made from those scores. Third, Messick states that the modern concept of validity is one which is evolutionary and based in the scientific method. As such, the question of validity is a relative one, based on existing evidence, which is continually evolving based on changes in personal and social values.

The modern concept of validity, specifically its unified characterization as construct validity, necessitated a reorganization of the classical validity types and the evidence used to support them. In the new conceptualization, each of the classical types of validity is recast, and connections between them and the overarching question of construct validation are made explicit. The classical concept of content validity is expanded to incorporate both the representativeness and relevance of the content of the instrument to the construct which is being evaluated. The key threats to construct validity which are posed by the content of the instrument are based on potential construct under-representation or construct-irrelevant variability. Construct under-representation results when important aspects of the construct being measured are omitted from the instrument. While Messick (1995) describes a small number of modern approaches to the development of items to represent the entire universe of the construct, and to selecting samples of items to comprise the domain of the measure, none of these methods differs significantly from the classical evidence of content validity. Sources of evidence which are cited as relevant to content representativeness include reviews of test specifications, comparison between the test domain and the domains of similar instruments, and expert judgment (AERA, 1999). The concepts of construct irrelevant variation, specifically the construct-irrelevant difficulty and construct-irrelevant easiness associated with item and response formats, do represent a change from classical approaches. Variability in scores which is due to variables other than the underlying construct being measured is considered evidence of poor construct validity. Construct-irrelevant difficulty results from variables not related to the construct that make the test more difficult, or the item more difficult to endorse, for some persons. Construct-irrelevant easiness refers to the a situation in which clues contained in items, or item formats, encourage correct responses or item endorsements which are unrelated to the underlying construct. Both types of construct-irrelevant variability can be conceptualized as contamination, or systematic error in measurement (Haladyna & Downing, 2004). Sources of evidence reflecting construct-irrelevant variability might include evidence of response biases (AERA, APA, & NCME, 1999), comparability of parallel test forms (Gallagher, Bennet, & Cahalan, 2000), and analyses of differential item functioning (Haladyna & Downing, 2004). The effect of response processes on scores can also be explored through the examination of relationships with measures of social desirability, think aloud protocols, response latencies, examination of the relationships between performance on different sections of the instrument, or differences in responses based on participant subgroup membership (AERA, APA, & NCME, 1999).

In the evaluation of content as a part of validation, Messick also suggests that item technical quality be included. The technical quality of the item/response refers generally to its clarity and suitability for use. Where items are worded ambiguously, or where response options are not clearly stated, an increase in the likelihood that responses may exhibit variability irrelevant to the construct being measured is likely. Sources of evidence regarding the technical quality of items and responses is generally obtained through expert evaluations of readability, item and response clarity, clarity of the instructions for item completion, and cognitive demand (Messick, 1995).

Messick's description of the unified theory of validity also involves a substantive component of construct validity. This aspect of validity is characterizes as the link between expert judgment of item content and observed response characteristics. The extent to which the theory underlying the construct explains which items perform as intended, an which do not, is the degree of substantive score validity. In order to investigate substantive validity, Messick suggests that the item pool be extended to include items based on competing theories, or items external to the construct of interest, and that evidence of discriminant and convergent validity be used to determine test content. Though the concept of substantive validity was discussed by Messick (1995) as a technique for instrument construction, a variety of evaluative methods, including MTMM, have been suggested as possible sources of substantive validity evidence. More recently, think aloud procedures have been suggested been utilized to evaluate if examinees are actually engaging in the intended responses processes (Gadermann, Guhn, & Zumbo, 2011; May & Warren, 2001).

Next, the modern conceptualization of validity looks to the internal structure of the instrument, and to the relationships between scores on the construct of interest and other variables for evidence of construct validity. With respect to the structural aspect of validity, the construct validity of scores is supported to the extent that the interrelationships among items, and the relationships between individual items and the total score, reflect the theorized structure of the underlying construct. For example, if the construct being measured is theorized to be unidimensional, evidence of strong item intercorrelations and item-total correlations would support score validity. Additionally, if the underlying construct is comprised of Guttman type items of increasing difficulty, then empirical evidence of the conformance of observed scores to this pattern provides evidence of construct validity (AERA, APA, & NCME, 1999). Evidence related to internal structure can take a variety of forms including correlational anaylysis and factor analysis (Eakman, & Eklund, 2011; Mutto, Lawoko, Ovuga, & Bangdiwala, 2010; Schell, 2012).

Among the aspects of validity described in Messick's chapter on unified validity, the generalizability of score interpretation across various groups and constructs plays an important role. Generalizability refers to the ability to apply research findings resulting from the study of a sample to the larger population from which the sample is drawn. The four classes of threats to external validity are the same as those described by Cook and Campbell (1979). First, statistical conclusion validity refers to the probability that the results of statistical hypothesis tests were negatively affected by sampling error. Internal validity is based on the strength of the evidence of the causal relationships between predictor and outcome variables, and the adequacy of controls over the effects of rival hypotheses. Construct validity in this context is referred to as the degree to which a causal relationship generalized across alternative measures of the cause and effect relationships. Finally, external validity is referred to as the degree of support for the interpretation of causal relationships across other groups of subjects and settings. As stated by Borsboom, Mellenbergh, and van Heerden (2004), the degree to which variation in the underlying construct causes variable in the score on the measure is a simple way of conceptualizing validity as a causal relationship. Thus, even where scores on a construct are not used for the purpose of making causal claims, the assumed causal relationship the underlying construct being measured and the score resulting from that measurement can be evaluated for genealizability. Instruction regarding acceptable evidence of generalizability is limited to meta-analytic techniques in the current APA standard (1999). In such studies, it can be possible to control for variation in context, and to estimate the degree to which the scores on the instrument have been replicated in different settings and populations.

Next, Messick (1995) describes the consequential aspect of score validity. This aspect of the validation processes refers to the implications of score interpretations, and their suitability for use. In order to fully evaluate if an instrument performs as intended, Messick argues that the appropriateness of what follows from its use must considered. It is not difficult to image both positive and adverse consequences from test-taking. Inextricably linked to values, the question of consequential validation provides evidence of whether the scores resulting from measurement should be used for a specific purpose, particularly in light of any adverse consequences that are observed. If scores resulting from measurement are shown to vary as a result of construct relevant sources, and adverse consequences of the use of the scores is observed, then the question of whether or not to use the scores is a policy question. However, if variability in scores is due to construct-irrelevant variability, then adverse consequences are a direct result of invalidity and as such become part of the validation process. For example, in the educational context, the use of high-stakes tests has ramifications for teachers and students. Test scores can result in curriculum changes and even school closures. Examination of the consequential aspects of score validity can include potential consequences, as well as consequences resulting from current use. Consequences from measurement can derive from the value implications of the construct or instrument labels used, or from the social consequences of their use (Hubley & Zumbo, 2011). Methods used to investigate the consequences of measurement may be qualitative or quantitative. In a study of the consequences of use of an evaluation of student teachers, researchers conducted interviews with both supervisors and teachers to discover the impact of evaluation on professional practice (Montecinos, Ritterhaussen, Cristina Solis, Contreras, & Contreras, 2010). Alternatively, in a study of the consequences of a variety of measures of oral fluency, researchers utilized methods related to measure sensitivity and selectivity (i.e., false positive, false negatives) to evaluate the consequences of score use (Valencia et al., 2010).

The unitarian view of validity is more comprehensive in scope than the prior trinitarian view. While frameworks of validation focused on content, construct, and criterion validity may seem more straightforward, they ignore the potential consequences of score use. Within the context of the current study, where the usefulness of the proposed measure of HIV transmission risk as a measure of intervention effectiveness is the subject of the investigation, the trinitarian view of validity does not provide an adequate framework for validation. The consequences of score use are inseparable from the investigation of their validity in this case. As such, the trinitarian concept of validity was observed to be better suited for this analysis.

Current APA Standards

In the current APA Standards for Educational and Psychological Testing (1999), the definition of validity is given as, "the degree to which evidence and theory support the interpretation of test scores entailed by proposed uses of tests" (p. 9). The standards further state that, "the process of validation involves accumulative evidence to provide a sound scientific basis," for the interpretation of scores resulting from test administration. The emphasis on the connection between the validity and scores, rather than validity and tests, as well as the inclusion of the proposed use of the scores in the definition, reflect the adoption of a unified concept of validity. Based on these definitions, the process of validating scores for use is based on the gathering of evidence and the construction of argument. Necessarily, the nature of that evidence and argument is specific not only to the scores, but to their intended use.

Summary

The goal of the first section of this review was to provide a brief description of the history of HIV and HIV prevention efforts in the United States. Second, a brief review of the importance of serostatus disclosure to HIV prevention efforts is undertaken. In this section, the primary focus is on the potential impact of disclosure on HIV transmission risk. In the third section, the review seeks to briefly describe and review examples of the basic types of sexual risk behavior measures which are currently being used by professionals in the study of sexual behavior and the challenges associated with the use of the measures, particularly in their use as measures of HIV prevention efforts. This section also describes briefly the types of statistical analyses commonly used to analyze data on sexual risk behavior. In the final section, the topic of validity is discussed. Emphasis in this section is placed on the meaning of validity and the analytical techniques used to explore it.

Conclusion

The impact of HIV and its transmission across the world has been well-documented. In the US, that impact has been deeply felt in the MSM community. Efforts toward developing, administering, and evaluating high quality interventions which can reduce the risk of transmission are underway. However, the effectiveness of those efforts is limited by the ability of researchers to adequately measure HIV transmission risk. A review of the relevant literature has found that, despite call for improved measures, no consensus on the best way to measure risk has been reached. In interventions involving MSM, the problem is particularly acute. Recently, interventions designed to increase serostatus disclosure in HIV positive MSM have shown promise in reducing the risk of HIV transmission. However, sexual risk reduction behavior in MSM involves complex seroadaptive strategies. Interventions involving MSM must be evaluated with measures which are sensitive to these behavioral strategies.

While both the MOR (Susser, Desvarieux, & Wittkowski, 1998) and Pinkerton and Galletly (2007) models of HIV transmission risk represent viable alternatives for the quantification of risk, neither approach has been widely adopted. Additionally, both procedures have attributes which, without further development, limit their usefulness in the study of the effect of disclosure on HIV transmission risk in MSM. The ordinal scale proposed by Osmond et al. (2007), with minor adaptation, has promise as a measure of HIV transmission risk in MSM.

The literature review also revealed a shift in the conception of validity from the classical trinitarian involving three distinct types of validity, to the unitarian view in which all aspects of validation reflect on the quality of construct validity. The seminal chapter on validity by Messick (1995) in the 3rd edition of *Educational Measurement*, was drawn upon extensively to explain the unitarian view, and to explore its utility as a framework for the validation of the ordinal measure of HIV transmission risk which is the subject of this study. The unitarian view was found to be more comprehensive. The emphasis placed on the importance of the consequences of score interpretation and use in this framework emerged as a particularly relevant strength. Given that the study seeks to evaluate this ordinal measure as an outcome measure which could be used to determine the effectiveness of behavioral interventions to reduce transmission risk, it was determined that unitarian validity was a more suitable framework for validation in this case.

CHAPTER 3: METHODOLOGY

Introduction

In prior sections, the purpose and background information required to undertake the study were explored. This chapter, comprised of four sections, focuses on the research process and methods which were employed to address the study's goals. First, the context of the proposed study is described. The data used for the study were gathered as part of a larger study. Therefore, it is the context of the larger study that is described. This description includes a brief summary of the study's purpose, the procedures for recruitment and retention, research design, data collection methods, and the instrumentation used to gather data. After the context of the larger study has been explained, procedures used for scale development and coding of the modified ordinal measure of HIV transmission risk are described. The third section begins with a brief description of data preparation, and descriptive analyses conducted in the current study. A review of the goals of the current study and the analytical methods used to achieve each of the goals completes this section. In the final section of this chapter, a summary of the approach taken to validation is provided.

Context of the Larger Study

Purpose. The current study utilizes data collected as a part of a larger study of the effect of an intervention designed to assist HIV-positive MSM in disclosing their serostatus to their sexual partners. The specific aims of the larger study focus on testing the effectiveness of an intervention designed to assist HIV positive MSM in disclosing their serostatus to casual sexual partners. When the larger study is complete, the sample will consist of 300 HIV-positive adult MSM living in the Columbus, Ohio area who are at least 18 years of age, sexually active with 2 or more partners in the last 12 months, and who were interested in learning more about disclosure of their serostatus to sexual partners. In addition, only participants who can speak and understand English and who planned on living in the Columbus area for at least 1 year are included in the sample. Exclusion criteria included women, men who exclusively have sex with women, those who could not speak and understand English, those who are not sexually active or who are behaviorally monogamous, those who didn’t plan on being in the Columbus area for one year, and children under the age of 18. The design of the larger study was longitudinal and involved data collection at multiple time points. At the time of the current study, the larger study is still in progress. Therefore, the current study is limited to data collected from the first 145 participants at the baseline observation.

Recruitment and Retention. Participants in the larger study were recruited in five ways. First, individuals were recruited through advertising efforts with local AIDS Service Organizations (ASO’s). Second, recruitment materials were made available at various HIV-related venues and forums held throughout the year in Central Ohio (e.g., OAC sponsored activities). Third, recruitment materials were displayed at local establishments which gear their services to the gay community. Fourth, advertisements were placed in a variety of local newspapers. Fifth, participants were recruited directly from venues in which solicitation for casual sex is more likely to occur including local bathhouses and websites known for facilitating casual hook ups. Recruited participants were then randomly assigned to one of two groups. Those in the experimental group received a 4-session intervention on serostatus disclosure to sexual partners entitled, "Exposing Yourself" (EY). Those in the attention-control group received standard care using a case management approach embedded within the evidence-based Comprehensive Risk Counseling Services (CRCS) intervention (CDC, 2012b).

Research Design. All activities associated with the larger study were conducted at a private research facility at The Ohio State University. Men recruited to the project were screened for eligibility prior to enrolling. The baseline ACASI questionnaire was administered to those who were eligible, who enrolled, and who provided documented consent to participate. After the baseline assessment, men were randomly assigned to either the experimental (EY) group, or the attention-control group (CRCS). An overview of the design of the larger study is provided in the diagram in Figure 1, called a CONSORT (Consolidated Standards of Reporting Trials) diagram.

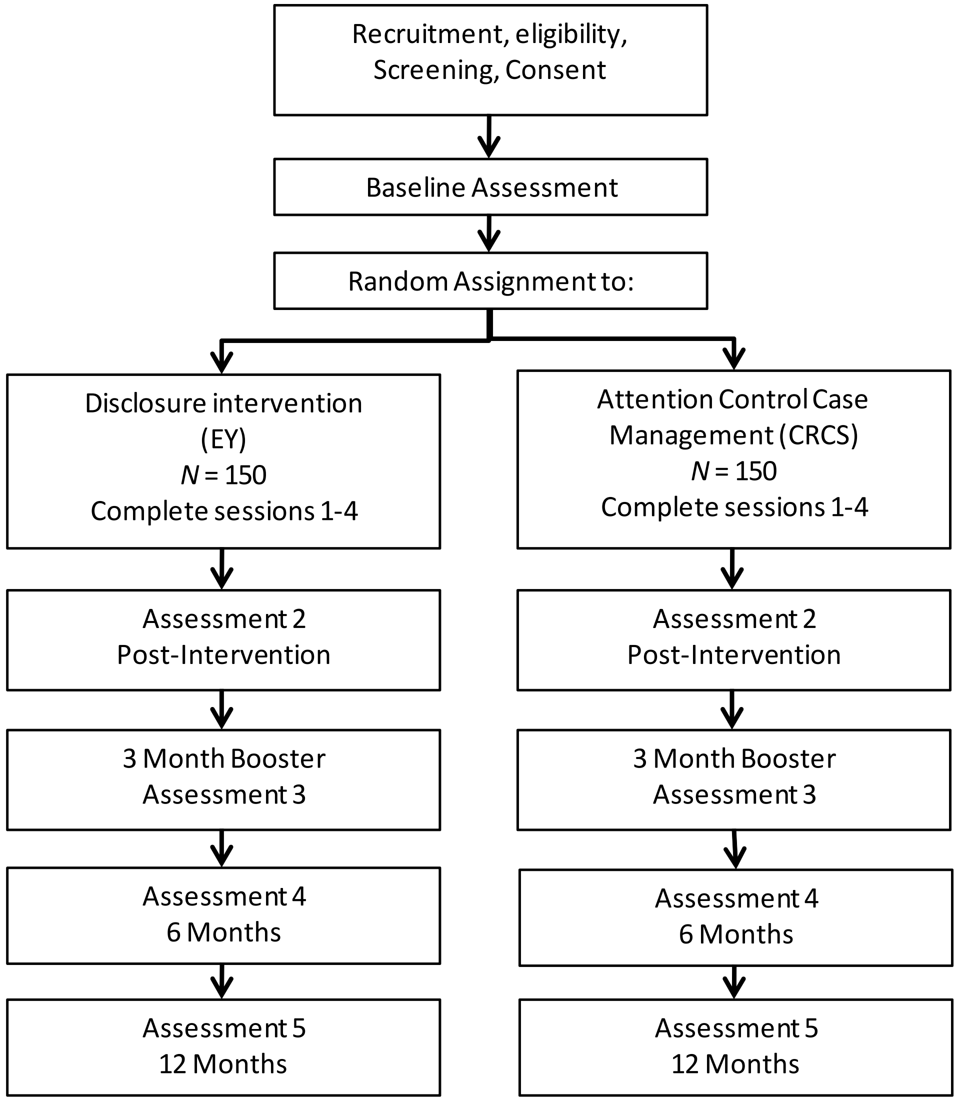


Figure 1. CONSORT Diagram of the Larger Study

Data Collection. As a part of the larger study, the full instrument was encoded into a custom-built questionnaire designed as an audio computer-assisted self-interview (ACASI) including both visual and audio prompts. The ACASI format was chosen as a means of reducing reporting bias and improving the quality of the self-report data. The use of computer-aided interviewing is recommended in any situation wherein the participant is asked to provide sensitive information (DesJarlais et al., 1999; Perlis, Des Jarlais, Friedman, Arasteh, & Turner, 2004). Using ACASI, the participant is able to self-administer the questionnaire at his own pace and in a private environment. Item visual and audio prompts are standardized to ensure that participants hear and see each item in the same way. Audio-assisted systems like ACASI are also effective regardless of the participants' reading abilities. ACASI has been demonstrated to be associated with greater reporting of potentially stigmatized drug, sex and HIV risk behaviors (DesJarlais et al., 1999; Perlis et al., 2004) and has been accepted and preferred as a method of data collection for future interviews by participants of HIV risk related studies (Perlis et al., 2004).

The ACASI questionnaire was customized for each observation to spread the participant burden across assessment points. Items administered to the EY and CRCS groups were identical with the exception of a subset of disclosure-related items. These items involved the strategies used for disclosure that formed an integral part of the intervention. In order to avoid contamination of the CRCS group, these items were administered only to the intervention group.

Instrumentation. Measures used in the larger study were developed by the research team under the leadership of the study's Principle Investigator (PI). Team members included experts in a variety of areas including HIV and serostatus disclosure among MSM, as well as graduate and undergraduate students with interests in HIV in the

MSM community. Additional expertise in instrument content and construction was provided by the measurement development team of the Nutrition-Infection Unit at Tufts University. The team at Tufts was responsible for the development of the Audio-Computer Assisted Self-Interviewing (ACASI) instrument used in data collection.

Scale and item development was conducted in an iterative fashion, beginning with a review of measures used in prior studies, and including both revision and editing phases. While attention was paid to sexual risk and disclosure behavior in the global sense, emphasis in the development process was placed on encounter-specific measures. Regular meetings with the research team over several months resulted in a collection of approximately 450 items and 33 scales which were administered to participants over 5 separate observations. While 18 of the scales adopted for use were based on instruments developed by others, 15 of the 33 scales, including the items used to assess sexual risk behavior and disclosure, were developed by the research team.

A series of count measures were developed to measure global sexual risk and disclosure behavior during the period prior to each observation. In all cases, participants were asked to provide the frequency of the targeted behavior during the prior 30 days. Global measures of sexual behaviors were comprised of a series of 12 self-report items. These items each focused on a single sexual behavior and included questions such as the number of sexual partners, the number of sexual encounters, and the frequency of anal sex (insertive and receptive) with and without a condom. Global measures of disclosure focused on the disclosure behavior of the participant rather than the partner. For example, participants were asked the number of partners who were aware of their serostatus prior to sex, and the number of partners to whom they disclosed during the prior 30 days period. Lists of these global measures can be found in Appendices A and B.

Measures of encounter-specific sexual risk and disclosure behavior were the focus of the measurement development process. Flow charts describing the items and order of item delivery to the participant are provided in Appendices C and D. Encounter-specific measures included questions on select partner characteristics (e.g., nature of relationship, serostatus) and characteristics of the encounter (e.g., type of sex, protection used). Participants were asked about condom use and substance use in conjunction with each sexual encounter and provided information about the perceived safety of each encounter. With respect to disclosure, participants reported how and when they disclosed their serostatus in connection with the encounter, reasons for not disclosing, and regret associated with the disclosure or non-disclosure.

At each observation, participants were asked to provide encounter-specific information about sexual encounters occurring in the prior 30 day period. Participants were asked to report on their most recent encounters, with a maximum of 5 encounters reported. Care was given in the development of the ACASI instrument to defining the word “encounter”. Instructions included some brief examples to guide participant response, and those instructions are reproduced in Appendix E. For the purpose of stimulating recollection, a calendar was presented to the participants, who were asked to indicate the dates of specific encounters. Once participants identified encounters, they were asked to supply the following information on each encounter: type of partner; nature of sexual behavior engaged (e.g., refusal of sex, receptive or insertive oral sex, receptive or insertive anal sex); condom use; whether disclosure occurred; and knowledge of partner's HIV status. Obstacles to disclosure (e.g., fear of rejection or violence) and difficulties with protection (e.g., condom availability or failure) were also be measured.

At baseline, a variety of other measures were administered to participants along with measures of disclosure and sexual activity. Scores on these measures will be used as required to investigate the validity of the ordinal measure of HIV transmission risk which is the subject of this study. Demographic information including age, race, and education, as well as descriptors such as income, relationship status (single, dating, or partnered) were collected. Trait variables associated with sexual mental health were also measured including openness, outness, depression, sexual compulsiveness, and internalized stigma. Openness was measured with a 10-item scale called the Opener Scale (Miller, Berg, & Archer, 1983) which evaluates the participants' willingness to disclose personal information to family, friends, or sexual partners. Outness was measured with an 11-item self-report questionnaire (Frost & Meyer, 2009) designed to measure the degree to which men are open about their sexual orientation to family, friends, and co-workers. The 20-item Depressed Mood Scale (Radloff, 1977) was used to evaluate participant depressive symptoms. Sexual compulsiveness was assessed with the 13-item Compulsive Behavior Inventory (Coleman, Miner, Ohlerking, & Raymond, 2001). Internalized stigma associated with HIV was evaluated with the 40-item HIV Stigma Scale (Berger, Ferrans, & Lashley, 2001). Other general measures included in the baseline assessment were measures of substance use, social support, and sexual communication. Substance use was evaluated with a 7-item scale designed to screen HIV positive persons for substance and alcohol abuse problems (Whetten et al., 2005). Procidano and Heller's (1983) scales of Perceived Social Support from Friends (20-item) and Family (20-item) were administered at baseline. The Assertive Sexual Communication Scale by Quina, Harlow, Morokoff, and Burkholder (2000) and the Health Protective Sexual Communication Scale by Catania (1998) were used to evaluate participants' sexual communication. Self-efficacy and outcome expectancy for disclosure were assessed using a series of scales developed by Semple, Patterson, and Grant (2004). The scales include items assessing self-efficacy for disclosure and outcome expectancy for disclosure. Similar scales were also used to evaluate self-efficacy and outcome expectancy related to condom use and condom use negotiation. Other disclosure-related scales administered at baseline included scales for disclosure attitudes, intentions, and behaviors. The "HIV Disclosure Attitudes" scale is a 13-item measure that assesses general attitudes related to disclosure; the "HIV Disclosure Intention" scale is a 13-item measure that assesses the participants' disclosure intentions about past, present, and future sex partners. The "HIV Disclosure Behaviors" scale is a 13-item measure that assesses the degree to which disclosure has occurred to a variety of different sex partner types. Disclosure-related regret was also evaluated with a 7-item scale. Each of the disclosure-related instruments was developed by the PI of the larger study.

Scale Development and Coding

For purposes of this analysis, the data obtained at the encounter level were recoded utilizing a modified version of the hierarchical risk scale suggested by Osmond and colleagues (2007). The six categories of the original measure incorporated sexual activity and partner status in recognition of the importance of serostatus to MSM sexual behavior. The categories include: (1) no intercourse, oral or anal, (2) no anal intercourse, (3) anal intercourse with 100% condom use, (4) anal intercourse without 100% condom use, but only with partners thought to be HIV positive, (5) unprotected anal intercourse where the insertive partner was HIV- or status unknown, and (6) unprotected anal intercourse where the receptive partner was HIV- or status unknown.

Given that the data from this study were collected from a sample of HIV-positive MSM, HIV risk transmission was operationalized to reflect the risk that the participant could transmit the HIV virus to a partner. For the purposes of this study, the Osmond scale was modified to fully incorporate the potential use of seropositioning as a protective strategy among MSM. Two scale items, XEF_E1_13 and XEF_E1_15, contained information about the positioning of the partners when condoms were used, and provided additional information regarding the use of seropositioning. The addition of these items to the scale resulted in the formation of 9 categories of risk as follows: (0) No sexual encounter, (1) Sexual encounter with no oral or anal intercourse, (2) Sexual encounter with oral intercourse, no anal intercourse, (3) Sexual encounter with anal intercourse, consistent condom use, where partner is also positive, (4) Sexual encounter with anal intercourse, unprotected, where partner is also positive, (5) Sexual encounter with *receptive* anal intercourse, consistent condom use, partner status is negative or unknown, (6) Sexual encounter with *insertive* anal intercourse, consistent condom use, partner status is negative or unknown, (7) Sexual encounter with *receptive* anal intercourse, unprotected, where partner status is negative or unknown, (8) Sexual encounter with *insertive* anal intercourse, unprotected, where partner status is negative or unknown. Based on these items, information on whether or not the partners may have used serosorting and/or seropositioning as a protective measure against HIV transmission can be evaluated. A diagram of the hierarchy of risk depicted by this ordinal scale is provided in Figure 2.

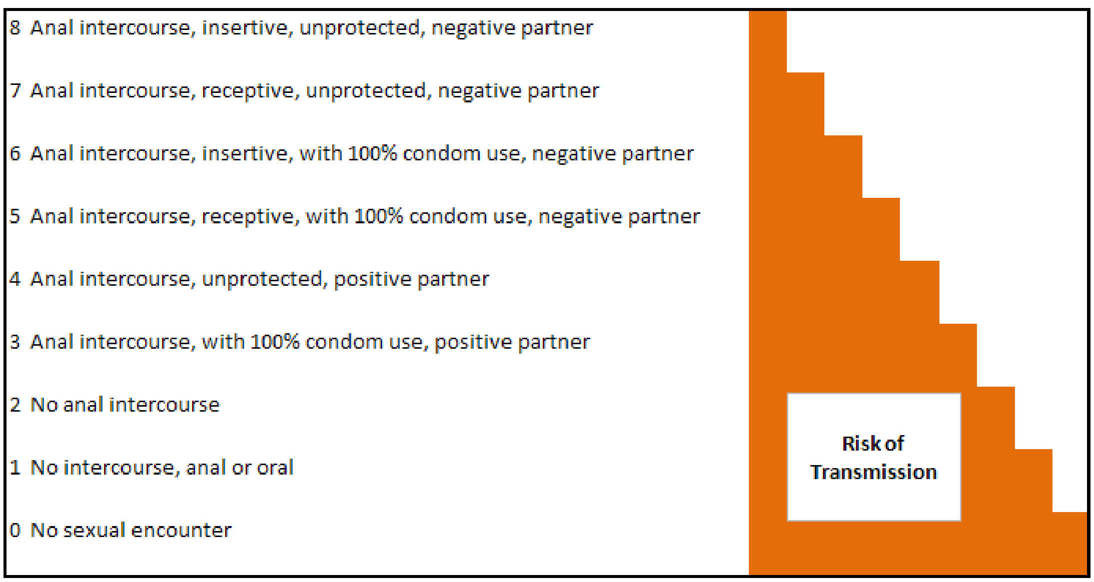


Figure 2. Hierarchy of Transmission Risk, Ordinal Scale

A total of 48 items were available for analysis and for potential inclusion in the proposed scale. Details of each item, including variable name, label, and scoring options are presented in Appendix F. Of the 48 items, participant scores on nine select items were required to create the proposed scale. The items used in scale construction are shown in Table 1.

Table 1. Items Used to Construct the Ordinal Measure of HIV Transmission Risk

Item	Description
XEF_05	Is this partner HIV positive?
XEF_12	Did this encounter involve anal intercourse?
XEF_13	During anal intercourse, were you ever the bottom (was it receptive)?
XEF_14	While you were the bottom, did you always use a condom?
XEF_15	During anal intercourse, were you ever the top (was it insertive)?
XEF_16	While you were the top, did you always use a condom?
XEF_18	Why did you consider this a safe sexual encounter? I only had oral sex
XEF_30	Why did you consider this encounter unsafe? I went down on him without a condom
XEF_31	Why did you consider this encounter unsafe? He went down on me without a condom

Participants who reported no sexual encounters were coded at the lowest level of risk. Among those reporting at least one encounter, item XEF_12 was used to distinguish between encounters involving anal intercourse and those that did not. In cases where the encounter involved both oral and anal intercourse, the risk of the encounter was coded based on the higher-risk anal intercourse reported. For participants who did not engage in anal intercourse, items XEF_18, XEF_30, and XEF_31 were used to determine if the encounter involved oral sex. Where neither oral nor anal sex was involved (i.e., the encounter involved sexual behavior which did not include intercourse), the encounter was placed in the second risk category. Encounters involving only oral sex were assigned to the third risk category. For those encounters involving anal intercourse, items XEF_E1_13, XEF_E1_14, XEF_E1_15, and XEF_E1_16 were used to determine whether the encounter involved an insertive or receptive role for the positive MSM participant, and whether or not condoms were consistently used in the encounter. Where the encounter involved both insertive and receptive roles, the higher-risk insertive role was coded. Item XEF_EX_05 ("Is this partner HIV positive?") was used to ascertain the partner's serostatus. Options for the item include, *'yes,' 'no,'* and '*I don't know.'* Responses to this item were recoded as '*yes'* or '*no or unknown'* to characterize the risk of transmission. This coding strategy was based on the assumption that the risk of transmission to or from partners of unknown status is the same as the risk of transmission to a negative partner.

The nine categories of the ordinal scale were created based on the sexual behavior reported by each participant at each sexual encountrer. The lowest risk category, *'0 No sexual encounter'*, was used only in person-based analyses to distinguish between participants who engaged in sexual encounters and those who did not. The remaining eight categories were used to operationalize risk on an encounter-specific basis. Categories 1 and 2 were used for encounters which didn't involve anal intercourse. Partner serostatus was not considered to significantly affect transmission risk in encounters where neither or nor anal sex was involved. While partner serostatus can impact the risk of transmission in oral sex, the degree of risk cannot be properly assessed without information on positioning (i.e., insertive or receptive). Since positioning information was not available in the data set, no distinction was made on the ordinal scale based on partner serostatus. Categories 3 and 4 were used to depict sexual behavior in seroconcordant couples. While there is no risk of primary transmission in such encounters, the potential risk of super-infection (i.e., infection with an alternate form of the virus) was considered in scale construction. In seroconcordant encounters between MSM, a higher risk was assigned to those where condoms were not used. The remaining four risk categories were used to assign risk levels to serodiscordant encounters. The highest risk category (8) was assigned to serodiscordant encounters involving unprotected anal intercourse (UAI), where the positive partner was *insertive*. In encounters involving serodiscordant UAI, where the positive partner was *receptive*, the lower risk category (7) was assigned. Risk scores of 6 and 5 were assigned to serodiscordant encounters, both receptive and insertive, where condoms were consistently used. From a seroadaptive standpoint, it is important to note certain comparisons between risk categories that provide insight into MSM risk reduction strategies. First, comparisons between categories 7 and 8 are consistent with seropositioning behavior. With this strategy, participants having serodiscordant anal intercourse choose the lowest risk positioning to minimize the risk of transmission. Comparison between encounters rated as 4 and those rated as 7 or 8 provide insight into serosorting behaviors. Those employing this strategy could choose to have unprotected sex with a positive partner rather than a negative partner. Finally, comparisons between encounters rated as 7 or 8, and those rated as 5 or 6, provide insight into condom use as a risk reduction strategy. Finally, comparisons between encounters rated as 4 and those rated as 7 or 8 and those with other ratings provide insight into condom use as a risk reduction strategy.

Data Set Preparation

The data set used for the study was assembled from data collected at the baseline observation of the larger study of HIV disclosure. In the larger study, participant level data were collected using the ACASI instrument, and existed in the comma-delimited format. Individual participant files were imported and concatenated in SPSS. Participant responses in the resulting data set were first reviewed to identify violations of scoring and/or item logic. Examples of violations of scoring, including invalid responses, were rare given the automated system for data collection. Violations of item logic were more frequent. As with any data collection effort, participants must be allowed to provide information which is valid to them. However, limitations in the construction of the ACASI instrument made it impossible to mistake-proof participant response. As a result, in some cases participant responses in one section of the questionnaire did not agree with responses in another section (i.e., global vs. encounter-level measures of sexual behavior). In areas relevant to the study, a full investigation of potential inconsistencies and a full accounting of the results of the investigation were made. Logical violations which were observed were noted, and decisions made regarding data reconciliation, disposition and cleaning were documented and presented along with the results of the analysis.

Descriptive Analysis

A thorough explanation of the validity of participant scores on a scale must begin with a thorough understanding of the sample from which those scores are obtained. To this end, a series of descriptive analyses were conducted on the study participants. First, a thorough description of participant characteristics, including demographics, gender and sexuality, and serostatus was undertaken. Sexual and disclosure-related behavior was then summarized using descriptive statistics of global measures. Participants' scores on the full complement of scales used at baseline were then analyzed. Cronbach's alpha was used to evaluate internal consistency reliability and in cases where reliability was below .8, inter-item correlations were used to explore scale properties.

Next, participants' encounter-specific information was summarized. The number of encounters reported by each participant, and summary statistics for characteristics of partners and encounter contexts were explored. Where it was deemed relevant, analysis was conducted by summing results across encounters and by reporting encounters individually by reported order (e.g., 1^{st} encounter reported, 2^{nd} encounter reported). Chi-square statistics were used to evaluate potential effects of reporting order, or of the number of encounters reported as this is a proxy for the degree of participant sexual activity. Partner serostatus information and information related to disclosure within the encounter context were also summarized. Focus was placed on the timing of disclosure by both the participant and the partner in each encounter.

Review of Goals and Statistical Analysis

Analyses conducted in support of the study goals are detailed below. All analyses were conducted using one of three statistical software packages. Descriptive analyses were conducting using SPSS (19.0). Hierarchical models were constructed using HLM (6.0). Rasch-based analysis was conducted using WINSTEPS (3.72.3). This section is organized by the relevant aspect of validity being investigated. In both classical- and Rasch-based analysis, evidence resulting from a given analysis may be relevant to a number of the aspects of a unified validity. As a result, some repetition in the methods used for analysis across sections may be noticed. The model for classical-based analysis is based on the work of Allen and Yen (2002). For an example of the Rasch-based approach, the article by Fendrich, Smith, Pollack, and Mackesy-Amiti (2009) should be consulted. It is important to note that, for the classical-based analysis, items were assessed both individually and in composite scale form. In the Rasch-based approach, items were regarded individually as dichotomies, and the potential for item additivity was explored through the application of the Rasch model.

The modern approach to validitation, based on the unified concept, was used to investigate the construct validity of the ordinal measure of HIV transmission risk. Each aspect of the unified concept described by Messick (1995) was regarded individually, with the goal of obtaining analytical evidence supporting validity. As was described in the literature review, the validation of the scores obtained from measurement cannot be separated from the intended use and interpretation of those scores. Therefore, for the purposes of this investigation, the scores on the ordinal measure of HIV transmission risk were regarded as indicators of encounter-specific sexual risk behavior in HIV positive MSM. Given that, in the larger study, the data collected on sexual behavior was intended for use as an outcome measure in the evaluation of the effectiveness of a disclosure intervention in reducing HIV transmission risk, this intended use was also assumed for the validation procedure. For each study goal, the methodology used to explore that aspect of validity, along with the justification for the choice of method, is provided. In some instances, one analytical method provided evidence related to more than one aspect of validity. To clarify the connection between method and validity evidence, a table of chosen methodologies, the aspects of validity for which they are relevant, and the section where the results of the analysis are reported, is provided in Appendix G.

Content Representativeness. The goal of the evaluation of content representativeness was to ascertain the extent to which the ordinal scale adequately defines the boundaries of the domain of sexual risk behavior in HIV-positive MSM, and to assess the appropriateness of the hierarchy of risk which the ordinal scale imposes on the construct of HIV transmission risk. Evidence of content representativeness is generally obtained from the judgment of experts (Messick, 1995). The process for this evaluation was largely judgmental, and proceeded from the review of current literature on the relative risk of transmission associated with various sexual behaviors. The structure of the ordinal measure of HIV transmission risk was then compared to published transmission rates as a way of validity the hierarchical position of categories, and to establish scale limitations.

Content Relevance. The evaluation of the relevance of the participants' scores on the ordinal measure to the construct of HIV transmission risk requires was based on an examination of potential sources of construct-irrelevant variation. This type of irrelevant variability can arise from interactions between characteristics of the participant (e.g., depression, HIV stigma) and self-reported sexual activities, and can result in over-endorsement of risk behaviors, under-endorsement of risk behaviors, or non-response. Based on prior work from Catania, Gibson, Chitwood, and Coates (1990), while little is known about the effect of participant demographic characteristics (e.g., gender, race, ethnicity) on the validity of self-reported sexual risk behavior, other characteristics including cognitive deficits, emotional issues, self-esteem, and motivation have been shown to play a role. Variation arising from characteristics of the measurement itself, including language used, the item structure, burden placed on the participant, and the order in which items are presented might also affect participants' self-reports. Finally, the mode of measurement administration has also been shown to affect data quality. Particularly in cases where the subject matter of the measurement is of a sensitive nature, methods which maximize participant privacy and confidentiality, such as self-administered questionnaires, have been found to encourage participant response.

The process used to evaluate the content relevance of the ordinal measure of HIV transmission risk involved 3 steps. First, missing responses to items on the instrument were analyzed. This analysis was conducted in order to evaluate the use of patterns of non-response as indicators of participant burden, motivation, or fatigue. Two types of missingness were explored, including participant skips (i.e., items where the participant chose not to respond) and N/A missingness. With respected to skipped items, missingness was explored as an indicator of participant fatigue, or as an indicator of social response bias. Given that some of the items on the questionnaire involved sensitive information regarding sexual behavior, it was possible that participants who were reluctant to answer these questions could have chosen to skip them. This behavior could result in under-reporting of risk behaviors, and lower scores on the ordinal measure of HIV transmission risk. The designation of N/A missingness was the result of the structure of the ACASI instrument. Based on participants' responses to certain branching items, follow-up questions were regarded as not applicable, and were designated as N/A missingness. The relevance of these patterns of missingness arises from the nature of instrument structure. For example, if a participant responded "no" to an item asking if a sexual encounter involved anal intercourse, then follow-up items asking about position during anal intercourse and condom use were not delivered. In the case of some instruments, this type of missingness might have no larger significance. However, in the case of the ACASI instrument used in this study, the encounter-specific information which was used in the construction of the ordinal measure of HIV transmission risk was affected by this branching structure. The participant completing the questionnaire was asked to answer the same list of questions for up to five separate encounters. This section of the instrument represented a total of 94 possible items per encounter, and increased the burden on the participant significantly. Given the repetitive nature of the item delivery, it was possible for the participant to "learn" the structure of the survey, and begin to choose responses deliberately in order to reduce the burden of response. The possibility for this type of behavior was increased given that the choice to skip an item resulted in the participant receiving a prompt to reconsider. This significantly slowed response to the questionnaire. For participants with low motivation to complete the instrument, or experiencing fatigue as a result of burden, the exploitation of the instruments' branching structure as a way to reduce burden could have been perceived as a more effective way of shortening the assessment period. For this reason, an evaluation of this type of missingness was undertaken, with the goal of discovery of construct-irrelevant difficulty (i.e., failure to endorse an item for a reason not related to the construct). Fisher's exact tests were used to evaluate patterns of missingess between encounters. Second, the ACASI instrument was reviewed for indicators particular to the participant which could be evaluated as potential sources of content irrelevant variability. A list of indicators was generated, and variables from this list were use in an analysis of the relationship between participant characteristics and missingness explored in the first section of this analysis. Significant relationships between these selected characteristics and either skip-related or N/A-related missingness were evaluated as indicators of the nature of the construct-irrelevant variability. For example, a relationship was found between participants' scores on HIV stigma and missingness could represent evidence of construct-irrelevant variation due to social response bias. Hierarchical generalized linear models were then constructed to explore the likelihood of missingness based on select participant characteristics including HIV stigma, depression, motivation, openness, and attitudes about disclosure. For consistency, all hierarchical modeling was conducted in HLM version 6.0. Third, selected participant characteristics and missingness were analyzed as predictors of participants' scores on the ordinal measure of HIV transmission risk. The results of this analysis were reported in the section exploring construct validity in connection with external factors.

Technical Quality. The goal of the evaluation of technical quality was to identify potential sources of ambiguity in the items used to construct the ordinal measure. In some cases, the ambiguity of an item, or of the response alternatives provided, may result in missingness related to skips. In other cases, this ambiguity may lead to logical conflicts between participant answers to different questions or to measurement error leading to higher or lower scores on the ordinal measure of risk. Without the participants' direct feedback about the thought processes which were engaged in while answering the items on the scale, the ability to assess technical quality is limited. Based on available data, four techniques were employed to gather evidence regarding the technical quality of the instrument. First, the items used to construct the ordinal measure was directly assessed using a procedure for evaluating item readability known as the SMOG (Simple Measure of Gobbledygook) score (McLaughlin, 1969). Readability was considered an important issue for technical quality, as low literacy levels among participants can lead to response inaccuracies and missing values to written items (Schroder, Carey, & Vanable, 2003). The SMOG score is a widely used measure of readability which provides an estimate of the years of education required to understand the language of a written passage. SMOG score procedure can be used with language in paragraph form or in the form of a written test item. The original SMOG procedure requires a sample of writing that is 30 sentences or items in length. Given that fewer than 30 items were required to construct the ordinal measure, a modified procedure suggested by ReadabilityFormalas.com (2012) was employed. Additionally, an updated form of the SMOG procedure which was modified by McLaughlin (2008) to increase the precision of the estimate was used. Next, participant measures which were indicative of participant education level, socio-economic status, and computer literacy were selected from among the demographic variables for investigation. Relationships between these characteristics and participant- skipped responses were then evaluated for evidence that participants' may have experienced confusion or ambiguity in the questions or item responses, causing them to skip the response. Because the number of skips observed on the items involved in the construction of the ordinal measure of HIV transmission risk was small, the focus of the investigation was broadened to include a larger number of items all related to the encounter-specific measure of sexual and disclosure behavior. The items which were used to construct the ordinal measure were included among those in the larger group. Focus was placed on the items in this section of the instrument because participants were asked repeated questions regarding their last five sexual encounters. While response latencies were not available to verify this, it was expected that the cognitive demand in this area of the instrument could be higher for two reasons. First, participants were asked to identify a list of encounters, and then to provided details of each one. This required the participant to create a hint or prompt to remind them of which encounter they were currently reporting on. In the event that participants became confused about the encounter being reported, an increase in skipped responses might be observed. Second, the wording of items used to evaluate sexual and disclosure behavior was of great concern in the item development process. Difficulties were encountered by the research team as they attempted to develop items which could capture the complexity of MSM sexual and disclosure behavior. For example, when developing items for use in exploring the timing of disclosure, great care was required to ascertain whether disclosure had occurred prior to having sex, prior to agreeing to have sex, or in connection with a prior encounter. Concern for the clarity of this section was great, so it was decided to include each of the items in this section in the analysis. While a relationship between participant characteristics and this larger group of items would not bear directly on the ordinal transmission scores, it could provide evidence of ambiguity or confusion related to the encounter-specific measure of sexual and disclosure behavior. Third, relationships between scores on the ordinal measure of HIV transmission risk and the set of participant characteristics related to level of education, socio-economic status, and computer literacy were evaluated. Though it is possible that participant education level, socioeconomic status, or computer literacy could be related to sexual risk behavior, it is also possible that significant relationships between those characteristics and ordinal scores could be due to construct-irrelevant variability due to problems with technical quality. The fourth analysis which was conducted to gather evidence of technical quality involved the comparison of participant reporting of sexual behavior in the global and encounter-specific sections of the ACASI instrument. While both sections of the instrument measured sexual and disclosure behavior, the encounter-specific section of the instrument was considered to be more demanding. Therefore, an examination of concurrence in reporting between the two sections was undertaken to identify potential sources of invalidity related to the technical quality of the items in the encounter-specific section. Results of this analysis are reported in the section concerning the aspect of substantive validity.

Substantive Aspect of Validity. The key question relevant to the investigation of the substantive aspect of construct validity is, “Do the scores on the ordinal measure of HIV transmission risk mean what they are intended to mean?” In the investigation of quality, investigators seek to understand the underlying cognitive processes which are engaged in by the participants during the response process, and to determine if these processes are those which are targeted. Four different analyses were conducted in order to explore this. First, comparisons were conducted between participant responses on the global and encounter-specific sections of the ACASI instrument. These sections involved similar questions about sexual behavior and disclosure, but on the global side of the instrument participants were asked to provide counts of specific sexual and disclosure behaviors over the prior 30 day period. In the encounter-specific section, participants are asked to recall behaviors associated with their last five sexual encounters. A high degree of concurrence between self-reports on both sections of the instrument would suggest that participants were actively engaged in recall when responding to the items, and that their responses on the ordinal measure were reflective of their most recent sexual behaviors. If reports on the two sides of the instrument did not agree, it might suggest that participants were not actually engaging in recall, or that the recall processes on the two sides of the instrument were actually different. Based on limitations in the design of the ACASI instrument, it was not possible to compare specific sexual behaviors (e.g., unprotected anal intercourse) or disclosure behaviors (e.g., sex without disclosure) across the two sections. However, it was possible to determine if the number of reported sexual encounters was consistently reported in the global and encounter-specific formats. The result of this comparison was a classification of participants in terms of response consistency (i.e., over-reported on the global section, consistent on both sections, over-reported on the encounter-specific section).

Second, relationships between participant demographic characteristics, and reporting consistency were examined. This investigation sought to find shed light on the source of any reporting inconsistencies which were detected. Where participant demographic variables were found to explain reporting consistency, it could suggest that participant cognitive processes were not consistent across those in the sample. This heterogeneity in cognitive processes across participants could signal the presence of differential item functioning. Chi-square tests were used to evaluate the potential relationships between inconsistency in response and select participant characteristics including demographic variables and scale scores.

In the third analysis, groups based on response consistency were compared on a variety of scale scores, including the ordinal scale of HIV risk behavior. Evidence of significant relationships between response inconsistency and participant scores on other constructs, and on scores on the ordinal measure of HIV transmission risk, could suggest potential sources of cognitive process variability (i.e., participant characteristics related to cognitive processing) or evidence that cognitive variability was a source of potential invalidity (i.e., response consistency related to HIV transmission risk). Simple analysis of variance, with LSD post hoc procedures was used to evaluate response consistency group differences on a variety of scale scores. Kruskal-Wallis analysis was used to evaluate the potential relationship between response inconsistency and participants' scores on the ordinal measure of HIV transmission risk. Significant relationships in this analysis guided the selection of predictor variables to be used in a multinomial regression model of response consistency predicted by participant characteristics.

The fourth analysis conducted to investigate the substantive aspect of construct validity was the multi-trait multi-method matrix. Results of this analysis are presented in the section on validity related to external factors.

Structural Aspect of Validity. The investigation of structural validity of participants' scores on the ordinal measure was conducted to evaluate the extent to which those scores accurately represent the construct of HIV transmission risk. A variety of methods to establish the relationship between the individual items used to construct the scores on the ordinal measure, and the resulting ordinal score were used in this inquiry. These methods can conveniently be divided into two general groups. The first group of methods was based in classical test theory and included and examination of response frequencies, inter-item correlations, item-total correlations, and internal consistency. The second group of methods was based in item-response theory and employed the Rasch model to evaluate the dimensionality of the scale of HIV risk. Rasch analysis was chosen as a tool for the investigation of the structure of scores on the ordinal measure because it permits the examination of the structure of the individual items used in the construction of the ordinal measure of HIV transmission risk, but from the perspective of the ease or difficulty of endorsement of each of the items. The Rasch model assumes that the items on a scale are related to each other in a Guttman-like pattern. For example, in a multi-item scale designed to measure test anxiety, items such as, "I feel nervous before a test," might be easier to endorse by the persons being measured than an item such as, “I have panic attacks before a test.” In a Guttman scale, agreement with the more difficulty to endorse items implies agreement with the simpler items to endorse. This assumption is useful in the construction of psychometric instruments, and the application of the Rasch model to the scores obtained from instrument provides an indication of the degree to which responses match a Guttman-like pattern. By designing items result in Guttman response patterns, psychometricians are able to create total scale scores with high construct validity. However, within the context of the current study where the construct being measured is behavioral (i.e., encounter-specific sexual behavior) participants' reports of specific sexual activities may not follow the same Guttman-like pattern. By applying the Rasch model to the data as a part of the validation process, and examination of the pattern of responses from the perspective of scaling can be undertaken. As a result of the Rasch analysis, an additional analysis exploring potentially different patterns of sexual behavior based on partner seroconcordance/discordance was conducted.

External Factors. Two general strategies were employed to establish the trait and nomological validity of the participants' scores on the ordinal measure of HIV risk transmission. First, a multi-trait multi-method analysis was conducted. The multi-trait multi-method (MTMM) matrix provides a means of assessing both the convergent and discriminant validity of a measure. The MTMM matrix accomplishes this through the construction of a specific pattern of bivariate correlations. In order to construct the matrix, two participant characteristics were chosen (i.e., HIV transmission risk, environmental risk) along with two scaling methods (i.e., constructed ordinal measures, tied rank scores) that could be applied to scaling of both characteristics. Environmental risk was chosen as alternative measure because it was also measured on an encounter-specific basis, and because the issues of environmental context are important to the topic of disclosure. The tied rank scoring procedure was used as the alternative measurement/scaling technique because, when the procedure is applied to dichotomous items, the rank ordering of scores using tied ranks is similar to the use of an additive or total score, but without the assumptions required for additivity (Cliff & Keats, 2003). After a summary of the scaling procedures, and results of scale construction, the MTMM matrix was constructed and the pattern of inter-correlation s among the subject scores was examined. The argument for validity was supported when two scores on HIV transmission risk (i.e., ordinal measure, tied rank score) were highly correlated, and when the correlation between scores on the ordinal measure of HIV transmission and both environmental scores were not correlated.

Additional evidence related to convergent and divergent validity was obtained through the analysis of a series of multi-level ordinal logistic regression models. The goal of modeling scores on the ordinal measure of HIV transmission risk was first to establish the significance of the within-person homogeneity in sexual risk by running an empty model and using the resulting variance estimates to compute the intra-class correlation coefficient. A series of subsequent models were run for the purpose of exploring the relationships between participant and encounter characteristics on the participants' scores on the ordinal risk measure. The validity of those scores was supported to the extent that ordinal risk scores converged with variables that should be related to transmission, and diverged from variables that should not be related to transmission. Prior to model construction a series of correlational analyses were conducted to evaluate relationships between participant demographics, scale scores, and response consistency. Significant relationships found in this analysis were used to support the inclusion of select variables as predictors in the model of ordinal HIV transmission risk. Following the modeling procedure, the procedure recommended by O'Connell (2006) for the investigation of the proportional odds assumption was applied. Results of this analysis were examined as a means of exploring the relationship between the structure of the construct (i.e., HIV transmission risk) and the structure of the ordinal scores resulting from measurement.

Potential for Generalizability. The interpretation of the participants' scores on the ordinal measure of HIV transmission risk is generalizable to the extent that that interpretation can be applied to contexts and samples (Messick, 1995). The question of generalizability involves both the items used to construct the ordinal measure and the measure itself. Evidence related to generalizability of score interpretations was based on analyses already conducted in conjunction with other study goals. Specifically, analyses which incorporated participant demographic information, focused on missingness, or which compared participant responses on the two sections of the ACASI instrument (i.e., global, encounter-specific) to evaluate response consistency were re-evaluated from the perspective of generalizability.

Consequential Validity. The relative consequential validity of the participants' scores on the ordinal measure of HIV transmission risk was demonstrated by comparing results on this measure to the alternative forms available including the global counts measures of risk, the tied rank total scores, and the dichotomous measures of unprotected receptive and insertive anal intercourse. The goal of these analyses was to evaluate the sensitivity, specificity, and utility of the ordinal measure relative to other available scoring mechanisms. These analyses were conducted on data from the first reported sexual encounter. Only the 128 participants reporting one or more encounters were included. Special attention in the analysis was paid to the evaluation of the effect of disclosure on HIV transmission risk.

To facilitate these comparisons, participants' scores on a variety of alternative measures were computed. Alternative measures of risk from the global portion of the ACASI instrument included several dichotomized measures indicating whether or not the participant reported one or more encounters involving unprotected anal intercourse, unprotected receptive anal intercourse, or unprotected insertive anal intercourse. These dichotomies were created based on participant response to count items (i.e., How many of these sexual encounters involved receptive anal sex (you were the bottom) without a condom?). Next, a total of these unprotected risk factors was computed. The range of this scale was from 0 to 5, with the maximum score assigned to participants who reported engaging in one or more encounters involving unprotected receptive anal, insertive anal, receptive oral, insertive oral, and vaginal intercourse. Count variables were also created which reflected the number of encounters in the last 30 days which involved unprotected insertive anal and receptive anal intercourse, as well as a total of all encounters involving unprotected anal intercourse.

Analyses to establish the relative sensitivity and specificity of the encounter-specific measures, including the ordinal measure of HIV transmission risk, were then conducted. The goal of this analysis was to evaluate the extent to which these alternative measures effectively “diagnosed” participant risk behavior. To facilitate comparison among the large number of measures, the ordinal measure of HIV transmission risk was regarded as the “gold standard” to which all others were compared. Three levels of risk comparison were of interest in the analysis. First, a positive “diagnosis” was assigned to participants who reported unprotected insertive anal intercourse with a discordant partner (i.e., a risk score of 8). These positives were compared to those who exhibited behavior consistent with *seropositioning* (i.e., a risk score of 7). Second, a positive “diagnosis” was assigned to participants reporting any unprotected anal intercourse with a discordant partner (i.e., risk score of 7 or 8). These positives were compared to those who exhibited behavior consistent with *serosorting* (i.e., a risk score of 4). In the final comparison, a “diagnosis” was assigned to those who participated in any unprotected anal intercourse (i.e., a risk score of 4, 7, or 8). These positives were compared to those who reported any other type of sexual activity.

It is important to consider, when examining the consequential validity of the scores resulting from measurement, that values play a significant role in both score interpretation, and in the interpretation of validity evidence. As such, values were incorporated into the procedures used to explore consequential validity. Controversies surrounding safe sex in the MSM community informed this incorporation. Specifically, it is controversial to accept seroadaptive behaviors including serosorting and seropositioning as protective. Some argue that seroadaptive strategies may not be reliable, or that HIV positive MSM remain at risk for superinfection and transmission of other STDs when engaging in any form of unprotected sex (Golden et al., 2008; Truong et al., 2006). However, for the purposes of validation for a measure that will be used to evaluate the effectiveness of an intervention designed to increase serostatus disclosure as a way of reducing HIV transmission risk, it was important to evaluate whether or not increased condom use, as well as increased use of seroadaptive strategies, could be detected.

Summary

This chapter provides a summary of several key aspects of the current study which must be understood prior to reviewing the results. First, the context of study was explored. This involved a description of key aspects of the larger study from which the data were obtained. The data were excerpted from a study of the effectiveness of an intervention designed to reduce the risk of HIV transmission in MSM through increased disclosure to sex partners. The ordinal measure which is the subject of this study is evaluated as a potential outcome measure to be used in assessing intervention effectiveness. Next, the procedures used in the construction of the ordinal measure of were explored. The items used in scale construction and the coding strategies used to impose an ordinal structure on the sexual behavior data were described. Connections between the scale and seroadaptive strategies used by MSM were then made to demonstrate the potential usefulness of the measure in evaluating intervention effectiveness. In the third section, data set preparation and the procedures use in descriptive analysis of the data were described. Limitations imposed by the structure and content of the instrument used to collect data, as well as procedures used to account for these limitations, were explained. In the final section of this chapter, the goals of the proposed study were reviewed within the context of the unified concept of validity described in Chapter 2. The methods and statistical analyses proposed to meet each of the study's goals were described. While the procedures proposed to evaluate validity are not exhaustive, an attempt was made to utilize methodologies that were appropriate for the structure of instrument and the data which were obtained from its administration. Some methods used in the analysis provided evidence which is relevant to more than one aspect of the construct validity. However, for the purposes of clarity, an attempt was made to associate specific methods with specific aspects of validity. In Chapter 4, as the results of the analyses undertaken are reported, an attempt will be made to synthesize the evidence which reflects on each aspect of validity.

CHAPTER 4: RESULTS

This chapter presents the evidence of validity obtained through a variety of analyses applied to the data obtained during the baseline observation of a randomized controlled trial of an intervention designed to assist HIV positive MSM in disclosing their serostatus to their casual sexual partners. The unified approach to validity, described by Messick (1995) in his seminal chapter on validity in *Educational Measurement* is used as the theoretical basis of this investigation. Prior to the presentation of results in support of the study goals, a comprehensive descriptive analysis was conducted on both participants and sexual encounters. As a part of the baseline assessment, participants were asked to provide demographic information, as well as to respond to a variety of scale measures. Additionally, participants provided information on their last five sexual encounters. The nested structure of these data necessitates that certain conventions be adopted to ensure clarity in the presentation of analytical results. When referring to the participants in the study, a subscript will be added the symbol for the frequency (n_p). When referring to the encounters in the study, a different subscript will be affixed (n_e). This convention will be utilized throughout the chapter to provide consistency. Following the descriptive analysis, results of individual procedures conducted to explore the various aspects of validity which are described in the study goals are presented in order, by the goal number. Summaries of the evidence obtained are provided at the end of each of these sections, as well as at the end of the chapter.

Participant Characteristics

The sample providing data for the study was comprised of 145 MSM living in a large Midwestern city. These men were participants in a larger study of the effectiveness of an intervention designed to assist HIV-positive MSM in disclosing their serostatus to their sexual partners. The demographic characteristics of the participants who provided data for this investigation of validity are provided in Table 2. On average, the men in the sample were 38.8 years of age (*SD* = 11.0). In terms of race and ethnicity, most of the men described themselves as either white (n_p = 73, 50.3) or Black/African American (n_p = 65, 44.8%). The majority of the participants (n_p = 124, 85.5%) were non-Hispanic. The level of education achieved by the participants was high, with 93 (64.1%) reporting attending at least some college.

Twenty-six participants (17.9%) reported earning a bachelor's degree, and 7 participants reported earning a post-graduate degree. Almost one-third of the participants (n_p = 43, 29.7%) reported earning less than $500 per month, and another third (n_p = 44, 30.3%) reported earning more than $500 but less than $1000 per month. Based on the US Poverty Guidelines published by the United States Department of Health and Human Services (2012), individuals earning less than $10,890 annually ($907.5/month) are living in poverty. Even without consideration of the participants' family sizes and number of dependents, the income levels reported by the study participants suggest that many of them were living in poverty. Additionally, more than half of the participants (n_p = 93, 64.1%) reported being unemployed at the time of data collection.

Table 2. Participant Demographic Characteristics (n_p =145)

Characteristic	n_p	%
Age (*M, SD)*	(38.8, 11.0)	
Race[a]		
American Indian / Alaska Native	9	6.2
Asian	0	0.0
Black or African American	65	44.8
Native Hawaiian / Pacific Islander	1	0.7
White	73	50.3
Other / Mixed	16	11.0
Ethnicity		
Hispanic / Latino	7	4.8
Non-Hispanic	124	85.5
Don't Know	14	9.7
Highest grade completed		
8th grade or less	2	1.4
Some high school	15	10.3
Finished high school / GED	35	24.1
Some college	60	41.4
Bachelor's degree	26	17.9
Post-graduate degree	7	4.8
Monthly income		
$0 to $500	43	29.7
$501 to $1000	44	30.3
$1001 to $1500	29	20.0
$1500 to $2000	14	9.7
Over $2000	15	10.3
Currently employed		
Yes	52	35.9
No	93	64.1

[a] Since participants could select more than one race, percentages across racial groups do not sum to 100.

A summary of participant gender and sexuality is provided in Table 3. All of the participants identified their gender at birth as being male, with one participant identifying currently as transgender. Participants predominantly identified their sexuality as gay (n_p = 111, 76.6%). The majority of participants (n_p = 105, 72.4%) also indicated that they had sex with only men.

Table 3. Participant Gender and Sexuality (n_p =145)

Characteristic	n_p	%
Birth gender		
Male	145	100.0
Female	0	0.0
Current gender		
Male	144	99.3
Female	0	0.0
Transgender	1	0.7
Sexual identity		
Gay	111	76.6
Bisexual	33	22.8
Straight / Heterosexual	1	0.7
Sexual partners		
Only men	105	72.4
Mostly men	23	15.9
Men and women	14	9.7
Mostly women	3	2.1
Only women	0	0.0
Current relationship status		
Single (not dating)	44	30.3
Single (dating)	49	33.8
Committed (monogamous)	36	24.8
Committed (non-monogamous)	16	11.0

However, more than one-fourth of the participants (n = 33, 22.8%) identified themselves as bisexual and one identified as straight or heterosexual. Forty participants (27.7%) indicated that they had some female sexual partners. Twenty-three participants (15.9%) reported mostly male sexual partners, 14 (9.7%) reported male and female sexual partners, and 3 (2.1%) reported mostly female sexual partners. The participants' current relationship status was largely single (n = 93, 64.1%), with 49 (33.8%) indicating that they were currently dating, and 44 (30.3%) indicating that they were both single and not dating. The remained of the participants (n_p = 52, 35.8%) indicated that they were currently in a committed or partnered relationship. Thirty-six participants (24.8%) indicated that their partnered relationship was monogamous, and the remaining 16 (11.0%) indicated that their partnered relationship was non-monogamous.

Participants also provided information about serostatus at baseline. These data are summarized in Table 4, and include the elapsed time since diagnosis with HIV and their most recent viral load information. The time since diagnosis varied widely across the participant group. More than half of the participants (n_p = 84, 57.9%) had been diagnosed for at least 5 years. Forty-four participants (30.3%) had been diagnosed for a period greater than 1 year but less than 5 years and 17 participants (11.7%) had been diagnosed for less than a year. The data for time since diagnosis were normally distributed ($\chi^2(2) = 0.40$, p = .808). The majority of participants (n = 77, 53.1%) also reported an undetectable viral load. Of the remaining 68 participants, 35 reported that they did not know their current viral load, and 3 indicated that they had never had a viral load taken. Only 30 participants (20.7%) reported a detectable viral load.

Table 4. Participant Serostatus Information (n_p = 145)

Characteristic	n_p	%
Time since diagnosis		
Less than 1 year	17	11.7
At least 1 year, but less than 5	44	30.3
At least 5 years, but less than 10	22	15.2
At least 10 years, but less than 20	33	22.8
20 years or longer	29	20.0
Most recent viral load		
Undetectable	77	53.1
I have never had a viral load taken	3	2.1
I don't know	35	24.1
1 to 5000	10	6.9
5001 to 50,000	12	8.3
More than 50,000	7	4.8
Skipped	1	0.7
Contracted HIV through unprotected sex		
Yes	114	78.6
No	6	4.1
Missing	25	17.2
Believes they were infected intentionally		
Yes	66	45.5
No	79	54.5

It is interesting to note that most of the participants (n_p = 114, 78.6%) reported contracting HIV through unprotected sex, and almost half (n_p = 66, 45.5%) indicated feeling that they had been infected by an HIV-positive person who did not disclose his or her serostatus prior to intercourse.

Participant Global Sexual Behavior

Participant sexual activity for the 30 days prior to baseline is summarized in Table 5. Participants were asked to provide summary counts of sexual partners and sexual encounters during the prior 30 day period. Though the inclusion criteria for the larger study required that participants be sexually active prior to enrollment, 12 participants (8.3%) reported having no sexual partners in the 30 days prior to baseline. Based on their report of no sexual partners, these 12 participants were not asked to report the number of sexual encounters they had in the prior 30 days. Thus, these participants were assigned a *not applicable* response to that item. Three participants who reported having one or more sexual partners also reported having no sexual encounters. This suggests that some ambiguity surrounding the definitions of the terms *sexual partner* and *sexual encounter* may have resulted in some inconsistency in the participant responses to questions regarding sexual behavior. The topic of logical inconsistency in response will be explored more deeply in the sections on content relevance, item technical quality, and substantive validity. More than one-third of participants (n_p = 51, 35.2%) reported having only 1 sexual partner in the prior 30 day period while more than half (n_p = 82, 56.6%) reported having 2 or more partners. Eleven participants (7.6%) reported having more than 5 partners during the prior 30 day period. The number of sexual encounters during the same period which were reported by participants varied widely. Twenty-one participants (14.5%) reported having only one sexual encounter, 76 (52.4%) reported having two to five encounters, and 33 (22.8%) reported having more than 5 encounters.

Table 5. Participant Global Sexual Activity, Last 30 Days (n_p =145)

Characteristic	n_p	%
Number of partners		
0	12	8.3
1	51	35.2
2 to 5	71	49.0
More than 5	11	7.6
Number of encounters		
0	3	2.1
1	21	14.5
2 to 5	76	52.4
More than 5	33	22.8
Not applicable	12	8.3

The largest number of sexual partners reported by any participant was 28, and the largest number of reported sexual encounters was 100. The distributions of both the number of sexual partners ($\chi^2(2) = 174.5$, $p = .000$) and the number of sexual encounters ($\chi^2(2) = 220.8$, $p = .000$) were positively skewed and severely leptokurtic.

Details of specific sexual behaviors during the prior 30 day period were also reported by participants. These data are summarized in Tables 6 and 7. After being asked to recount the number of sexual encounters, participants were asked to detail the number of encounters involving receptive and insertive intercourse, both anal and oral. Additionally, participants reported the number of these encounters which were protected and unprotected. About half of the participants (n_p = 72, 49.7%) reported at least one encounter involving protected insertive anal intercourse. A lower proportion (n_p = 63, 43.4%) reported one or more encounters involving protected receptive anal

Table 6. Participant Global Reports of Anal Intercourse, Last 30 Days (n_p =145)

Characteristic	n_p	%
Number of encounters, *protected* insertive anal intercourse		
0	73	50.3
1	17	11.7
2 to 5	28	19.3
More than 5	7	4.8
Not applicable	15	10.3
Skip	5	3.4
Number of encounters, *unprotected* insertive anal intercourse		
0	68	46.9
1	22	15.2
2 to 5	27	18.6
More than 5	9	6.2
Not applicable	15	10.3
Skip	4	2.8
Number of encounters, *protected* receptive anal intercourse		
0	69	47.6
1	28	19.3
2 to 5	22	15.2
More than 5	8	5.5
Not applicable	15	10.3
Skip	3	2.1
Number of encounters, *unprotected* receptive anal intercourse		
0	72	49.7
1	21	14.5
2 to 5	24	16.6
More than 5	10	6.9
Not applicable	15	10.3
Skip	3	2.1

intercourse. Approximately half of the participants reported one or more unprotected encounters involving insertive (n_p = 77, 53.1%) or receptive (n_p = 73, 50.3%) anal intercourse. Reports suggest that condom use associated with oral intercourse was rare regardless whether the participant's role was insertive or receptive.

Table 7. Participant Global Reports of Oral Intercourse, Last 30 Days (n_p =145)

Characteristic	n_p	%
Number of encounters, giving *protected* oral intercourse		
0	107	73.8
1	6	4.1
2 to 5	11	7.6
More than 5	2	1.4
Not applicable	15	10.3
Skip	4	2.8
Number of encounters, giving un*protected* oral intercourse		
0	24	16.6
1	27	18.6
2 to 5	54	37.2
More than 5	24	16.6
Not applicable	15	10.3
Skip	1	0.7
Number of encounters, receiving *protected* oral intercourse		
0	112	77.2
1	5	3.4
2 to 5	10	6.9
More than 5	3	2.1
Not applicable	15	10.3
Number of encounters, receiving un*protected* oral intercourse		
0	47	32.4
1	18	12.4
2 to 5	44	30.3
More than 5	21	14.5
Not applicable	15	10.3

Only 38 (26.2%) participants reported giving oral intercourse to a partner with a condom used. The same number (n_p = 38, 26.2%) reported receiving oral intercourse from a partner with a condom used. Unprotected oral intercourse was more common, with 121 participants (83.4%) reporting 1 or more encounters where they gave oral intercourse to a partner without a condom and 98 participants (67.6%) reporting one or more encounters where they received oral sex from a partner without a condom. Not unexpectedly, the distributions of the count variables were both positively skewed and leptokurtic. It is important to note that, for the 15 participants who reported either no sexual partners or no sexual encounters in the prior 30 day period, these items were regarded as *not applicable*. As the ordinal measure of HIV transmission risk is applied to encounter-specific data, these participants were regarded as having no partners, and no risk of transmission.

Table 8. Participant Global Reports of Non-Disclosure, Last 30 Days (n_p =145)

Characteristic	n_p	%
Number of partners who don't know serostatus		
0	83	57.2
1	32	22.1
2 to 5	13	9.0
More than 5	5	3.4
Not applicable	12	8.3
Number of partners who know now, but didn't know before sex		
0	104	71.7
1	7	4.8
2 to 5	4	2.8
More than 5	0	0.0
Not applicable	30	20.7

Participants were asked to provide global counts of serostatus disclosure during the 30 days prior to the baseline observation. Data on nondisclosure are summarized in Table 8. More than half of the participants in the sample indicated that they had no sexual partners who were not aware of the HIV-positive status. Thirty-two participants (22.1%) reported that they had not yet disclosed to one of their partners, and 18 participants (12.4%) reported more than one partner who did not know their serostatus.

Table 9. Participant Global Reports of Disclosure, Last 30 Days (n_p =145)

Characteristic	n_p	%
Number of partners disclosed to in the last 30 days		
0	28	19.3
1	38	26.2
2 to 5	43	29.7
More than 5	6	4.1
Not applicable	30	20.7
Number of times rejected after disclosure		
0	109	75.2
1	9	6.2
2 to 5	19	13.1
More than 5	8	5.5
Number of times verbally or physically abused after disclosure		
0	135	93.1
1	5	3.4
2 to 5	5	3.4
More than 5	0	0.0

Data provided by participants on disclosure are summarized in Table 9. Almost 20% of the participants reported that they did not disclose to any participants during the period. While this result could indicate that some participants are having sex without disclosing their serostatus, it is not necessarily so. These men may not have disclosed because they were having sex with a partner who was already aware of their serostatus. Additionally, these men may have been engaging in lower risk sexual activities. Among the remaining participants, 87 (60%) reported disclosing to one or more partners during the 30 day period. This question was considered *not applicable* to the 30 (20.7%) participants who reported no partners, no encounters, or no partners who were aware of their serostatus. Participants also reported global reactions to disclosure including violence, abuse, and rejection. Almost one-fourth (n_p = 35, 24.1%) of the participants reported being rejected for sex one or more times after disclosing their serostatus, and 6.8% reported being verbally or physically abused.

Participant Scale Scores

The men in the study were asked to complete a variety of other baseline scales measuring aspects related to disclosure, communication, mental health, and social support. Participant scores on these scales are summarized in Table 10. With respect to disclosure, participants provided scores on scales of behaviors, attitudes, intentions, self-efficacy, outcome expectancy and regret. With the exception of two scales, total scores were computed only for participants who provided answers for each scale item. Based on the branching structure of the disclosure and substance use scales, the resulting patterned missingness made the use of total scores impossible. To overcome this problem, the mean item score was reported for the scales measuring disclosure behaviors, attitudes, and intentions.

In the case of the substance use scale, participants who reported not drinking at all were not presented with the items asking about alcohol consumption (n_p = 15). For these participants, a score of zero was imputed for those items.

Table 10. Participant Scale Scores

Scale	n_p	No. Items	Scale (Min-Max)	*M*	*Median*	*SD*	*α*
Disclosure scale[a]							
Disclosure behavior	144	14	1 - 5	3.7	4.1	1.3	.957
Disclosure attitude	145	14	1 - 4	3.3	3.4	0.6	.950
Disclosure intention	145	14	1 - 4	3.4	3.5	0.6	.953
Self-efficacy							
Condom	145	3	3 - 12	10.8	12.0	1.7	.723
Disclosure	145	3	3 - 12	9.3	10.0	2.5	.856
Negotiation	145	3	3 - 12	10.3	11.0	1.9	.768
Outcome expectancy							
Condom	145	5	5 - 20	14.3	14.0	3.3	.711
Disclosure	143	4	4 - 16	11.6	11.0	2.2	.509
Negotiation	144	5	5 - 20	16.5	17.0	3.0	.811
Assertive sexual communication	145	6	6 - 30	23.5	24.0	4.8	.857
Health protective communication[a]	144	8	1 - 4	2.0	1.9	0.6	.801
Compulsivity	145	13	13 - 65	32.8	33.0	10.6	.908
Openness	145	10	0 - 40	31.4	31.0	5.1	.873
Outness	143	5	0 - 15	11.2	12.0	3.5	.790
Substance use	145	7	5 - 35	15.3	14.0	5.8	.769
Stigma	143	40	40 - 160	102.1	102.0	22.5	.960
Regret of prior disclosure	145	7	0 - 35	11.6	10.0	7.7	.750
Depression	145	20	0 - 60	21.6	22.0	13.1	.933
Social support from family	144	20	20 - 80	52.2	52.5	13.9	.952
Social support from friends	143	20	20 - 80	60.8	60.0	9.5	.929

[a] Score provided is the mean item score

This made it possible to report a total substance use score for these participants. Internal consistency reliability was computed for each of the scales administered. Generally, observed reliability was high across the scales, with Cronbach's alpha coefficients ranging from .71 to .96. Score reliability on the disclosure outcome expectancy scale was below generally acceptable levels ($\alpha = .51$). The poor internal consistency of this measure could suggest that the scale of outcome expectancy, when applied to disclosure, is not unidimensional.

Table 11. Item Scores and Intercorrelations for Disclosure Outcome Expectancy

	Item	n_p	*M*	*SD*	Intercorrelations[a] 1	2	3
1.	I believe that my partner(s) will reject me if I tell him/her that I am HIV positive.	145	2.4	0.9			
2.	I believe that disclosing my HIV status to my sexual partner(s) will increase my sexual pleasure.	143	2.4	1	.23**		
3.	I will become better at disclosing my HIV status if I practice what I will say and do.	145	3.4	0.8	.03	.25**	
4.	I will feel good about myself if I disclose my HIV+ status to all my sexual partners.	145	3.5	0.7	.05	.25**	.47**

*p < .05. **p < .01. ***p < .001.
[a] Correlations reported are Spearman's Rho coefficients.

Descriptive statistics and inter-item correlation coefficients were computed for each of the items on this scale. Results of this analysis are provided in Table 11. Inter-item correlation coefficients were positive and statistically significant with two exceptions involving the item which deals with potential partner rejection in the case of disclosure. The relationship between the expectancy of partner rejection and the expectancy of improving disclosure skills with practice ($r_s(143) = .03$, $p = .705$) was not significant. This was also true of the relationship between the expectancy of partner rejection and the expectancy for feeling good about oneself after disclosure ($r_s(143) = .05$, $p = .524$). The items dealing with expectancies of improving disclosure skills with practice and feeling good about oneself with disclosure were significantly and directly related ($r_s(143) = .47$, $p = .000$).

Encounter-Specific Measures

The men participating in the larger study were also asked to report specific details of their last five sexual encounters. The applicable encounters were limited to those that occurred during the 30 days prior to the baseline assessment. Data summarizing the encounters reported by participants are provided in Table 12. Of the 145 participants included in the sample, 128 (27.3%) reported at least one sexual encounter in the past 30 days. The remaining 17 participants provided no details of their sexual activities. In the reported results for the participant global sexual behaviors, it was reported that 15 men indicated having so sexual partners in the prior 30 days, and 3 men reported at least one partner but no sexual encounters. Crosstabs of the global and encounter-specific measures indicated relatively low levels of agreement between the two reports. Of the 17 participants who did not provide details of at least one sexual encounter, 4 had reported one sexual partner in the prior 30 days, and 1 had reported two partners.

Table 12. Number of Participants Reporting Encounters, by Reporting Order

Reported Order	n_p	%
1st encounter	128	27.3
2nd encounter	111	23.7
3rd encounter	91	19.4
4th encounter	73	15.6
5th encounter	66	14.1

Additionally, 2 of these participants reported two or more sexual encounters in the global measures section, yet failed to provide details of any sexual encounters in this section. Details of these logical inconsistencies are covered in greater detail in the section on content relevance, technical quality, and substantive validity. Among those in the sample, 111 (23.7%) reported at least two sexual encounters, 91 (19.4) reported at least three sexual encounters, 73 (15.6%) reported at least four sexual encounters, and 66 (14.1%) reported details on five sexual encounters.

Overall, the participants in the sample provided detailed descriptions of 469 sexual encounters. Details of partner characteristics, encounter context, and specific sexual behaviors are summarized the following section. Given that some participants provided information on more than one sexual encounter, the encounter-specific data are hierarchically structured, with encounters nested within participants. It is this hierarchical structure that requires the use of multilevel analysis when analyzing encounter-specific behaviors in conjunction with this study's goals for the evaluation of validity. In order to describe encounter-based data, tables will be used that provide summary statistics across all encounters. Additionally, tables depicting these statistics by encounter (i.e., 1st encounter, 2nd encounter, etc.) are provided.

Beginning with partner characteristics, Table 13 summarizes information about partner type across encounters and by encounter number respectively. Across encounters, 159 (33.9%) of all reported encounters involved a main partner. Various casual partners including friends with benefits, hookups, and fuck buddies accounted for 245 (52.2%) of reported encounters. The remaining encounters involved anonymous partners (n_e = 33, 7.0%), exchange partners (n_e = 4, 0.9%), or partners with an unknown relationship to the participant (n_e = 28, 19.3%). When the data are examined by encounter, the proportion of encounters involving main partners appears relatively constant, varying from 32.4% to 38.4%. The percentage of casual partners including friends with benefits, hookups, and fuck buddies tended to decrease as the number of the encounter increased. The proportion of casual partners at the first reported encounter was 57%. By the fifth encounter, this proportion had dropped to 45.5%. The proportion of anonymous/exchange partners was lowest in the first encounter at 5.5%, increasing to 9.1% by the 5th encounter. These results suggest that participants who report more sexual activity tend to report that activity with main partners or anonymous partners. However, a chi-square test of the partner type by encounter number was not significant ($\chi^2(24)$ = 20.4, p = .672). Participants also reported the serostatus of their partners in each encounter. Overall, 197 encounters (42.0%) involved a positive partner, 159 (33.9%) a negative partner, and 113 (24.1%) involved a partner with an unknown serostatus. The relative proportions of concordant and discordant encounters remained relatively stable across the number of the reported encounter ($\chi^2(8) = 6.1$, $p = .639$).

Encounter context was also available for the 469 reported encounters. Results of descriptive analysis for the context variables are provided in Table 14. Approximately one-third of all encounters ($n_e = 158$, 33.7%) involved partners that the participant had met online. An additional 22.4% ($n_e = 105$) involved partners met through a friend, and 15.1% ($n_e = 71$) of encounters involved a partner met in a bar or club.

Table 13. Partner characteristics, By the Number of the Encounter

	Encounter Number										Total Encounters	
	1		2		3		4		5			
	n_p = 128		n_p = 111		n_p = 91		n_p = 73		n_p = 66		n_e = 469	
Characteristic	n_p	%	n_p	%	n_p	%	n_p	%	n_p	%	n_e	%
Relationship type												
Main partner	43	33.6	36	32.4	27	29.7	28	38.4	25	37.9	159	33.9
Anonymous partner	5	3.9	7	6.3	10	11.0	5	6.8	6	9.1	33	7.0
Friend with benefits	18	14.1	21	18.9	13	14.3	7	9.6	7	10.6	66	14.1
Exchange partner	2	1.6	1	0.9	0	0.0	1	1.4	0	0.0	4	0.9
Hookup	30	23.4	26	23.4	20	22.0	14	19.2	7	10.6	97	20.7
Fuck buddy	25	19.5	14	12.6	15	16.5	12	16.4	16	24.2	82	17.5
Other	5	3.9	6	5.4	4	4.4	4	5.5	5	7.6	24	5.1
Skip	0	0.0	0	0.0	2	2.2	2	2.7	0	0.0	4	0.9
Partner serostatus												
Positive	53	41.4	45	40.5	41	45.1	29	39.7	29	43.9	197	42.0
Negative	44	34.4	41	36.9	24	26.4	30	41.1	20	30.3	159	33.9
Unknown	41	32.0	25	22.5	26	28.6	14	19.2	17	25.8	113	24.1

NOTE: n_p refers to the number of participants, n_e refers to the number of encounters

Table 14. Encounter Context, By the Number of the Encounter

	Encounter Number										Total	
	1		2		3		4		5		Encounters	
	n_p = 128		n_p = 111		n_p = 91		n_p = 73		n_p = 66		n_e = 469	
Characteristic	n_p	%	n_p	%	n_p	%	n_p	%	n_p	%	n_e	%
Where met partner												
Online	40	31.3	34	30.6	34	37.4	27	37.0	23	34.8	158	33.7
Bar or club	17	13.3	21	18.9	15	16.5	12	16.4	6	9.1	71	15.1
Through a friend	29	22.7	25	22.5	21	23.1	14	19.2	16	24.2	105	22.4
Bathhouse	6	4.7	5	4.5	3	3.3	2	2.7	2	3.0	18	3.8
Public sex environment	3	2.3	2	1.8	1	1.1	0	0.0	2	3.0	8	1.7
Other	33	25.8	23	20.7	16	17.6	17	23.3	17	25.8	106	22.6
Skip	0	0.0	1	0.9	1	1.1	1	1.4	0	0.0	3	0.6
Where sex occurred												
Their place	39	30.5	36	32.4	31	34.1	17	23.3	18	27.3	141	30.1
My place	52	40.6	41	36.9	36	39.6	35	47.9	30	45.5	194	41.4
Our place	16	12.5	12	10.8	11	12.1	13	17.8	10	15.2	62	13.2
Bathhouse	5	3.9	7	6.3	3	3.3	2	2.7	2	3.0	19	4.1
Public sex environment	3	2.3	4	3.6	3	3.3	1	1.4	1	1.5	12	2.6
Other	13	10.2	11	9.9	6	6.6	5	6.8	5	7.6	40	8.5
Skip	0	0.0	0	0.0	1	1.1	0	0.0	0	0.0	1	0.2

NOTE: n_p refers to the number of participants, n_e refers to the number of encounters

The proportion of encounters (n_e = 106, 22.6%) involving a partner that the participant met in an unreported manner (i.e., other, skip) was unexpectedly high. The proportion of encounters with partners met in bathhouses or public sex environments (n_e = 26, 5.5%) was low. When examined by encounter, the proportions of partners met in different contexts remained relatively stable ($\chi^2(20) = 10.7$, $p = .953$) from the first reported encounter through the fifth reported encounter. Across all encounters, the proportion of encounters which took place in contexts associated with high risk behavior (i.e.,bathhouses, public sex environments) was low (n_e = 31, 6.7%). Crosstabulations of the place in which the encounter occurred by the encounter number suggest that this proportion was relatively stable (χ^2 (20) = 10.3, $p = .962$) from the first through the fifth encounter.

Participants were also asked to provide information on the use of drugs and alcohol in conjunction with sexual encounters. Details of their reports are provided in Table 15. Overall, 140 (29.9%) encounters involved the self-reported use of alcohol. Of the encounters involving alcohol use, the average level of intoxication reported was 5.4 on a scale of 1 to 10. Fewer encounters involved (n_e = 88, 18.8%) reported drug use. However, the average level of intoxication in encounters involving drug use was somewhat slightly higher at 6.9 on a scale of 1 to 10. When the use of alcohol and drugs was evaluated by the number of the encounter, it appeared that the proportion of encounters involving substance use declined from the first through the fifth encounter.

The decrease in the proportion involving alcohol use was particularly notable, decreasing from 38.3% at the first encounter to 18.2% at the fifth encounter.

The observed difference in the proportion of encounters involving alcohol use was statistically significant ($\chi^2(4) = 13.5$, $p = .009$) based on the encounter number, but the observed difference in the proportion of encounters involving drug use was not ($\chi^2(4) = 1.8$, $p = .773$). It is important to remember the structured nature of the data at this point. The observed decrease in alcohol use across encounters could be the result of lower sexual activity reported by participants who drink.

Information regarding the serostatus of participants' sexual partners at the encounter-specific level is summarized in Tables 16. Seroconcordant encounters (n_e = 197, 42%) were more likely than serodiscordant encounters (n_e = 159, 33.9%). However, in considering the risk of HIV transmission, it is important to consider the impact of unknown partner status. If those encounters involving partners of unknown status (n_e = 111, 23.7%) are added to those involving negative partners, these potentially risky encounters (n_e = 270) represent 57.6% of the total number of reported encounters. Examination of partner serostatus by encounter indicates that the proportion of seroconcordant encounters remained consistent regardless of the number of encounter reported ($\chi^2(12) = 9.8$, $p = .652$). This suggests that sexual active partners were not more likely to engage in serodiscordant sex than were those who were less sexually active.

Disclosure of serostatus by the partner is summarized in Tables 17 and 18. Among seroconcordant encounters (n_e = 197) partner disclosure occurred at that encounter 57.4% of the time (n_e = 113).

Table 15. Substance Use Behaviors, By the Number of the Encounter

	Encounter Number										Total	
	1		2		3		4		5		Encounters	
	n_p = 128		n_p = 111		n_p = 91		n_p = 73		n_p = 66		n_e = 469	
Characteristic	n_p	%	n_p	%	n_p	%	n_p	%	n_p	%	n_e	%
Alcohol use	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
How drunk were you feeling (*n* =140, scale of 1-10)	6.9	2.6	6.8	2.5	6.3	2.5	6.3	3.0	5.3	2.6	5.4	2.5
Drinking alcohol before sex	n_p	%	n_p	%	n_p	%	n_p	%	n_p	%	n_e	%
Yes	49	38.3	40	36.0	24	27.0	15	21.1	12	18.2	140	29.9
No	79	61.7	71	64.0	65	73.0	56	78.9	54	81.8	329	70.1
Drug use	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
How high were you feeling (*n* =140, scale of 1-10)	7.2	2.9	7.4	3.1	6.2	2.3	6.0	2.4	6.8	3.3	6.9	2.5
Using drugs before sex	n_p	%	n_p	%	n_p	%	n_p	%	n_p	%	n_e	%
Yes	26	20.3	23	20.7	18	20.2	10	14.1	11	16.7	88	18.8
No	102	79.7	88	79.3	71	79.8	61	85.9	55	83.3	381	81.2

NOTE: n_p refers to the number of participants, n_e refers to the number of encounters

Table 16. Partner Serostatus, By the Number of the Encounter

	Encounter Number										Total	
	1		2		3		4		5		Encounters	
	n_p = 128		n_p = 111		n_p = 91		n_p = 73		n_p = 66		n_e = 469	
Characteristic	n_p	*%*	n_p	*%*	n_p	*%*	n_p	*%*	n_p	*%*	n_e	*%*
Is this partner positive												
Yes	53	41.4	45	40.5	41	45.1	29	39.7	29	43.9	197	42.0
No	44	34.4	41	36.9	24	26.4	30	41.1	20	30.3	159	33.9
Don't know	31	24.2	25	22.5	25	27.5	13	17.8	17	25.8	111	23.7
Skip	0	0.0	0	0.0	1	1.1	1	1.4	0	0.0	2	0.4

NOTE: n_p refers to the number of participants, n_e refers to the number of encounters

Table 17. Partner Disclosure Behaviors (Positive Partners), By Encounter

	Encounter Number										Total Encounters	
	1		2		3		4		5			
	$n_p = 128$		$n_p = 111$		$n_p = 91$		$n_p = 73$		$n_p = 66$		$n_e = 469$	
Characteristic	n_p	%	n_p	%	n_p	%	n_p	%	n_p	%	n_e	%
Partner told you he/she is positive at this encounter	$n_p = 53$		$n_p = 45$		$n_p = 41$		$n_p = 29$		$n_p = 29$		$n_e = 197$	
Yes	39	73.6	27	60.0	18	43.9	14	48.3	15	51.7	113	57.4
No	13	24.5	15	33.3	22	53.7	14	48.3	13	44.8	77	39.1
Skip	1	1.9	3	6.7	1	2.4	1	3.4	1	3.4	7	3.6
When did this partner tell you he/she is positive	$n_p = 39$		$n_p = 27$		$n_p = 18$		$n_p = 14$		$n_p = 15$		$n_e = 113$	
Before sex	36	92.3	26	96.3	16	88.9	13	92.9	15	100.0	106	93.8
During sex	1	2.6	0	0.0	1	5.6	1	7.1	0	0.0	3	2.7
After sex	2	5.1	1	3.7	1	5.6	0	0.0	0	0.0	4	3.5
If not at this encounter, how did you know	$n_p = 13$		$n_p = 15$		$n_p = 22$		$n_p = 14$		$n_p = 13$		$n_e = 77$	
Someone told me	0	0.0	1	6.7	1	4.5	1	7.1	0	0.0	3	3.9
I read his/her online profile	3	23.1	4	26.7	4	18.2	3	21.4	3	23.1	17	22.1
Told me at a previous encounter	9	69.2	6	40.0	14	63.6	8	57.1	8	61.5	45	58.4
Other	1	7.7	4	26.7	3	13.6	2	14.3	2	15.4	12	15.6

NOTE: n_p refers to the number of participants, n_e refers to the number of encounters

Where disclosure did occur at this encounter, it largely occurred before sex (n_e = 106, 93.8%). The predominant reason that disclosure did not occur at this encounter was the participants' prior knowledge of the partner's serostatus based on a prior disclosure (n_e = 45, 58.4%). In 17 cases (22.1%), the partner's status was obtained from an online profile prior to the encounter.

In encounters where the participant reported a serodiscordant partner (Table 18; n_e = 159), partner disclosure occurred at that encounter in 45.3% (n_e = 72) cases. Where disclosure did occur at this encounter it generally occurred before sex (n_e = 68, 94.4%). The predominant reason that disclosure did not occur at this encounter was prior knowledge of the partner's negative serostatus (n_e = 63, 72.4). Participants reported 14 encounters in which the partner's negative serostatus was assumed with disclosure (16.1%). In the remaining encounters, partner serostatus was obtained through online profiles (n_e = 4, 4.6%) or other methods (n_e = 6, 6.9%).

Table 18. Partner Disclosure Behaviors (Negative Partners), By Encounter

	Encounter Number										Total Encounters	
	1		2		3		4		5			
	n_p = 128		n_p = 111		n_p = 91		n_p = 73		n_p = 66		n_e = 469	
Characteristic	n_p	%	n_p	%	n_p	%	n_p	%	n_p	%	n_e	%
Partner told you he/she is negative at this encounter	n_p = 44		n_p = 41		n_p = 24		n_p = 30		n_p = 20		n_e = 159	
Yes	20	37.7	19	42.2	11	26.8	14	48.3	8	27.6	72	45.3
No	24	45.3	22	48.9	13	31.7	16	55.2	12	41.4	87	54.7
When did this partner tell you he/she is negative	n_p = 20		n_p = 19		n_p = 11		n_p = 14		n_p = 8		n_e = 72	
Before sex	17	85.0	18	94.7	11	100.0	14	100.0	8	100.0	68	94.4
During sex	1	5.0	1	5.3	0	0.0	0	0.0	0	0.0	2	2.8
After sex	1	5.0	0	0.0	0	0.0	0	0.0	0	0.0	2	2.8
Skip	1	5.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
If not at this encounter, how did you know	n_p = 24		n_p = 22		n_p = 13		n_p = 16		n_p = 12		n_e = 87	
I read online profile	2	8.3	0	0.0	1	7.7	14	87.5	1	8.3	4	4.6
Told me at a previous encounter	15	62.5	16	72.7	10	76.9	0	0.0	8	66.7	63	72.4
I assumed he/she is negative	5	20.8	5	22.7	2	15.4	1	6.3	1	8.3	14	16.1
Other	2	8.3	1	4.5	0	0.0	1	6.3	2	16.7	6	6.9

NOTE: n_p refers to the number of participants, n_e refers to the number of encounters

Participant encounter-specific self-reports also included information regarding participant serostatus disclosure. Table 19 provides a summary of participant disclosure across encounters. In the majority of cases, participants reported that they did not disclose serostatus at this encounter (n_e = 294, 62.7%). Serostatus was disclosed in 170 encounters (36.2%), and information on disclosure was not available for 5 encounters (1.1%).

Table 19. Participant Disclosure, Summing the Last 5 Encounters (n_e =469)

Characteristic	n_e	%
Did you disclose your serostatus to this partner at this encounter		
Yes	170	36.2
No	294	62.7
Skip	5	1.1

Table 20 provides details concerning specific participant disclosure-related behaviors including the timing and method of disclosure, as well as information regarding disclosure regret. In most cases (n_e = 145, 85.3%) disclosure was made personally by the participant, rather than through a friend or other person. However, 25 participants (14.7%) used some other method for disclosing. Only a few disclosure were reported to occur during or after sex (n_e = 3, 1.8%). Regret associated with disclosure was relatively rare, with only 15

Table 20. Participant Disclosure Behaviors, By the Number of the Encounter

Characteristic	Encounter Number 1		2		3		4		5		Total Encounters	
	n_p = 128		n_p = 111		n_p = 91		n_p = 73		n_p = 66		n_e = 469	
	n_p	%	n_p	%	n_p	%	n_p	%	n_p	%	n_e	%
How did you disclose to this partner	n_p = 54		n_p = 41		n_p = 29		n_p = 24		n_p = 22		n_e = 170	
I told this partner directly	44	81.5	38	92.7	24	82.8	20	83.3	19	86.4	145	85.3
Other	10	18.5	3	7.3	5	17.2	4	16.7	3	13.6	25	14.7
When did you tell this partner you are positive	n_p = 54		n_p = 41		n_p = 29		n_p = 24		n_p = 22		n_e = 170	
Before sex	53	98.1	39	95.1	28	96.6	24	100.0	22	100.0	166	97.6
During sex	1	1.9	1	2.4	1	3.4	0	0.0	0	0.0	2	1.2
After sex	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	1	0.6
Skip	0	0.0	1	2.4	0	0.0	0	0.0	0	0.0	1	0.6
Do you regret disclosing to this partner	n_p = 54		n_p = 41		n_p = 29		n_p = 24		n_p = 22		n_e = 170	
Yes	1	1.9	2	4.9	4	13.8	4	16.7	4	18.2	15	8.8
No	53	98.1	38	92.7	25	86.2	20	83.3	18	81.8	154	90.6
Skip	0	0.0	1	2.4	0	0.0	0	0.0	0	0.0	1	0.6

NOTE: n_p refers to the number of participants, n_e refers to the number of encounters

participants (8.8%) indicating that they regretted disclosing their serostatus to their sexual partner and 154 participants (90.6%) reporting no disclosure-related regret. Reasons for regret were not provided.

Encounters involving participant non-disclosure are summarized in Table 21. The majority of these encounters (n_e = 204, 69.4%) involved a partner who already knew the participant's serostatus, making disclosure unnecessary. However, in the remaining encounters where disclosure did not occur (n_e = 90, 30.6%), participants were not asked to provide a specific reason for non-disclosure. Regret was relatively common with non-disclosure. In 48.9% (n_e = 44) of encounters where disclosure did not occur, participants reported feelings of regret.

Table 21. Participant Non-Disclosure Behaviors, By the Number of the Encounter

	Encounter Number										Total Encounters	
	1		2		3		4		5			
	$n_p = 128$		$n_p = 111$		$n_p = 91$		$n_p = 73$		$n_p = 66$		$n_e = 469$	
Characteristic	n_p	%	n_p	%	n_p	%	n_p	%	n_p	%	n_e	%
Reason for not disclosing at this encounter	$n_p = 73$		$n_p = 69$		$n_p = 61$		$n_p = 48$		$n_p = 43$		$n_e = 294$	
They already knew, so I didn't need to	48	65.8	53	76.8	39	63.9	34	70.8	30	69.8	204	69.4
Other	25	34.2	16	23.2	22	36.1	14	29.2	13	30.2	90	30.6
Do you regret your decision not to disclose to this partner	$n_p = 25$		$n_p = 16$		$n_p = 22$		$n_p = 14$		$n_p = 13$		$n_e = 90$	
Yes	14	56.0	9	56.3	10	45.5	5	35.7	6	46.2	44	48.9
No	11	44.0	7	43.8	12	54.5	9	64.3	7	53.8	46	51.1

NOTE: n_p refers to the number of participants, n_e refers to the number of encounters

Specific sexual behaviors were also reported for the participants' last five sexual encounters. Analysis presented in Table 22 summarizes participants' self-reported behaviors including anal intercourse (insertive and receptive) and condom use. Across the 469 reported encounters, a majority (n_e = 307, 65.5%) involved anal intercourse. Among these, receptive intercourse (n_e = 204, 66.4%) was somewhat more common than insertive intercourse (n_e = 166, 54.1%). This relationship was consistent across the first four encounters. However, at the fifth encounter insertive anal intercourse (n_p = 27, 61.4%) was more slightly common than receptive anal intercourse (n_p = 25, 56.8%). However, this observed variability was not statistically significant ($\chi^2(4) = 3.5$, $p = .485$). Additionally, across encounters, receptive intercourse was more commonly associated with consistent condom (n_e = 83, 40.7%) use than was insertive intercourse (n_e = 55, 33.1%) across all encounters.

Table 22. Anal Intercourse, By the Encounter Number

	Encounter Number										Total Encounters	
	1		2		3		4		5			
	n_p = 128		n_p = 111		n_p = 91		n_p = 73		n_p = 66		n_e = 469	
Characteristic	n_p	%	n_p	%	n_p	%	n_p	%	n_p	%	n_e	%
Did this encounter involve anal intercourse?	n_p = 128		n_p = 111		n_p = 91		n_p = 73		n_p = 66		n_e = 469	
Yes	83	64.8	77	69.4	57	62.6	46	63.0	44	66.7	307	65.5
No	45	35.2	34	30.6	34	37.4	27	37.0	22	33.3	162	34.5
During anal intercourse, were you the bottom (receptive)?	n_p = 83		n_p = 77		n_p = 57		n_p = 46		n_p = 44		n_e = 307	
Yes	59	71.1	52	67.5	39	68.4	29	63.0	25	56.8	204	66.4
No	23	27.7	22	28.6	17	29.8	15	32.6	19	43.2	96	31.3
Skip	1	1.2	3	3.9	1	1.8	2	4.3	0	0.0	7	2.3
While you were the bottom, did you always use a condom?	n_p = 59		n_p = 52		n_p = 17		n_p = 15		n_p = 19		n_e = 162	
Yes	27	45.8	22	42.3	16	41.0	10	34.5	8	32.0	83	40.7
No	32	54.2	30	57.7	23	59.0	19	65.5	17	68.0	121	59.3

NOTE: n_p refers to the number of participants, n_e refers to the number of encounters

Continued

Table 22. Continued

	Encounter Number										Total	
	1		2		3		4		5		Encounters	
	n_p = 128		n_p = 111		n_p = 91		n_p = 73		n_p = 66		n_e = 469	
Characteristic	n_p	%	n_p	%	n_p	%	n_p	%	n_p	%	n_e	%
During anal intercourse, were you the top (insertive)?	n_p = 83		n_p = 77		n_p = 57		n_p = 46		n_p = 44		n_e = 307	
Yes	41	49.4	42	54.5	30	52.6	26	56.5	27	61.4	166	54.1
No	40	48.2	32	41.6	26	45.6	19	41.3	17	38.6	134	43.6
Skip	2	2.4	3	3.9	1	1.8	1	2.2	0	0.0	7	2.3
While you were the top, did you always use a condom?	n_p = 41		n_p = 42		n_p = 30		n_p = 26		n_p = 27		n_e = 166	
Yes	16	39.0	14	33.3	8	26.7	11	42.3	6	22.2	55	33.1
No	25	61.0	28	66.7	22	73.3	15	57.7	21	77.8	111	66.9

NOTE: n_p refers to the number of participants, n_e refers to the number of encounters

However, a close examination of Table 22 shows that the probability of engaging in receptive anal intercourse appeared to increase from encounter one (n_p = 23, 27.7%) through encounter five (n_p = 19, 43.2%). The probability of unprotected insertive intercourse also appeared to increase from the first encounter (n_p = 25, 62.5%) to the fifth (n_p = 21, 77.8%). However, neither the relationship between the number of the reported encounter and unprotected receptive intercourse ($\chi^2(4)$ = 2.4, p = .661) or unprotected insertive intercourse ($\chi^2(4)$ = 5.3, p = .258) was statistically significant.

Participant perceptions of sexual risk were also evaluated on an encounter-specific basis. Of the 469 reported encounters, participants perceived 322 (68.7%) to be safe and the remaining 147 (31.3%) to be unsafe. Reasons for regarding an encounter as safe are summarized in Table 23.

Table 23. Reasons for Considering the Encounter Safe, By Encounter

Characteristic	Encounter Number 1 $n_p = 128$		2 $n_p = 111$		3 $n_p = 91$		4 $n_p = 73$		5 $n_p = 66$		Total Encounters $n_e = 469$	
	n_p	%	n_p	%	n_p	%	n_p	%	n_p	%	n_e	%
Did you consider this a safe encounter?	$n_p = 128$		$n_p = 111$		$n_p = 91$		$n_p = 73$		$n_p = 66$		$n_e = 469$	
Yes	90	70.3	79	71.2	60	65.9	53	72.6	40	60.6	322	68.7
No	38	29.7	32	28.8	31	34.1	20	27.4	26	39.4	147	31.3
Reasons for considering the encounter safe (choose all that apply)[a]	$n_p = 90$		$n_p = 79$		$n_p = 60$		$n_p = 53$		$n_p = 40$		$n_e = 322$	
I only had oral sex	27	30.0	21	26.6	23	38.3	18	34.0	11	27.5	100	31.1
I wasn't the top	16	17.8	16	20.3	11	18.3	7	13.2	7	17.5	57	17.7
I used a condom	29	32.2	28	35.4	19	31.7	20	37.7	15	37.5	111	34.5
My partner used a condom	25	27.8	16	20.3	9	15.0	7	13.2	6	15.0	63	19.6
I didn't ejaculate	17	18.9	13	16.5	10	16.7	8	15.1	4	10.0	52	16.1
My viral load was low	16	17.8	16	20.3	14	23.3	11	20.8	6	15.0	63	19.6
My partner was positive	18	20.0	18	22.8	11	18.3	8	15.1	8	20.0	63	19.6
I'm on meds	22	24.4	23	29.1	16	26.7	14	26.4	10	25.0	85	26.4
Other	19	21.1	12	15.2	13	44.8	12	22.6	9	22.5	65	20.2

[a] Since participants could select more than one reason, percentages across options do not sum to 100.

Participants were permitted to select multiple reasons for perceiving an encounter as safe. The most frequently endorsed reasons were use of a condom (n_e = 111, 34.5%) and only having oral intercourse (n_e = 100, 31.1%). Almost 20% of participants cited condom use by a partner as a reason for perceived safety. Interestingly, close to 20% of encounters involved considerations about the use of HIV medication (n_e = 85, 26.4%) or viral load (n_e = 63, 19.6%) as a reason for perceived safety.

A similar proportion of encounters involved a perception of safety based on a partner's positive serostatus (n_e = 63, 19.6%). Table 24 provides summary statistics across encounters for the remaining 147 (31.3%) encounters that were regarded as unsafe by participants. The most frequently endorsed reasons for regarding the encounter as unsafe each involved the failure to use condoms in conjunction with a specific behavior such as receiving anal sex (n_e = 77, 52.4%), giving anal sex (n_e = 69, 46.9%), receiving oral sex (n_e = 64, 43.5%), or giving oral sex (n_e = 70, 47.6%). It is interesting to note that oral sex was given as a reason for perceived safety as well as for perceived risk. Partner serostatus was also a factor, with increased perceived risk was associated with partner seronegative status in 26 (17.7%) encounters.

Table 24. Reasons for Considering the Encounter Unsafe, By Encounter

	Encounter Number										Total Encounters	
	1		2		3		4		5			
	n_p = 128		n_p = 111		n_p = 91		n_p = 73		n_p = 66		n_e = 469	
Characteristic	n_p	%	n_p	%	n_p	%	n_p	%	n_p	%	n_e	%
Reasons for considering the encounter unsafe (choose all that apply)[a]	n_p = 38		n_p = 32		n_p = 31		n_p = 20		n_p = 26		n_e = 147	
I received anal sex without a condom	18	47.4	16	50.0	18	58.1	13	65.0	12	46.2	77	52.4
I gave anal sex without a condom	15	39.5	17	53.1	13	41.9	10	50.0	14	53.8	69	46.9
I went down on him without a condom	20	52.6	13	40.6	16	51.6	8	40.0	13	50.0	70	47.6
He went down on me without a condom	13	34.2	15	46.9	15	48.4	8	40.0	13	50.0	64	43.5
The condom broke or came off	1	2.6	0	0.0	1	3.2	0	0.0	0	0.0	2	1.4
My viral load was high	3	7.9	0	0.0	1	3.2	0	0.0	0	0.0	4	2.7
My partner was negative	8	21.1	7	21.9	3	9.7	4	20.0	4	15.4	26	17.7
I wasn't taking meds	3	7.9	1	3.1	2	6.5	1	5.0	1	3.8	8	5.4
Other	5	13.2	7	21.9	3	9.7	2	10.0	5	19.2	22	15.0

[a] Since participants could select more than one reason, percentages across options do not sum to 100.

The Ordinal Measure of HIV Transmission Risk

The goal of this study was to explore the validity of an ordinal measure of HIV transmission risk, adapted from the measure proposed by Osmond and colleagues (2007). The participants in the larger study provided data on a total of 469 sexual encounters. Data from these encounters were then used to construct the ordinal measure of HIV transmission risk which is the subject of this study.

Table 25. Items Used to Construct the Ordinal Scale

Item	Description
XEF_05	Is this partner HIV positive?
XEF_12	Did this encounter involve anal intercourse?
XEF_13	During anal intercourse, were you ever the bottom (was it receptive)?
XEF_14	While you were the bottom, did you always use a condom?
XEF_15	During anal intercourse, were you ever the top (was it insertive)?
XEF_16	While you were the top, did you always use a condom?
XEF_18	Why did you consider this a safe sexual encounter? I only had oral sex
XEF_30	Why did you consider this encounter unsafe? I went down on him without a condom
XEF_31	Why did you consider this encounter unsafe? He went down on me without a condom

The procedures used in coding and constructing the ordinal measure were described in Chapter 3 of this dissertation. The developed scale was constructed using participants' responses to the 9 items shown in Table 25. Responses to these items were combined to create eight categories of participant risk ranging from 1 to 8. The lowest category of risk was assigned to reported encounters that did not involved either oral or anal intercourse. The highest risk ranking was assigned to encounters involving unprotected insertive (on the part of the participant) anal intercourse with a serodiscordant partner. Table 27 summarizes the scores assigned to individual sexual encounters based on the ordinal scale.

Table 26. Levels of Ordinal HIV Transmission Risk

Label	Level of risk
1	Sexual encounter, no oral or anal intercourse
2	Sexual encounter, oral sex only
3	Sexual encounter, positive partner, 100% condom use
4	Sexual encounter, positive partner, without consistent condom use
5	Sexual encounter, negative or unknown partner, 100% condom use, receptive
6	Sexual encounter, negative or unknown partner, 100% condom use, insertive
7	Sexual encounter, negative or unknown partner, without consistent condom use, receptive
8	Sexual encounter, negative or unknown partner, without consistent condom use, insertive

The ordinal scores assigned to each of the participants' reported sexual encounters are reported by encounter number in Table 27. The lowest risk of transmission was assigned to encounters involving no anal or oral intercourse (n_e = 49, 10.4%). While these encounters could involve a variety of activities associated with at least some level of transmission risk (e.g., rimming, fisting), participants were not asked to provide details of these activities. Next in order of risk were encounters involving oral intercourse but not anal intercourse (n_e = 113, 24.1%). Given the very low risk of transmission associated with these sexual activities, partner serostatus and condom use were not considered in the ordering of risk at the first two levels. This represents a potential limitation to the sensitivity of the ordinal measure. Levels 3 and 4 of the scale were used to characterize the potential for superinfection. In both cases, anal intercourse was involved. However, encounters involving consistent condom use (n_e = 36) were assigned a lower level of transmission risk, and a higher risk of superinfection was assigned to seroconcordant encounters involving anal intercourse without a condom (n_e = 107, 22.8%).

Table 27. Level of Ordinal HIV Transmission Risk, By Encounter

Level of Risk	*Encounter Number* 1 n_p = 128		2 n_p = 111		3 n_p = 91		4 n_p = 73		5 n_p = 66		Total Encounters n_e = 469	
	n	*%*	*n*	*%*	*n*	*%*	*n*	*%*	*n*	*%*	*n*	*%*
1	12	9.4	12	10.8	8	8.8	9	12.3	8	12.1	49	10.4
2	33	25.8	22	19.8	26	28.6	18	24.7	14	21.2	113	24.1
3	13	10.2	9	8.1	5	5.5	6	8.2	3	4.5	36	7.7
4	25	19.5	25	22.5	24	26.4	15	20.5	18	27.3	107	22.8
5	18	14.1	13	11.7	10	11.0	7	9.6	6	9.1	54	11.5
6	7	5.5	8	7.2	5	5.5	6	8.2	4	6.1	30	6.4
7	10	7.8	11	9.9	7	7.7	7	9.6	6	9.1	41	8.7
8	10	7.8	11	9.9	6	6.6	5	6.8	7	10.6	39	8.3

NOTE: n_p refers to the number of participants, n_e refers to the number of encounters

The highest risk of HIV transmission was assigned to serodiscordant encounters involving anal intercourse. Given the effectiveness of condoms in preventing transmission, encounters involving condom use were assigned lower levels of risk than those that were unprotected. Where anal sex was protected and receptive on the part of the positive partner (n_e = 54, 11.5%), a risk level of 5 was assigned. Discordant encounters involving protected anal intercourse which was insertive on the part of the positive partner (n_e = 30, 6.4%) were assigned a risk level of 6. In unprotected serodiscordant encounters, a risk level of 7 was assigned in cases where the positive partner was receptive (n_e = 41, 8.7%). The highest level of risk was assigned to discordant encounters involving unprotected anal intercourse in which the insertive partner was HIV positive (n_e = 39, 8.3%).

After construction of the measure, the procedures described in Chapter 3 for the exploration of validity were applied to the scores on the ordinal measure of HIV transmission risk. In each of 8 separate sections, result reporting will begin with a recapitulation of the procedures use to investigate the relevant aspect of validity. Next, results for those planned analyses will be reported. Finally, each section will conclude with a synthesis of the evidence relating to that aspect of validity.

Goal 1: Evaluation of Content Representativeness

The goal of the evaluation of content representativeness was to explore the extant research on HIV transmission, and to determine the extent to which the ordinal measure of HIV transmission represents the domain of sexual risk behavior in HIV-positive MSM. The ordinal nature of the proposed measure also required that the ranking of categories from low risk of transmission to high risk of transmission be logically supported. The process for this evaluation was largely judgmental, and proceeded from the review of current literature on the relative risk of transmission associated with various sexual behaviors. What follows is a summary of those findings.

Unprotected anal intercourse, both receptive and insertive, between serodiscordant partners is associated with the greatest risk of HIV transmission. Insertive anal sex on the part of the positive partner, with ejaculation, carried the highest risk of transmission (1 in 70, 1.43%). Without ejaculation, the risk of transmission was cut by more than 50% (1 in 154, .65%). Receptive anal sex on the part of the positive partner was associated with a lower risk of transmission. In this case, circumcision of the negative partner was the major factor affecting transmission. In cases where the partner was uncircumcised the risk of transmission was 1 in 161 (.62%). Where the partner was circumcised the risk dropped considerably to 1 in 909, or 11% (NAM, 2012). The use of a condom in either insertive or receptive anal intercourse has been shown to reduce transmission dramatically. Condoms have been found to be 90-95% effective in preventing transmission if consistently used (Pinkerton & Abramson, 1997). Oral intercourse is also considered to play an important role in HIV transmission. Among MSM, oral sex may involve contact between mouth and anus (*anilingus*) or contact between the mouth and penis (*fellatio*). The risk of HIV transmission is thought to be higher with fellatio than with anilingus, and the risk of both types of oral sex is considered to be much lower than risks associated with anal intercourse (NAM, 2012). While cases of HIV transmission through oral sex have been documented (Vittinghoff et al., 1997), the transmission risk through oral-genital contact is thought be very low (Campo et al., 2006). Factors including oral ulcers, bleeding gums, and the presence of other sexually transmitted diseases are thought to increase the risk of oral transmission (CDC, 2009). Specific rates of transmission resulting from other sexual behaviors are also unavailable in the literature. Additionally, the popular messages about the safety of practices including mutual masturbation, French kissing, fisting, rimming, and use of sex toys are mixed. In acts that do not involve any chance of contact with semen, urine, excrement, or blood (e.g., French kissing, mutual masturbation) the risk of transmission is thought to be negligible (NAM, 2012; The Swiss AIDS Federation, 2012). Acts involving potential contact with bodily fluids (e.g., rimming, fisting, shared sex toys) are portrayed as risky by some sources (Visibility Campaign, 2012) and not risky by others (NAM, 2012; The Swiss AIDS Federation, 2012). Overall, these activities are considered to involve a lower risk of HIV transmission than activities involving intercourse.

Finally, debate continues regarding the significance of the risk of superinfection (i.e., re-infection of a positive person with an additional strain of the HIV virus) among seroconcordant partners engaging in unprotected anal intercourse, as well as the implications for health in the event that it actually occurs (Allen, 2002; CDC, 2009b; Cheonis, 2005; Marcus, McConnell, & Grant, 2011; Project Inform, 2004). Based on existing evidence, the incidence of superinfection is estimated to be 5% within 6 to 12 months of initial infection. Additionally, greater virulence of the superinfecting strain can lead to poorer health outcomes (Smith et al., 2004).

Descriptive results, shown in Table 28, for the risk of HIV transmission across encounters support the content representativeness of the ordinal measure as encounters were found in each of the categories of risk identified. The categories as characterized also make it possible to examine the potential use of risk some risk reduction strategies including avoidance of anal intercourse, avoidance of oral and intercourse, serosorting, and seropositioning. The frequency of unprotected anal intercourse sex was greater in seroconcordant encounters. Among encounters that involved anal intercourse and were serodisconcordant (n_e = 164), 84 (51.20%) were described as protected. In encounters where partners were concordant (n_e = 143), only 36 (25.2%) involved consistent condom use. While participants were not asked to report the use of serosorting as a risk reduction strategy, this pattern of behavior is consistent with its use. Among unprotected encounters involving serodiscordant partners (n_e = 80), the observed frequencies of receptive (n_e = 41, 51.3%) and insertive (n_e = 39, 48.7%) positioning do not support the use of seropositioning as a strategy. While the proportion of risk varied among the encounters, no significant differences in the proportions were observed ($\chi^2(28) = 15.8$, $p = .972$).

Limitations of the content representativeness of the ordinal measure arise from the ACASI instrument used in data collection and in the extremely low rates of transmission involved in a variety of sexual activities common with MSM. Given the much higher rates of HIV transmission associated with anal intercourse, and with the potential use of serosorting and seropositioning as risk reduction strategies, emphasis in instrument development was placed on encounters involving anal intercourse. The larger study involved an RCT of an intervention to promote serostatus disclosure among sex partners. Data collection focused on gathering information of sufficient detail to detect shifts in risk behavior as a result of the intervention. Shifts among very low risk behaviors (e.g., rimming, mutual masturbation) were not significant to the question of intervention effectiveness. Additionally, difficulties in establishing definitive estimates of risks among these behaviors made ranking impossible. By omitting items related to these behaviors, increased burden on the participant from the need to respond to a large number of items related to low risk behavior was also avoided. Larger risk changes associated with the avoidance of anal intercourse, serosorting and seropositioning were of greater importance and were therefore prioritized. Given the documented importance of withdrawal prior to ejaculation and partner circumcision to the risk of transmission, items related to these factors should be considered for addition to the ordinal scale.

Based on the review of the literature regarding HIV transmission risk, the content representativeness of the proposed ordinal measure is supported. Anal intercourse is the sexual activity associated with the highest risk of HIV transmission. Further distinctions in the level of risk are also apparent based on partner serostatus (i.e., concordant, discordant), sexual positioning (i.e., insertive, receptive), and condom use. Oral intercourse represents a lower risk of transmission. As with anal intercourse, partner serostatus, sexual positioning, and condom use play a significant role in transmission. Still lower levels of risk are associated with a variety of other activities that do not involved intercourse. The ordinal measure of risk is able to capture the full range of sexual activities which are accompanied by a risk of transmission. The sensitivity of the scale it the lower end of the risk continuum is not as great as the sensitivity at the high end of the scale. This is due to the lack of information available on lower level risk activities. This weakness of the scale represents a limit to its usefulness as a measure of the effectiveness of risk reduction interventions in groups of MSM that do not engage in higher risk activities (i.e., anal intercourse, oral intercourse). It is also apparent that, with the exception of strategies based on withdrawal prior to ejaculation, the ordinal measure also adequately represents the domain of seroadaptive behaviors.

Goal 2: Evaluation of Content Relevance

The goal of the evaluation of content relevance was to explore participant responses and non-responses to the items which were used to construct the ordinal measure for evidence of variability due to participant characteristics, characteristics of items, or characteristics of the measurement process. Evidence of such relationships is suggestive of construct-irrelevant variability, and poses a threat to the valid interpretation of scores on the measure. A 3-step process was employed to investigate this aspect of validity. First, patterns of participant non-response (i.e., skip, N/A) across reported encounters were explored. Second, relationships between participant characteristics and non-response were explored. Third, participant scores on the ordinal measure were regressed on missingness, and on select participant characteristics. Details of this analysis are reported in the section on external factors. However, results of that analysis are synthesized in the summary of content relevance which concludes this section.

First, patterns of missingness associated with each scale item were examined. A total of 469 encounters were reported, thus a total of 469 valid responses were possible for each item used to construct the ordinal measure. Responses were considered valid in all cases except those where the participant was not presented with the question as a result of instrument branching (*N/A*), or where the participant elected to skip an item (*Skip*). The designation *N/A* was used to characterize pattern missing data related to the branching structure of the ACASI instrument. For example, responses to item XEF_13 which asked the participant to elaborate on the position taken during anal intercourse were only given to participants who had answered affirmatively to item XEF_12 which asked whether the encounter had involved anal intercourse. Thus, with respect to items where *N/A* related missingness was observed, the issue related to potential content irrelevance was restricted to the item which lead to the branching. Missing items identified as *Skips* occurred as the result of direct participant action. In order to skip an item, the ACASI instrument required that the participant press a *Skip* button. Following the skip, a recorded message was delivered to the participant reminding them of the confidentiality and importance of their responses. The participant was then prompted to either return to the item to respond, or to continue the skip.

Table 28. N/A Responses, by Item and Encounter

	Encounter Number										Total Encounters	
	1		2		3		4		5			
	n_p = 128		n_p = 111		n_p = 91		n_p = 73		n_p = 66		n_e = 469	
Item	n_p	%	n_p	%	n_p	%	n_p	%	n_p	%	n_e	%
XEF_05	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
XEF_12	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
XEF_13	45	35.2	34	34.0	34	37.4	27	37.0	22	33.3	162	34.5
XEF_14	69	53.9	59	53.2	52	57.1	44	60.3	41	62.1	265	56.5
XEF_15	45	35.2	34	30.6	34	37.4	27	37.0	22	33.3	162	34.5
XEF_16	87	68.0	69	62.2	61	67.0	47	64.4	39	59.1	303	64.6
XEF_18	38	29.7	32	28.8	31	34.1	20	27.4	26	39.4	147	31.3
XEF_30	90	70.3	79	71.2	60	65.9	53	72.6	40	60.6	322	68.7
XEF_31	90	70.3	79	71.2	60	65.9	53	72.6	40	60.6	322	68.7

NOTE: n_p refers to the number of participants, n_e refers to the number of encounters

Table 28 provides a summary of the missingness related to *N/A* responses by encounter number, as well as the total *N/A* responses across all encounters. Decreasing or increasing patterns of responses across reported encounters could suggest a relationship between the likelihood of responding positively to a branching question. For example, an increase in the percentage of *N/A* responses across reported encounters could suggest that participants began to respond negatively to a question as a result of fatigue (i.e., as a way of avoiding additional burden). A decreasing percentage of *N/A* responses across encounters might suggest that those who were more sexually active were less likely to be inhibited by social response acceptability than were participants who were less sexually active. Examination of the data in the table did indicate some variability in the proportion of *N/A* responses by encounter. An increase in the percentage of *N/A* responses was observed in item XEF_14. To ascertain the significance of the relationship between *N/A* response and encounter number, a set of multi-level logistic models using full penalized quasi-likelihood estimation and robust standard errors using the Huber/White correction was constructed. For each of the items use to construct the ordinal measure, a model with *N/A* response to that item as the outcome (i.e., 1 = N/A, 0 = Other) and encounter number as the predictor, was applied to the data set. In all each case, encounter number was not a significant predictor of *N/A* response.

A review of the frequencies of missingness among the items used to construct the ordinal measure suggests that participants were unlikely to skip an item. In only 2 of the

469 encounters (0.4%) did participants choose not to indicate the serostatus of their partners. In 7 of the 469 encounters (1.5%), participants chose not to provide specific information about the position they assumed during anal intercourse. Table 30 provides a summary of missingness related to participants' skipped responses. Overall, the frequency of skipping in these data was low. However, an increase or predominance of skipping behavior in the latter reported encounters might suggest a relationship between fatigue and missingness. Additionally, a predominance of skips in the early encounters might suggest that those who were less sexually active were more affected by a desire to provide socially acceptable responses. Examination of the data in this table did not indicate any such pattern in the skipped responses. As a result of these very low levels of missingness related to participants skips, no analysis of potential relationships with irrelevant variables was undertaken.

Table 29. Skipped Items, By Item and Encounter

	Encounter Number										Total Encounters	
	1 n_p = 128		2 n_p = 111		3 n_p = 91		4 n_p = 73		5 n_p = 66		n_e = 469	
Item	n_p	%	n_p	%	n_p	%	n_p	%	n_p	%	n_e	%
XEF_05	0	0.0	0	0.0	1	1.1	1	1.4	0	0.0	2	0.4
XEF_12	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
XEF_13	1	0.8	3	2.7	1	1.1	2	2.7	0	0.0	7	1.5
XEF_14	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
XEF_15	2	1.6	3	2.7	1	1.1	1	1.4	0	0.0	7	1.5
XEF_16	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
XEF_18	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
XEF_30	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
XEF_31	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0

NOTE: n_p refers to the number of participants, n_e refers to the number of encounters

Relationships between the likelihood of *N/A* missingness and selected participant characteristics were evaluated using hierarchical linear modeling. A series of models were run to explore potentially irrelevant variability. For each analysis, a two-level hierarchial logistic model was run using full penalized quasi-likelihood estimation and robust standard errors using the Huber/White correction. A separate analysis was run for each item used to construct the ordinal measure. In each case, the dependent variable in the model was a nominal indicator of missingess on that item. Analyses were not run for the item involving partner's serostatus (Item XEF_05) and anal intercourse (Item XEF_12) as no *N/A* missingness was reported on those items. Participant characteristics targeted for analysis include HIV stigma, depression, motivation, openness, and attitudes about disclosure. Participants' scores on these variables were added at level 2 of the model as predictors of the model intercept. In such a model, a regression coefficient significantly different from zero for any of the predictors would suggest a relationship between that variable and the log odds of *N/A* missingness for that item.

Results of these analyses are summarized in Table 30. Participant self-reported readiness for safer sex was significantly related to *N/A* missingness on item XEF_18. The odds ratio (0.84) was below 1 (95% CI 0.728, 0.958) indicating that the odds of *N/A* missingness on that item decreased as motivation to have safe sex increased. Missingness due to *N/A* on this item was due to the participants' response to a prior item (XEF_17).

Table 30. N/A Responses[a] , Predicted by Participant Characteristics

Parameter[b]	XEF_13[c] *Coeff (SE)*	XEF_14 *Coeff (SE)*	XEF_16 *Coeff (SE)*	XEF_18[d] *Coeff (SE)*
	Item Number			
Level 1 (sexual encounter)				
Intercept	0.53 (1.67)	0.97 (1.63)	3.23* (1.58)	2.91 (1.67)
Level 2 (participant characteristics)				
Attitudes about disclosure	0.28 (0.33)	0.41 (0.33)	0.08 (0.35)	-0.08 (0.34)
Openness	-0.01 (0.03)	-0.03 (0.04)	-0.01 (0.04)	-0.03 (0.04)
HIV-related stigma	-0.02 (0.01)	-0.00 (0.01)	-0.02† (0.01)	-0.01 (0.01)
Depression	-0.01 (0.02)	-0.01 (0.02)	-0.01 (0.02)	0.01 (0.02)
Readiness for safer sex	-0.07 (0.07)	-0.12 (0.08)	-0.01 (0.09)	-0.18* (0.07)
Readiness to disclose	0.04 (0.07)	-0.01 (0.08)	-0.02 (0.08)	-0.06 (0.08)
Random Parameters (Variance Estimates)				
τ_{00}	1.45*** (0.38)	2.04*** (0.46)	2.26*** (0.52)	1.77*** (0.44)

*p < .05 **p < .01 ***p < .001

[a] Full Penalized Quasi-likelihood estimation, Bernoulli distribution for the outcome variable

[b] Unit-specific estimates with Huber-White robust standard errors

[c] Results for items XEF_13 and XEF_15 were identical

[d] Results for items XEF_18, XEF_30, and XEF_31 were identical

[†] Coefficient was significant at the alpha = 0.1 level

If the participant responded “no” when asked if the encounter was safe, or skipped that item, the value of *N/A* was assigned to XEF_18 through XEF_27 which detailed reasons associated with safe encounters. Additionally, participant HIV-related stigma approached significance as a predictor of *N/A* missingness in connection with item XEF_16. This suggests that participants with higher levels of HIV-related stigma were more likely to endorse the prior item (XEF_15) which asked the participant if anal intercourse had been insertive in that encounter.

Evidence related to content relevance supports the validity of the ordinal measure. First, very low frequency of participant skipped responses suggests non-response as the result of lack of motivation to respond or excessive fatigue is not a threat to valid score interpretation. Multilevel analyses of participant *N/A* responses to items used in the construction of the ordinal measure showed no significant relationships with encounter number. Additionally, *N/A* response was found to be unrelated to a variety of participant characteristics which could also be related to sexual behavior and disclosure behavior including attitudes about disclosure, openness, depression, and readiness to disclose. In all cases, significant relationships could suggest construct-irrelevant variability in responses to component items. Those who expressed higher readiness to have safe sex were either more likely to report safer sex in their encounters (i.e., more likely to receive item XEF_18). Given that these participants were reporting more safe sex in a manner consistent with their reported motivations, this relationship actually provides additional evidence of response validity. Participants' levels of internalized HIV-related stigma were negatively related to *N/A* response to item XEF_16. Though the relationship was not statistically significant at the alpha = .05 level, further investigation of this relationship could be warranted. In the investigation of the validity in relationship to external factors (see section entitled *Goal 6: Evaluation of Validity Considering External Factors*), levels of the participants' internalized HIV-related stigma were also found to be negatively related to scores on the ordinal measure of HIV transmission risk. However, given that *N/A* missingness on item XEF_16 results in a higher ordinal score (i.e., the participant is more likely to report insertive anal intercourse), participant lack of motivation to respond or fatigue as evidenced by *N/A* missingness to item XEF_16 is not a likely reason for the relationship between internalized HIV-related stigma and the ordinal measure. Therefore, from the perspective of content relevance, no threat to validity due to internalized HIV-related stigma was apparent.

Goal 3: Evaluation of technical quality

The goal of the evaluation of technical quality is to identify potential sources of ambiguity in the items used to construct the ordinal measure. Poor item construction, in addition to low literacy levels in participants, can lead to ambiguity. In some cases, the ambiguity of an item, or of the response alternatives provided, may result in missing values or logically inconsistent responses. Given that participant ability to read and understand the items on the instrument was not directly observable, several indirect methods were chosen to conduct the investigation of technical quality. In each case, the investigation focused on attributes of the item (i.e., readability) or participant (e.g., literacy) that could result in ambiguity in item meaning. These methods included an evaluation of the readability of each of the items on the scale, a review of the participant responses to detect patterns of missingness and other signs of ambiguity, and an exploration of the relationship between participant responses and participant educational preparation, computer skills, and employment. Relationships between scale scores and these construct irrelevant variables could suggest issues related to technical quality.

Table 31. Item Readability

Item	Description	No. of Polysyllabic Words
XEF_05	Is this partner HIV positive?[b]	2
XEF_12	Did this encounter involve anal intercourse?	2
XEF_13	During anal intercourse, were you ever the bottom (was it receptive)?[c]	2
XEF_14	While you were the bottom, did you always use a condom?	0
XEF_15	During anal intercourse, were you ever the top (was it insertive)?[d]	2
XEF_16	While you were the top, did you always use a condom?	0
XEF_18	Why did you consider this a safe sexual encounter? I only had oral sex	2
XEF_30	Why did you consider this encounter unsafe? I went down on him without a condom	2
XEF_31	Why did you consider this encounter unsafe? He went down on me without a condom	2
Total Number of Items		9
Total Number of Polysyllabic Words		14
Average Polysyllabic Words per Item		1.6
Scaling the Result to Accommodate Fewer than 30 sentences (multiply by 21)		33.6
Adding the Total Number of Polysyllabic Words		47.6
Final SMOG Score		10

An evaluation of item readability was conducted on the 9 items used to construct the ordinal measure (See Table 31). The frequency of polysyllabic words (i.e., words containing three or more symbols) contained in the items was used to assess readability. This approach, known as the SMOG readability formula (McLaughlin, 1969), results in a grade level estimate of reading difficulty. In the case of this 9 item group, the readability is 10, suggesting that the text is written at a tenth grade reading level. Review of the participant demographics (Table 1) indicates that 2 participants (1.4%) had not completed the 8th grade, an additional 15 (10.3%) indicating that they had only completed some high school, and 35 (24.1%) had completed high school or a GED. Given that some of these students may not be reading at a tenth-grade level, ambiguity based on readability is a concern.

The key polysyllabic words which add to the SMOG count of the items include the words *HIV, positive, encounter, intercourse, receptive, insertive, consider*. The greatest concern related to ambiguity arises from the use of the words *encounter, intercourse, receptive,* and *insertive,* as the participant must have a clear understanding of these potentially ambiguous terms in order to correctly report sexual risk behavior. Prior research on sexual behavior in persons living with HIV suggests that the use of scientific terms to relate to sexual behaviors can result in increased reliability (Catania, 1990). However, concern about the ability of participants to understand the term *insertive anal intercourse*, and to properly determine which partner in the encounter was being referred to, were great enough that the popular phrases *were you the top* and *were you the bottom* were used in addition to the scientific terms. Additionally, branching structure in the instrument itself was used to prevent logical inconsistencies. Participants were prevented from answering *while you were the bottom, did you always use a condom?* if they had answered *no* when asked if they had ever been the bottom. A review of participant skipped responses for these items (see Table 27) suggests that the highest proportion of skips was related to items XEF_13 and XEF_15 which ask about the insertive and receptive positions during anal intercourse. For each of these items, 7 of 169 responses (4.1%) resulted in skips. Skips on these items could be due to a variety of issues including the reading level or ambiguity of these terms. They could also be due to a problem occurring at item XEF_12 where the participant is asked if anal intercourse was involved in the encounter. The type of anal intercourse that was being referred to by the instrument involves a penis inserted into an anus. The participant may have thought when answering XEF_12 that the definition included other activities involving the anus such as postillionage (i.e., inserting a finger into the anus), rimming, or fisting. If they answered If they answered "yes", they may have then been faced with two options (i.e., were you ever the top, were you ever the bottom*)* that they did feel applied to the actual activity engaged in, resulting in a skipped response.

A correlational analysis was conducted to explore any relationships that might exist between measures of participant education, income, or facility with technology and signs of ambiguity including skips. For this analysis, the number of skips was extended to include a total of 100 items related to sexual activity and disclosure. Skip frequencies are summarized in Table 33. Across the 100 selected items, 19.2% (*n* = 90) of encounter reports included one or more skipped items. Among those with skips, most included 2, 3, or 4 skipped items in total. The frequency of skipped items did not appear to increase across encounters.

Table 32. Frequency of Skipped Responses[a] , By Encounter

Number of Skipped Responses in the Encounter Report	*Encounter Number*									
	1 (*n* = 128)		2 (*n* = 111)		3 (*n* = 91)		4 (*n* = 73)		5 (*n* = 66)	
	n	*%*	*n*	*%*	*n*	*%*	*n*	*%*	*n*	*%*
0	103	80.5	91	82.0	73	80.2	57	78.1	55	83.3
1	5	3.9	0	0.0	0	0.0	2	2.7	0	0.0
2	10	7.8	10	9.0	10	11.0	6	8.2	7	10.6
3	4	3.1	5	4.5	3	3.3	3	4.1	1	1.5
4	5	3.9	3	2.7	4	4.4	4	5.5	2	3.0
5	1	0.8	1	0.9	1	1.1	1	1.4	1	1.5
6	0	0.0	1	0.9	0	0.0	0	0.0	0	0.0

[a]Summary of skipped responses across 100 items related to sexual behavior

Participant characteristics which might be related to ambiguity were then selected, and correlations between these characteristics and the number of skipped responses in each reported encounter (i.e., among the 100 selected items related to sexual behavior) were examined. Results of this analysis are shown in Table 33. No significant negative relationships were found between the number of skips on items relating to sexual behavior and any of the selected participant characteristics. This was true, regardless of the number of the reported encounter. The observed lack of relationship supports the technical quality of the items used to evaluate participant sexual behavior in each reported encounter.

Table 33. Correlations With Number of Skipped Responses, By Encounter

	Encounter Number				
	1 (n = 128)	2 (n = 111)	3 (n = 91)	4 (n = 73)	5 (n = 66)
Participant Characteristic	r_s	r_s	r_s	r_s	r_s
Highest Grade Completed	-.11	-.09	-.07	-.13	-.01
Monthly Income	-.13	-.18	-.15	-.14	-.11
Comfort With Using the Internet	-.12	-.07	-.09	-.22	-.09

*p < .05. **p < .01. ***p < .001.
[a] Correlations reported are Spearman's Rho coefficients

Taking a more indirect approach, item technical quality was evaluated by examining the correlations between the participant characteristics which might be related to ambiguity or literacy (i.e., highest grade completed, monthly income, comfort using the internet) and the participants' scores on the ordinal measure. A separate analysis was conducted for each of the five reported sexual encounters. As was the case with the number of skips, no significant relationships between participant characteristics and participant scores on the ordinal measure of HIV transmission risk were observed. Significant relationships could suggest that some type of construct irrelevant variability in the ordinal score due to ambiguity could be present. Absence of any such relationship supports not only the technical quality of the items used to measure encounter-specific sexual behavior, but the validity of the measurement process.

The investigation of item technical quality supported the quality of the items used to evaluate participant encounter-specific sexual behavior in general, as well as the quality of the specific items used to create the ordinal measure of HIV transmission risk. While the reading level of the items used to construct the measure was relatively high (i.e., 10th grade level), the vast majority of the participants (*n* = 143, 98.6%) reported having completed at least some high school education, and 88.3% (*n* = 127) reported having completed high school. However, no direct measures of participants' interpretation s of the meanings of the items, nor of participants' literacy levels, were available. Future studies should consider employing more direct measures of technical quality and item ambiguity in order to ensure that participants' interpretations of item meaning are clear and unambiguous. A lack of observed relationship between skips (i.e., evidence of ambiguity) and participant characteristics, as well as a lack of observed relationships between skips and participants' scores on the ordinal measure of HIV transmission risk, suggest that item technical quality was sufficient to ensure that participants' interpretations of item meaning were consistent.

Goal 4: Evaluation of Substantive Validity

The investigation of the substantive validity of the participants' scores on the ordinal measure of HIV transmission risk sought to ascertain the degree to which participants engaged in the cognitive processes targeted by the ACASI instrument. When responding to the encounter-specific section of the instrument, participants were asked to report the details of the last five sexual encounters experienced during the prior 30 day period. Participants were provided with a calendar of the prior 30 days, and asked to indicate the date on which each of the encounters occurred. As a part of the process, participants were also asked to assign a temporary name or nickname to the partner for that encounter. This temporary name was displayed on the ACASI screen while the participant responded to the items for that encounter as a way of helping the participant recall accurately.

In order to establish the degree to which the participants' responses to the 9 items used to create the ordinal measure of transmission risk demonstrated substantive validity, an analysis was first conducted comparing the encounter-specific responses to the participant's responses on the global portion of the ACASI instrument. A high degree of concurrence between self-reports on both sections of the instrument would suggest that participants were actively engaged in recall when responding to the items, and that their scores on the ordinal measure were reflective of their most recent sexual behaviors. Response consistency was evaluated at a basic level, beginning with agreement in the number of sexual encounters reported. Table 34 provides a summary of the findings of this investigation. Reports were considered consistent when one of two situations was observed. First, if the number of sexual encounters reported in response to the question "How many sexual encounters have you had over the last 30 days?" was equivalent to the number of sexual encounters reported in the encounter specific section, the report was considered consistent. Second, if the number of reports in the global section was greater than five, and five encounters were reported in the encounter-specific section, the report was considered consistent. The latter condition was the result if a participant had more sexual encounters in the last 30 days than were possible to report in the encounter-specific section. Reports were considered inconsistent if the number of global encounters was lower than the number of encounter-specific encounters (i.e., over-reporting on the encounter-specific section). If the number of reported encounter-specific encounters was less than the maximum allowable (i.e., less than 5) and the global encounters exceeded the number of encounter-specific encounters, this was also regarded as response inconsistency (i.e., over-reporting on the global section).

Of the 145 participants in the study, 12 (8.3%) participants reported no sexual encounters on either the global or the encounter-specific section. Another 84 (57.9%) participants' reports of sexual activity were consistent between the global section and the encounter section of the instrument. Fifty-six participants reported the exact number of encounters in both sections, and the remaining 29 reported more than 5 encounters in the global section, and reported all 5 allowable encounters in the encounter-specific section. The remaining participant responses were considered inconsistent.

Table 34. Number of Encounters[d] , Cross-Tab By Section

How many sexual encounters have you had over the last 30 days?	Number of Specific Encounters Reported[a,b,c]					
	0	1	2	3	4	5
0	**3**					
1		**10**	3			8
2	*1*	*3*	**13**	3		3
3	*1*		*1*	**10**	2	3
4		*1*	*1*	*2*	**4**	8
5		*2*	*1*		*1*	**16**
6		*1*		*1*		**3**
7						**1**
8						**2**
9						**1**
10				*2*		**2**
12						**2**
14						**1**
15						**3**
20						**2**
22						**1**
25						**3**
30						**4**
40						**2**
50			*1*			
100						**1**

[a] Bolded frequencies represent participants reporting consistently across sections (*n* = 85, 58.6%). [b] Italicized frequencies indicate participants who reported more encounters on the global counts than in the encounter-specific section (*n* = 18, 12.4%). [c] Underlined frequencies indicate participants who reported more encounters in the encounter-specific section than in the global section (*n* = 30, 20.7%). [d] Participants who reported no activity on either section are omitted (*n* = 12, 8.3%).

Again, two types of inconsistent reports were possible. First, 19 (13.1%) participants reported a greater number of encounters in the global section than they did in the encounter-specific section. Seven of these participants were fairly close to consistency, reporting only one more encounter in the global side than they did on the encounter-specific side. The others' reports were more inconsistent. The most inconsistent report was provided on one participant who reported 50 encounters on the global side, but only 2 in the encounter-specific section. The second type of inconsistency arose when participants reported fewer encounters in the global section than they reported in the encounter-specific section. This type of inconsistency accounted for 30 (20.7%) participants' reports. Of these, about half (n = 16, 53.3%) reported only one more encounter in the encounter-specific section than they did in the global section.

In order to evaluate potential relationships between response consistency and participant demographics, a series of chi-square tests of independence were conducted. A significant result on any of these tests would suggest that participant race, sexuality, or sexual behavior could be related to the observed pattern of reporting. Based on the results of these tests, no significant relationships were found between response consistency and participant gay sexual identity/straight or bisexual sexual identity ($\chi^2(2) = 0.3$, p = .878), or monogamy/non-monogamy ($\chi^2(2) = 1.5$, p = .470). However, a significant relationship was found between consistency and participant minority/non-minority status ($\chi^2(2) = 7.2$, p = .027). Participants who identified their ethnicity as Caucasian were more likely to provide consistent reports across the sections of the instrument than were those identified with one or more non-majority ethnicity. This finding is of concern, as it might suggest cultural bias or differential item functioning associated with the ACASI instrument. The relationship between consistency and the participant's scale scores was also investigated. One-way analysis of variance was used to compare the means scores of the three consistency groups (i.e., over-reported global, consistent, over-reported encounter-specific) on scales for disclosure attitudes, self-efficacy, outcome expectancy, sexual communication, health communication, openness, outness, substance use, HIV stigma, disclosure regret, depression, and social support. Significant differences were found between the groups' mean scores on disclosure regret ($F(2, 130) = 3.68$, $p = .028$). Post-hoc comparisons using the LSD procedure indicated that those who over-reported in the global section reported significantly higher disclosure regret ($t(1) = 2.54$, $p = .012$) than did those who reported consistently. No significant difference in disclosure regret was found between those who over-reported in the global section and those who over-reported in the encounter-specific section ($t(1) = 1.11$, $p = .136$). This finding is important because, in the evaluation of the effectiveness of the disclosure intervention, under-reporting in the encounter-specific section limits the ability to evaluate the effect of the disclosure on those who might potentially be having the most difficulty disclosing. Significant differences were also found between the consistency groups on health communication scores ($F(2, 130) = 4.79$, $p = .010$). In this case, those who over-reported in the encounter-specific section reported significantly higher health communication scores ($t(1) = 2.95$, $p = .004$) than those who reported consistently. No significant difference was found between those who over-reported in the encounter-specific section and those who over-reported in the global section ($t(1) = 0.79$, $p = .492$). This suggests that those who are more likely to communicate with their sex partners about sex are more likely to over-report in the encounter-specific section.

Finally, Kruskal-Wallis analysis of variance tests were used to evaluate the potential relationship between reporting consistency and HIV transmission risk as measured with the ordinal scale. If an observed relationship between consistency in reporting and the ordinal score was observed, it could suggest the presence of some type of construct-irrelevant variability. First, individual analyses were conducted by encounter. A significant result on any of these tests would suggest that response consistency was responsible for some portion of the variability among participants in their sexual risk scores. No significant relationship was found between response consistency and HIV transmission risk in the first ($\chi^2(2) = 1.5$, $p = .472$), second ($\chi^2(2) = 1.0$, $p = .681$), third ($\chi^2(2) = 1.5$, $p = .473$), fourth ($\chi^2(2) = 1.4$, $p = .487$), or fifth ($\chi^2(2) = 0.3$, $p = .570$) reported encounter. These results suggest that no significant relationship existed between response consistency and the ordinal measure. An additional analysis using a multilevel cumulative proportional odds model was then conducted. The purpose of this analysis was to evaluate the effect of response consistency on the odds of success in each individual sexual encounter; in this case, success was defined as a score at or below a given level of HIV transmission risk on the ordinal scale. For this analysis, response consistency was dummy-coded, with consistent responders acting as the reference group. With HIV transmission risk as the dependent variable at level 1 of the model, response consistency was added at level of 2 of the model as a participant-specific variable predicting the intercept. No significant difference in the ordinal measure of HIV transmission risk was found between those who over-reported in the global section ($t(125) = 1.02$, $p = .310$) or those who over-reported in the encounter-specific section ($t(125) = 1.34$, $p = .185$) and those who reported consistently.

Table 35. Multinomial Regression of Reporting Consistency ($n_p = 133$)[a]

Characteristic	*Coeff (SE)*	95% CI for Odds Ratio Lower	Odds Ratio	Upper
Over-reporting in the Global Section vs. Consistent Reporting				
Intercept	-3.77*** (1.03)			
Health Protective Communication	0.65 (0.48)	0.75	1.91	4.84
HIV Regret	0.08* (0.07)	1.01	1.09	1.16
Over-reporting in the Encounter-Specific Section vs. Consistent Reporting				
Intercept	-3.47*** (0.89)			
Health Protective Communication	1.09** (0.41)	2.96	1.33	6.61
HIV Regret	0.02 (0.03)	1.02	0.96	1.09

Note: $R^2 = .105$ (Cox & Snell), .125 (Nagelkerke). Model X^2 (4) =14.58, $p = .006$, *$p < .05$ **$p < .01$ ***$p < .001$.

[a] Analysis includes only those participants reporting sexual encounters on both global and encounter-specific sections.

Additional analyses using a single-level multinomial logistic regression were employed to explore participant characteristics which might be significantly related to either over-reporting on the global side of the instrument, or over-reporting on the encounter side of the instrument. The consistency of response was recoded for this analysis, with the reference category assigned to those who reported consistently. Variables considered as potential predictors of inconsistency included minority ethnicity, health protected communication, and HIV regret. The purpose of these analyses was to evaluate the effect of these variables on the likelihood of over-reporting on either the global and encounter-specific sides of the ACASI instrument. Given that the analysis was exploratory, a backward elimination strategy was used for model building. Results of this analysis are shown in Table 35.

Minority ethnicity was not a significant predictor of either type of inconsistency and was removed from the model. The effect of the removal of this variable was not statistically significant ($\chi^2(2) = 3.0$, $p = .225$). Relative to those who reported consistently, the odds of over-reporting on the global side were positively related to HIV regret (Wald statistic (1) = 5.3, $p = .021$). The log odds of over-reporting in the global section increased by 0.082 for every one unit increase in HIV regret score. However, this was not true of those who over-reported on the encounter-specific side of the instrument. The odds of over-reporting in that section of the instrument were positively related to health protective communication (Wald statistic (1) = 7.0, $p = .008$). The log odds of over-reporting in the encounter-specific section of the ACASI instrument increased by 1.086 for every one unit increase in the participants' health protective communication score.

Evidence of the substantive validity of participants' scores on the ordinal measure of HIV transmission risk was mixed. The level of inconsistency observed in participants' responses to the global and encounter-specific sections of the instrument were relatively high. Approximately one-third (*n* = 48, 33.1%) of participant over-reported on either the global or encounter-specific section of the ACASI instrument. The lack of concurrence between the two reports suggests that participants may not have been engaging in a consistent recall process when answering the items in each section. Another 8.3% of participants (*n* = 12) did not report any sexual encounters in the last 30 days on either section of the instrument. While their responses were consistent, they did indicate a potential problem. The inclusion criteria for the study specified that participants must be sexually active within the 30 days prior to baseline. These participants had answered affirmatively that they were sexually active during recruitment, but then answered negatively when asked about sexual activity during the baseline assessment. It is not possible to ascertain which of the participants' responses were valid, but it is possible that their responses to the ACASI instrument did not reflect the targeted recall process. It is also important to note that the nature of the items on the two sections of the instrument did not permit a more thorough investigation of consistency (e.g., consistency in reports of anal intercourse, condom use, etc.). The relationship between participant minority status and response consistency was not observed in the multilevel analysis, suggesting that there was no interaction between threats to substantive validity and ethnicity. However, HIV regret and health communication were significantly related to response inconsistency. This finding could suggest a significant interaction between participant characteristics and the consistency of the recall process.

Goal 5: Evaluation of Structural Validity

The investigation of structural validity of participants' scores on the ordinal measure was conducted to evaluate the extent to which those scores accurately represent the construct of HIV transmission risk. A variety of methods were used to establish the relationship between the individual items used to construct the scores on the ordinal measure and the resulting ordinal score. These methods can conveniently be divided into two general groups. The first group of methods including response frequencies, inter-item correlations, item-total correlations, and internal consistency reliability are based in classical test theory. The second group of methods was based in item-response theory and employed the Rasch model to evaluate the dimensionality of the scale of HIV risk.

Beginning with techniques based in classical test theory, response frequencies for each of the items used to construct the ordinal measure were computed (Table 32). The original items were recoded prior to complete this analysis, with the purpose of creating a series of individual, independent, dichotomous risk factors. In each case, a positive coded response to the item was assigned to behaviors that represented increased risk. For example, item XEF_05 regarding partner serostatus was recoded (XEF_05_R1) such that a response of 1 indicated an HIV negative partner, or a partner whose serostatus in unknown. In either case, this represents an increased risk of HIV transmission over the reference category, which in this case was comprised of HIV positive partners. Items asking about the types of intercourse reported (i.e., receptive anal, insertive anal, oral) and condom use were combined to provide indicators of unprotected receptive anal (XEF_13_R2) and unprotected insertive anal (XEF_15_R2) intercourse. All reported sexual encounters (*n* = 469) were coded as positive for some type of sexual activity (XEF_12_R1). In Table 36, the frequencies of these risk behaviors encounters are reported in descending order of prevalence, by encounter. Sex with a partner whose serostatus was negative/unknown was the most frequently occurring risk factor across encounters (*n* = 270, 57.6%). Receptive anal intercourse was the next frequently reported risk factor and involved 204 (43.5%) encounters. Oral sex was the next most frequently occurring risk behavior (*n* = 179, 38.2%), followed by insertive anal intercourse (*n* = 166, 35.4%). The least frequent behaviors reported were unprotected receptive anal intercourse (*n* = 121, 25.8%) and unprotected insertive anal intercourse (*n* = 111, 23.7%). The observed frequencies of these sexual behaviors suggest that frequencies of observed behavior in MSM is patterned differently than the severity of risk represented by the ordinal measure of transmission risk.

Table 36. Frequency of Dichotomous Risk Behaviors, By Encounter

	Encounter Number										Total	
	1		2		3		4		5		Encounters	
	n_p = 128		n_p = 111		n_p = 91		n_p = 73		n_p = 66		n_e = 469	
Characteristic	n_p	%	n_p	%	n_p	%	n_p	%	n_p	%	n_p	%
Did this encounter involve sexual activity?[a]												
Yes	128	100.0	111	100.0	91	100.0	73	100.0	66	100.0	469	100.0
No	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
Did this encounter involve a partner with a *negative or unknown* serostatus?												
Yes	75	58.6	66	59.5	49	53.8	43	58.9	37	56.1	270	57.6
No	53	41.4	45	40.5	41	45.1	29	39.7	29	43.9	197	42.0
Skip	0	0.0	0	0.0	1	1.1	1	1.4	0	0.0	2	0.4
Did this encounter involve *receptive* anal intercourse?												
Yes	59	46.1	52	46.8	39	42.9	29	39.7	25	37.9	204	43.5
No	69	53.9	59	53.2	52	57.1	44	60.3	41	62.1	265	56.5
Did this encounter involve *oral* sex?												
Yes	50	39.1	37	33.3	41	45.1	26	35.6	25	37.9	179	38.2
No	78	60.9	74	66.7	50	54.9	47	64.4	47	71.2	290	61.8

Continued

Table 36. Continued

Characteristic	Encounter Number 1 n_p = 128		2 n_p = 111		3 n_p = 91		4 n_p = 73		5 n_p = 66		Total Encounters n_e = 469	
	n_p	%	n_p	%	n_p	%	n_p	%	n_p	%	n_e	%
Did this encounter involve *insertive* anal intercourse?												
Yes	41	32.0	42	37.8	30	33.0	26	35.6	27	40.9	166	35.4
No	87	68.0	69	62.2	61	67.0	47	64.4	39	59.1	303	64.6
Did this encounter involve *unprotected receptive* anal intercourse												
Yes	32	25.0	30	27.0	23	25.3	19	26.0	17	25.8	121	25.8
No	96	75.0	81	73.0	68	74.7	54	74.0	49	74.2	348	74.2
Did this encounter involve *unprotected insertive* anal intercourse?												
Yes	25	19.5	28	25.2	22	24.2	15	20.5	21	31.8	111	23.7
No	103	80.5	83	74.8	69	75.8	58	79.5	45	68.2	358	76.3

Specifically, lower risk behaviors such as oral intercourse are reported less frequently than higher risk activities such as receptive anal intercourse. It is important to note that this might have been the result of the lack of focus in the ACASI instrument on oral sex. Chi-square tests of independence were conducted to explore potential relationships between then number of the encounter (i.e., an indicator of sexual activity) and the proportion of encounters involving each risk factor. A significant relationship in any of these tests would suggest that participants who were more sexually active (i.e., greater number of encounters reported) reported different sexual risk behaviors than those who were less sexually active. No significant relationship was found between the encounter number and partner negative/unknown serostatus ($\chi^2(4) = .8$, $p = .943$), receptive anal intercourse ($\chi^2(4) = 2.1$, $p = .709$), oral sex ($\chi^2(4) = 3.2$, $p = .529$), insertive anal intercourse ($\chi(4) = 2.04$, $p = .729$), unprotected receptive anal intercourse ($\chi^2(4) = .2$, $p = .997$), or unprotected insertive anal intercourse ($\chi^2(4) = 4.2$, $p = .380$).

Further examination of these response patterns was undertaken to explore the structure of the HIV transmission risk construct. Inter-item correlations between the dichotomous items used to construct the ordinal measure were evaluated as were item-total correlations. Inter-item correlations for the first reported encounter (Table 37) provided interesting insights into risk behavior among the MSM in the sample.

Table 37. Summary of Item Intercorrelations[a], Encounter 1 Only

Item	*1*	*2*	*3*	*4*	*5*	*6*
Did this encounter involve…						
1. XEF_05_R1 ...a partner with a negative or unknown serostatus?						
2. XEF_13_R1 ...receptive anal intercourse?	-.05					
3. XEF_13_R2 ...unprotected receptive anal intercourse	-.17 *	.62 **				
4. XEF_15_R1 ...insertive anal intercourse?	-.27 **	-.03	.11			
5. XEF_15_R2 …unprotected insertive anal intercourse?	-.23 *	.02	.26 **	.72 **		
6. XEF_12_R3 ...oral sex?	-.04	-.29 **	-.06	-.31 **	-.11	
7. UI ...either no condom use or inconsistent condom use[b]	-.19 *	.26 **	.66 **	.27 **	.56 **	.10

**p* < .05 ***p < .01* ****p < .001*

[a] Correlations reported are Tau-B coefficients.

[b] This variable was not used in scale construction, but is reported to illustrate the dependence between other items.

Sex with a partner of negative or unknown serostatus was negatively associated with insertive (τ_B = -.27, p = .002) and unprotected receptive anal intercourse (τ_B = -.23, p = .011), however it was not significantly related to protected receptive anal intercourse (τ_B = -.05, p = .573). This pattern is consistent with the use of seroadaptive behaviors, as MSM may choose behaviors with lower risk of transmission when engaging in serodiscordant intercourse. Condom use was positively associated with both receptive (τ_B = .26, p = .004) and insertive (τ_B = .27, p = .002) anal intercourse, but was not associated with oral sex (τ_B = .10, p = .256). This response pattern could suggest that MSM are more likely to use a condom as the risk of transmission increases. Condom use (i.e., a dichotomized item reflecting any reported unprotected intercourse) was negatively associated with serodiscordance (τ_B = -.19, p = .036), suggesting that MSM may be more likely to engage in unprotected sex with seroconcordant partners (i.e., serosorting).

Non-parametric correlations between the items used to construct the ordinal measure and participants' ordinal scores on the measure itself were then computed using the Somer's *D* coefficient. The results of this analysis, by encounter number, are provided in Table 38. Positive correlations between risk factors and the overall score are indicative of a correspondence between the HIV transmission risk behavior, and the ordinal scaling of transmission risk. Significant positive relationships between participants' scores on each item and the ordinal measure of HIV transmission risk were observed in most cases. However, item-total correlations between oral sex and the

Table 38. Correlations[a] Between Ordinal Scores and Items, By Encounter

Item No.	Characteristic	Encounter Number					Across All Encounters
		1 (n_p = 128)	2 (n_p = 111)	3 (n_p = 91)	4 (n_p = 73)	5 (n_p = 66)	(n_e = 469)
XEF_05_R1	Partner negative serostatus	.29 **	.41 ***	.25 *	.25	.37 **	.32 ***
XEF_13_R1	Receptive anal intercourse	.62 ***	.64 ***	.59 ***	.61 ***	.57 ***	.61 ***
XEF_13_R2	Unprotected receptive anal intercourse	.54 ***	.59 ***	.48 ***	.59 ***	.50 ***	.54 ***
XEF_15_R1	Insertive anal intercourse	.45 ***	.40 ***	.47 ***	.45 ***	.52 ***	.46 ***
XEF_15_R2	Unprotected insertive anal intercourse	.49 ***	.39 **	.40 **	.41 **	.47 ***	.43 ***
XEF_12_R3	Oral sex	-.39 ***	-.48 ***	-.50 ***	-.38 **	-.27	-.41 ***

*p < .05. **p < .01. ***p < .001.
[a] Correlations reported are asymmetric Somers' D coefficients.

participants' scores on the ordinal measure of HIV transmission risk were negative in all cases except with the fifth reported encounter. This suggests that those engaging in oral sex could be less likely to engage in the higher risk behaviors (i.e., anal intercourse) that result in higher ordinal risk scores.

Internal consistency reliability of items used to construct the scale was also evaluated. Corrected point-biserial item-total correlations for each item, as well as Cronbach’s alpha reliability coefficients are reported in Table 39. The point-biserial correlation is a parametric statistic which relates participants’ item scores to a simple scale total computed by adding the item scores. The use of a parametric correlation coefficient with an ordinal variable is not typical; however, descriptive analyses of participants’ summed scale scores revealed that these summed scores were normally distributed. Examination of these item-total correlations revealed negative relationships between participants’ summed scale scores and both partner negative serostatus and oral sex. Cronbach’s alpha, the measure of internal consistency reliability, was also very low for each of the reported encounters. Negative correlations between the item score and the sum of the items scores suggests that participants engaging in those specific risk behaviors having lower summed scores. For example, encounters involving partners with a negative or unknown serostatus are less likely to involve unprotected anal intercourse than are encounters involving lower risk positive partners. In item analysis, this score pattern is observed when the scale is not strictly additive.

Table 39. Corrected Item-Total Correlations, By Encounter

Item No.	Characteristic	Encounter Number 1[a] (n_p = 128)	2[b] (n_p = 111)	3[c] (n_p = 91)	4[d] (n_p = 73)	5[e] (n_p = 66)	Across All Encounters [f] (n_e = 469)
XEF_05_R1	Partner negative serostatus	-.29	-.24	-.41	-.31	-.22	-.30
XEF_13_R1	Receptive anal intercourse	.10	.21	.15	.25	.23	.18
XEF_13_R2	Unprotected receptive anal intercourse	.40	.40	.42	.48	.44	.42
XEF_15_R1	Insertive anal intercourse	.03	.15	.24	.10	.13	.12
XEF_15_R2	Unprotected insertive anal intercourse	.31	.32	.39	.43	.28	.34
XEF_12_R3	Oral sex	-.32	-.29	-.35	-.24	-.15	-.28

[a]Cronbach's α = -.02. [b]Cronbach's α = .14. [c]Cronbach's α = .03. [d]Cronbach's α = .17. [e]Cronbach's α = .22.
[f]Cronbach's α = .09.

Methods based in item-response theory were then used to evaluate the dimensionality of the items used to construct the ordinal scale. The purpose of this investigation was to assess the relatively difficulty with which each of the sexual behaviors was endorsed by participants and, if taken together, the underlying items represented a Guttman-like scale of risk behavior. Prior to conducting the analysis, the structure of several of the original scale items had to be changed slightly to eliminate dependencies between items. First, participants were asked whether or not the encounter involved anal intercourse. Had they answered "no" to that item, they would not have received items on positioning and condom use during anal intercourse. Second, where anal intercourse and been receptive, a positive response meant that they were presented with an item asking about condom use in that position. A similar item on insertive anal intercourse resulted in the same branching. In order to eliminate dependencies, the information obtained from these items was recoded as shown in Table 40. Examination of these frequencies suggests potential problems with response patterns across participants. In a Guttman-based scale, items endorsed by the largest percentage of participants are considered "easy" items. Item difficulty increases as the frequency of endorsement decreases. In this framework, the easiest item to endorse was serodiscordant partnering, followed by receptive anal intercourse. As a scale of HIV transmission risk, this suggests that the least risky behavior a participant can engage in serodiscordant partnering. Next in order of increasing risk was receptive intercourse, followed by unprotected sex (i.e., inconsistent condom use, no condom use).

Table 40. Modified Dichotomous Risk Behaviors[a]

Item	Behavior	n_p [b]	%
XEF_05_R1	Did this encounter involve a partner with a *negative or unknown* serostatus?		
	Yes	75	58.6
	No	53	41.4
XEF_13_R1	Did this encounter involve *receptive* anal intercourse?		
	Yes	69	53.9
	No	59	46.1
UI_ALT	Did this encounter involved unprotected sex?		
	Yes	56	43.8
	No	72	56.3
XEF_12_R3	Did this encounter involve *oral* sex?		
	Yes	50	39.1
	No	78	60.9
XEF_15_R1	Did this encounter involve *insertive* anal intercourse?		
	Yes	41	32.0
	No	87	68.0

[a] Results are arranged in descending order by prevalence
[b] Results are reported for encounter 1 only (n_p = 128)

The pattern of response was further analyzed through of the observed data to the Rasch model. Output from this analysis is shown in Figure 3. The item misfit order table provides a summary of each of the items used in the scale, the number of endorsements for each, and the logit estimation of item difficulty.

TABLE 10.1 REED_DISS_ENCOUNTER ONE ONLY_022512.s ZOU884WS.TXT Feb 25 18:37 2012
INPUT: 128 PERSON 5 ITEM REPORTED: 128 PERSON 5 ITEM 2 CATS WINSTEPS 3.73

PERSON: REAL SEP.: .00 REL.: .00 ... ITEM: REAL SEP.: 1.85 REL.: .77

ITEM STATISTICS: MISFIT ORDER

ENTRY	TOTAL	TOTAL		MODEL	INFIT		OUTFIT		PT-MEASURE		EXACT MATCH		
NUMBER	SCORE	COUNT	MEASURE	S.E.	MNSQ	ZSTD	MNSQ	ZSTD	CORR.	EXP.	OBS%	EXP%	ITEM
4	50	128	.23	.20	1.15	1.9	1.23	2.2	A .24	.40	61.6	68.4	XF_12_R3
1	75	128	-.71	.20	1.14	1.9	1.22	2.1	B .25	.39	64.8	67.3	XF_05_R1
3	41	128	.59	.21	1.07	.8	1.12	1.0	C .31	.39	69.6	71.8	XF_15_R1
2	59	128	-.11	.19	.94	-.9	.91	-1.1	b .46	.40	65.6	66.4	XF_13_R1
5	56	128	.00	.19	.68	-5.2	.63	-4.9	a .71	.40	87.2	67.1	UI_ALT
MEAN	56.2	128.0	.00	.20	1.00	-.3	1.02	-.2			69.8	68.2	
S.D.	11.2	.0	.43	.00	.17	2.7	.23	2.7			9.1	1.9	

Figure 3. Output from Rasch Analysis of Scale Items

With a logit difficulty of -0.71, serodiscordant partnering (shown as item XF_05_R1 in the table) is the easiest item for participants to endorse, followed by receptive anal intercourse, unprotected intercourse, and oral sex in increasing order of difficulty. The relative difficulty in endorsing the lower risk activity (i.e., oral sex) could be due either to the lack of focus on this activity on the ACASI instrument, or due to the lower frequency of this activity among the reported encounters. The most difficult item to endorse was insertive anal intercourse. This arrangement of items is not totally inconsistent with the hierarchy of risk established by the ordinal measure. In both cases, insertive anal intercourse is the most difficult/riskiest item. However, when the conditions of dependency were removed from the items in order to facilitate application of the Rasch-based model, the result was a scale that reflects behavioral frequency rather than HIV transmission risk. Further examination of the figure indicates that the fit of items to the Rasch model is poor. Items in the figure are listed in order of misfit, with items XF_12_R3 (i.e. oral sex) and XF_05_R1 (i.e., serodiscordant partner) showing the most misfit. The reason for the misfit on these items was explored through inspection of poor-fitting responses. Among those with the most unexpected responses were several participants reporting insertive anal intercourse and oral sex (i.e., the two most difficult items to endorse) yet responding that the partner was not serodiscordant (i.e., the easiest item to endorse). Other participant response demonstrating ill fit included a group of 10 participants who reported oral sex (i.e., the most difficult item to endorse) but reported no other risk factors. The results of the Rasch analysis indicate the items used to in the construction of the ordinal measure of HIV risk, as coded, are not additive. In other words, endorsement of the most "difficult" (i.e., endorsed by fewest persons) items does not equate to the greatest risk of HIV transmission. It also points to problems in using behavioral frequencies as a proxy for risk. At least in situations in which the item responses do not conform to the model, use of such a score would result in improper ranking of the risk of HIV transmission in each encounter. The misfit of these items also contributes to poor item and person reliability. The Rasch item reliability (.77) is an indicator of the adequacy of the sample size and range of item difficulty. Values close to 1.0 suggest that estimates of item difficulty and standard errors are stable. In this case, the reliability is within the acceptable range. However, person reliability, analogous to internal consistency reliability, was estimated at .00. This is indicative of a severe misfit of the item to the Rasch model and/or multi-dimensionality (i.e., the items on the "scale" are not all measuring the same thing).

To further explore the potential interactions between items, item ordered was examined separately for seroconcordant and serodiscordant encounters, and series of layered chi-squares tests were run to evaluate the statistical significance of observed differences. Results of the descriptive analysis are provided in Table 41. Response frequencies for the item on condom use (UI_ALT) demonstrate an apparent relationship between serotatus and condom use.

Table 41. Dichotomous Risk Behaviors by Partner Serostatus[a]

Item	Behavior	Seroconcordant (n_p = 53)		Serodiscordant (n_p = 75)	
		n_p	%	n_p	%
XEF_13_R1	Did this encounter involve *receptive* anal intercourse?				
	Yes	27	50.9	33	44.0
	No	26	49.1	42	56.0
XEF_13_R2	Did this encounter involve *unprotected receptive* anal intercourse				
	Yes	18	34.0	14	18.7
	No	35	66.0	61	81.3
UI_ALT	Did this encounter involved unprotected sex?				
	Yes	29	54.7	27	36.0
	No	24	45.3	48	64.0
XEF_12_R3	Did this encounter involve *oral* sex?				
	Yes	22	41.5	28	37.3
	No	31	58.5	47	62.7
XEF_15_R1	Did this encounter involve *insertive* anal intercourse?				
	Yes	25	47.2	16	21.3
	No	28	52.8	59	78.7
XEF_15_R2	Did this encounter involve *unprotected insertive* anal intercourse?				
	Yes	16	30.2	9	12.0
	No	37	69.8	66	88.0

[a] Results reported are for encounter 1 only (n_p = 128)

Specifically, unprotected sex or inconsistent condom use was more likely than protected sex in seroconcordant encounters (consistent with serosorting) and protected sex is more likely in serodiscordant encounters. Chi-square analysis of the relationships between unprotected anal intercourse (insertive and receptive) and partner serostatus were then conducted. Several significant findings were noted in these analyses. In serodiscordant encounters, unprotected sex was less likely to occur ($\chi^2(1) = 4.4$, $p = .036$) than in concordant encounters. Additionally, insertive anal intercourse ($\chi^2(1) = 9.5$, $p = .002$), and specifically unprotected insertive anal intercourse ($\chi^2(1) = 6.5$, $p = .011$) were less likely in serodiscordant encounters. With respect to receptive anal intercourse, there was no observed relationship between the serostatus of the partner and the occurrence of receptive anal intercourse ($\chi^2(1) = 0.3$, $p = .572$). However, unprotected receptive anal intercourse was less likely in serodiscordant encounters ($\chi^2(1) = 3.9$, $p = .049$). These interrelationships are consistent with seroadaptive behavior among MSM. This underscores a potential problem in using a total sum score of risk behaviors as a scale of HIV transmission risk, as some risk behaviors like partner serodiscordance are actually related to lower risk sexual behaviors. Clearly a scaling mechanism which addresses this problem is needed.

Evidence of structural validity supported the use of the ordinal scale for HIV transmission risk. Support for the validity of the scaled ordinal score arose, in part, from the lack of additivity observed in the items which were used to create the scale. Frequencies of item responses showed higher reported frequencies for higher risk behaviors (e.g., anal intercourse) than for lower risk behaviors (e.g., oral sex). This observation underscored the potential problem in using sum scores, based in the assumption of Guttman response scale, of sexual risk behaviors on an encounter-specific measure of risk. In each encounter, MSM may engage in high risk behaviors such as insertive anal intercourse without having also engaged in lower risk activities like oral sex. This pattern of behavior results in poor fit of the items to the Guttman scale. While the observed "out of risk order" position of oral sex could have been the result of a lack of focus on low risk behaviors in the ACASI instrument, further evidence of non-additivity was also observed. Specifically, participants were more likely to report safer sex with riskier partners (e.g., use of a condom with a partner who was serodiscordant) and risker sex with safe partners (e.g., insertive anal intercourse with concordant partners). As a result of this pattern in responses, it cannot be assumed that the responses are independent. Therefore, simple sum scores cannot be used to characterize encounter-specific risk, as they mask complex interactions between certain risk behaviors. In contrast, the structure of the constructed ordinal measure is based on actual risk levels supported in the literature, and is not influenced by the frequency of any given behavior, or by the dependencies between behaviors. As such, the ordinal measure appears to be a better choice for modeling encounter-specific HIV transmission risk in MSM.

Goal 6: Evaluation of Validity Considering External Factors

Two strategies were employed to establish the trait and nomological validity of the participants' scores on the ordinal measure of HIV risk transmission. First, a mulit-trait multi-method analysis was conducted. The multi-trait multi-method matrix provides a means of assessing both the convergent and discriminant validity of a measure. The MTMM matrix accomplishes this analysis through the use of a specific pattern of bivariate correlations. The two traits selected for use in this analysis were HIV transmission risk and environmental risk. As has been discussed, the ordinal measure of

HIV transmission risk was a single scale, scored from 0 to 8 as shown in Table 25. The environmental risk scale was comprised of items that describe environmental risk factors for unsafe sex. This trait was chosen because it also involved encounter-specific responses and, as was the case with the ordinal measure of HIV transmission, it involves a behavioral construct. Four risk items were selected for the environmental scale. The first item was scored 1 if the participant indicated that he had met his partner in a high risk venue (i.e., online, bar, bathhouse, public sex environment). The second item was scored positively if the participant indicating use drugs and/or alcohol in conjunction with the encounter. Third, a risk score of 1 was assigned if the partner was identified as a high risk partner (i.e., anonymous partner, exchange partner, or hookup). The final risk factor used to construct the scale was an indicator that the sexual encounter took place in a high risk place (e.g., bathhouse, public sex environment). Frequencies for the environmental risk factors involved in the first reported encounter are shown in Table 42.

Two scaling methods were selected for use in the analysis. First, the method used to construct the ordinal measure of HIV transmission risk was applied to the set of environmental risk factors to create an ordinal measure of environmental risk. Scale construction resulted in a 5 category scale. The first category was assigned a score of 0, and was assigned to any encounter in which none of the 4 risk factors were present. To construct the remaining categories, risk factors were arranged in ascending order, with the most commonly reported risks incorporated first. A risk factor of 1 was assigned to any encounter in which the partner was met in a high risk place (e.g., online, at a bar, bathhouse, or public sex environment. A score of 2 on the risk scale was assigned to encounters which also involved alcohol or drug use.

Table 42. Dichotomous Environmental Risk Factors[a]

Item	Behavior	n_p[b]	%
XEF_09_R7	Did you meet this partner online, at a bar, at a bathhouse, or in a public sex environment?		
	Yes	66	51.6
	No	62	48.4
XEF_41_R7	Did this encounter involve drugs and/or alcohol?		
	Yes	63	49.2
	No	65	50.8
XEF_01_R7	Did this encounter involve a high-risk sex partner?		
	Yes	37	28.9
	No	91	71.1
XEF_10_R7	Did you have sex with partner in a bathhouse or public sex environment?		
	Yes	8	6.3
	No	120	93.8

[a] Results are reported in descending order by prevalence
[b] Results are reported for encounter 1 only (n_p = 128)

If the partner was also identified as high risk (e.g., anonymous partner, hookup, exchange partner) a score of 3 was assigned. In the highest risk category, participants also reported having sex in a high risk place such as a bathhouse or public sex environment. The resulting scale scores for the first reported encounter are shown in Table 43. It is important to note that this ordinal scale was constructed based on observed frequencies in this sample and, unlike the ordinal measure of HIV transmission risk, is not actually based on transmission rates.

Table 43. Level of Ordinal Environmental Risk

Label	Level of risk	n_p [a]	%
0	No risk factor	62	48.4
1	Met partner in a high risk venue	36	28.1
2	Met partner in a high risk venue, and alcohol/drugs were involved	15	11.7
3	Met partner in a high risk venue, alcohol/drugs were involved, partner is high risk	13	10.2
4	Met partner in a high risk venue, alcohol/drugs were involved, partner is high risk, and had sex in a high risk environment	2	1.6

[a] Results reported are for encounter 1 only (n_p=128)

In the second scoring method, a tied rank score was calculated for each trait based on the scale items associated with HIV transmission risk and environmental risk. Tied-rank scores were computed for each item used to construct the ordinal measures for both traits. Total scores per encounter were computed by adding the tied rank scores for each scale item. These computations were conducted on the first encounter reported by each participant (n = 128). Tied ranks for each scale item on the HIV transmission risk scale are shown in Table 44. The mean score for the tied rank scale of HIV transmission risk was 322.5 (*SD* = 66.2). The minimum score assigned to any of the first reported encounters was 182.0 and the maximum score was 502.0.

Table 44. Tied Rank Scores, HIV Risk Transmission Total

Item Number	Item	Tied Rank Score	n_p[a]	%
	Did this encounter involve…			
XEF_05_R1	...a partner with a negative or unknown serostatus?			
	Yes	332.5	75	58.6
	No	99.0	53	41.4
XEF_13_R1	...receptive anal intercourse?			
	Yes	367.5	59	46.1
	No	133.0	69	53.9
XEF_13_R2	...unprotected receptive anal intercourse			
	Yes	409.0	32	25.0
	No	174.5	96	75.0
XEF_15_R1	...insertive anal intercourse?			
	Yes	386.5	41	32.0
	No	152.0	87	68.0
XEF_15_R2	…unprotected insertive anal intercourse?			
	Yes	414.0	25	80.5
	No	179.5	103	19.5
XEF_12_R3	...oral sex?			
	Yes	380.0	50	39.1
	No	145.5	78	60.9

[a] Results reported are for encounter 1 only (n_p = 128)

Tied ranks for each scale item on the environmental risk scale are shown in Table 45. The mean score on this scale was 955.87 (*SD* = 227.4). The minimum score assigned to any encounter was 637.0 and the maximum was 1575.0.

Table 45. Tied Rank Scores, Environmental Scale Total

Item Number	Item	Tied Rank Score	n_p[a]	%
XEF_09_R7	Did you meet this partner online, at a bar, at a bathhouse, or in a public sex environment?			
	Yes	342.0	66	51.6
	No	107.5	62	48.4
XEF_41_R7	Did this encounter involve drugs and/or alcohol?			
	Yes	376.5	63	49.2
	No	142.0	65	50.8
XEF_01_R7	Did this encounter involve a high-risk sex partner?			
	Yes	402.5	37	28.9
	No	168.0	91	71.1
XEF_10_R7	Did you have sex with partner in a bathhouse or public sex environment?			
	Yes	454.0	8	6.3
	No	120.0	120	93.8

[a] Results reported are for encounter 1 only (n_p = 128)

Once these scores were calculated a set of correlations were used to compare the convergent and divergent validity of the ordinal measure of HIV transmission risk. Convergent validity would be supported by significant correlations between the ordinal measure is correlated and the tied rank score for the same trait. Divergent validity of the measure would be supported by the lack of correlation between the ordinal measure and each measure of environmental risk. The results of this analysis are provided in Table 46. As was expected, the correlation between the ordinal and tied rank scores for HIV transmission risk was positive and strong ($r(126) = .63$, $p = .000$) as was the correlation between the two measures of environmental risk (*r*(126) = .82, *p* = .000). The correlations between the ordinal scores on the two traits (*r*(126) = -.02, *p* = .796) and the tied ranks scores on the two traits (*r*(126) = -.14, *p* = .124) were not significant.

Table 46. Multi-Trait Multi-Method Matrix, Encounter 1 Only (n_p = 128)

Scale	Trait	Pearson Correlation				
		1		*2*		*3*
1. Ordinal Measure	HIV Transmission Risk					
2. Ordinal Measure	Environmental Risk	-.02				
3. Tied Rank Score[a]	HIV Transmission Risk	.65	***	-.11		
4. Tied Rank Score[b]	Environmental Risk	.03		.82	***	-.05

**p* < .05. ***p* < .01. ****p* < .001.

[a]Cronbach's alpha = -.018

[b]Cronbach's alpha = .242

Additional evidence related to convergent and divergent validity was obtained through the analysis of a series of multi-level ordinal logistic regression models. To begin the modeling process, a series of bivariate correlations were run between the ordinal measure of HIV transmission risk for each reported encounter and selected participants traits (e.g., openness, outness, HIV stigma). Results of these bivariate analyses are found in Table 47.

Scores on several disclosure-related measures including disclosure behaviors, attitudes, intentions, disclosure self-efficacy, disclosure outcome expectancy, and disclosure-related regret were among those scales which were significantly related to participants' ordinal risk scores. Other scale scores which were significantly correlated to transmission risk included HIV-related stigma, depression, substance use, condom outcome expectancy, health protective sexual communication, and assertive sexual communication. Inter-correlations among these scales were then examined as a means of selecting variables for inclusion in the contextual model. High inter-correlations were observed between disclosure behaviors, attitudes, and intentions. Based on the pattern of these relationships, disclosure attitude scores were chosen to add to the contextual model. Bivariate relationships between select participant characteristics and the ordinal measure of HIV transmission risk were also explored. Results of these analyses are provided in Table 48. Among the variables explored, only global reports of multiple partners, unprotected anal intercourse, and sex without disclosure were significantly related to transmission risk.

Table 47. Correlations Between Scale Scores and Risk, By Encounter

	Encounter				
	1 (n_p = 128)	2 (n_p = 111)	3 (n_p = 91)	4 (n_p = 73)	5 (n_p = 66)
Disclosure Behavior	.012	-.152	-.170	-.164	-.336 **
Disclosure Attitudes	-.005	-.276 **	-.206	-.170	-.329 **
Disclosure Intentions	.001	-.307 **	-.283 **	-.143	-.410 **
Condom Self-Efficacy	-.083	-.067	.014	.095	-.134
Disclosure Self-Efficacy	-.131	-.284 **	-.197	-.153	-.421 **
Negotiation Self-Efficacy	.024	-.162	.008	.102	-.144
Condom Outcome Expectancy	-.165 *	-.056	-.009	-.062	-.132
Disclosure Outcome Expectancy	-.036	-.231 *	-.149	-.158	-.361 **
Negotiation Outcome Expectancy	-.004	-.134	-.075	-.044	-.059
Assertive Sexual Communication	.113	.012	-.235 *	-.194	-.209

*p < .05. **p < .01. ***p < .001.

Continued

Table 47. Continued

	Encounter				
	1 (n_p = 128)	2 (n_p = 111)	3 (n_p = 91)	4 (n_p = 73)	5 (n_p = 66)
Health Protective Communication	.104	-.116	-.071	.058	-.305 *
Sexual Compulsiveness	-.130	-.065	.159	.034	.207
Openness	.149	.077	-.073	.044	.008
Outness	.030	.069	-.006	.094	.030
Substance Use	.050	.195 *	.225 *	.158	.224
HIV-Related Stigma	.060	.095	.287 **	.232	.383 **
HIV Disclosure Regret	.042	-.005	.043	.132	.334 **
Depression	-.002	.124	.242 *	.099	.291 *
Social Support from Family	.008	-.117	-.030	-.103	.062
Social Support from Friends	.056	-.042	-.078	-.096	-.075

*p < .05. **p < .01. ***p < .001.

Table 48. Correlations[a] Between Participant Characteristics and Risk, By Encounter

	Encounter 1 (n_p = 128)	Encounter 2 (n_p = 111)	Encounter 3 (n_p = 91)	Encounter 4 (n_p = 73)	Encounter 5 (n_p = 66)
CON_D1A Participant over-reported in the global section	-.099	.048	-.101	-.105	.082
CON_D1B Participant over-reported in the encounter-specific section	-.016	-.089	-.062	-.086	-.070
MIN_1 Minority ethnicity	.045	.107	.048	.226	.061
XID_1 Sexual identification as straight or bisexual	-.028	-.136	-.197	-.196	-.023
XPRT_1R Has sex with females	-.057	-.047	-.169	-.200	.022
RLG_1 Religiosity	.004	-.087	-.093	-.011	-.036
RSTA_1A Non-monogamous	.021	.101	.031	-.064	.210
MLTPRT Reported more than one partner in the last 30 days	.217 **	.054	.078	-.054	.211
UAI_1 Global report of 1 or more encounters with unprotected anal intercourse	.329 **	.019	.143	-.033	.083
NDIS_1 Global report of one or more partners without disclosure	.164 *	-.009	.088	.060	.268 *

**p* < .05. ***p* < .01. ****p* < .001.

[a] Correlations are Spearman's rho correlations

Based on the results of the exploratory bivariate analysis, a series of multilevel proportional odds models for ordinal response variables were constructed to evaluate the contextual effects of participant variables on the ordinal measure of HIV transmission risk. Select encounter-level variables, including disclosure and environmental risk, were also included in the analysis. The purpose of these analyses was to explore convergent and divergent validity in the multilevel environment. The first step in the modeling process was to run an empty model (Model 1) to evaluate the structure of the participant data. Specifically, the empty model was created as a mean of obtaining the variance estimates necessary to compute the intraclass correlation coefficient. In the process of modeling, 5 of the 128 participants who had reported 1 or more sexual encounters were eliminated from the analysis as a result of missing scale scores. The final sample used in the analysis was composed of 123 participants, and 446 encounters. The estimate of τ_{00} (3.596) was used to compute the intraclass correlations coefficient (.52) as recommended by O'Connell et al. (2008). This value for the ICC suggests a high degree of clustering in the data stemming from variability between participants and supports the use of the multilevel modeling. The coefficients for the intercept and thresholds in the empty ordinal model were also examined. Each coefficient was statistically significant.

The first set of proportional odds models were developed to explore the contextual effects of participant characteristics (Model 2A), participant disclosure-related scale scores (Model 2B), and other participant scale scores (Model 2C) on the variability in HIV transmission risk among participants. Results of these analyses are shown in

Table 49. All scale scores were grand-mean centered during modeling. Among these models, only HIV-related stigma significantly predicted participants' HIV transmission risk. The odds ratio of 0.94 suggests that increases in HIV-related stigma resulted in a significant decrease in the odds of scoring at or below each ordinal threshold, and an increase in the odds of scoring above each threshold. This result is consistent with the idea that internalized stigma is positively related to sexual risk taking. The next proportional odds model (Model 3) explored the effects of encounter-level variables on HIV transmission risk. Variables including disclosure, partner type (e.g., hook-up, fuck-buddy), use of drugs or alcohol, encounter location (e.g., bathhouse, home), and the way in which the partner was met (e.g., online, at a bar) were used to predict sexual risk taking at the encounter level. Results of these analyses are shown in Table 51. No significant relationships were found between any of the encounter-level predictors and the ordinal measure of HIV transmission risk.

Table 49. Participant-Level Predictors of HIV Transmission Risk [a]

	Model 1	Model 2A	Model 2B	Model 2C
	Coeff (SE) OR	Coeff (SE) OR	Coeff (SE) OR	Coeff (SE) OR
Fixed effects				
Model for the intercept (β_0)				
Intercept (γ_{00})	-3.17 (0.32)*** 0.04	-2.66 (0.57)*** 0.07	-3.20 (0.32)*** 0.04	-3.21 (0.32)*** 0.04
Participant characteristics				
Multiple partners[b]		0.13 (0.48) 1.14		
Global report of unprotected anal intercourse[c]		-0.74 (0.51) 0.47		
Global report of non-disclosure[d]		-0.43 (0.46) 0.65		

*p < .05. **p < .01. ***p < .001.

[a]Restricted maximum likelihood estimation, unit specific estimates with robust standard errors

[b]Reference category is 1 partner or less reported

[c]Reference category is no global reports of unprotected anal intercourse

[d]Reference category is no global reports of non-disclosure

Continued

Table 49. Continued

	Model 1 Coeff (SE) OR	Model 2A Coeff (SE) OR	Model 2B Coeff (SE) OR	Model 2C Coeff (SE) OR
Disclosure-related scales				
Disclosure attitudes[e]			0.31 (0.35) 1.37	
Disclosure self-efficacy[e]			0.12 (0.09) 1.12	
Disclosure outcome expectancy[e]			0.12 (0.10) 1.13	
Disclosure regret[e]			0.00 (0.03) 1.00	

*p < .05. **p < .01. ***p < .001.

[a]Restricted maximum likelihood estimation, unit specific estimates with robust standard errors

[b]Reference category is 1 partner or less reported

[c]Reference category is no global reports of unprotected anal intercourse

[d]Reference category is no global reports of non-disclosure

Continued

Table 49. Continued

	Model 1 Coeff (SE) OR	Model 2A Coeff (SE) OR	Model 2B Coeff (SE) OR	Model 2C Coeff (SE) OR
Other scale scores				
Condom outcome expectancy[e]				0.05 (0.06) 1.05
Assertive sexual communication[e]				-0.00 (0.04) 1.00
Health protective sexual communication[e]				0.25 (0.39) 1.28
Substance use[e]				-.06 (0.04) 0.94
HIV stigma[e]				-0.06 (0.04)* 0.94
Depression[e]				-0.00 (0.02) 1.00

*p < .05. **p < .01. ***p < .001.

[a]Restricted maximum likelihood estimation, unit specific estimates with robust standard errors

[b]Reference category is 1 partner or less reported

[c]Reference category is no global reports of unprotected anal intercourse

[d]Reference category is no global reports of non-disclosure

Continued

Table 49. Continued

	Model 1	Model 2A	Model 2B	Model 2C
	Coeff (SE)	Coeff (SE)	Coeff (SE)	Coeff (SE)
For thresholds:	OR	OR	OR	OR
δ2	2.22	2.24	2.23	2.23
	(0.26)***	(0.26)***	(0.26)***	(0.26)***
	9.18	9.38	9.29	9.32
δ3	2.67	2.69	2.68	2.69
	(0.27)***	(0.28)***	(0.28)***	(0.27)***
	14.14	14.76	14.62	14.69
δ4	4.05	4.08	4.07	4.08
	(0.32)***	(0.32)***	(0.32)***	(0.32)***
	51.21	58.93	58.74	58.24
δ5	4.93	4.97	4.97	4.98
	(0.35)***	(0.35)***	(0.35)***	(0.35)***
	139.07	143.81	144.31	146.12
δ6	5.58	5.62	5.62	5.64
	(0.39)***	(0.39)***	(0.39)***	(0.39)***
	265.68	276.92	278.32	282.43
δ7	6.74	6.79	6.79	6.82
	(0.49)***	(0.49)***	(0.49)***	(0.49)***
	846.12	890.552	896.66	920.08

*p < .05. **p < .01. ***p < .001.

[a]Restricted maximum likelihood estimation, unit specific estimates with robust standard errors

[b]Reference category is 1 partner or less reported

[c]Reference category is no global reports of unprotected anal intercourse

[d]Reference category is no global reports of non-disclosure

Continued

Table 49. Continued

	Model 1 Coeff (SE) OR	Model 2A Coeff (SE) OR	Model 2B Coeff (SE) OR	Model 2C Coeff (SE) OR
Random effects				
Variance in intercepts ($\tau 00$)	3.59***	3.67***	3.42***	3.40***

*p < .05. **p < .01. ***p < .001.

[a]Restricted maximum likelihood estimation, unit specific estimates with robust standard errors

[b]Reference category is 1 partner or less reported

[c]Reference category is no global reports of unprotected anal intercourse

[d]Reference category is no global reports of non-disclosure

Table 50. Encounter-Level Predictors of HIV Transmission Risk[a]

Fixed effects	Model 3 Coeff (SE) OR
Model for the intercept (β_0)	
Intercept (γ_{00})	-3.06 (0.54)*** 0.05
Encounter-level predictors	
Disclosure[b]	0.02 (0.45) 1.02
High risk partner[c]	-0.29 (0.28) 0.75
Use of drugs or alcohol[d]	-0.41 (0.28) 0.66
Sex in a high risk location[e]	0.87 (0.60) 2.38
Met partner in a high risk way[f]	0.11 (0.31) 1.11

*p < .05. **p < .01. ***p < .001.

[a]Restricted maximum likelihood estimation, unit specific estimates with robust standard errors

[b]Reference category is non-disclosure

[c]Reference category is not high risk partner

[d]Reference category is no drug or alcohol use

[e]Reference category is not high-risk location

[f]Reference category is not high risk meeting

Continued

Table 50. Continued

For thresholds:	Model 3 Coeff (SE) OR
δ2	2.23 (0.26)*** 9.34
δ3	2.68 (0.27)*** 14.68
δ4	4.08 (0.31)*** 59.02
δ5	4.98 (0.34)*** 144.83
δ6	5.62 (0.39)*** 276.91
δ7	6.78 (0.48)*** 882.19
Random effects	
Variance in intercepts (τ00)	3.55***

*p < .05. **p < .01. ***p < .001.

[a]Restricted maximum likelihood estimation, unit specific estimates with robust standard errors

[b]Reference category is non-disclosure

[c]Reference category is not high risk partner

[d]Reference category is no drug or alcohol use

[e]Reference category is not high-risk location
[f]Reference category is not high risk meeting

The lack of significant relationships with other important predictors of sexual risk taking was notable in the prior models, raising concern about possible violations of the proportional odds assumption. In order to evaluate the tenability of the assumption, as series of 7 logistic regression models were constructed using each of the 7 thresholds associated with the ordinal measure as the dependent variable. Given the purpose of the larger study (i.e., a randomized controlled trial of a disclosure intervention), disclosure-related scale scores and encounter-level variables were used as predictors in each model. Results of these analyses are provided in Table 51. In order for the proportional odds assumption to be supported, the effect of predictor values should be consistent across each of the 7 threshold points. With respect to the participant-level predictors this is generally the case. However, in the model predicting HIV transmission risk scores less than 7, disclosure outcome expectancy is a significant predictor of HIV transmission risk. The odds ratio (1.32) suggests that those with higher disclosure outcome expectancy are more likely to score in the lower risk category. In other words, increased scores in disclosure outcome expectancy reduce the probability of engaging in unprotected receptive or insertive anal intercourse with a discordant partner. In the model predicting HIV transmission risk scores less than 8, disclosure regret is a significant predictor of transmission risk. The odds ratio (1.08) suggests that those with higher score on the HIV disclosure regret scale have a higher probability of scoring 7 or less than of scoring 8 on the ordinal measure.

Table 51. Cumulative Analyses[a] for an Ordinal Model of Risk Category

Fixed effects	$R_{ij}<2$ Coeff (SE) OR	$R_{ij}<3$ Coeff (SE) OR	$R_{ij}<4$ Coeff (SE) OR	$R_{ij}<5$ Coeff (SE) OR	$R_{ij}<6$ Coeff (SE) OR	$R_{ij}<7$ Coeff (SE) OR	$R_{ij}<8$ Coeff (SE) OR
Model for the intercept (β_0)							
Intercept (γ_{00})	-2.68 (0.72)*** 0.07	-0.02 (0.45) 0.98	-0.02 (0.45) 0.98	0.67 (0.53) 1.94	1.81 (0.60)** 6.11	1.66 (0.64)* 5.27	2.29 (0.69)** 9.85
Disclosure-related scales							
Attitudes[e]	0.95 (0.45) 2.58	0.15 (0.32) 1.16	0.38 (0.33) 1.46	0.34 (0.44) 1.40	0.31 (0.43) 1.37	0.14 (0.43) 1.15	-0.26 (0.48) 0.77
Self-efficacy[e]	-0.13 (0.11) 0.88	0.13 (0.08) 1.14	0.06 (0.08) 1.07	0.10 (0.10) 1.10	0.10 (0.09) 1.11	0.06 (0.11) 1.06	0.17 (0.13) 1.18
Outcome expectancy[e]	0.06 (0.15) 1.06	0.07 (0.10) 1.07	0.08 (0.10) 1.08	0.02 (0.12) 1.02	0.08 (0.12) 1.08	0.27 (0.11)* 1.32	0.13 (0.14) 1.14
Regret[e]	0.02 (0.03) 1.02	-0.03 (0.02) 0.97	-0.02 (0.02) 0.98	-0.02 (0.03) 0.98	-0.03 (0.03) 0.97	0.02 (0.03) 1.02	0.08 (0.04)* 1.08

*p < .05. **p < .01. ***p < .001.

[a]Restricted maximum likelihood estimation, unit specific estimates with robust standard errors

[b]Reference category is non-disclosure

[c]Reference category is not high risk partner

[d]Reference category is no drug or alcohol use

[e]Reference category is not high-risk location

[f]Reference category is not high risk meeting

Continued

Table 51. Continued

Encounter-level predictors	R_{ij}<2 Coeff (SE) OR	R_{ij}<3 Coeff (SE) OR	R_{ij}<4 Coeff (SE) OR	R_{ij}<5 Coeff (SE) OR	R_{ij}<6 Coeff (SE) OR	R_{ij}<7 Coeff (SE) OR	R_{ij}<8 Coeff (SE) OR
Disclosure[b]	0.45 (0.72) 1.02	-0.95 (0.47) 0.38	-0.56 (0.45) 0.57	0.34 (0.45) 1.41	0.05(0.52) 1.05	0.32 (0.56) 1.37	1.51 (0.61)* 4.56
High risk partner[c]	-0.49 (0.51) 0.75	-0.32 (0.33)* 0.73	-0.29 (0.34) 0.75	-0.71 (0.34)* 0.49	-0.59 (0.36) 0.55	0.17 (0.34) 1.19	0.27 (0.50) 1.31
Use of drugs or alcohol[d]	-0.31 (0.45) 0.66	0.04 (0.30) 1.05	0.13 (0.29) 1.15	-0.40 (0.33) 0.67	-0.41 (0.35) 0.66	-0.17 (0.37) 0.84	-0.85 (0.44) 0.43
Sex in a high risk location[e]	1.41 (0.73) 2.38	1.03 (0.61) 2.80	0.92 (0.59) 2.52	1.20 (0.56)* 3.33	0.38 (0.55) 1.46	-0.08 (0.68) 0.92	-0.84 (0.95) 0.43
Met in a high risk way[f]	-0.53 (0.51) 1.11	-0.01 (0.31) 0.99	0.03 (0.30) 1.02	0.30 (0.37) 1.34	0.11 (0.40) 1.11	0.40 (0.43) 1.49	0.35 (0.55) 1.42
Randome effects, variance in intercepts (τ00)	2.97***	1.75***	1.79***	2.73***	2.53***	2.60***	3.18***

*p < .05. **p < .01. ***p < .001.

[a]Restricted maximum likelihood estimation, unit specific estimates with robust standard errors

[b]Reference category is non-disclosure

[c]Reference category is not high risk partner

[d]Reference category is no drug or alcohol use

[e]Reference category is not high-risk location

[f]Reference category is not high risk meeting

Specifically, increased regret of disclosure was associated with a lower probability of engaging in unprotected insertive anal intercourse with a discordant partner. Inconsistency in the coefficients of the encounter-level predictors was also observed. Sex with a high risk partner was negatively associated with scoring less than 5 on the ordinal measure. The odds ratio (0.49) suggests that encounters involving sex with a high-risk partner (e.g., anonymous partner, trick, hookup) were more likely to be scored at risk level of 5 or above. Given that all scores above 4 are associated with discordant partnering, this is indicative of a greater probability of engaging in discordant sex when the partner is from a high risk category. Sex in a high risk location (e.g., bathhouse, park, rest stop) was also related to scoring less than 5 on the ordinal measure. The odds ratio (3.53) suggests that persons engaging in sex in high risk environments are more likely to report serconcordance in the encounter. Finally, the significant odds ratio for disclosure (4.56) indicates that those who disclose at the encounter, or who have disclosed to the partner prior to the encounter, are less likely to score an 8 in HIV transmission risk (i.e., unprotected insertive anal sex with a discordant partner) than to score a 7 or less. The variability observed in the coefficient associated with the encounter-level variable disclosure across thresholds of the ordinal measure did not support the tenability of the proportional odds assumption relevant to the prior analyses.

As a result of the investigation of validity related to external factors, the validity of the participants' scores on the ordinal measure of HIV transmission risk was supported. The multi-trait multi-method matrix demonstrated a greater correlation between parallel measures of the same trait, than between unrelated characteristics measured using the same type of measure. Initial models predicting ordinal risk of HIV transmission with a variety of participant and encounter characteristics thought to be related to risk did not support those expected relationships. However, the assumption of proportional odds associated with the models used was violated. Several characteristics, including condom outcome expectancy, serostatus disclosure regret, and disclosure itself were found to have different effects on the ordinal score across the different thresholds of the ordinal scale. These observed patterns were consistent with the use of seroadaptive behaviors in MSM, and support the use of a measure which is sensitive to those behaviors.

Goal 7: Evaluation of the Potential for Generalizability

The interpretations of participants' scores on the ordinal measure of HIV transmission risk are generalizable to the extent that they can be applied to other times, other parallel tests, and to other HIV positive MSM. This question of generalizability involves both the items used to construct the ordinal measure and the measure itself. Evidence related to generalizability of score interpretations was based on analyses already conducted in conjunction with other study goals and will be summarized here.

Evidence supporting the generalizability of the scores on the ordinal measure of HIV transmission risk was found in the analysis of content relevance. Very low observed frequencies of missingness (Table 30) due to participant item skipping suggest that the instrument is easy to complete, and that scores obtained on the constructed ordinal measure of risk should be representative of participant behavior. Additionally, the lack of an observed relationship between missingness from either skips or N/A responses and sexual activity (Tables 18 and 29) is consistent with the comparable reporting of participants who are highly sexually active and those who are not. From the section on technical quality, the SMOG count of the instrument (10.2) supported the use of the instrument for adult HIV positive MSM with at least a 10^{th} grade education. The sample which was used in this study was comprised in large part of participants with at least some high school (*n* = 143, 98.6%), therefore the SMOG count was considered acceptable. However, the items on the ACASI instrument might not be appropriate for participants with low levels of literacy. Additionally, no significant relationships were found between partner education level and ordinal HIV transmission risk or item missingness (See Table 33). The analysis of substantive validity also provided evidence of the potential for generalizability. While some response inconsistency between global reports of sexual behavior and encounter-specific reports of sexual behavior were observed, no significant relationships were found between over-reporting on either the global or the encounter-specific side of the instrument and scores on the ordinal measure of HIV transmission risk.

Evidence supporting the generalizability of the ordinal measure of HIV transmission risk was found in the results from analyses associated with several other aspects of the unified validity. However, not all of the results supported the argument for generalizability. The lack of consistency between participant responses on the global and encounter-specific sections of the ACASI instrument suggests that some type of construct irrelevant variability in item interpretation is present. In another analysis associated with the substantive aspect of validity, a relationship was found between participant ethnicity and the consistency in response across the global and encounter-specific sections of the ACASI instrument. Though the relationship between minority status and response consistency was not significant in the multilevel analysis, further study of the interactions between language, ethnicity, and meaning should be conducted.

Goal 8: Evaluation of Consequential Validity

The relative consequential validity of the participants' scores on the ordinal measure of HIV transmission risk was demonstrated by comparing results on this measure to the alternative forms available including the global counts measures of risk, the tied rank total scores, and the dichotomous measures of unprotected receptive and insertive anal intercourse. The goal of these analyses was to evaluate the sensitivity, specificity, and utility of the ordinal measure, relative to other available scoring mechanisms, as a measure transmission risk, and as an outcome measure for the evaluation of intervention effectiveness. These analyses were conducted on data from the first reported sexual encounter. Only participants reporting one or more encounters (n_p = 128) were included. Special attention in the analysis was paid to the evaluation of the effect of disclosure on HIV transmission risk.

To facilitate these comparisons, participants' scores on a variety of alternative measures were computed. Alternative measures of risk from the global portion of the ACASI instrument included several dichotomized measures indicating whether or not the participant reported one or more encounters involving unprotected anal intercourse, unprotected receptive anal intercourse, or unprotected insertive anal intercourse. These dichotomies were created based on participant response to count items (i.e., *How many of* *these sexual encounters involved receptive anal sex (you were the bottom) without a condom?).* Next, a total of these unprotected risk factors was computed. The range of this scale was from 0 to 5, with the maximum score assigned to participants who reported engaging in one or more encounters involving unprotected receptive anal, insertive anal, receptive oral, insertive oral, and vaginal intercourse.

Table 52. Risk Scores Using Global Measures

Measure	n_p[a]	%
Dichotomous		
Had unprotected anal intercourse	78	60.9
Had unprotected receptive anal intercourse	55	43.0
Had unprotected insertive anal intercourse	58	45.3
Total number of unprotected risk factors reported		
0	13	10.2
1	21	16.4
2	33	25.8
3	30	23.4
4	29	22.7
5	2	1.6
Counts	*M*	*SD*
Number of encounters with unprotected insertive anal intercourse	1.78	5.01
Number of encounters with unprotected receptive anal intercourse	2.44	7.05
Total number of encounters with unprotected anal intercourse	4.126	10.9

[a]Results are reported for encounter 1 only

Count variables were also created which reflected the number of encounters in the last 30 days which involved unprotected insertive anal and receptive anal intercourse, as well as a total of all encounters involving unprotected anal intercourse. A summary of participants' scores on these measures is provided in Table 52.

Alternative risk scores were also computed on the encounter-specific portion of the ACASI instrument. Dichotomous measures of unprotected anal intercourse regardless of position, unprotected receptive anal intercourse, and unprotected insertive anal intercourse were established. Additionally, tied rank scores based on participants' responses to 5 encounter-specific items were computed. Tied ranks were then summed to obtaine a participant ranking across the 5 risk factors.

Table 53. Risk Scores Using Encounter-Specific Measures

Measure	n_p[a]	%
Dichotomous		
Had unprotected anal intercourse	56	43.8
Had unprotected receptive anal intercourse	32	25.0
Had unprotected insertive anal intercourse	25	19.5
Counts		
Total of tied rank scores		
Rank 1 (score = 182)	2	1.6
Rank 2 (score = 246)	29	22.7
Rank 3 (score = 310)	54	42.2
Rank 4 (score = 374)	31	24.2
Rank 5 (score = 438)	12	9.4

[a]Results are reported for encounter 1 only

Risk factors included in this analysis included partner negative/unknown serostatus, unprotected intercourse (i.e., no condom use, inconsistent condom use), oral sex, receptive anal intercourse, and insertive anal intercourse. A summary of participants' scores on the alternative encounter-specific measures is provided in Table 53.

Relationships between these alternative measures and the ordinal measure of HIV transmission risk were then explored. Results of this analysis are provided in Table 54. Strong relationships between the alternative measure and the ordinal measure of HIV transmission risk suggest that the two scales are measuring a similar construct.

Table 54. Correlations[ab] Between Risk Measures

Alternative measure	r_s	
Global		
Dichotomous		
Had unprotected anal intercourse	.14	
Had unprotected receptive anal intercourse	.10	
Had unprotected insertive anal intercourse	.08	
Total number of unprotected risk factors reported	.09	
Counts		
Number of encounters with unprotected receptive anal intercourse	.11	
Number of encounters with unprotected insertive anal intercourse	.10	
Total number of encounters with unprotected anal intercourse	.17	
Encounter-specific		
Dichotomous		
Had unprotected anal intercourse	.40	**
Had unprotected receptive anal intercourse	.41	**
Had unprotected insertive anal intercourse	.34	**
Counts		
Total of tied rank scores	.60	***

p = .006, *p < .05 **p < .01 ***p < .001

[a] Coefficients are Spearman's rho correlations

[b]Results are reported for encounter 1 only (n_p = 128)

Notably, none of the global dichotomies, nor the scores on the global total risk scale, were significantly related to the ordinal measure. This was also true of the global counts of unprotected intercourse. Relationships with alternative measures on the encounter-specific section of the instrument were both positive and significant. The nature of these relationships is consistent with the fact that the ordinal scale was constructed from participants' responses to these items. Interestingly, the correlation between the encounter-specific tied rank scores and the ordinal measure were strong ($r_s(126) = .60$, $p = .000$).

Analyses to establish the relative sensitivity and specificity of the encounter-specific measures, including the ordinal measure of HIV transmission risk, were then conducted. The goal of this analysis was to evaluate the extent to which these alternative measures effectively "diagnosed" participant risk behavior. To facilitate comparison among the large number of measures, the ordinal measure of HIV transmission risk was regarded as the "gold standard" to which all others were compared. Three levels of risk comparison were of interest in the analysis. First, a positive "diagnosis" was assigned to participants who reported unprotected insertive anal intercourse with a discordant partner (i.e., a risk score of 8; $n = 10$). These positives were compared to those who exhibited behavior consistent with *seropositioning* (i.e., a risk score of 7; $n_p = 10$). Second, a positive "diagnosis" was assigned to participants reporting any unprotected anal intercourse with a discordant partner (i.e., a risk score of 7 or 8; $n_p = 20$). These positives were compared to those who exhibited behavior consistent with *serosorting* (i.e., a risk score of 4; $n_p = 25$). In the final comparison, a "diagnosis" was assigned to those who participated in any unprotected anal intercourse (i.e., a risk score of 4, 7, or 8; $n_p = 45$).

These positives were compared to those who reported any other type of sexual activity (n_p = 83). Results of these analyses are provided in Table 55.

A computed sensitivity close to zero suggests that the measure does not effectively identify participants who are positive. Low specificity suggests that the measure does not effectively identify participants who are negative. The predictive value for a positive result reflects the proportion of participants who are identified by the measure as positive who are actually positive. For the predictive value for a negative result, the value reflects the proportion of participants who are actually negative. In the detection of seropositioning and serosorting, it was expected that the relative performance of the dichotomous measures would be poor, as they do not contain serostatus information. In the case of the third comparison, the performance of these items is limited only by the positioning information they contain. For example, all those flagged for having unprotected insertive anal intercourse were correctly diagnosed as having unprotected anal intercourse. However those who were diagnosed as negative on that measure may not have truly been negative, as they could have engaged in unprotected receptive anal intercourse. Generally, the performance of the dichotomous indicators improved as the nature of the diagnosis became more general.

Table 55. Relative Performance of Alternative Measures[a]

Alternative Measure	Sensitivity	Specificity	Predictive Value for a Positive Result	Predictive Value for a Negative Result
Comparison 1 (Seropositioning vs. No seropositioning)				
Dichotomous				
Had unprotected anal intercourse	1.00	.00	.50	
Had unprotected insertive anal intercourse	.90	1.00	1.00	.91
Had unprotected receptive anal intercourse	.30	.00	.23	.00
Counts				
Total of tied rank scores (Highest rank vs. others)	.00	.92	.00	.58
Comparison 2 (Serosorting vs. No serosorting)				
Dichotomous				
Had unprotected anal intercourse	1.00	.00	.44	
Had unprotected insertive anal intercourse	.65	.28	.42	.50
Had unprotected receptive anal intercourse	.45	.36	.36	.45
Counts				
Total of tied rank scores (Highest two ranks vs. others)	.25	.48	.28	.44

[a]Results reported are for encounter 1 only (n_p = 128)

Continued

Table 55. Continued

Alternative Measure	Sensitivity	Specificity	Predictive Value for a Positive Result	Predictive Value for a Negative Result
Comparison 3 (Unprotected Anal Intercourse vs. Other)				
Dichotomous				
Had unprotected anal intercourse	1.00	1.00	1.00	1.00
Had unprotected insertive anal intercourse	.56	1.00	1.00	.81
Had unprotected receptive anal intercourse	.70	1.00	1.00	.85
Counts				
Total of tied rank scores (Highest three ranks vs. others)	.71	.35	.37	.69

[a]Results reported are for encounter 1 only (n_p = 128)

Examination of the performance of the total scores for the tied ranks was poor for each of the three comparisons. This score ranks participants in descending order, from those who engaged in all of the risk behaviors, to those who did not engage in any of the risk behaviors. Differences in the performance of this measure relative to the ordinal measure stem from the structure of the construct. As was discussed in the results for structural validity, participant scoring on these items reflects the frequency of the reported behavior, rather than the risk of HIV transmission. Higher rank scores are assigned to encounters involving activities that are less frequent, rather than to activities which carry a greater risk of transmission. Sensitivity and specificity results shown reflect the consequences of using such a measure as an indicator of risk reduction, given that the two underlying structures do not match. In the case of the sample used in this study, only 37% of those ranking at level 3or higher (n_p = 97, 76%) were actually engaging in unprotected sex, and none of those receiving the highest risk ranking (level 5; n_p = 12, 9.4%) actually engaged in the activity associated with the highest risk of HIV transmission (i.e., unprotected insertive anal intercourse with a serdiscordant partner).

The investigation of consequential validity was based on the appropriateness of the use of scores as an outcome measure in the test of a HIV transmission risk-reduction intervention specific to HIV positive MSM. Of particular interest was the ability of the measure to detect seroadaptive changes in sexual behavior which could be adopted by MSM as a result of the intervention. A variety of alternative measures similar to those typically used to evaluate intervention effectiveness were computed, and their sensitivity and specificity compared to the proposed ordinal measure of HIV transmission risk.

Alternative measures constructed from data collected in the global section of the ACASI instrument did not correlate well with scores on the ordinal measure, and were dropped from the analysis. Poor performance in the encounter-specific dichotomous measures, and in the encounter-specific tied rank scores, suggest that the proposed ordinal measure of HIV transmission risk is preferable to these alternatives as a measure of risk reduction.

Summary

Evidence of validity for the ordinal measure of HIV risk transmission was gathered from a variety of analyses applied to the data set from a larger study of the effectiveness of an intervention designed to assist HIV positive MSM in disclosing their serostatus to their casual sexual partners. The data set included data on the participants in the study (n_p = 145) and on each participants' most recent sexual encounters (i.e., last 5 encounters in the prior 30 day period). The goal of this investigation was to ascertain the extent to which the construct validity of the scores obtained on the proposed ordinal measure, and the appropriateness of their use as an outcome measure for the test of intervention effectiveness, were supported by the procedures employed to develop the ACASI measure used to collect the data, as well as the data set obtained at the baseline observation. The theoretical approach to validation was informed by the unified conception of validity (Messick, 1995).

Evidence of validity obtained from the investigation support the use of the proposed ordinal measure of HIV transmission risk as a measure of risk in HIV positive MSM. Particular strengths of the measure include its sensitivity to seroadaptive behaviors (e.g., serosorting and seropositioning) that might be employed by MSM as a way of reducing the risk of transmission. Failure to capture these types of risk reduction behaviors, as evidenced by the reduced sensitivity of commonly used measures of risk, could negatively impact the ability of researchers to evaluate the effectiveness of interventions designed to reduce HIV transmission risk in MSM populations. The constructed measure was also found to perform better as a measure of encounter-specific risk than alternative proxy measures based on the sum of observed risk factors. Due to the fact that sexual behavior at the encounter level might not conform to the assumption of additivity which apply to such measures, a constructed measure, such as the one proposed, can be used to overcome this issue.

Opportunities to improve the validity of the measure were also identified. These opportunities generally pertain to improved item construction. In the analysis of substantive validity, considerable differences were found between participant-reports of sexual activity on the global and encounter-specific sections of the ACASI instrument. These inconsistencies resulted in both over-reporting on the global section, and over-reporting on the encounter-specific section. The ability to investigate this observed response behavior was limited by the nature of the ACASI instrument. The possibility exists that the recall process used by participants to report sexual activity was different in the two sections, resulting in different reported activities.

CHAPTER 5: DISCUSSION

The validity of a measure such as the ordinal measure of HIV transmission risk which was the subject of this study is not an “all or nothing” concept. In his seminal chapter on validity, Samuel Messick (1995) wrote, “It is important to note that validity is a measure of degree, not all or none.” (p. 13). In fact, the validity I refer to is not a quality of the measure itself, but rather of the scores which result from its administration. The data obtained from measurement are only valid insofar as they support inference, action, and decision-making (Messick, 1993). Our skills in measurement, our underestanding of the validity the resulting data, and our abilities to draw inferences from those data must remain the subjects of study and of continuous improvement. Yet it is more than a desire to improve that drives the effort to improve measurement. Valid measurement is a prerequisite to any meaningful quantitative research, and invalid measurement is a potential root cause for the failure of many important research efforts.

Purpose and Goals

The specific measurement-related problem addressed by this study was the lack of available measures of sufficient precision to accurately reflect the seroadaptive behaviors of HIV positive MSM, and the resulting difficulty in evaluating the effectiveness of interventions designed to decrease the risk of HIV transmission. The potential usefulness of participants' scores obtained from the ordinal measure of HIV transmission risk, proposed by Osmond (2007) and modified for use in this study, was evaluated through a series of analyses conducted to provide evidence relevant to the argument for score validity. Eight goals were established for this undertaking - each focused on a singular aspect of construct validity. First, evidence of content representativeness was sought. This evaluation sought to determine whether the constructed scale accurately reflected a continuum of HIV transmission risk, and to ascertain whether the entire range of risk was adequately reflected by participant scores. Second, the content relevance of the scores was examined. Third, the technical quality of the obtained measurement was evaluated. Next, the substantive validity of the scores obtained from the measure was explored. The fifth goal involved a study of the structural validity of the resulting risk scores. In the next portion of the study, the validity of the scores relevant to external factors was investigated. The seventh goal was to evaluate the potential for generalizability of the scores to other contexts and samples. The final goal of the study was to explore the consequential validity of the participant risk scores obtained through measurement.

Discussion of Content and Technical Quality

The results of the study were promising in many ways. Beginning with the content of the measure, it was possible to craft a measure of HIV transmission risk in MSM from relatively few standard items concerning sexual behavior. Scores on the scale were constructed from existing data, demonstrating the usefulness of the constructed scale as a way of improving data analysis, even in cases where data have already been collected. With further refinement, the single ordinal measure could be used in data collection to replace a series of simpler measures, thus reducing participant burden. The measure was suitable for use in an ACASI environment, was encounter-specific, and the resulting scale scores made it possible to discriminate between high risk and low risk encounters. Additionally, scores on the ordinal measure made it possible to evaluate the rather complex behavioral concept referred to as seroadaptation. Specifically, it was possible to discern when seroadaptive behaviors including seropositioning and serosorting may have been employed by MSM to reduce the risk of transmission. The examination of construct relevance was very promising. Despite the large number of items involved in the baseline assessment, very few participant skipped responses were observed. The construction of the ACASI instrument made it possible to reduce participant burden by employing a branching structure. The nature of missingness arising from this structure was also explored. Only participant motivation to engage in protected sex was significantly related to the likelihood of response. Not surprisingly, participants who reported higher motivation to practice safe sex were less likely to report perceived unsafe behavior. While this pattern of response is consistent with consistency between participant motivation and behavior, it could also be explained by a desire to report acceptable behavior. It is also interesting to note that, despite concerns for the effect of participant burden on response, the relationship between degree of sexual activity, as evidenced by the number of encounters reported, and participant response missingness was not statistically significant.

The investigation of technical quality yielded some interesting information regarding the readability of the measure. Use of a standard readability assessment suggested that a items used to construct the ordinal measure were written at a 10th-grade reading level. This raises concern about item wording effects on the validity of scores on the ordinal measure. Significant relationships found between risk scores and participant characteristics including income and comfort using the internet add to the concern that construct-irrelevant variability resulting from item wording might be present in the resulting risk scores.

Discussion of Substantive and Structural Validity

The measure was acceptable and understandable to a wide variety of MSM, and participant non-response issues were rare. The ranked scores obtained from the ordinal of the measure provided information beyond the dichotomous scores of risk behavior which are typically used in research on HIV transmission. The resulting ranks were suitable for contrasting seroadaptive behaviors including seropositioning and serosorting. The investigation of substantive validity pointed to a high degree of concurrence between reports of sexual activity on the global and encounter-specific sections of the instrument. However, the investigation of this concurrence was limited by the characteristics of the instrument, and of the items it contained. Only a limited number of encounters could be compared. Participants were only able to enter specific information on the last five sexual encounters. For those who were more sexually active, there was no way to definitively evaluate concurrence between the two forms of reporting. Additionally, global reports were not partner specific, and encounters were not identifiable by partner, so it was not possible to evaluate reports of disclosure on the two sections of the instrument. Response inconsistencies were found to be related to HIV disclosure regret and health protective communication. The underlying cause of these inconsistencies is not immediately apparent. These relationships could be spurious, in which case further investigation is needed to establish the source of the apparently construct-irrelevant variability in reporting. If they are not spurious, they could be a sign that the items on the instrument are biased. These findings have implications for research, as the validity of the scores on the ordinal measure of HIV transmission risk is contingent on the accurate reporting of sexual activity in individual encounters. If participants either over-report or under-report sexual activity, the data obtained from measurement will not be valid for use in determining the effectiveness of risk reduction interventions.

As a part of the investigation of structural validity, the scores obtained on the ordinal measure of HIV transmission risk were compared to the structure of the items used to construct the measure. The results of this analysis were informative. When the items were examined individually, and ranked in order by "difficulty" (i.e., the probability of endorsement) the resulting hierarch was noticeably dissimilar than the hierarchy imposed by the ordinal measure. The two strategies for scale construction resulted in very different outcomes. If the items are recoded to remove dependencies (i.e., unprotected sex with a negative partner was recoded to an item for unprotected sex and an item for partner serostatus), the resulting structure was not Guttman-based. Participants who did not take risks associated with the "easiest" items (i.e., having a serodiscordant partner), endorsed items which were much more unusual (i.e., having oral sex). The result of these patterns of response was a poor fit of the items to the one-parameter Rasch model . This underscores the importance of the conceptual definition of HIV transmission risk. If, as in Fendrich et al. (2009) the construct is operationalized as a continuum of risk behaviors, then the use of the probability of endorsement could be appropriate. Assuming that those activities that are engaged in by very few men are indicators of greater risk of transmission (e.g., fisting, rimming, sex parties), then the use of the Rasch approach to model risk is consistent. However, if those activities that are engaged in by few are actually only indicative of sexual preferences and do not equate to a greater risk of transmitting the virus (i.e., they are not accompanied by unprotected anal intercourse with a discordant partner), then their results in a disordered risk ranking. Those with the highest scores might not have engaged in activities that pose the greatest threat for transmission of the virus. This is particularly relevant to measures of individual encounters where sexual activity can be highly targeted or limited based on the partner or context. Participants may engage only in the highest risk activities (e.g., unprotected receptive anal intercourse) with certain partners, and only low risk activities with other partners. This is particularly problematic when the concept of seroadaptation is considered. When decisions about sexual activities are actually based on partner serostatus, then the probability of endorsement of any of the individual items on the risk scale is dependent, rather than independent, of partner serostatus. Conversely, if the risk of HIV transmission is conceptualized in terms of the probability of the spread of infection, and the hierarchy of risk is imposed on the specific activities engage in (i.e., as with the ordinal measure of HIV transmission risk), the structure of the construct is maintained in the resulting scores. Those with the highest scores have engaged in the highest risk activities. This is a strength of the ordinal measure and an argument supporting the validity of the scores obtained from its administration. The measure ensures the proper ordering of encounters based on the risk of disease transmission. Additionally, it accommodates the dependencies or interactions that could exist between sexual behavior and context. By incorporating the partner serotatus into the measure, the potential interaction between seroconcordance and sexual activity (i.e., seroadaptation) can be modeled.

Discussion of Validity Considering External Factors

The relationship between scores on the ordinal measure of HIV transmission risk and participant scores on a variety of other scales was use to evaluate their trait and nomological validity. Patterns of correlations between the scores on the measure of transmission risk were appropriately divergent with scores of environmental risk. Results were mixed for the variety of other measures explored. For example, prior research suggests that sexual risk-taking could be related to disclosure (Serovich et al., 2009). This relationship was supported in encounter specific analyses, but disclosure-related variables, and other encounter-specific measures of sexual context, were not significantly related to sexual risk taking in the multilevel proportional odds model. Participant characteristics which were significant in the encounter-specific analysis (e.g., multiple partners) were also found to be unrelated to transmission risk in the multilevel proportional odds model. This could be due in part to the results of the test of proportional odds. Examination of a series of regression conducted at each of the 7 thresholds of risk indicated that the relationship of several covariates, including disclosure outcome expectancy and partner type, were not consistent. This suggests that the salient features of the partner, participant, and the context of the sexual encounter that are important in determining sexual behavior could be different depending on the level of HIV transmission risk.

Discussion of Generalizability and Consequential Validity

The potential for generalizability of the scores on the ordinal measure of HIV transmission risk to other contexts and populations was generally supported. Overall, response missingness was limited, particularly on the items required to construct the ordinal measure. Observed response inconsistencies between the global and encounter-specific sections of the ACASI instrument were of concern, however, no relationships were found between response inconsistency and HIV transmission risk. The observed relationship between response consistency and reported ethnicity was potentially indicative of some construct-irrelevant variation due to language that could be consistent with differential item functioning. These issues were specific to the items used to construct the ordinal measure and, as such, may affect the resulting ordinal scores.

With respect to consequential validity, the scores on the ordinal measure were compared to participants' scores on a variety of alternative measures. Correlations between ordinal encounter-specific scores and global measures of sexual risk were positive, but not statistically significant. A number of factors, including response inconsistency between the two sections of the instrument, could account for the lack of relationship. Only data from the first reported encounter were used. As such, the measures on the encounter-specific section measure sexual risk and those on the global side measure risk prevalence. Correlations with individual, encounter-specific items, as well as the encounter-specific tied rank score were stronger.

The potential consequences of using the alternative measures of HIV transmission risk which are commonly used in the literature suggests that, if seroadaptive behaviors are considered, exisiting measures of HIV transmission risk can provide a biased view. In the first measure comparision, the ordinal measure of HIV transmission risk was superior to the other measures in the ability to determine whether or not seropositioning had occurred. Not surprisingly, the failure of the variety of dichotomous measures to consider partner serostatus in the determination of risk resulted in predictive values for both positive and negative results which were, in some cases, very low. Even the tied rank score, which incorporated each of the risk factors involved in the construction of the ordinal measure, was complete ineffective in distinguishing between seropositioners and others. Low scores on any of these measures can be regarded as measurement error, particularly in the case that these scores are being used to make decisions about HIV transmission risk. Where the negative predictive value is low (e.g., tied ranks scores PPV = .58 in Comparison 1), the ability to correctly determine risk category is severely impaired. In the example given, of those who engaged in unprotected intercourse with a discordant partner (*n* = 20) only about one-fourth were corrected identified as seropositioners. The ability to correctly identify persons using serosorting to reduce HIV transmission risk was also low among the alternative measures. As a result of these differences, the use of one of the alternative measures as an outcome measure in the study of intervention effectiveness could lead to a serious distortion of the results.

Implications

Results of the study are promising for the refinement of measurements of HIV transmission risk, and for the understanding of seroadaptive behavior in MSM. For researchers who are seeking to demonstrate the effectiveness of interventions designed to reduce HIV transmission risk, the ordinal measure provides a means for detecting qualitative shifts in sexual activity which can be critical to the question of effectiveness. Though some might not agree that serosorting (i.e., unprotected anal intercourse, but only with seroconcordant partners) is a safe health practice, most would agree that it drastically reduces the risk of HIV transmission. If an intervention with MSM results in lowering risk by increasing serosorting, the ordinal measure of HIV risk transmission makes it possible to detect this shift. Failure to incorporate partner serostatus in the identified outcomes could lead to incorrect decisions regarding intervention use and support. As the fight against the spread of HIV continues, researchers must employ more sophisticated concepts about sexual behavior. The validation of a measure which enables researchers to capture important sources of variability in sexual behavior, and to model the relationships of potential covariates, could lead to important new discoveries about MSM as well as about the risk of HIV transmission.

Limitations and Further Research

This study was conducted on data obtained from volunteers. Though the descriptive analysis of the sample demographic characteristics indicates that the sample was diverse in many ways, self- selection bias could have resulted in a sample which is not representative of the target population of sexually active MSM. Of particular concern is the behavior of highly sexually active men who engage in risky sexual behavior. If those participants are under-represented in the sample because of an unwillingness to participants in research on disclosure, this could have affected the generalizability of this study's results. Data available for the analysis arose from the measure designed to evaluate transmission risk, not measurement validity. A limited number of items or scales were available to assess the convergent validity of the HIV transmission risk scores. Given that no gold standards exist for measuring HIV transmission risk or risky sexual behavior in MSM, and the inability to ascertain the relative correspondence of the information recollected by the participants and their actual sexual behaviors, the ability to develop evidence of construct validity was severely limited. The validity investigation was also limited by the structure and content of the ACASI instrument and the lack of process information (e.g., think aloud interviews, response latencies).

The modification of the original risk measure proposed by Osmond et al. (2007) was based on expert judgment. While the ordering of the categories of risk is generally supported by the extant research on HIV transmission, the relative reduction in risk in moving from a higher to lower risk category is widely variable, and difficult to substantiate. Further research is needed to determine whether the 8 proposed categories of risk are enough (i.e., should other activities be included in the hierarchy) or too much (i.e., is there a significant reduction in HIV transmission risk in moving from protected sex with a positive partner to oral sex) to adequately discriminate among sexual encounters. Further research should also include an investigation of the proportional odds assumption in conjunction with the effect of important encounter- and participant-level covariates. The usefulness of other models for ordinal outcomes, including partial proportional odds and nonproportional odds models should also be explored.

REFERENCES

Agresti, A. (2007). *An introduction to categorical data analysis*. Hoboken, NJ: Wiley & Sons.

Allen, J., Timmerman, L., Ksobiech, K., Valde, K., Gallagher, E. B., Hookham, L., Bradford, L., & Emmers-Sommer, T. (2008). Persons living with HIV: Disclosure to sexual partners. *Communication Research Reports, 25*(3), 192-199. doi:10.1080/08824090802237592

Allen, M. J., & Yen, W. M. (2002). *Introduction to measurement theory*. Prospect Heights, IL: Waveland Press.

Allen, T. M. (2002). HIVsuperinfection and immune control: Implications for vaccine development? *PRN Notebook, 7*(4), 1-7. Retrieved from http://www.prn.org/images/pdfs/allen_todd_v7n4_393.pdf

American Educational Research Association, Committee on Test Standards, and National Council on Measurements Used in Education. (1955). Technical recommendations for achievement tests. Washington, DC: AERA.

American Educational Research Association, American Psychological Association, & National Council of Measurement in Education (1999). *Standards for educational and psychological testing*. Washington, DC: AERA.

Amirkhanian, Y. A., Kelly, J. A., Kuznetsova, A. V., DiFranceisco, W. J., Musatov, V. B., & Pirogov, D. G. (2011). People with HIV in HAART-era Russia: Transmission risk behavior prevalence, anti-retroviral medication-taking, and psychosocial distress. *AIDS and Behavior, 15*(4), 767-777. doi:

Anastasi, A. (1954). *Psychological Testing*. New York: MacMillan.

AVERT (2011). History of AIDs up to 1986. Retrieved from http://www.avert.org/aids-history-86.htm

Barash, E. & Golden, M. (2010). Awareness and use of HIV pre-exposure prophylaxis among attendees of a Seattle Pride event and sexually transmitted disease clinic. *AIDS Patient Care and STDs, 24*(11), 689-691. doi:10.1089/apc.2010.0173

Berger, B. E., Ferrans, C. E., & Lashley, F. R. (2001). Measuring stigma in people with HIV: Psychometric assessment of the HIV stigma scale. *Research in Nursing and Health, 24*, 518-529.

Biglan, A., Metzler, C. W., Wirt, R., Ary, D., Noell, J., Ochs, L., French, C., & Hood, D. (1990). Social and behavioral factors associated with high-risk sexual behavior among adolescents. *Journal of Behavioral Medicine, 13,* 245-261.

Bird, J., Fingerhut, D., & McKirnan, D. (2011). Ethnic differences in HIV disclosure and sexual risk. *AIDS Care, 23*(4), 444-448. doi: 10.1080/09540121.2010.507757

Boffey, P. M. (1986). Surgeon general urges frank talk to young on AIDS. New York Times, October 23, 1986. Retrieved from http://www.nytimes.com/1986/10/23/us/surgeon-general-urges-frank-talk-to-young-on-aids.html.

Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2004). The concept of validity. *Psychological Review, 11,* 1061-1071.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 59,* 81-105.

Campo, J., Perea, M. A., del Romero, J., Cano, J., Hernando, V., & Bascones, A. (2006). Oral transmission of HIV, reality or fiction? An update. *Oral Diseases, 12,* 219-228. doi: 10.1111/j.1601-0825.2005.01187.x

Carpenter, K. M., Stoner, S. A., Mikko, A. N., Dhanak, L. P., & Parsons, J. T. (2010). Efficacy of a web-based intervention to reduce sexual risk in men who have sex with men. *AIDS and Behavior, 14*(3), 549-557. doi: 10.1007/s10461-009-9578-2

Catania, J. A., Gibson, D. R., Chitwood, D. D., & Coates, T. J. (1990). Methodological problems in AIDS behavioral research: Influences on measurement error and participation bias in studies of sexual behavior. *Psychological Bulletin, 108*(3), 339-362.

Catania, J. A. (1998). Health protective communication scale. In C. M. Davis, W. L. Yarber, R. Bauserman, G. Schreer, & S. L. Davis (Eds.), *Handbook of Sexuality-Related Measures* (p. 544-547). Thousand Oaks, CA: Sage Publications.

Catania, J. A., Osmond, D., Neilands, T. B., Canchola, J., Gregorich, S., & Shiboski, S. (2005). Commentary on Schroder et al. *Annals of Behavioral Medicine, 29*(2), 86-95.

Centers for Disease Control (1988). Update: Acquired immunodeficiency syndrome and human immunodeficiency virus infection among health-care workers. *Morbidity and Mortality Weekly Report*, *37*, 229-239.

Centers for Disease Control (2011a). Retrieved from http://www.cdc.gov/hiv/topics/funding/state-awards/2009/index.htm

Centers for Disease Control (2011b). Retrieved from http://www.cdc.gov/nchhstp/newsroom/docs/FastFacts-MSM-FINAL508COMP.pdf.

Centers for Disease Control (2011c). Retrieved from http://www.cdc.gov/hiv/topics/msm/challenges_partner.htm

Centers for Disease Control (2011d). Retrieved from http://www.cdc.gov/mmwr/preview/mmwrhtml/mm5521a4.htm

Centers for Disease Control (2011e). Retrieved from http://www.cdc.gov/hiv/topics/prev_prog/acdp/.

Centers for Disease Control (2011f). Retrieved from http://www.cdc.gov/hiv/topics/research/prs/resources/factsheets/healthy-relationship.htm

Centers for Disease Control (2011g). Retrieved from http://www.cdc.gov/hiv/topics/msm/index.htm.

Centers for Disease Control (2011h). CDC Fact Sheet, HIV and AIDS among gay and bisexual men. Retrieved from http://www.cdc.gov/nchhstp/newsroom/docs/fastfacts-msm-final508comp.pdf

Centers for Disease Control (2012a). Prevention Programs. Retrieved from http://www.cdc.gov/hiv/topics/prev_prog/index.htm

Centers for Disease Control (2012b). Comprehensive Risk Counseling Services. Retrieved from http://www.cdc.gov/hiv/topics/prev_prog/CRCS/index.htm

Chen, X., Lunn, S., Deveaux, L., Li, X., Brathwaite, N., Cottrell, L., & Stanton, B. (2009). A cluster randomized controlled trial of an adolescent HIV prevention program among Bahamian youth: Effect at 12 months post-intervention. *AIDS and Behavior, 13,* 499-508.

Chiasson, M. A., Shaw, F. S., Humberstone, M., Hirshfield, S., & Hartel, D. (2009). Increased HIV disclosure three months after an online video intervention for men who have sex with men (MSM). *AIDS Care, 21*(9), 1081-1089. doi: 10.1080/09540120902730013

Cheonis, N. (2005). Dual HIV infection. Retrieved from http://www.thebody.com/content/art2513.html

Ciccarone, D. H., Kanouse, D. E., Collins, R. L., Miu, A., Chen, J. L., Morton, S. C., & Stall, R. (2003). Sex without disclosure of positive HIV serostatus in a US probability sample of persons receiving medical care for HIV infection. *American Journal of Public Health, 93*(6), 949-954. doi:

Cliff, N., & Keats, J. A. (2003). *Ordinal measurement in the behavioralsSciences.* Mahwah, New Jersey: Lawrence Erlbaum Associates.

Coleman, E., Miner, M., Ohlerking, F., & Raymond, N. (2001). Compulsive sexual behavior inventory: A preliminary study of reliability and validity. *Journal of Sex & Marital Therapy, 27*, 325-332.

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Boston, MA: Houghton Mifflin Company.

Coyle, K. K., Kirby, D. B., Robin, L. H., Banspach, S. W., Baumler, E., & Glassman, J. R. (2006). ALL4YOU! A randomized trial of an HIV, other STDs, and pregnancy prevention intervention for alternative school students. *AIDS Education and Prevention, 18*(3), 187-203.

Cronbach, L. J. (1949). *Essentials of psychological testing*. New York: Harper.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52,* 281-302.

Derlega, V. J., Winstead, B. A., Greene, K., Serovich, J., & Elwood, W. N. (2004). Reasons for disclosure /nondisclosure in close relationships: Testing a model of HIV-disclosure decision making. *Journal of Social and Clinical Psychology, 23*(6), 747-767.

Des Jarlais, D. C., Paone, D., Milliken, J., Turner, C. F., Miller, H., Gribble, J., Shi, Q., Hagan, H., & Friedman, S. R. (1999). Audio-computer interviewing to measure risk behaviour for HIV among injecting drug users: a quasi-randomized trial. *Lancet, 535(9165),* 1657-1662.

Driskell, J. R., Salomon, E., Mayer, K., Capistrant, B., & Safren, S. (2008). Barriers and facilitators of HIV disclosure: Perspectives from HIV-infected men who have sex with men. *Journal of HIV/AIDS & Social Services*, *7*(2), 135-156. doi: 10.1080/15381500802006573

Eakman, A. M., & Eklund, M. (2011). Reliability and structural validity of an assessment of occupational value. *Scandinavian Journal of Occupational Therapy, 18*(3), 231-240.

Eaton, L. A., Cherry, C., Cain, D., & Pope, H. (2011). A novel approach to prevention for at-risk HIV-negative men who have sex with men: Creating a teachable moment to promote informed sexual decision-making. *American Journal of Public Health, 101*(3), 539-545. doi: 10.2105/AJPH.2010.191791.

Eaton, L. A., Kalichman, S. C., & Cherry, C. (2010). Sexual partner selection and HIV risk reduction among Black and White men who have sex with men. *American Journal of Public Health, 100*(3), 503-509.

Eaton, L. A., Kalichman, S. C., O'Connell, D. A., & Karchner, W. D. (2009). A strategy for selecting sexual partners believed to pose little/no risks for HIV: Serosorting and its implications for HIV transmission. *AIDS Care, 21*(10), 1278-1288. doi: 10.1080/09540120902803208

Eaton, L. A., West, T. V., Kenny, D. A., & Kalichman, S. C. (2009). HIV transmission risk among HIV seroconcordant and serodiscordant couples: Dyadic processes of partner selection. *AIDS and Behavior, 13*(2), 185-195. doi: 10.1007/s10461-008-9480-3

Fendrich, M., Mackesy-Amiti, M. E., Johnson, T. P., & Pollack, L. M. (2010). Sexual risk behavior and drug use in two Chicago samples of men who have sex with men: 1997 vs. 2002. *Journal of Urban Health: Bulletin of the New York Academy of Medicine, 87*(3), 452-466. doi: 10.1007/s11524-009-9432-x

Fendrich, M., Smith, E. V., Pollack, L. M., & Mackesy-Amiti, M. E. (2009). Measuring sexual risk for HIV: A Rasch scaling approach. *Archive of Sexual Behavior, 38,* 922-935.

Fox, J., & Fidler, S. (2010). Sexual transmission of HIV-1. *Antiviral Research, 85*, 276-285.

Frost, D. M., & Meyer, I. H. (2009). Internalized homophobia and relationship quality among lesbians, gay men, and bisexuals. *Journal of Counseling Psychology, 56,* 97-109.

Gadermann, A., Guhn, M., & Zumbo, B. (2011). Investigating the substantive aspect of construct validity for the Satisfaction with Life Scale Adapted for Children: A focus on cognitive processes. *Social Indicators Research, 100*(1), 37-60.

Gallagher, A., Bennet, R. E., & Cahalan, C. (2000). Detecting construct-irrelevant variance in an open-ended, computerized mathematics task. Retrieved from http://www.ets.org/research/policy_research_reports/rr-00-18

Galletly, C. L., DiFranceisco, W., & Pinkerton, S. D. (2008). HIV positive persons' awareness and understanding of their state's criminal HIV disclosure law. *AIDS and Behavior, 13*(6), 1262-1269.

Gaskins, S., Foster, P. P, Sowell, R., Lewis, T., & Parton, J. (2011). Reasons for HIV disclosure and non-disclosure: An exploratory study of rural African American men. *Issues in Mental Health Nursing, 32*, 367-373.

Global Working Group (2010). *Proven HIV Prevention Strategies*. Retrieved from http://www.kff.org/hivaids/upload/050106_HIVPreventionStrategies.pdf.

Golden, M. R., Stekler, J., Hughes, J. P., & Wood, R. W. (2008). HIV serosorting in men who have sex with men: Is it safe? *Journal of Acquired Immune Deficiency Syndrome, 49*(2), 212-218.

Gorbach, P. M., Galea, J. T., Amani, B., Shin, A., Celum, C., Kerndt, P., & Golden, M. R. (2004). Don't ask, don't tell: Patterns of HIV disclosure among HIV positive men who have sex with men with recent STI practising high risk behaviour in Los Angeles and Seattle. *Sexually Transmitted Infection, 80,* 512-517. doi: 10.1136/sti.2004.010918

Gore-Felton, C., Power, R., Duran, R., Sarlati, S., Palmer, L., Koopman, C., Israelski, D. M., Porter, J., & Spiegel, D. (2008). Psychosocial factors associated with disclosure of HIV positive serostatus: Understanding the influence of depression and social support. *Journal of HIV/AIDS and Social Services, 7*(1), 83-98.

Gray, B. T. (January, 1997). Controversies regarding the nature of score validity: Still crazy after all these years. Unpublished paper presented at the Annual Meeting of the Southwest Educational Research Association, Austin, TX.

Grossman, C. I., Forsyth, A., Purcell, D. W., Allison, S., Toledo, C., & Gordon, C. M. (2011). Advancing novel HIV prevention intervention research with MSM - meeting report. *Public Health Reports, 126*, 472-479.

Guion, R. M. (1980). On trinitarian doctrines of validity. *Professional Psychology, 11*(3), 385-398.

Guzman, R., Colfax, G. N., Wheeler, S., Mansergh, G., Marks, G., Rader, M., & Buchbinder, S. (2005). Negotiated safety relationships and sexual behavior among a diverse sample of HIV-negative men who have sex with men. *Journal of Acquired Immune Deficiency Syndrome, 38*(1), 82-86.

Haladyna, T. M., & Downing, S. M. (2004). Construct-irrelevant variance in high-stakes testing. *Educational Measurement: Issues and Practice, 23*(1), 17–27.

Halkitis, P. N., Moeller, R. W., & Pollock, J. A. (2008). Sexual practices of gay, bisexual, and other nonidentified MSM attending New York City gyms: Patterns of serosorting, strategic positioning, and context selection. *Journal of Sex Research, 45*(3), 253-261. doi: 10.1080/00224490802204456

Halkitis, P. N., Mukherjee, P, P., & Palamar, J. J. (2009). Lontiduinal modeling of methamphetamine use and sexual risk behaviors in gay and bisexual men. *AIDS and Behavior, 13*(4), 783-791.

Hedeker, D., & Gibbons, R. D. (1994). A random-effects ordinal regression model for multilevel analysis. *Biometrics, 50*, 933-944.

Hedeker, D. (2007). Multilevel models for ordinal and nominal variables. In Jan de Leeuw & E. Meijer (Eds.), *Handbook of Multilevel Analysis*. New York: Springer.

Herbst, J. H., Sherba, R. T., Crepaz, N., DeLuca, J. B., Zohrabyan, L., Stall, R. D., Lyles, C. M., & The AIDS Research Synthesis Team (2005). A meta-analytic review of HIV behavioral interventions for reducing sexual risk behavior of men who have sex with men. *Journal of Acquired Immune Deficiency Syndromes, 39*(2), 228-241.

Hien, D. A., Campbell, A. N. C., Killen, T., Mei-Chen, H., Hanswer, C., Huiping, J., Hatch-Maillette, M., Miel, G. M., Cohen, L. R., Gan, W., Resko, S. M., DiBono, M., Wells, E. A., & Nunes, E. V. (2009). The impact of trauma-focused group therapy upon HIV sexual risk behaviors in the NIDA clinical trials network “Women and Trauma” multi-site study. *AIDS and Behavior, 14*, 421-430.

Highlaymen, L. (1/16/09). Serosorting and other non-condom HIV prevention strategies for men who have sex with men. Retrieved from http://www.hivandhepatitis.com/recent/2009/011609_d.html

Holt, M., Rawstorne, P., Worth, H., Bittman, M., Wilkinson, J., & Kippax, S. (2011). Predictors of HIV disclosure among untested, HIV-negative and HIV-positive Australian men who had anal intercourse with their most recent casual male sex partner. *AIDS and Behavior, 15*, 1128-1139.

Hong, D. S., Goldstein, R. B., Rotheram-Borus, M. J., Wong, L. F., Gore-Felton, C., & The NIMH Health Living Trial Group (2006). Perceived partner serostatus attribution of responsibility for prevention of HIV transmission, and sexual risk behavior with “main” partner among adults living with HIV. *AIDS Education and Prevention*, *18*(2), 150-162.

Huang, J., Jacobs, D. F., & Derevensky, J. L. (2010). Sexual risk-taking behaviors, gambling, and heavy drinking among U. S. college athletes. *Archives of Sexual Behavior, 39*, 706-713.

Hubley, A. M., & Zumbo, B. D. (2011). Validity and the consequences of test interpretation and use. *Social Indicators Research, 103*, 219-230.

Janssen, R. S., & Valdiserri, R. O. (2004). HIV prevention in the United States: Increasing emphasis on working with those living with HIV. *Journal of Acquired Immune Deficiency Syndrome, 37*, S119-S121.

Kaiser Family Foundation (2009). 2009 survey of Americans on HIV/AIDS: Summary of findings on the domestic epidemic. Retrieved from http://www.kff.org/kaiserpolls/upload/7889.pdf

Kalichman, S. C., Cherry, C., White, D., Jones, M., Grebler, T., Kalichman, M. O., Detorio, M., Caliendo, A. M., & Schinazi, R. F. (2011). Sexual HIV transmission and antiretroviral therapy: A prospective cohort study of behavioral risk factors among men and women living with HIV. *Annals of Behavioral Medicine, 42,* 111-119.

Kalichman, S. C., Rompa, D., Cage, M., DiFonzo, K., Simpson, D., Austin, J., Luke, W., Buckles, J., Kyomugisha, F., Benotsch, E., Pinkerton, S., & Graham, J. (2001). Effectiveness of an intervention to reduce HIV transmission risks in HIV-positive people. *American Journal of Preventive Medicine, 21*(2), 84-92.

Kane, M. (2009). Validating the interpretations and uses of test scores. In R. W. Lissitz (Ed.), *The concept of validity* (pp. 39-64). Charlotte, NC: Information Age.

Klitzman, R., Exner, T., Correale, J., Kirshenbaum, S. B., Remien, R., Ehrhardt, A. A., Lightfoot, M., Catz, S. L., Weinhardt, L. S., Johnson, M. O., Morin, S. F., Rotheram-Borus, M. J., Kelly, J. A., & Charlebois, E. (2007). It's not just what you say: Relationships of HIV dislosure and risk reduction among MSM in the post-HAART era. *AIDS Care, 19*(6), 749-756. doi: 10.1080/09540120600983971

Leigh, B. C. & Stall, R. (1993). Substance use and risky sexual behavior for exposure to HIV: Issues in methodology, interpretation, and prevention. *American Psychologist, 48*(10), 1035-1045.

Marcus, J. L., McConnell, J. J., & Grant, R. M. (2005). HIV superinfection: What clinicians and patients should know: Behavioral implications/risk. *Medscape HIV/AIDS, 11*(1). Retrieved from http://www.medscape.com/viewarticle/504811_3

Marks, G., Millett, G. A., Bingham, T., Lauby, J., Murrili, C. S., & Stueve, A. (2010). Prevalence and protective value of serosorting and strategic positioning among Black and Latino men who have sex with men. *Sexually Transmitted Diseases, 37*(5), 325-327.

Mattson, C. L., Campbell, R. T., Karabatsos, G., Agot, K., Ndinya-Achola, J. O., Moses, S., & Bailey, R. C. (2010). Scaling sexual behavior or "sexual risk propensity" among men at risk for HIV in Kisumu, Kenya. *AIDS and Behavior, 14*(1), 162-172.

May, L. A., & Warren, S. (2001). Measuring quality of life of persons with spinal chord injury: Substantive and structural validity. *Quality of Life Research, 10*(6), 503-515.

McConnell, J. J., Bragg, L., Shiboski, S., & Grant, R. M. (2010). Sexual seroadaptation: Lessons for prevention and sex research from a cohort of HIV-positive men who have sex with men. *Plos One, 5*(1). Retrieved from http://www.plosone.org/article/info:doi%2F10.1371%2Fjournal.pone.0008831

McCready, K. C., & Halkitis, P. N. (2008). HIV serostatus disclosure to sexual partners among HIV-positive methamphetamine-using, gay, bisexual, and other men who have sex with men. *AIDS Education and Prevention, 20*(1), 15-29.

McCusker, J., Stoddard, A. M., Zapka, J. G., Zorn, M., & Mayer, K. H. (1989). Predictors of AIDS-preventive behavior among homosexually active men: A longitudinal study. *AIDS, 3*, 443-448.

McFarland, W., Chen, Y. H., Raymond. H. F., Nguyen, B., Colfax, G., Mehrtens, J., Robertson, T., Stall, R., Levine, D., Truong, H. H. (2011). HIV seroadaptation among individuals, within sexual dyads, and by sexual episodes, men who have sex with men, San Francisco, 2008. *AIDS Care, 23*(3), 261-268. doi: 10.1080/09540121.2010.507748

McLaughlin, G. H. (1969). SMOG grading: A new readability formula. *Journal of Reading, 12*(8), 639-646.

McLaughlin, G. H. (2008). SMOG: Simple GobbledyGook. Retrieved from http://www.harrymclaughlin.com/SMOG.htm

Merson, M. H., O'Malley, J., Serwadda, D., & Apisuk, C. (2008). The history and challenge of HIV prevention. *The Lancet*, *372*, 475-488.

Messick, S. (1995). The validity of psychological assessment: Validation of inferences from person's responses and performances as scientific inquiry into score meaning. *American Psychologist, 50*(9), 741-749.

Miller, L. C., Berg, J. H., & Archer, R. L. (1983). Openers: Individuals who elicit intimate self-disclosure. *Journal of Personality and Social Psychology, 44*(6), 1234-1244.

Montecinos, C., Rittershaussen, S., Cristina Solis, M., Contreras, I., & Contreras, C. (2010). Standards-based performance assessment for the evaluation of student teachers: A consequential validity study. *Asia-Pacific Journal of Teacher Education, 38*(4), 285-300.

Morin, S. F., Myers, J. J., Shade, S. B., Koester, K., Maiorana, A., & Rose, C. D. (2007). Predicting HIV transmission risk among HIV-infected patients seen in clinical settings. *AIDS and Behavior, 11*(1), S6-S16.

Morin, S. F., Shade, S. B., Steward, W. T., Carrico, A. W., Remien, R. H., Rotheram-Borus, M. J., Kelly, J. A., Charlebois, E. D., Johnson, M. O., Chesney, M. A., & The Health Living Project Team (2008). A behavioral intervention reduces HIV transmission risk by promoting sustained serosorting practices among HIV-infected men who have sex with men. *Journal of Acquired Immune Deficiency Syndrome, 49*(5), 544-551.

Mullen, P. D., Ramirez, G., Strouse, D., Hedges, L. V., & Sogolow, E. (2002). Meta-analysis of the effects of behavioral HIV prevention interventions on the sexual risk behavior of sexually experienced adolescents in controlled studies in the United States. *Journal of Acquired Immune Deficiency Syndromes, 30,* S94-S105.

Mutto, M., Lawoko, S., Ovugo, O., & Bangdiwala, S. (2010). Structural validity and reliability of the Integrated Conflict Scale. *International Journal of Injury Contol & Safety Promotion, 17*(3), 141-144.

Myers, J. J., Shade, S. B., Rose, C. D., Koester, K., Maiorana, A., Malitz, F. A., Bie, J., Kangu-Dufour, M., & Morin, S. F. (2010). Interventions delivered in clinical settings are effective in reducing risk of HIV transmission among people living with HIV: Results from the Health Resources and Services Administratation (HRSA)'s Special Projects of National Significance Initiative. *AIDS and Behavior, 14*, 483-492. doi: 10.1007/s10461-010-9679-y

Nagin, D. S. (1999). Analyzing developmental trajectories: A semiparametric, group-based approach. *Psychological Methods, 4*(2), 139–157. doi:10.1037/1082-989X.4.2.139.

NAM (2012). HIV risk levels for the insertive and receptive partner in different types of sexual intercourse. Retrieved from http://www.aidsmap.com/HIV-risk-levels-for-the-insertive-and-receptive-partner-in-different-types-of-sexual-intercourse/page/1443490/#item1443493

Nappi, C. M., Thakral, C., Kapungu, C., Donenberg, G. R., DiClemente, R., Brown, L., & Project STYLE Study Group (2009). Parental monitoring as a moderator of the effect of family sexual communication on sexual risk behavior among adolescents in psychiatric care. *AIDS and Behavior, 13*, 1012-1020.

National Institutes of Mental Health Multisite HIV/STD Prevention Trial for African American Couples Group (2010). Risky sexual behavior and correlates of STD prevalence among African American HIV serodiscordant couples. *AIDS and Behavior, 14*, 1023-1031.

Neumann, M. S., Johnson, W. D., Semaan, S., Flores, S. A., Peersman, G., Hedges, L. V., & Sogolow, E. (2002). Review and meta-analysis of HIV prevention intervention research for heterosexual adult populations in the United States. *Journal of Acquired Immune Deficiency Syndrome*, 30, S106-S117.

Niccolai, L. M., King, E., D'Entremont, D., & Pritchett, E. N. (2006). Disclosure of HIV serostatus to sex partners: A new approach to measurement. *Sexually Transmitted Diseases, 33*(2), 102-105). doi: 10.1097/01.olq.0000194591.97159.66

Noar, S. M. (2008). Behavioral interventions to reduce HIV-related sexual risk behavior: Review and synthesis of meta-analytic evidence. *AIDS and Behavior, 12*(3), 335-353.

O'Connell, A. A. (2006). *Logistic regression models for ordinal response variables*. Thousand Oaks, CA: Sage.

O'Connell, A. A., Goldstein, J., Rogers, H. J., & Peng, C. Y. J. (2008). Multilevel logistic models for dichotomous and ordinal data. In A. A. O'Connell & B. McCoach (Eds.), *Multilevel modeling of educational data*. Charlotte, NC: Information Age Publishing, Inc.

OhioDepartment of Health (2010). Current HIV infection annual surveillance tables. Retrieved from http://www.odh.ohio.gov/ASSETS/6F1EEA61EA1342E7A074CF38A2984D47/Franklin09.pdf

Operario, D., Smith, C. D., Arnold, E., & Kegeles, S. (2010). The Bruthas Project: Evaluation of a community-based HIV prevention intervention for African American men who have sex with men and women. *AIDS Education and Prevention, 22*(1), 37-48.

Osmond, D. H., Pollack, L. M., Paul, J. P., & Catania, J. A. (2007). Changes in prevalence of HIV infection and sexual risk behavior in men who have sex with men in San Francisco: 1997-2002. *American Journal of Public Health, 97*(9), 1677-1683.

Parsons, J. T., Scrimshaw, E. W., Wolitski, R. J., Halkitis, P. N., Purcell, D. W., Hoff, C. C., et al. (2005). Sexual harm reduction practices of HIV-seropositive gay and bisexual men: Serosorting, strategic positioning, and withdrawal before ejaculation. *AIDS*, *19*(1), S13–S25.

Patterson, T. L., & Strathdee, S. A. (2005). From Don Giovanni to Magic Johnson: Methodological conundrums in the measurement of sexual risk behavior. *Annals of Behavioral Medicine*, *29*(2), 83-85.

Pearson, C. R., Kurth, A. E, Cassels, S., Martin, D. P., Simoni, J. M., Hoff, P., Matediana, E., & Gloyd, S. (2007). Modeling HIV transmission risk among Mozambicans prior to their initiating highly active retroviral therapy. *AIDS Care, 19*(5), 594-604. doi: 10.1080/09540120701203337

Perlis, T. E., Des Jarlais, D. C., Friedman, S. R., Arasteh, K., & Turner, C. F. (2004). Audio-computerized self-interviewing versus face-to-face interviewing for research data collection at drug abuse treatment programs. *Addiction, 99*, 885-896.

Philip, S. S., Yu, X., Donnell, D., Vittinghoff, E., & Buchbinder, S. (2010). Serosorting is associated with a decreased risk of HIV seroconversion in the EXPLORE study cohort. PLoS ONE, *5*(9): e12662. doi: 10.1371/journal/pone.0012662.

Pilcher, C. D., Fiscus, S.A., Nguyen, T. Q., Foust, E., Wolf, L., Williams, D., Ashby, R., O'Dowd, J. O., McPherson, J. T., Stalzer, B., Hightow, L., Miller, W. C., Eron, J. J., Cohen, M. S., Leone, P. A. (2005). Detection of acute infections during HIV testing in North Carolina. *New England Journal of Medicine, 44*(2), 196-202. doi: 10.1056/NEJMoa042291

Pinkerton, S. D., & Abramson, P. R. (1993). Evaluating the risks: A Bernoulli process model of HIV infection and risk reduction. *Evaluation Review, 17,* 504-528.

Pinkerton, S. D., & Abramson, P. R. (1994). An alternative model of the reproductive rate of HIV infection. Formulation, evaluation, and implications for risk reduction interventions. *Evaluation Review, 18,* 371-388.

Pinkerton, S. D., & Abramson, P. R. (1997). Effectiveness of condoms in preventing HIV transmission. *Social Science and Medicine, 44*(99), 1303-1312. doi: 10.1016/S0277-9536(96)00258-4

Pinkerton, S. D., & Galletly, C. L. (2007). Reducing HIV transmission risk by increasing serostatus disclosure: A mathematical modeling analysis. *AIDS and Behavior, 11*(5), 698-705.

Poppen, P. J., Reisen, C. A., Zea, M. C., Bianchi, F. T., & Echeverry, J. J. (2005). Serostatus disclosure, seroconcordance, partner relationship, and unprotected anal intercourse among HIV-positive Latino men who have sex with men. *AIDS Education and Prevention, 17*(3), 227-237.

Prochaska, J. O., & DiClemente, C. C. (1992). Stages of change in the modification of problem behaviors. *Progress in behavior modification*, 28, 184-218. doi:

Procidano, M. E., & Heller, K. (1983). Measures of perceived social support from friends and from family: Three validation studies. *American Journal of Community Psychology, 11*, 1-24.

Project Inform (2004). Re-infection: Is it a concern for people living with HIV? Retrieved from http://www.thebody.com/content/art5042.html

Quina, K., Harlow, L. L., Morokoff, P. J., & Burkholder, G. (2000). Sexual communication in relationships: When words speak louder than actions. *Sex Roles, 42*(7/8), 523-549.

Radloff, L. S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*, *1*, 385-401.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*. Thousand Oaks, CA: Sage.

Reisner, S. L., Mimiaga, M. J., Skeer, M., & Mayer, K. H. (2009). Beyond anal sex: Sexual practices associated with HIV risk reduction among men who have sex with men in Boston, Massachussetts. *AIDS Patient Care and STDs, 23*(7), 545-550. doi:10.1089=apc.2008.0249

Readability Formulas.com (2012). The SMOG Readability Formula. Retrieved from http://www.readabilityformulas.com/smog-readability-formula.php

Rosser, B. R. S., Horvath, K. J., Hatfield, L. A., Peterson, J. L., Jacoby, S., & Stately, A. (2008). Predictors of HIV disclosure to secondary partners and sexual risk behavior among a high-risk sample of HIV-positive MSM: Results from six epicenters in the US. *AIDS Care, 20*(8), 925-930. doi: 10.1080/09540120701767265

Rosser, B. R. S., Oakes, J. M., Horvath, K. J., Konstan, J. A., Danilenko, G. P., & Peterson, J. L. (2009). HIV sexual risk behavior by men who use the internet to seek sex with men: Results of the men's INTernet Sex Study-II (MINTS_II). *AIDS and Behavior*, *13*, 488-498. doi: 10.1007/s10461-009-9524-3

Schell, K. (2012). The Error-Oriented Motivation Scale: An examination of structural and convergent validity. *Personality and Individual Differences, 52*(3), 352-356.

Schroder, K. E. E., Carey, M. P., & Vanable, P. A. (2003). Methodological challenges in research on sexual risk behavior: I. Item content, scaling, and data analytical options. *Annals of Behavioral Medicine*, *26*(2), 76-103.

Schroder, K. E. E., Carey, M. P., & Vanable, P. A. (2005). Methodological challenges in research on sexual risk behavior: Response to commentary. *Annals of Behavioral Medicine, 29*(2), 96-99.

Semaan, S., Des Jarlais, D. C., Sogolow, E., Johnson, W. D., Hedges, L. V., Ramirez, G., Flores, S. A., Norman, L., Sweat, M. D., & Needle, R. (2002). A meta-analysis of the effect of HIV prevention interventions on the sex behaviors of drug users in the United States. *Journal of Acquired Immune Deficiency Syndrome,* 30, S73-S93. doi: 10.1097/01.QAI.0000017613.32101.90

Semple, S. J., Patterson, T. L., & Grant, I. (2004). Psychosocial characteristics and sexual risk behavior of HIV+ men who have anonymous sex partners. *Psychology and Health, 19*(1), 71-87.

Serovich, J. M., & Mosack, K. E. (2003). Reasons for disclosure or nondisclosure to casual sexual partners. *AIDS Education and Prevention, 15*(1), 70-80.

Serovich, J. M., Reed, S., Grafsky, E. L., & Andrist, D. (2009). An intervention to assist men who have sex with men disclose their serostatus to causal sex partners: Results from a pilot study. *AIDS Education and Prevention, 21*(3), 207-219.

Sheon, N., & Lee, S. (2009). Sero-skeptics: Discussions between test counselors and their clients about sexual partner HIV status disclosure. *AIDS Care, 21*(2), 133-139. doi: 10.1080/09540120801932181

Sincolfi, D. E., & Moeller, R. W. (2007). Serosorting. Retrieved from http://www.thebody.com/content/art40765.html#benefits

Smith, D. M., Wong, J. K., Daar, E. S., Richman, D. D., & Little, S. J. (2004). Incidence of HIV superinfection following primary infection. *Journal of the American Medical Association, 292*(10), 1177-1178. doi:10.1001/jama.292.10.1177

Snowden, J. M., Raymond, H. F., & McFarland, W. (2009). Prevalence of seroadaptive behaviours of men who have sex with men, San Francisco, 2004. *Sexually Transmitted Infections, 85*(6), 469-476. doi: 10.1136/sti.2009.036269

Steward, W. T., Remien, R. H., Higgins, J. A., Dubrow, R., Pinkerton, S. D., Sikkema, K. J., Truong, H. M., Johnson, M. O., Hirsch, J., Brooks, R. A., & Morin, S. F. (2009). Behavior change following diagnosis with acute/early HIV infection - a move to serosorting with other HIV-infected invididuals. The NIMH Multisite Acute HIV Infection Study: III. *AIDs and Behavior, 13*, 1054-1060. doi: 10.1007/s10461-009-9582-6

Susser, E., Desvarieux, M., & Wittkowski, K. M. (1998). Reporting sexual risk behavior for HIV: A practical risk index and a method for improving risk indices. *American Journal of Public Health, 88*(4), 671-674.

Tate, J., Singh, K., Ndubani, P., Kamwanga, J., & Buckner, B. (2010). Measurement of HIV prevention indicators: A comparison of the PLACE method and a household survey in Zambia. *AIDS and Behavior, 14*, 209-217. doi: 10.1007/s10461-008-9505-y

The Center for HIV Law and Policy (2012). Criminal Law. Retrieved from http://www.hivlawandpolicy.org/resourceCategories/view/2

The Global HIV Working Group (2010). Proven HIV prevention strategies. Retrieved from http://www.kff.org/hivaids/upload/050106_HIVPreventionStrategies.pdf

The Swiss AIDS Federation (2012). Risk and protection. Retrieved from http://www.aids.ch/e/fragen/risiko.php#FAQ4

Tieu, H. V., Xu, G., Bonner, S., Spikes, P., Egan, J. E., Goodman, K., Stewart, K., & Koblin, B. A. (2011). Sexual partner characteristics, serodiscordant/serostatus unknown unprotected anal intercourse and disclosure among Human Immunodeficiency Virus-infected and uninfected Black men who have sex with men in New York City. *Sexually Transmitted Diseases*, *38*(6), 548-554. doi: 10.1097/OLQ.0b013e318203e2d7

Trochim, W. M. (2006). *The research methods knowledge base*, 2nd Edition. Retrieved from www.socialresearchmethods.net/kb/ (version current as of October 20, 2006).

Truong, H. H. M., Kellogg, T., Klausner, M. H., Dilley, J., Knapper, K., Chen, S., Prabhu, R., Grant, R. M., Louie, B., & McFarland, W. (2006). Increases in sexually transmitted infections and sexual risk behavior without a concurrent increase in HIV incidence among men who have sex with men in San Francisco: A suggestion of HIV serosorting? *Sexually Transmitted Infections*, *82*(6), 461-466. doi: 10.1136/sti.2006.019950

United States Department of Health and Human Services (2012). 2012 HHS poverty guidelines. Retrieved from http://aspe.hhs.gov/poverty/12poverty.shtml

Valencia, S. W., Smith, A. T., Reece, A. M., Li, M., Wixson, K. K., & Newman, H. (2010). Oral reading fluency assessment: Issues of construct, criterion, and consequential validity. *Reading Research Quarterly, 45* (3), 270-291.

Validity. (2012). *Merriam-Webster.com*. Retrieved April 20, 2012 from http://www.merriam-webster.com/dictionary/validity

Van Griensven, F. (2008). Non-condom use risk-reduction behaviours: Can they help to contain the spread of HIV infection among men who have sex with men? *AIDS, 23*, 253-255. doi: 10.1097/QAD.0b013e32831fb56e

Varghese, B., Maher, J. E., Peterman, T. A., Branson, B. M., & Steketee, R. W. (2002). Reducing the risk of sexual HIV transmission: Quantifying the per-act risk for HIV on the basis of choice of partner, sex act, and condom use. *Sexually Transmitted Diseases, 29*(1), 38-43.

Visibility Campaign (2012). Visibility Campaign's Sexual Diversity Handbook, Appendix 3: HIV/AIDS. Retrieved from http://www.stanford.edu/group/viva/facts/visibility.html#6

Vittinghoff, E., Douglas, J., Judson, F., McKirnan, D., MacQueen, K., & Buchbinder, S. P. (1999). Per-contact risk of human immunodeficiency virus transmission between male sexual partners. *American Journal of Epidemiology*, *150* (3), 306-311.

Wei, C., & Raymond, H. F. (2011). Preference for and maintenance of anal sex roles among men who have sex with men: Sociodemographic and behavioral correlates. *Archives of Sexual Behavior, 40*, 829-834. doi: 10.1007/s10508-010-9623-2

Wei, C., Raymond, H. F., Guadamuz, T. E., Stall, R., Colfax, G. N., Snowden, J. M., & McFarland, W. (2011). Racial and ethnic differences in seroadaptive and serodisclosure behaviors among men who have sex with men. *AIDS and Behavior, 15*(1), 22-29. doi:10.1007/s10461-010-9683-2

Weinhardt, L. S., Forsyth, A. D., Carey, M. P., Jaworski, B. C., & Durant, L. E. (1998). Reliability and validity of self-report measures of HIV-related sexual behavior: Progress since 1990 and recommendations for research and practice. *Archives of Sexual Behavior*, *27*(2), 155-180.

Whetten, K., Reif, S., Swartz, M., Stevens, R., Ostermann, J., Hanisch, L., Eron, J. J. (2005). A brief mental health and substance abuse screener for persons with HIV. *AIDS Patient Care*, *19*(2), 89-99.

Wilson, D. P., Regan, D. G., Heymer, K. J., Jin, F., Prestage, G. P., & Grulich, A. E. (2010). Serosorting may increase the risk of HIV acquisition among men who have sex with men. *Sexually Transmitted Diseases, 37*(1), 13-16. doi: 10.1097/OLQ.0b013e3181b35549

Wilton, L., Herbst, J. H., Coury-Doniger, P., Painter, T. M., English, G., Alvarez, M. E., Scahill, M., Roberson, M. A., Lucas, B., Johnson, W. D., & Carey, J. W. (2009). Efficacy of an HIV/STI prevention intervention for Black men who have sex with men: Findings from the *Many Men, Many Voices (3MV)* Project. *AIDS and Behavior, 13*(3), 532-544. doi: 10.1007/s10461-009-9529-y

Zablotska, I. B., Imrie, J., Prestage, G., Crawford, J., Rawstorne, P., Gulich, A., Jin, F., & Kippax, S. (2009). Gay men's current practice of HIV serconcordant unprotected anal intercourse: Serosorting or seroguessing? *AIDS Care, 21*(4), 501-510. doi: 10.1080/09540120802270292

Zea, M. C., Reisen, C. A., Poppen, P. J., & Bianchi, F. T. (2009). Unprotected anal intercourse among immigrant Latino MSM: The role of characteristics of the person and the sexual encounter. *AIDS and Behavior, 13,* 700-715. doi:

APPENDIX A: Global Measures of Sexual Behavior

We would like to ask you some questions about your *sexual partners* over the last 30 days.

XBC_	001	How many different sexual partners have you had over the last 30 days?

We would now like to ask you some questions about the *sexual encounters* you have had over the last 30 days.

XBC_	002	How many sexual encounters have you had in the last 30 days?
XBC_	003	How many of these sexual encounters involved insertive anal sex (you were the top) *with* a condom?
XBC_	004	How many of these sexual encounters involved insertive anal sex (you were the top) *without* a condom?
XBC_	005	How many of these sexual encounters involved receptive anal sex (you were the bottom) *with* a condom?
XBC_	006	How many of these sexual encounters involved receptive anal sex (you were the bottom) *without* a condom?
XBC_	007	How many of these sexual encounters involved you giving oral sex (you went down on him) *with* a condom?
XBC_	008	How many of these sexual encounters involved you giving oral sex (you went down on him) *without* a condom?
XBC_	009	How many of these sexual encounters involved you receiving oral sex (your partner went down on you) *with* a condom?
XBC_	010	How many of these sexual encounters involved you receiving oral sex (your partner went down on you) *without* a condom?
XBC_	011	How many of these sexual encounters involved vaginal sex *with* a condom?
XBC_	012	How many of these sexual encounters involved vaginal sex *without* a condom?

APPENDIX B: Global Measures of Disclosure

The following questions refer to your sexual partners during the last 30 days.

DBC_	001	Of these partners, how many know that you are HIV positive?
DBC_	002	How many of these partners knew you were HIV positive before you had sex with them?
DBC_	003	How many of these partners did you disclose to during the last 30 days?
DBC_	004	In the last 30 days, how many times were you rejected for sex after disclosing to a potential partner?
DBC_	005	In the last 30 days, how many times were you verbally or physically abused by a partner after disclosing?

APPENDIX C: Sexual Encounter Form

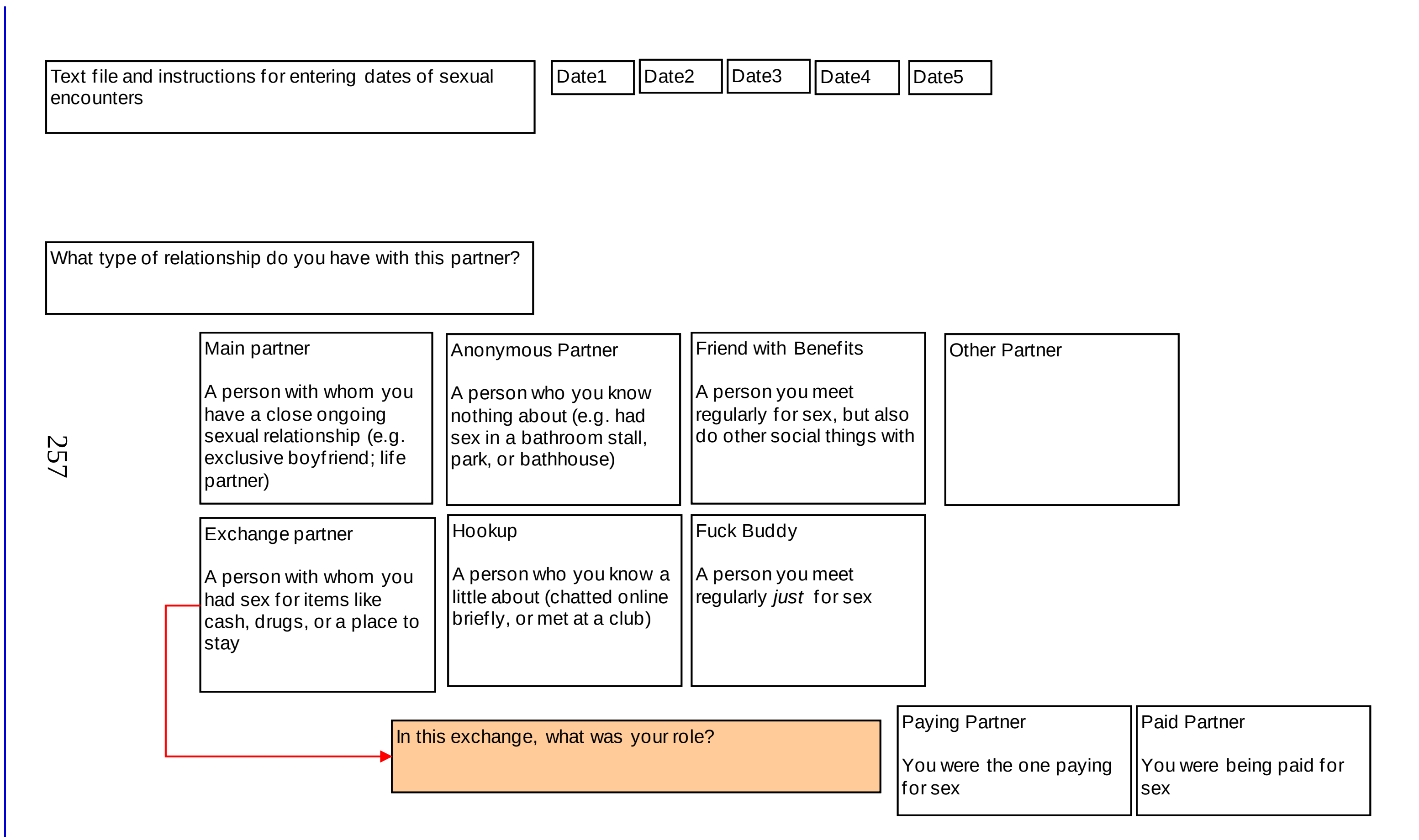
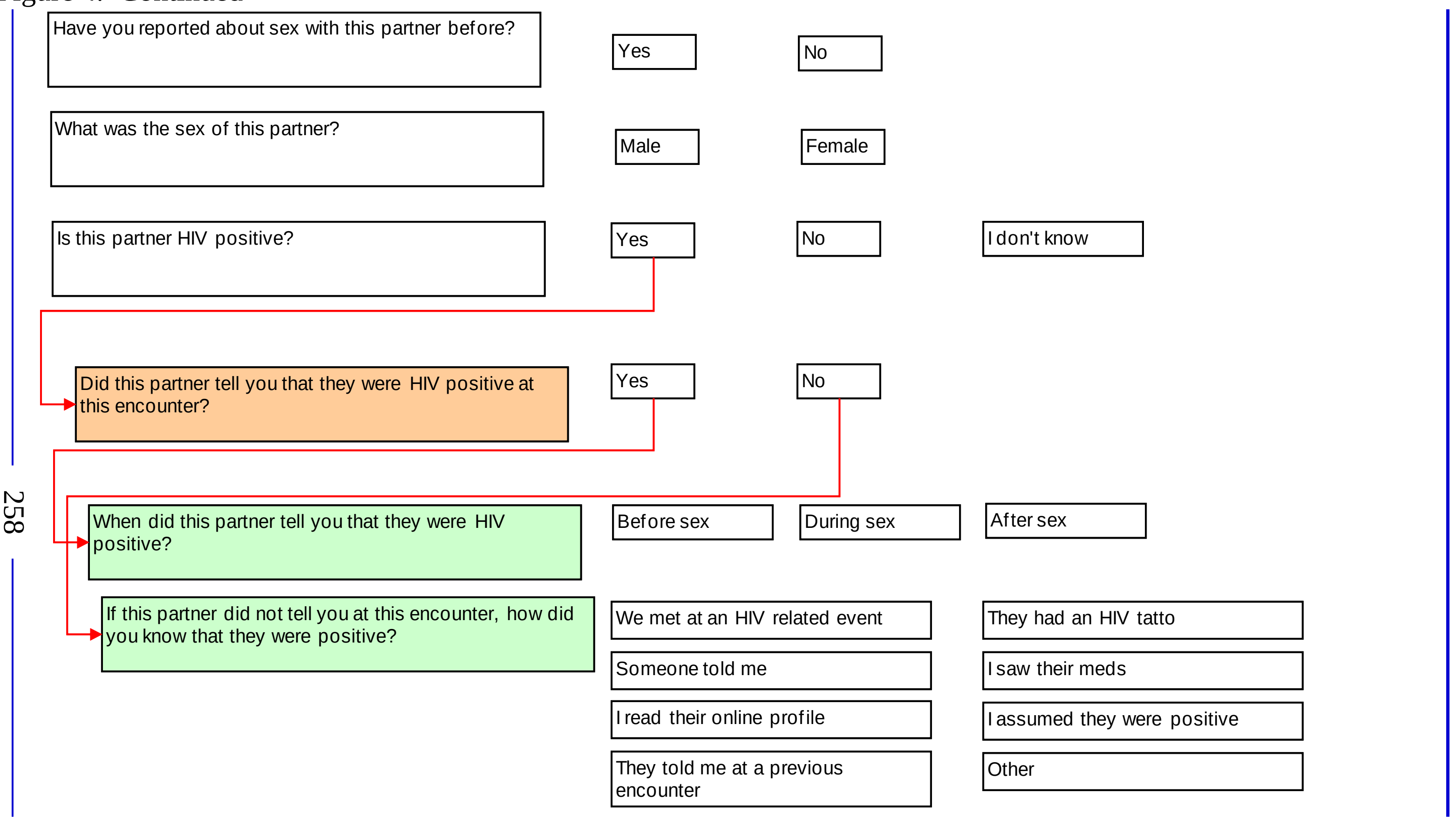


Figure 4. Sexual Encounter Form

Continued

Figure 4. Continued



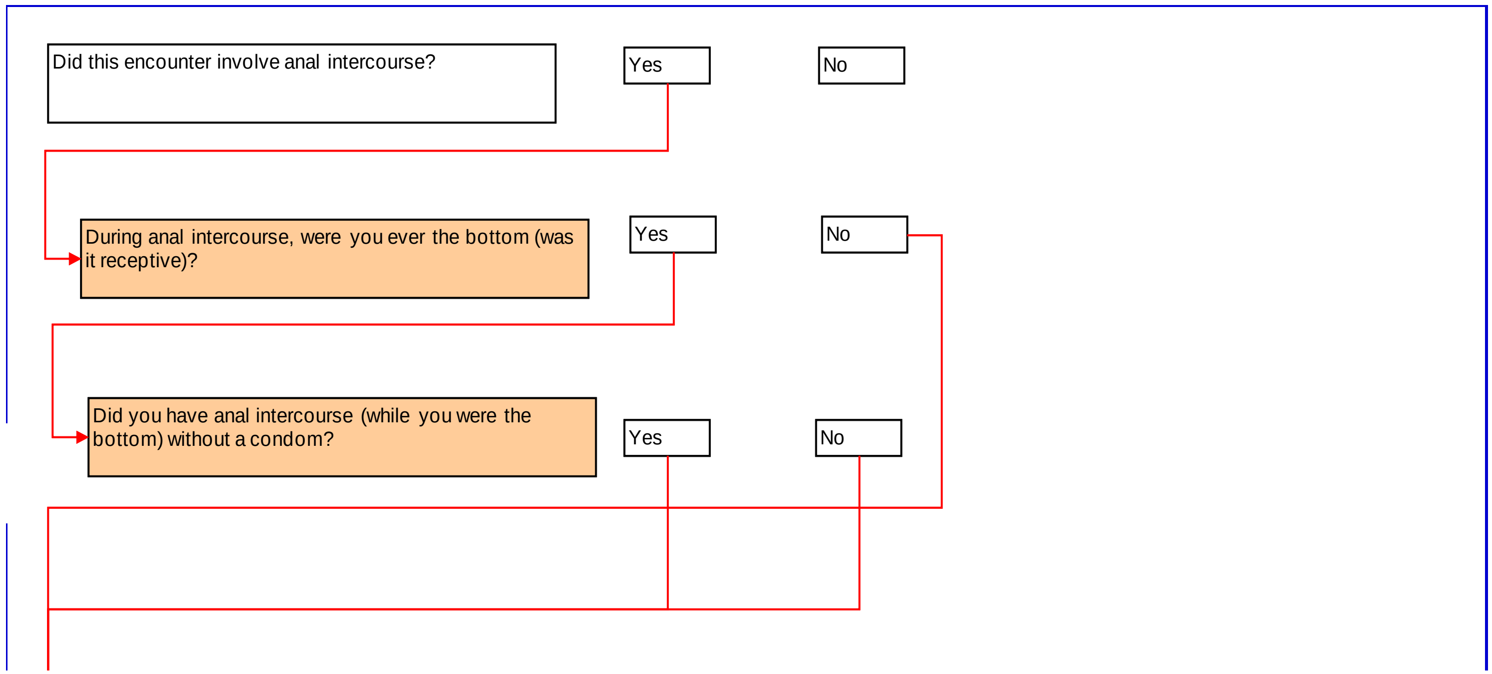
Continued

Figure 4. Continued

How did you meet this partner?	Public sex environment (Park, rest stop)	Bar or club
	Online	Bathhouse
	Through a friend	Other
Where did you have sex with this partner?	Their place	Bathhouse
	My place	Public sex environment (Park, rest stop)
		Other
Is there a chance that you will have sex with this partner again?	Yes	No

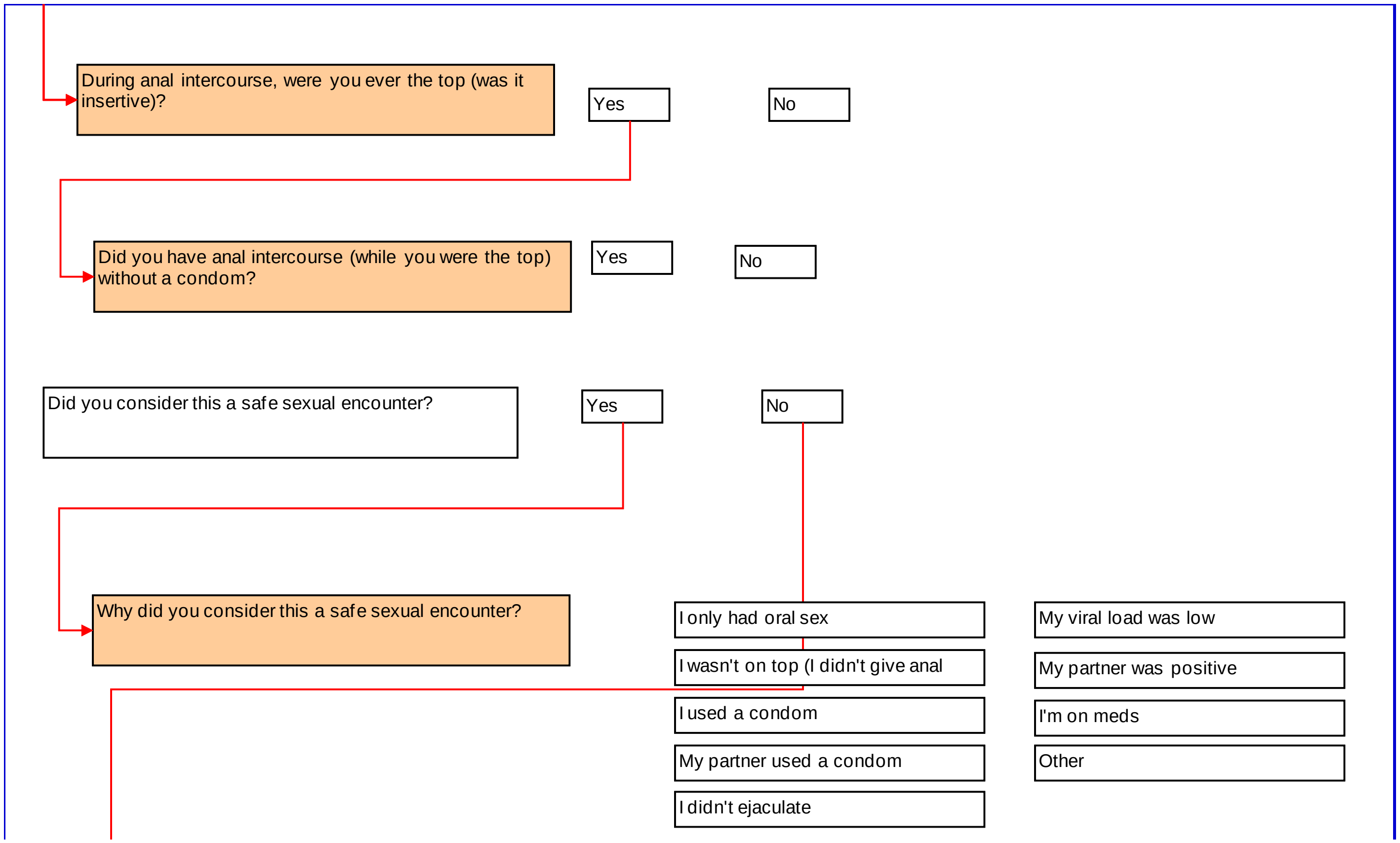
Continued

Figure 4. Continued



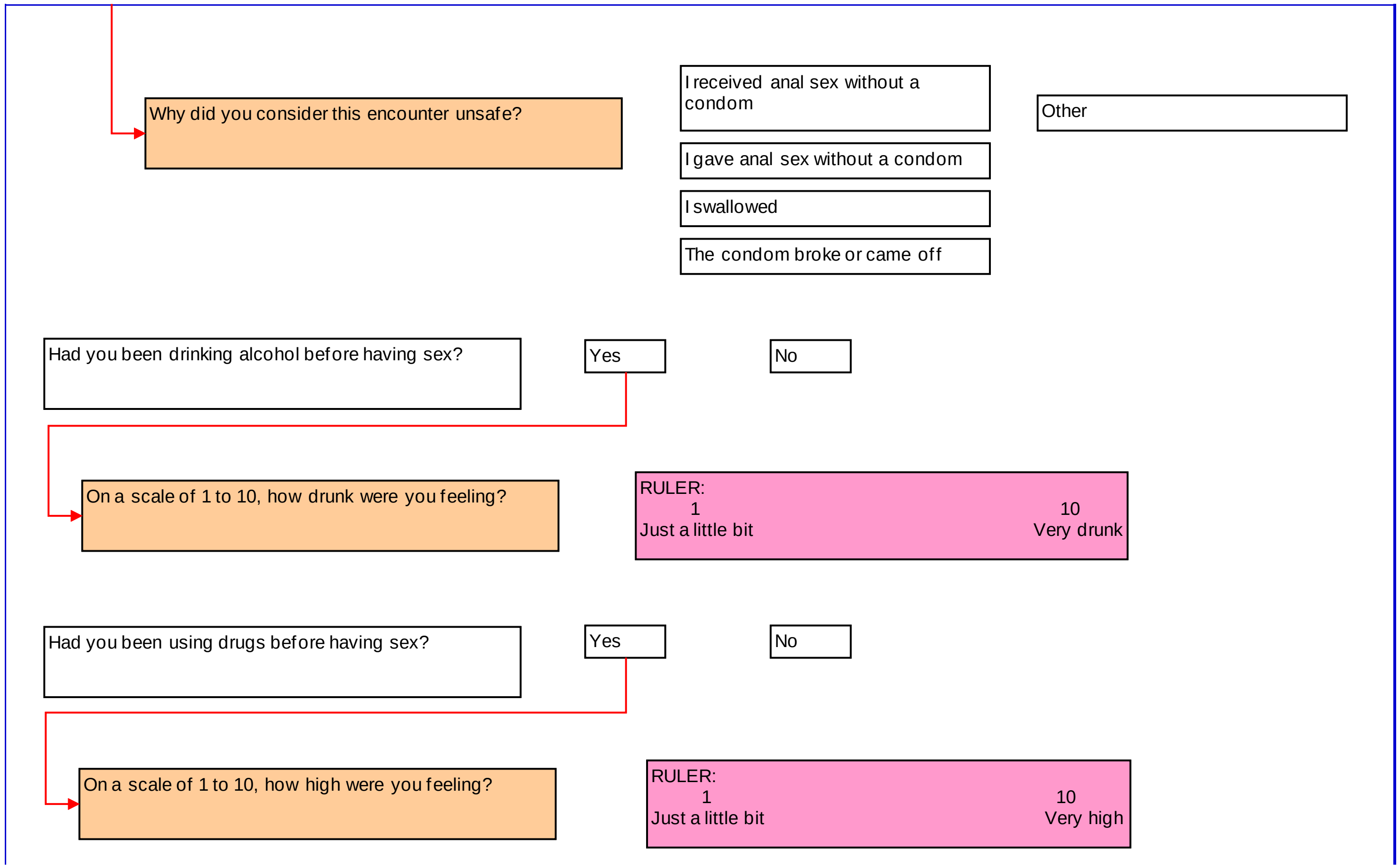
Continued

Figure 4. Continued



Continued

Figure 4. Continued



Continued

Figure 4. Continued

I asked my partner if he/she ever had an HIV test	Yes	No
I asked my partner about his/her previous partners	Yes	No
I asked my partner if he/she had ever had sex with someone who shoots drugs with a needle	Yes	No
Was there anything else going on that impacted your decision to have sex with this partner?	OPEN ENDED	

APPENDIX D: Disclosure Form

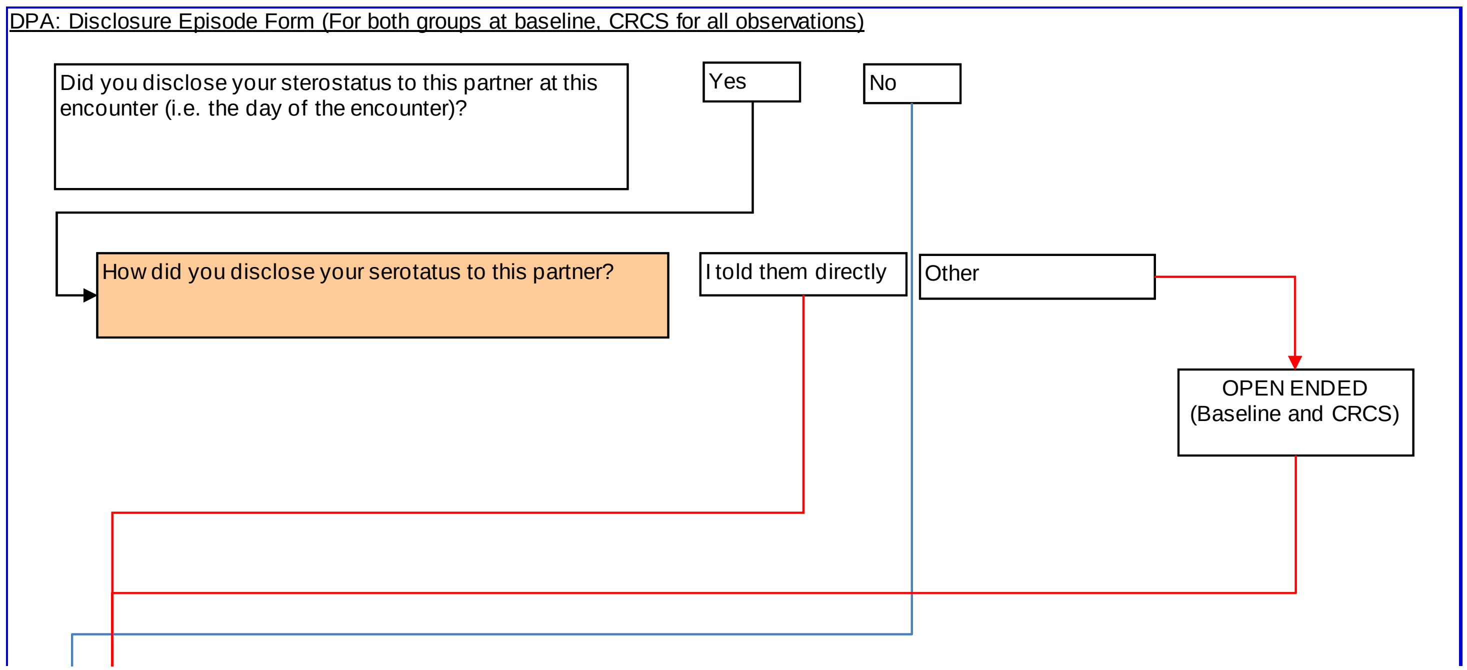
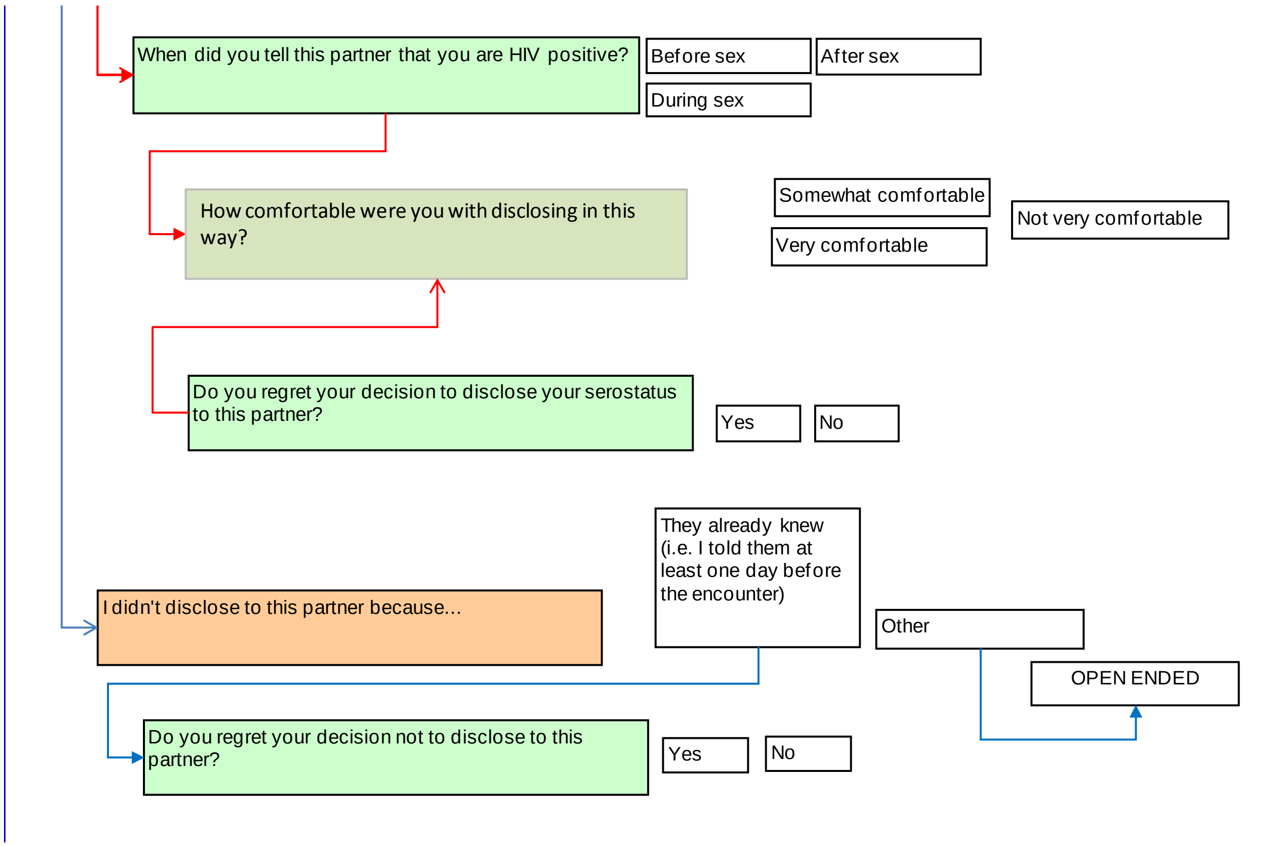


Figure 5. Disclosure Form

Continued

Figure 5. Continued



Continued

Figure 5. Continued

Many people think about different things before they decide to tell or not to tell their partners about their HIV status. When you thought about whether or not to tell this partner that you are HIV-positive, on a scale of 1 - 5, how important were the following issues in making that decision?

See list of Costs and Rewards

Was there anything else going on that impacted your decision about whether or not to disclose your HIV status to this partner?

OPEN ENDED

APPENDIX E: Instructions to Participants

Screen 1 Text

Now we would like to ask you some questions about your most recent sexual encounters.

The information that we ask you to provide may be personal, so before we start, we would like to take a moment to remind you of a few things.

First, the information you provide in response to these questions cannot be tied to your personal information.

Second, we will not be asking you to provide the names of any of your sexual partners or any information that would make it possible for anyone to identify them.

Third, you can skip any questions that you do not feel comfortable answering.

Screen 2 Text

Have you had any sexual encounters in the last 30 days?

MULTIPLE CHOICE (One Answer): **Yes** (= 1); **No** (= 0)
SKIP (= -9)

Figure 6. Instructions to Participants

Continued

Figure 6. Continued

Screen 3 Text

Next we would like you to select the dates of your last five sexual encounters during the last 30 days.

Use the calendar on the desk next to the computer to help you remember

While you enter your encounters, remember that:

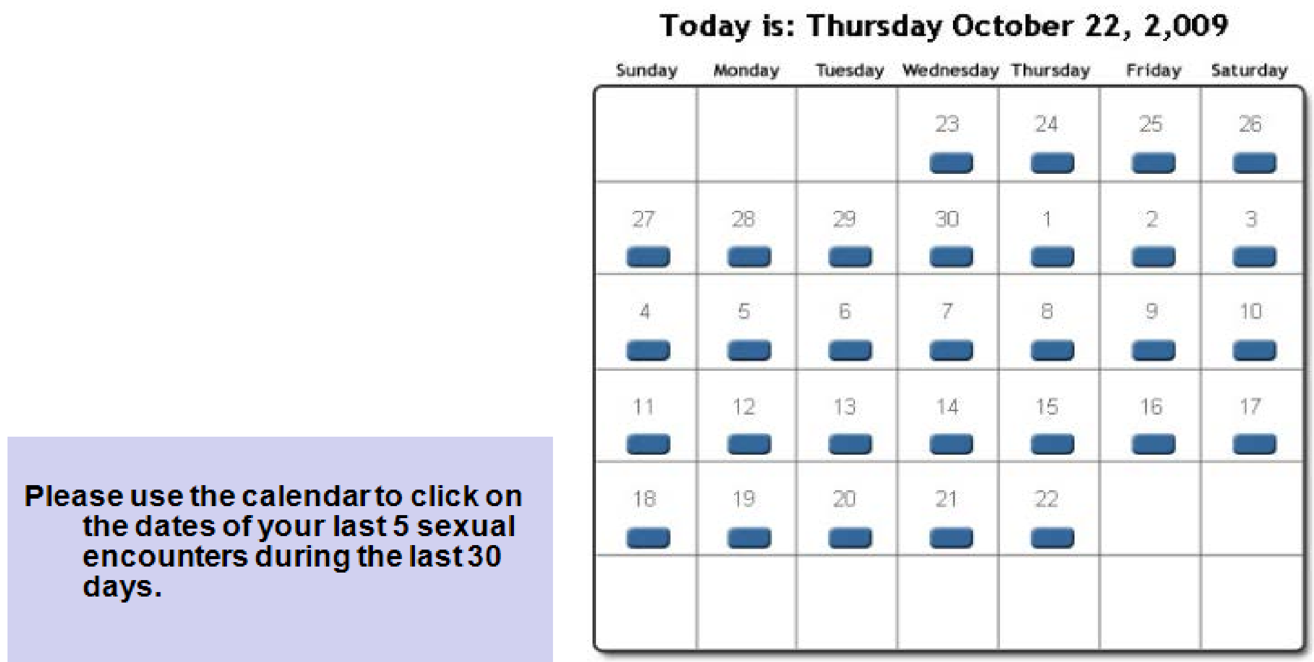
- If you had sex with 2 different partners on one day, this would count as two sexual encounters.
- If you had sex with one partner on two different days, this would also count as two sexual encounters.

When you are ready to proceed, click the forward arrow.

Continued

Figure 6. Continued

Screen 4 Text



Continued

Figure 6. Continued

Screen 5 Text

Please confirm the dates of your last sexual encounters.

For each, add a hint to help you remember the encounter

For example, you can type in “guy from the bar” or “guy I met at the party” to help you remember.

The hint you use will be erased when you are done reporting about the encounter.

APPENDIX F: Original Items for Use in Scale Construction

Table 56. Original Items for Use in Scale Construction

Item No.	Item Content	Response Option							
XEF_E1_01	XEF_E1_01 What type of relationship do you have with this partner?	1 '1 Main Partner - A person with whom you have a close ongoing sexual relationship'	2 '2 Anonymous Partner - A person who you know nothing about'	3 '3 Friend with Benefits - A person you meet regularly for sex, but also do other social things with'	4 '4 Exchange Partner - A person with whom you had sex for items like cash, drugs, or a place to stay'	5 '5 Hookup - A person who you know a little about , chatted with briefly online, or met at a club'	6 '6 Fuck Buddy - A person you meet regularly just for sex'	7 '7 Other Partner'	-7 '-7 N/A'
XEF_E1_02	XEF_E1_02 In this exchange, what was your role?	1 '1 Paying Partner (You were the one paying for sex)'	2 '2 Paid Partner (You were being paid for sex)'						
XEF_E1_03	XEF_E1_03 Have you ever told us about an encounter with this partner before now?	1 '1 Yes'	0 '0 No'						
XEF_E1_04	XEF_E1_04 What was the sex of this partner?	1 '1 Male'	0 '0 Female'						
XEF_E1_05	XEF_E1_05 Is this partner HIV positive?	1 '1 Yes'	0 '0 No'	2 '2 I Don’t Know'					
XEF_E1_06	XEF_E1_06 Did this partner tell you that they were HIV positive at this sexual encounter?	1 '1 Yes'	0 '0 No'						
XEF_E1_07	XEF_E1_07 When did this partner tell you that they were HIV positive?	1 '1 Before Sex'	2 '2 During Sex'	3 '3 After Sex'					
XEF_E1_08	XEF_E1_08 If this partner did not tell you at this sexual encounter, how did you know that they were HIV positive?	1 '1 We met at an HIV related event'	2 '2 Someone told me'	3 '3 I read their online profile'	4 '4 They told me at a previous encounter'	5 '5 They had an HIV tattoo'	6 '6 I saw their meds'	7 '7 I assumed they were positive'	8 '8 Other'

Continued

Table 56. Continued

Item No.	Item Content	Response Option					
XEF_E1_09	XEF_E1_09 How did you meet this partner?	1 '1 Online'	2 '2 Bar or Club'	3 '3 Through a Friend'	4 '4 Bathhouse'	5 '5 Public Sex Environment (Park, rest stop)'	6 '6 Other'
XEF_E1_10	XEF_E1_10 Where did you have sex with this partner?	1 '1 Their Place'	2 '2 My Place'	3 '3 Our Place'	4 '4 Bathhouse'	5 '5 Public Sex Environment (Park, rest stop)'	6 '6 Other'
XEF_E1_11	XEF_E1_11 Is there a chance that you will have sex with this partner again?	1 '1 Yes'	0 '0 No'				
XEF_E1_12	XEF_E1_12 Did this encounter involve anal intercourse?	1 '1 Yes'	0 '0 No'				
XEF_E1_13	XEF_E1_13 During anal intercourse, were you ever the bottom (was it receptive)?	1 '1 Yes'	0 '0 No'				
XEF_E1_14	XEF_E1_14 While you were the bottom, did you always use a condom?	1 '1 Yes'	0 '0 No'				
XEF_E1_15	XEF_E1_15 During anal intercourse, were you ever the top (was it insertive)?	1 '1 Yes'	0 '0 No'				
XEF_E1_16	XEF_E1_16 While you were the top, did you always use a condom?	1 '1 Yes'	0 '0 No'				

Continued

Table 56. Continued

Item No.	Item Content	Response Option	
XEF_E1_17	XEF_E1_17 Did you consider this a safe sexual encounter?	1 '1 Yes'	0 '0 No'
XEF_E1_18	XEF_E1_18 Why did you consider this a safe sexual encounter? I only had oral sex	1 '1 Yes'	0 '0 No'
XEF_E1_19	XEF_E1_19 Why did you consider this a safe sexual encounter? I wasn't the top (I didn't give anal sex)	1 '1 Yes'	0 '0 No'
XEF_E1_20	XEF_E1_20 Why did you consider this a safe sexual encounter? I used a condom	1 '1 Yes'	0 '0 No'
XEF_E1_21	XEF_E1_21 Why did you consider this a safe sexual encounter? My partner used a condom	1 '1 Yes'	0 '0 No'
XEF_E1_22	XEF_E1_22 Why did you consider this a safe sexual encounter? I didn't ejaculate	1 '1 Yes'	0 '0 No'
XEF_E1_23	XEF_E1_23 Why did you consider this a safe sexual encounter? My viral load was low	1 '1 Yes'	0 '0 No'
XEF_E1_24	XEF_E1_24 Why did you consider this a safe sexual encounter? My partner was positive	1 '1 Yes'	0 '0 No'
XEF_E1_25	XEF_E1_25 Why did you consider this a safe sexual encounter? I'm on meds	1 '1 Yes'	0 '0 No'
XEF_E1_26	XEF_E1_26 Why did you consider this a safe sexual encounter? Other	1 '1 Yes'	0 '0 No'
XEF_E1_27	XEF_E1_27 Why did you consider this a safe sexual encounter? Skip	1 '1 Yes'	0 '0 No'

Continued

Table 56. Continued

Item No.	Item Content	Response Option	
XEF_E1_28	XEF_E1_28 Why did you consider this encounter unsafe? I received anal sex without a condom	1 '1 Yes'	0 '0 No'
XEF_E1_29	XEF_E1_29 Why did you consider this encounter unsafe? I gave anal sex without a condom	1 '1 Yes'	0 '0 No'
XEF_E1_30	XEF_E1_30 Why did you consider this encounter unsafe? I went down on him without a condom	1 '1 Yes'	0 '0 No'
XEF_E1_31	XEF_E1_31 Why did you consider this encounter unsafe? He went down on me without a condom	1 '1 Yes'	0 '0 No'
XEF_E1_32	XEF_E1_32 Why did you consider this encounter unsafe? The condom broke or came off	1 '1 Yes'	0 '0 No'
XEF_E1_33	XEF_E1_33 Why did you consider this encounter unsafe? My viral load was high	1 '1 Yes'	0 '0 No'
XEF_E1_34	XEF_E1_34 Why did you consider this encounter unsafe? My partner was negative	1 '1 Yes'	0 '0 No'
XEF_E1_35	XEF_E1_35 Why did you consider this encounter unsafe? I wasn’t taking my meds	1 '1 Yes'	0 '0 No'
XEF_E1_36	XEF_E1_36 Why did you consider this encounter unsafe? Other	1 '1 Yes'	0 '0 No'

Continued

Table 56. Continued

Item No.	Item Content	Response Option	
XEF_E1_37	XEF_E1_37 Why did you consider this encounter unsafe? Skip	1 '1 Yes'	0 '0 No'
XEF_E1_38	XEF_E1_38 Had you been drinking alcohol before having sex?	1 '1 Yes'	0 '0 No'
XEF_E1_39	XEF_E1_39 On a scale of 1 to 10, how drunk were you feeling?		
XEF_E1_40	XEF_E1_40 Had you been using drugs before having sex?	1 '1 Yes'	0 '0 No'
XEF_E1_41	XEF_E1_41 On a scale of 1 to 10, how high were you feeling?		
XEF_E1_42	XEF_E1_42 Did you asked this partner if he/she ever had an HIV test?	1 '1 Yes'	0 '0 No'
XEF_E1_43	XEF_E1_43 Did you asked this partner about previous partners?	1 '1 Yes'	0 '0 No'
XEF_E1_44	XEF_E1_44 Did you asked this partner if he/she had ever had sex with someone who shoots drugs with a needle?	1 '1 Yes'	0 '0 No'
XEF_E1_45	XEF_E1_45 Was there anything else going on that impacted your decision to have sex with this partner?		
XEF_E1_46	XEF_E1_46 Did this partner tell you he/she was HIV negative at this sexual encounter?	1 '1 Yes'	0 '0 No'

Continued

Table 56. Continued

Item No.	Item Content	Response Option				
XEF_E1_47	XEF_E1_47 When did this partner tell you he/she was HIV negative?	1 '1 Before Sex'	2 '2 During Sex'	3 '3 After Sex'		
XEF_E1_48	XEF_E1_48 If this partner did not tell you at this sexual encounter, how did you know that he/she was HIV negative?	1 '1 Someone told me'	2 '2 I read their online profile'	3 '3 They told me at a previous encounter'	4 '4 I assumed they were negative'	5 '5 Other'

APPENDIX G: Table of Methodologies

Table 57. Table of Methodologies

Section Where Results are Reported	Technique	Content representativeness	Content relevance	Technical Quality	Substantive aspect	Structural aspect	Validity related to external factors	Possibility for generalizability	Consequential aspects of validity
		Relevant Aspect of Unified Validity							
I.	Review of current literature Evaluation of the ordering of risk categories	**X**							
II.	Descriptive analysis of missingness (skips, NA)		**X**						
II.	Relationships between participant characteristics (HIV stigma, openness, depression, self-esteem, and motivation) and		**X**						
II.	Patterns of missingness across encounters		**X**						

Continued

Table 57. Continued

Section Where Results are Reported	Technique	Content representativeness	Content relevance	Technical Quality	Substantive aspect	Structural aspect	Validity related to external factors	Possibiltiy for generalizability	Consequential aspects of validity
VII.	Relationships between HIV stigma, openness, depression, self-esteem, and motivation and ordinal scores		**X**						
III.	SMOG scores		**X**	**X**					
III.	Relationships between participant education, computer skills, and employment and missingness		**X**	**X**					
III.	Relationships between participant education, computer skills, and employment and ordinal scores		**X**	**X**					

Relevant Aspect of Unified Validity

Continued

Table 57. Continued

Section Where Results are Reported	Technique	Relevant Aspect of Unified Validity: Content representativeness	Content relevance	Technical Quality	Substantive aspect	Structural aspect	Validity related to external factors	Possibiltiy for generalizability	Consequential aspects of validity
IV.	Comparisons between global and encounter-specific sections of the ACASI instrument			**X**	**X**			**X**	
IV.	Relationship between participant demographic variables and response consistency				**X**			**X**	
IV.	Comparison of participant scale scores across categories of response consistency (over-report global, consistent, over-report encounter-specific) using ANOVA				**X**				
IV.	Relationship between ordinal scores of HIV transmission risk and response consistency category				**X**				
IV.	Multinomial logistic regression predicting response consistency with select participant characteristics				**X**			**X**	

Continued

Table 57. Continued

Section Where Results are Reported	Technique	Content representativeness	Content relevance	Technical Quality	Substantive aspect	Structural aspect	Validity related to external factors	Possibiltiy for generalizability	Consequential aspects of validity
		Relevant Aspect of Unified Validity							
VI.	Multi-trait multi-method matrix				X			X	
V.	Examination of response frequencies, item intercorrelations, item-total correlations, and internal consistency					X			
V.	Rasch modeling of item scores					X			
V.	Pattern responses based on seroconcordance/discordance					X			
VI.	Relationships between scores on the ordinal scale of HIV transmission risk and participant demographics, scale scores, and response consistency				X		X	X	

Continued

Table 57. Continued

Section Where Results are Reported	Technique	Content representativeness	Content relevance	Technical Quality	Substantive aspect	Structural aspect	Validity related to external factors	Possibiltiy for generalizability	Consequential aspects of validity
		Relevant Aspect of Unified Validity							
VI.	Ordinal regression of HIV tranmission risk on select variables				**X**		**X**		
VI.	Exporation of the proportional odds assumption					**X**			
VIII.	Comparison of the sensitivy and specificity of the ordinal risk measure and alterntive measures of risk								**X**