Development and Testing of a New Version of the Hypermasculinity Index

Jay Peters, Carey Nason, and Winston M. Turner

—————

The high prevalence and serious consequences associated with rape makes its prevention an important social work goal. Rape prevention necessitates understanding the attitudes and personality characteristics of actual and potential rapists. Within the research on attitudinal correlates of rape, hypermasculinity consistently emerges as one of the strongest predictors. The most commonly used measure of hypermasculinity, the Hypermasculinity Index (HMI), uses a forced-choice format that impairs its psychometric properties. This article presents the results of testing a revised version of the HMI using a phrase-completion response format. A convenience sample of undergraduate men ($N = 284$) from a rural New England university was used. Findings indicate that the new version yields more normally distributed data with a higher internal reliability coefficient. Even more important, the revised version greatly reduced social desirability bias and improved the ability to detect the underlying structure of hypermasculinity.

KEY WORDS: *hypermasculinity; macho; phrase completion; psychometric properties; rape*

Sexual assault of adult women by men is a widespread social problem. Lifetime prevalence of sexual assault in the United States is conservatively estimated between 12.65% (Kilpatrick, Acierno, Resnick, Saunders, & Best, 1997; Resnick, Kilpatrick, Dansky, Saunders, & Best, 1993) and 14.8% of women (Tjaden & Thoennes, 2000). Somewhat older and less conservative estimates of lifetime prevalence rates range from approximately one in four (Koss, Gidycz, & Wisniewski, 1987) to one in two (Randall & Haskell, 1995) adult women in the United States.

Recent estimates of the annual incidence of rape within the United States range from 95,136 (64.8 per 100,000 women; Federal Bureau of Investigation [FBI], 2002) to 126,500 (103 per 100,000 women; Catalano, 2004). On college campuses there is an average of 35 attempted or completed incidents of sexual assault per 1,000 college women each academic year (Fisher, Cullen, & Turner, 2000). Although the exact number of sexual assaults is unknown as a result of underreporting (Banyard et al., 2005), even these conservative estimates of the prevalence and incidence of sexual assault indicate that it is a widespread social problem. There is evidence that both men and women are capable of being sexually aggressive; however, this article examines only sexual assault of women by men because this focus "reflects the reality that men are responsible for the vast majority of sexual assaults" (Krahé, Waizenhöfer, & Möller, 2003, p. 219).

In addition to being widespread, sexual assault also produces serious short- and long-term consequences for the victim (Burgess & Holmstrom, 1974). These consequences include increased anxiety symptoms (Burgess & Holmstrom; Calhoun & Atkinson, 1991), including posttraumatric stress disorder (PTSD) (Calhoun & Resnick, 1993), clinically significant depression (Calhoun & Atkinson; Saunders, Villeponteaus, Lipovsky, Kilpatrick, & Veronen, 1992), sexual issues, and health problems (Calhoun & Atkinson). In addition to these largely individual problems, sexual assault by men can often result in the dissolution of long-term relationships and marriages (Davis, Taylor, & Bench, 1995; Monnier, Resnick, Kilpatrick, & Seals, 2002; Rodkin, Hunt, & Cowan, 1982) as well as lost work and educational productivity, all of which contribute to lowered economic status for rape victims and their children (Calhoun & Atkinson; Monnier et al.).

PREDICTING MALE OFFENDERS

Given the scope and seriousness of sexual assault by men, the need for sexual assault prevention is

obvious and fits well within the purposes of social work (Council on Social Work Education, 2001). However, to prevent rape it is important to improve the ability to predict which men, as a group, are most likely to assault women. This ability would better facilitate targeted education, intervention, and prevention resources for high-risk groups. Predictive ability, however, depends on a detailed understanding of the causes and correlates of sexual assault behavior in men. The literature regarding this important topic is growing (for example, Abbey, McAuslan, Zawacki, Clinton, & Buck, 2001; Aberle & Littlefield, 2001; Koss & Gaines, 1993; Loh, Gidycz, Lobo, & Luthra, 2005), but questions remain regarding what attitudes, beliefs, personality characteristics, and environmental factors are related to an increased likelihood of perpetrating rape. A measure that can reliably detect such tendencies would greatly enhance the ability to intervene and alter such behavior among groups identified as being at high risk.

The relationship between rape perpetration and attitudes such as endorsement of rape myths and stereotypical sex roles has been studied extensively beginning with Burt's (1980) landmark study. These attitudes are weakly correlated (at approximately the $r = .22$ level) with self-reported sexually coercive behavior in men (for a review of the literature see Check & Malamuth, 1985) or with intentions to commit sexual assault (Douglass, 2003). A meta-analysis of the literature on attitudinal predictors of assault behavior (Murnen & Byrne, 1991), found that the strongest predictor of sexual assault behavior ($r = .29$) was a measure of hypermasculinity, specifically, the Hypermasculinity Index (HMI) (Mosher & Sirkin, 1984). And, in a retrospective study, hypermasculinity was moderately correlated ($r = .40$) with the self-reported use of force to obtain sex (Mosher & Anderson, 1986). Therefore it appears that measures of hypermasculinity are particularly important.

Overview of the HMI

Theoretically, *hypermasculinity* results from gender role socialization (Crowell & Burgess, 1996) in which cultural expectations of maleness produce both a turning away from relational ways of being (Bergman, 1991) and also an adversarial relationship with females who are viewed as deficient, "other," and "dangerous" (Gilligan, 1982). Consequently, hypermasculinity is thought to be related to violence against women through a process in which women are both desired and feared (Dutton, 1998; Gilligan). A solution to the contradictory impulses of desire and fear is to exert physical and sexual power and control over the feared object. As a result, hypermasculinity has been associated with both sexual and physical violence against women (see Moore & Stuart, 2005, for a review of masculinity and partner violence and Byers, 1996, for a review of the theoretical literature on hypermasculinity and sexual assault). Furthermore, the fear itself is thought to be involved in the construction of the hypermasculine personality, resulting in three components of the construct: (1) danger as exciting, (2) fighting (coercive control), and (3) calloused sexuality (Mosher & Sirkin, 1984).

In addition to these correlational studies, empirical support for the relationship of hypermasculinity and sexual assault has also been consistently found in studies using multiple regression or structural equation and path analysis models (Bourg, 2001; Johnson & Knight, 2000; Koralewski & Conger, 1992). In fact, Bourg found that hypermasculinity was strongly related to each of the four components of a model of sexual assault of women, concluding therefore that "hypermasculinity can be thought of as the common thread which binds together the fundamental constructs of coerciveness against women" (p. 80). The only study in which hypermasculinity was not an important factor in predicting sexual assault on women was that of Loh and colleagues (2005), which used a different measure of hypermasculinity, the Hypergender Ideology Scale (Hamburger, Hogben, McGowan, & Dawson, 1996).

In addition to its relationship with sexual assault, hypermasculinity, as assessed with the HMI, was also correlated with alcohol use ($r = .28$), use of both drugs and alcohol ($r = .26$), delinquency ($r = .38$), early sexual intercourse ($r = .36$), aggression ($r = .47$) and aggression after drinking ($r = .65$) (Mosher & Anderson, 1986), and aggressive driving ($r = .37$) (Krahé & Fenske, 2002). In addition, the HMI was strongly correlated with personality traits such as impulsivity and lack of empathy (Mosher & Sirkin, 1984) and with attitudes such as rape myth acceptance ($r = .53$) and with negative views of women ($r = .47$) (Koralewski & Conger, 1992). These traits and attitudes, in turn, have also been linked with an increased likelihood that a man will perpetrate rape (Tyler, Hoyt, & Whitbeck, 1998). Taken together, these findings suggest that the HMI is a valid measure of a set of attitudes that have been shown to

be highly correlated with male sexually coercive behavior. As such, the measure has potential use in prevention efforts.

Although there is strong evidence of this instrument's predictive ability, psychometrically, the HMI has some serious limitations. Following the recommendation of Edwards (1953), Mosher and Sirkin (1984) used a forced-choice format to reduce social desirability response bias (Cronbach, 1946). In the HMI, respondents are instructed to circle one of two equally socially desirable or undesirable choices, thus compelling participants to choose between responding as a "macho man" or a "wimp" (Mosher & Sirkin, p. 160). For example, item 3 of the HMI asks participants to choose between the two statements: "Call me a name and I'll pretend not to hear you." or "Call me a name and I'll call you another." Similarly, item 18 forces a choice between "I win by not fighting" or "I fight to win."

Despite the cogent logic supporting its design, the forced-choice format results in unacceptably high levels of missing data (Edwards & Diers, 1962). Without a "Both," "Don't know," or "Not applicable" response option, many participants leave many questions blank. The forced-choice format also results in reduced variance in the data (Hamburger et al., 1996). That reduced variance then diminishes the HMI's ability to discriminate between subtle (though potentially important) differences in the level of endorsement of the underlying traits (for an example, see Stöber, Deete, & Musch, 2002). Consequently, any data analysis that is rooted in correlations will suffer from this lack of variance. As DeVellis (1991) noted, "A measure cannot covary if it does not vary" (p. 64). The lack of variation may then lead to incorrect conclusions. For example, Stöber and colleagues found that when comparing a continuous version of a scale with a dichotomous version, the continuous format was able to detect which participants had been instructed to "fake bad" and which had not; the dichotomously scored instrument was unable to detect them ($p < .001$). Similarly, the continuous version was able to detect a theoretically expected correlation between impression management and extraversion, whereas the dichotomous version was not. In both cases, the authors concluded that the use of the dichotomous version would lead to the erroneous conclusion that the scale was not affected by social desirability bias and it lacked convergent validity. Finally, the lack of variability in the dichotomous format data means

that analyses that use multiple regression are likely to be unstable and consequently may lead to erroneous conclusions (DeVellis). Thus, although Mosher and Sirkin's logic of choosing a forced-choice format was solid, the lack of variability in the final measure adversely affected the utility of the instrument.

An Alternative to Forced Choice

Fortunately, Hodge and Gillespie (2003) have devised a new "phrase-completion" format for survey questions that may provide a good alternative for the forced-choice format. Rather than using two forced-choice statements or a statement with a Likert-type response set of "strongly agree" to "strongly disagree," the phrase-completion format presents the reader with the stem of a sentence along with two possible endings for that stem. For example, we might have a stem "I think presidential debates…" and then two alternate endings of "…are boring" or "…are interesting." These two opposing phrases are arranged along a 10- or 11-point continuum as in the example below:

I think presidential debates

are boring.						are interesting.				
0	1	2	3	4	5	6	7	8	9	10

This question format preserves the polar extremes of the forced-choice format but still allows respondents to indicate their degree of endorsement of those extremes. In addition, the format allows the use of a 10- or 11-point response set (as opposed to a typical five-point scale for a Likert format), which helps maximize variability in the responses (Nunnally, 1970). For questions that contain obvious socially desirable answers (for example, "In my opinion…some women are good for only one thing/all women deserve the same respect as men"), this format allows the respondent to indicate very slight endorsement of the socially undesirable response. The format thus preserves the advantages of forced choice sought by Mosher and Sirkin (1984) for the HMI scale and adds an increased ability to detect subtle differences in endorsement of the underlying construct.

In their study comparing the properties of the phrase completion format to Likert-type formats, Hodge and Gillespie (2003) revised the frequently used six-item version of Allport and Ross's (1967) intrinsic measure of religion to fit the phrase-completion format. After administering the revised scale to 78 graduate social work students

(no demographic data given), they compared the phrase-completion results with results from numerous earlier studies and found that the phrase-completion format resulted in higher reliability coefficients ($\alpha = 0.95$ compared with low to mid 0.80s in earlier studies). In addition, when using the phrase-completion format, the factor loadings of items increased dramatically (from 0.57–0.80 to .79–0.93). This evidence of the increased power for both Cronbach's alpha and factor analysis (both of which are based on correlation), implies that regression analyses should also be improved with use of the phrase-completion format.

Because hypermasculinity appears to be one of the strongest predictors of sexual assault, an improved measure of this construct would be helpful in both research and prevention settings. Such a measure could be used in the context of basic research, needs assessment, evaluation of community change programs, and rape prevention programs. For example, an improved measure of hypermasculinity with a valid and stable factor structure would facilitate greater understanding of the precise elements of hypermasculinity that are more or less related to sexual assault by men. This understanding could then be integrated into community programming and program evaluations designed to reduce the "rape prone" culture (Palmer, 1989) that supports sexual assault of women.

Based on the centrality of hypermasculinity in sexual assault by men and the apparent psychometric superiority of the phrase-completion format compared with a forced-choice response set, the purpose of the present research was to develop and test an improved measure of hypermasculinity. We first created a new, phrase-completion version of the HMI, titled the HMI-R, and then compared it with Mosher and Sirkin's (1984) forced-choice version. We hypothesized that, when compared with the original forced-choice version, the HMI-R would demonstrate substantially greater variability in the data, have a higher internal reliability coefficient, and contain fewer items that were highly skewed.

METHOD

Participants

We first conducted a pilot study of the HMI-R with a convenience sample of students in an undergraduate social work class ($N = 32$). Students were randomly assigned to complete the phrase-completion and forced-choice versions of the instrument. The students then participated in a focus group discussion to identify possible problems with our revision of the scale or the new response format. The recommendations of those respondents were incorporated into a further revised scale.

A convenience sample ($N = 686$) of undergraduate students in five classes at a rural university in northern New England was then recruited for our study comparing the psychometric properties of the HMI and HMI-R. The classes were in business ($n = 134$, 19.5%), engineering ($n = 84$, 12.3%), human development ($n = 337$, 49.1%), sociology ($n = 101$, 14.7%), and women's studies ($n = 30$, 4.4%).

Within these classes, instruments were distributed so that every other person in a classroom row received the HMI or the HMI-R. This randomization was designed to eliminate between-group differences based on seating location. Data about race and ethnicity were not collected because 97% of the undergraduate student body was white. Therefore, no meaningful comparisons based on race or ethnicity would have been possible. In addition, the research goals were only to compare the psychometric properties of the original and revised versions of the HMI. We therefore reduced participant burden by eliminating any demographic questions except gender.

Of approximately 823 surveys that were distributed, 343 usable forced-choice surveys were returned and 341 usable phrase-completion surveys were returned, for a sample size of 686 and an approximate return rate of 83.4%. One participant from each group was eliminated because of obviously bad data (for example, circling all ones or 10s on the phrase completion format). Of this sample 315 or 45.9% were women and 284 or 41.4% were men. Three students (.4%) were transgender and 84 (12.3%) had missing data for gender. For the present study, only the 284 male responses were used. Of these 284 male respondents, 135 completed the HMI; 148 completed the HMI-R.

Instruments

To improve the clarity and ease of comprehending the items, we modified the original forced-choice version of the HMI by deleting three of the original 30 items and revising the wording of some of the remaining 27 items (see appendix). For example, because it does not distinguish between being a participant or bystander, we recast the original item, "I hope to forget past unpleasant experiences with

male aggression," to "I hope to forget past unpleasant experiences when I was in a physical fight." To reflect contemporary culture we reworded the item "I would rather be a famous scientist than a famous prizefighter" to "I would rather be a famous scientist than a famous WWF fighter." Other items were revised to achieve brevity and clarity. For instance, the item "There is no such thing as too big a risk, if the payoff is large enough" was revised to "When it comes to taking risks, I'm a high roller."

We next recast the slightly revised HMI into a phrase-completion format, the HMI-R. In most cases recasting items for the HMI-R was a straightforward process. For example, the pair of forced-choice items "I'd rather gamble than play it safe" or "I'd rather play it safe than gamble" simply became "I'd rather…" with two possible endings ("…play it safe than gamble" or "…gamble than play it safe."). In some cases, however, more complete reworking of the item was required. The original items in the forced-choice pair, "Fair is fair in love and war" and "All is fair in love and war" did not share a common sentence root. We therefore recast this item to "In love and war…" with the completing phrases, "you should still play by the rules" or "anything goes." Any time we made such changes to create a phrase completion item, we then also revised the original HMI wording so that we ended up with two identical scales with 27 items.

Finally, because we want to use the HMI-R in future research to measure internalization of hypermasculinity among women, we eliminated male pronouns. Thus the item "I like fast cars and fast women" was recast to "I like fast cars and fast lovers."

RESULTS

Scoring

The forced-choice response set of the HMI yields dichotomous data that were coded with the hypermasculine option as a one, and its opposite as a zero. As a result, the final (summed) score represents the total number of hypermasculine items endorsed by the respondent. These scores have a possible range of 0 to 27. The actual range in this sample was from 0 to 22.

The HMI-R used a continuous response format with the familiar anchors of 1 to 10. To compare HMI-R data with the HMI data, we recoded the HMI-R responses to a scale of 0 to 9 and then divided the total score for each individual by 9,

resulting in total HMI-R scores that ranged from 0 (least hypermasculine) to 27 (most hypermasculine). The actual range in this sample was from 1 to 22. We used a 10-point response set for the HMI-R rather than the 11-point scale recommended by Hodge and Gillespie (2003) because numerous pilot-test participants reported using the middle value of 5 to signify "no opinion" or "I don't know."

Descriptive Characteristics

For the HMI the mean for the entire sample was 9.32 (SD = 5.2) with a total variance of 26.7. In comparison, the HMI-R had a mean of 11.3 (SD = 3.5) with a total variance of 12.2 or less than one-half the variance of the forced-choice scale. Our first hypothesis was therefore not supported in that the HMI-R had substantially less variability than did the HMI. Examination of the left-hand tail of the distributions for the two scales (see Figure 1) indicates that part of this reduced variance may stem from the fact that nine respondents (3.17%) scored two or less on the HMI and only two respondents (1.4%) to the HMI-R scored two or less.

We next tested our second hypothesis, which stated that the internal consistency reliability for the HMI-R would be higher than that of the HMI. The Cronbach's alpha coefficient for the HMI-R was 0.90) compared with 0.79 for the HMI in the present study and to 0.89 reported by Mosher and Sirkin (1984), and 0.80 to 0.87 in subsequent studies (Burk, Burkhart, & Sikorski, 2004; and Sullivan & Mosher, 1990 respectively). The second hypothesis was therefore supported with the HMI-R demonstrating slightly to somewhat higher internal consistency reliability.

Our third hypothesis stated that when compared with the HMI, the HMI-R would contain fewer items with skewed distributions. This hypothesis was strongly supported with 10 items on the HMI evidencing almost no variability as indicated by a skew greater than 1.0. In contrast responses to the HMI-R resulted in only two items with skewed distributions. Thus the phrase-completion response format of the HMI-R resulted in data with substantially fewer skewed items.

Examination of two items from the HMI-R, one skewed and one not, further indicates ways in which this phrase-completion format is superior to the original. First as can be seen at the top of Figure 2, in responding to the item "Using drugs or alcohol to 'encourage' a woman to have sex with you is "(a)
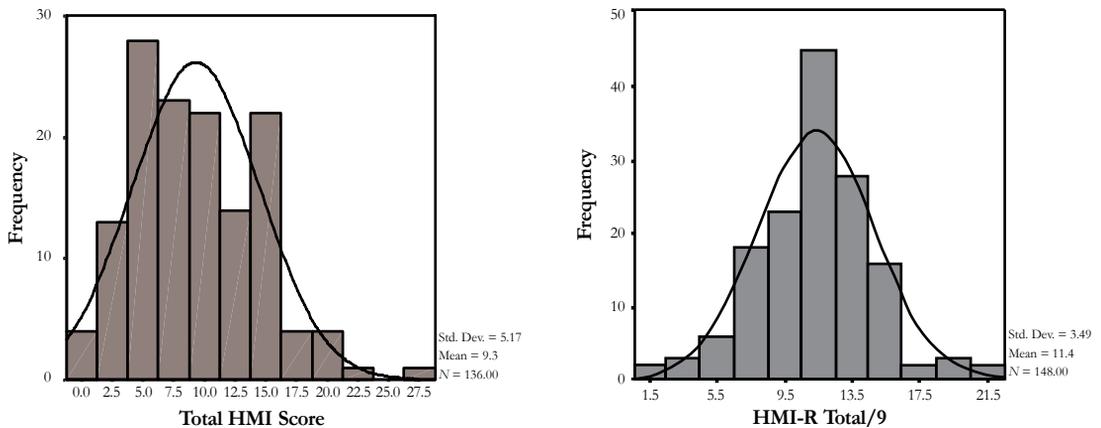
gross and unfair" or "(b) OK if you can get away with it," only 13 participants (9.6%) selected the socially undesirable answer when using the original HMI. In contrast, with the HMI-R, 76 or 51.8% of the participants endorsed some degree of the socially undesirable answer (see top pair of items in Figure 2). The lower two graphs in Figure 2 show that the item "So-called prick teasers should be (a) forgiven or (b) raped" was endorsed by only 5.9% (14) HMI participants. In contrast fully 53.6% or 98 of the HMI-R participants endorsed some degree of this rather extreme sentiment. Taken together, the substantial reduction in skewed items and increased variance for items with clearly socially correct responses indicates that the HMI-R has better psychometric properties when compared with the HMI.

Finally, given these results we compared the distribution of total HMI and HMI-R scores. As can be seen in Figure 1, the total HMI had a positive though still acceptable skew (0.55) and HMI-R had an almost perfectly normal distribution with skewness of −0.06. This difference in skew indicates that when using the HMI-R, participants were more willing to endorse greater amounts of the underlying construct of hypermasculinity.

Post Hoc Analysis

Two post hoc analyses were conducted, the first concerning missing data arose from our observations while conducting the usual data cleaning techniques (Tabachnick & Fidell, 2001) and the second regarding the factor structure of the HMI-R

followed from the finding reported earlier about the substantial increase in variability for many items on the HMI-R compared with the HMI.
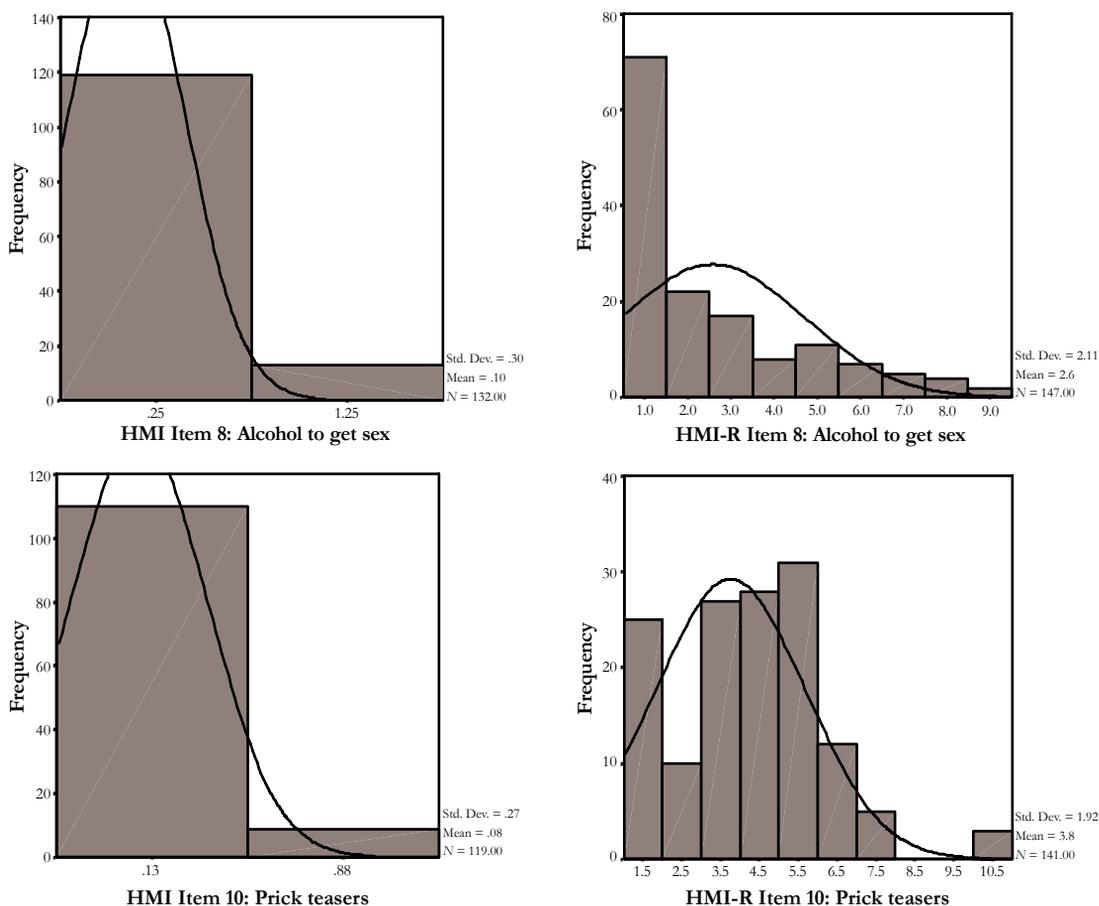
Missing Data

Although not included in our predictions, we noticed that there was a substantial difference in the amount of missing data between the two versions. Although the forced-choice format was recommended by Edwards (1953) as a way to reduce social desirable responding, he later noted that the format often results in unacceptable amounts of missing data (Edwards & Diers, 1962). In the present case, for the HMI there were a total of 161 items with missing data, compared with only 90 items for the HMI-R. Given the dichotomous data of the HMI, it is obviously impossible to impute values for missing data. Therefore, researchers using the HMI would have to eliminate many of the participants because of excessive missing data. This loss of participants leads to a loss of power to detect meaningful differences and increases the chances of making a Type II error. The reduced amount of missing data and the possibility of using imputation statistics (Tabachnick & Fidell, 2001) to correct for missing data when using the HMI-R both point to the superior psychometric properties of the HMI-R.

Factor Analysis

In constructing the HMI, Mosher and Sirkin (1984) selected items to reflect three theoretically derived aspects of hypermasculinity: Danger as

HMI Item 8: Alcohol to get sex
Std. Dev. = .30
Mean = .10
N = 132.00

HMI-R Item 8: Alcohol to get sex
Std. Dev. = 2.11
Mean = 2.6
N = 147.00

HMI Item 10: Prick teasers
Std. Dev. = .27
Mean = .08
N = 119.00

HMI-R Item 10: Prick teasers
Std. Dev. = 1.92
Mean = 3.8
N = 141.00

Note: HMI= Hypermasculinity Inventory; HMI-R = Hypermasculinity Index-Revised [Mosher, D. L., & Sirkin, M. (1984). Measuring a macho personality constellation. *Journal of Research in Personality, 18,* 150–163.]

exciting, physical fighting, and calloused attitudes. Despite the nominal level of measurement, Mosher and Sirkin conducted a primary axis factor analysis that revealed the presence of a single principle factor and eight uninterpretable factors. They concluded that the instrument was unifactoral. Nevertheless, they and most subsequent researchers have continued to use the three theoretical factors in their analysis of HMI data (for example, Bourg, 2001; Burk et al., 2004; Krahé & Fenske, 2002). The greater number of items with normal distributions yielded by the HMI-R would be expected to produce a more interpretable factor solution than did the HMI. We therefore conducted an exploratory factor analysis with varimax rotation. Examination of the scree plot indicated three of the seven factors with eigenvalues greater than 1.0 should be

retained. These factors accounted for 46.1% of the variation in scores and were easily interpretable. Consistent with the literature on hypermasculinity and machismo (Mosher & Anderson, 1986; Mosher & Sirkin), these factors included fighting, calloused sex, and danger as exciting. A total of eight items, however, did not load ($r > 0.40$) on any of the first three factors. In particular, four items, which on their face appeared related to calloused sexuality ("prick teasers deserve to be raped," "men need sex," "pick ups should expect to put out," and "all is fair in love and war"), loaded separately from the other calloused sexuality items. Given these results we next conducted a confirmatory restricted factor analysis, forcing the results on to three factors as indicated by the underlying theory of the HMI. The confirmatory factor analysis of the HMI-R

data accounted for 46.1% of the variance compared with 24% of the variance in Mosher and Sirkin's analysis of HMI data. In addition, all 27 of the HMI-R items loaded on the three expected factors related to danger as exciting, fighting, and calloused sexual attitudes (see Table 1). Only item number 9, "I like: a) fast cars and fast lovers or b) dependable cars and faithful lovers" loaded on both the calloused sex and danger as exciting factors. Given the content of the item, this dual loading makes sense and does not detract from the clear interpretation of the factors.

These factor analytic results support the construct validity of the HMI-R and demonstrate the psycho- metric advantages of the HMI-R which, unlike the original, detected the underlying three-factor structure of the construct of hypermasculinity, which explained almost twice the variance in scores.

DISCUSSION

The present study sought to extend the usefulness of the HMI as an explanatory or predictive instrument in research on attitudes held by men related to sexual assault of women by altering its dichotomous response format to a phrase-completion format (HMI-R). The HMI-R, like the forced-choice original, presents participants with a choice between two extreme possibilities but, unlike the original, allows them to select responses arranged physically and conceptually along the familiar "from 1 to 10" continuum.

Use of this phrase-completion format resulted in nine expected and unexpected positive results. Our first hypothesis was not supported: The HMI-R had lower variability than the original, although this lower variability was at least in part due to fewer extremely low or socially desirable scores on the HMI-R compared with the HMI.

As predicted in our second and third hypotheses, when compared with the forced-choice version, the HMI-R produced mean scores that were almost perfectly distributed along a normal curve. In addition, the HMI-R had only two items with unacceptable skewness, compared with 10 items for the HMI. Supporting our final expectation, the new phrase-completion version of the HMI had a greater internal consistency reliability than did the original forced-choice version.

Based on these promising results, we conducted additional analyses of the data. Specifically we looked at the distributions of the items with obviously high levels of social desirability. Because the goal of the HMI is to measure a socially undesirable trait, it must contain socially undesirable items. However, socially undesirable items are known to be psychometrically problematic because of their low rates of endorsement. With the HMI-R, almost 10 times the number of participants endorsed some degree of the socially undesirable response compared with the original HMI. This result suggests that when compared with the forced-choice response format, the phrase-completion format allows participants to more accurately report what are often considered socially undesirable attitudes or personality traits. The phrase-completion format

Table 1: Confirmatory Factor Analysis of HMI-R Data			
	Component		
Item Number and Description	**1**	**2**	**3**
12r Man needs sex	.703		
21r Pick-ups	.694		
8 Alcohol to get sex	.679		
25 Lesbians need respect/cock	.662		
10 Prick teasers	.638		
9r Fast cars and lovers[a]	.580		.431
22r Women deserve respect	.542		
5r Kind of parties I like	.429		
23 Consensual sex/ war of sex	.423		
24 Scientist or WWF wrestler			
18 Fighting is stupid		.770	
26r Fight if challenged		.763	
17 To fight or not		.725	
15) Fight to win/not fight		.652	
27r Fight if insulted		.614	
16r Fights are natural		.598	
3 Call me a name		.580	
11 Drink = mellow/trouble		.447	
7 Effeminate men deserve			
1 Dangerous experience			.781
14 Risk taker/play it safe			.757
6r Take risks			.678
2r Gamble/play it safe			.658
20 Drive safe/edge of danger			.580
19 If bored: TV or excitement?			.473
13r Drink relax			.446
4 In love and war			.416

Notes: Extraction method used was principal component analysis. Rotation method used was varimax with Kaiser normalization. HMI-R = Hypermasculinity Index-Revised.
[a]Rotation converged in five iterations.

therefore appears to help the researcher to detect subtle variations in the presence or absence of the underlying trait—variations that are masked by the forced-choice format.

Other unanticipated benefits of the phrase-completion version compared with the forced-choice version included substantially less missing data and an improved factor analytic solution that explained more of the variability in scores and provided a better fit of theory and data. Mosher and Sirkin (1984) constructed the HMI to tap into three aspects of hypermasculinity but did not find these three factors in their data. In fact they found a single factor and eight other uninterpretable factors. Our confirmatory analysis of the phrase-completion version produced a clear three-factor solution with factors precisely related to the three dimensions of calloused sexual attitudes, danger (risky behavior) and fighting, as originally hypothesized by Mosher and Sirkin. Given the differences in distributions on many of the items in the HMI, we suspect that Mosher and Sirkin's inability to detect the hypothesized three factors was at least partially due to the lack of variability and severe skewness of the majority of items in their original scale. Their use of factor analysis instead of the more appropriate latent class analysis statistic may have further contributed to their failure to find the desired factor structure. The ability of the phrase completion format version to reveal the latent structure of the instrument is an important advantage in establishing the construct validity of this measure of hypermasculinity.

Future Research

The results of this research indicate that further research on the HMI-R and on the use of the phrase-completion response format is warranted and likely to be productive. Regarding the HMI-R, replication of this study with more diverse samples is needed. Because we changed the wording of some items, validation of the HMI-R is also needed. Such validation studies might explore the relationship between hypermasculinity and other predictors of violence against women (see, for example, Moore & Stuart, 2005; Parrott & Zeichner, 2003). In addition, studies exploring the criterion validity of the HMI-R would be helpful.

In general, a good instrument, whether used to measure a personality trait, classroom learning, or social attitude, should be able to detect differences in that construct among study participants (DeVellis,

1991; Nunnally, 1970). The present study demonstrates that when compared with a forced-choice dichotomous response format, a phrase-completion response format improves a measure of hypermasculinity in its ability to detect even slight variations in the presence or absence of the underlying construct. This ability is especially important in measuring constructs with strong social desirability where participants are often unwilling to endorse an extremely unpopular or unflattering statement but are, as demonstrated in this study, willing to mildly endorse that statement when given that opportunity. The findings of this study indicate that the phrase-completion format appears especially well suited for research on topics likely to be subject to social desirability bias. Therefore, future research should compare a phrase-completion format with Likert-type formats in scales measuring attitudes and attributes such as rape or domestic violence myths, misogyny, sex-role stereotyping, and use of power and control in intimate relationships. If the results indicate that social desirability bias does not differ between randomly assigned groups but that the participants are more willing to endorse the socially undesirable attitudes and attributes when using the phrase-completion format, this would further support the superiority of the phrase-completion format for research in socially sensitive areas.

Limitations

The limitations of the present study should be noted and carefully weighed in interpreting the results. These limitations include its reliance on a nonrepresentative (predominantly white undergraduate student) sample of convenience, the lack of demographic information, and the absence of an independent measure of social desirability. A sample of college students was sufficient for comparing phrase-completion and forced-choice versions of the HMI, but exploration of the convergent and criterion validity of the HMI-R should use more diverse samples. In addition, future studies should more directly assess the effect of the phrase-completion response format on social desirability bias. **SWR**

REFERENCES

Abbey, A., McAuslan, P., Zawacki, T., Clinton, A. M., & Buck, P. O. (2001). Attitudinal, experiential, and situational predictors of sexual assault perpetration. *Journal of Interpersonal Violence*, *16*, 784–807.

Aberle, C. C., & Littlefield, R. P. (2001). Family functioning and sexual aggression in a sample of college men. *Journal of Interpersonal Violence*, *16*, 565–579.

Allport, G. W., & Ross, M. J. (1967). Personal religious orientation and prejudice. *Journal of Personality and Social Psychology*, *5*, 432–443.

Banyard, V. L., Plante, E. G., Cohn, E. S., Moorhead, C., Ward, S., & Walsh, W. (2005). Revisiting unwanted sexual experiences on campus: A 12-year follow-up. *Violence Against Women*, *11*, 426–446.

Bergman, S. J. (1991). Men's psychological development: A relational perspective [Working paper no. 48]. Wellesley, MA: Wellesley College, Stone Center.

Bourg, S. N. (2001). Sexual and physical aggression within a dating/acquaintance relationship: Testing models of perpetrator characteristics. *Dissertation Abstracts International*, *62*(1-B), 537.

Burgess, A. W., & Holmstrom, L. L. (1974). Rape Trauma syndrome. *American Journal of Psychiatry, 131,* 981–986.

Burk, L. R., Burkhart, B. R., & Sikorski, J. F. (2004). Construction and preliminary validation of the Auburn Differential Masculinity Inventory. *Psychology of Men & Masculinity*, *5*, 4-17.

Burt, M. R. (1980). Cultural myths and supports for rape. *Journal of Personality and Social Psychology*, *38*, 217–230.

Byers, E. S. (1996). How well does the traditional sexual script explain sexual coercion? Review of a program of research. *Journal of Psychology & Human Sexuality*, *8*(1), 7–25.

Calhoun, K. S., & Atkinson, B. M. (1991). *Treatment of rape victims: Facilitating psychosocial adjustment*. Elmsford, NY: Pergamon Press.

Calhoun, K. S., & Resnick, P. A. (1993). Post-traumatic stress disorder. In D. H. Barlow (Ed.), *Clinical handbook of psychological disorders* (pp. 48–98). New York: Guilford Press.

Catalano, S. M. (2004). *Criminal victimization, 2003*. Washington, DC: U. S. Department of Justice, Office of Justice Programs.

Check, J. V. P., & Malamuth, N. M. (1985). An empirical assessment of some feminist hypotheses about rape. *International Journal of Women's Studies*, *8*, 414–423.

Council on Social Work Education. (2001). *Educational policy and accreditation standards*. Alexandria VA: Author.

Cronbach, L. J. (1946). Response sets and test validity. *Educational and Psychological Measurement*, *6*, 475–493.

Crowell, N. A., & Burgess, A. W. (1996). *Understanding violence against women*. Washington, DC: National Academies Press.

Davis, R., Taylor, B., & Bench, S. (1995). Impact of sexual and nonsexual assault on secondary victims. *Violence and Victims*, *10*, 73–84.

DeVellis, R. F. (1991). *Scale development: Theory and applications*. Newbury Park, CA: Sage Publications.

Douglass, L. O. (2003). *The impact of two acquaintance rape prevention programs on cognitive elaboration in low and high hypermasculine men*. Dissertation Abstracts International, 63(8-B), p. 3975.

Dutton, D. G. (1998). *The abusive personality: Violence and control in intimate relationships.* New York: Guilford Press.

Edwards, A. L. (1953). The relationship between judged desirability of a trait and the probability that the trait will be endorsed. *Journal of Applied Psychology*, *37*, 90–93.

Edwards, A. L., & Diers, C. J. (1962). Social desirability and the factorial interpretation of the MMPI. *Educational and Psychological Measurement*, *22*, 501–509.

Federal Bureau of Investigation. (2002). *Crime in the United States 2002*. Washington, DC: Author.

Fisher, B., Cullen, F., & Turner, M. (2000). *The sexual victimization of college women*. Washington, DC: U.S.

Department of Justice, Office of Justice Programs, National Institute of Justice.

Gilligan, C. (1982). *In a different voice: Psychological theory and women's development*. Cambridge, MA: Harvard University Press.

Hamburger, M. E, Hogben, M., McGowan, S., & Dawson, L. J. (1996). Assessing hypergender ideologies: Development and initial validation of a gender-neutral measure of adherence to extreme gender role beliefs. *Journal of Research in Personality*, *30*, 157–178.

Hodge, D. R., & Gillespie, D. (2003). Phrase completions: An alternative to Likert scales. *Social Work Research*, *27*, 45–55.

Johnson, G. M., & Knight, R. A. (2000). Developmental antecedents of sexual coercion in juvenile sexual offenders. *Sexual Abuse*, *12*, 165–178.

Kilpatrick, D. G., Acierno, R., Resnick, H. S., Saunders, B. E., & Best, C. L. (1997). A 2-year longitudinal analysis of the relationship between violent assault and substance use in women. *Journal of Consulting and Clinical Psychology*, *65*, 834–847.

Koralewski, M. A., & Conger, J. C. (1992). The assessment of social skills among sexually coercive college males. *Journal of Sex Research*, *29*, 169–189.

Koss, M. P., & Gaines, J. A. (1993). The prediction of sexual aggression by alcohol use, athletic participation, and fraternity affiliation. *Journal of Interpersonal Violence*, *8*, 94–108.

Koss, M. P., Gidycz, C. A., & Wisniewski, N. (1987). The scope of rape: Incidence and prevalence of sexual aggression and victimization in a national sample of higher education students. *Journal of Consulting and Clinical Psychology*, *55*, 162–170.

Krahe, B., & Fenske, I. (2002). Predicting aggressive driving behavior: The role of macho personality, age, and power of car. *Aggressive Behavior*, *28*, 21–29.

Krahé, B., Waizenhöfer, E., & Möller, I. (2003). Women's sexual aggression against men: Prevalence and predictors. *Sex Roles*, *49*, 219–232.

Loh, C., Gidycz, C. A., Lobo, T. R., & Luthra, R. (2005). A prospective analysis of sexual assault perpetration: Risk factors related to perpetrator characteristics. *Journal of Interpersonal Violence*, *20*, 1325–1348.

Monnier, J., Resnick, J. S., Kilpatrick, D. G., & Seals, B. (2002). The relationship between distress and resource loss following rape. *Violence and Victims, 17,* 85–91.

Moore, T. M., & Stuart, G. L. (2005). A review of the literature on masculinity and partner violence. *Psychology of Men & Masculinity*, *6*, 46–61.

Mosher, D. L., & Anderson, R. D. (1986). Macho personality, sexual aggression, and reactions to guided imagery of realistic rape. *Journal of Research in Personality*, *20*, 77–94.

Mosher, D. L., & Sirkin, M. (1984). Measuring a macho personality constellation. *Journal of Research in Personality*, *18*, 150–163.

Murnen, S. K., & Byrne, D. (1991). Hyperfemininity: Measurement and initial validation of the construct. *Journal of Sex Research*, *28*, 479–489.

Nunnally, J. C., Jr. (1970). *Introduction to psychological measurement*. New York: McGraw-Hill.

Palmer, C. (1989). Is rape a cultural universal? A re-examination of the ethnographic data. *Ethnology*, *28*(1), 1–16.

Parrott, D. J., & Zeichner, A. (2003). Effects of hypermasculinity on physical aggression against women. *Psychology of Men & Masculinity*, *4*, 70–78.

Randall, M., & Haskell, L. (1995). Sexual violence in women's lives: Findings from the Women's Safety Project, a community-based survey. *Violence Against Women*, 1, 6–31.

Resnick, H.S., Kilpatrick, D. G., Dansky, B. S., Saunders, B. E., & Best, C. L. (1993). Prevalence of civilian trauma and posttraumatic stress disorder in a representative sample of women. *Journal of Consulting and Clinical Psychology*, *61*, 984–991.

Rodkin, L. I., Hunt, E. J., & Cowan, S. D. (1982). A men's support group for significant others of rape victims. *Journal of Marital & Family Therapy*, *8*, 91–97.

Saunders, B. E., Villeponteaus, L. A., Lipovsky, J. A., Kilpatrick, D. G., & Veronen, L. J. (1992). Child sexual abuse as a risk factor for mental disorders among women: A community survey. *Journal of Interpersonal Violence*, 7, 189–204.

Stöber, J., Deete, D. E., & Musch, J. (2002). Comparing continuous and dichotomous scoring of the Balanced Inventory of Desirable Responding. *Journal of Personality Assessment*, 78, 370–389.

Sullivan, J. P., & Mosher, D. L. (1990). Acceptance of guided imagery of marital rape as a function of macho personality. *Violence and Victims*, *5*, 275–286.

Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics* (4th ed.). Boston: Allyn & Bacon.

Tjaden, P., & Thoennes, N. (2000). *Full report of the prevalence, incidence, and consequences of violence against women: Findings from the National Violence Against Women Survey.* Washington, DC: U.S. Department of Justice and the Centers for Disease Control and Prevention.

Tyler, K. A., Hoyt, D. R., & Whitbeck, L. B. (1998). Coercive sexual strategies. *Violence and Victims*, *13*, 47–61.

Jay Peters, PhD, is assistant professor, School of Social Work, University of Maine, 5770 Social Work Building, Orono, ME 04469; e-mail: jpeters@maine.edu. *Carey Nason, PhD,* is coordinator, Safe Campus Project, and **Winston M. Turner, PhD,** is adjunct professor, School of Social Work, University of Maine, Orono. The authors thank Donald P. Mosher, PhD, and Mark Sirkin, PhD, who not only granted the authors permission to revise their Hypermasculinity Index (HMI) and to label it HMI-R, but also encouraged them in the endeavor.

Appendix A: Hypermasculinity Inventory-Revised

1. my knees feel weak and I shake all over.	2	3	4	5	6	7	8	9	10. I feel high.

1). After I've gone through a really dangerous experience

2). I'd rather	
gamble than play it safe.	play it safe than gamble.
3). Call me a name and	
I'll pretend not to hear you.	I'll call you another.
4). In love and war	
you should still play by the rules.	anything goes.
5). When I go to parties	
I like wild, uninhibited parties.	I like quiet parties with good conversations.
6). Some people have told me	
I take foolish risks.	I ought to take more chances.
7). So-called effeminate men	
are more artistic and sensitive.	deserve to be ridiculed.
8). Using drugs or alcohol to "encourage" a woman to have sex with you is	
gross and unfair.	OK if you can get away with it
9). I like	
fast cars and fast lovers.	dependable cars and faithful lovers.
10). So-called prick teasers	
should be forgiven.	should be raped.
11). When I have had a few drinks	
I mellow out.	I look for trouble.
12). Any man who is a man	
needs to have sex regularly.	can do without sex.
13). When I have a drink or two	
I feel ready for whatever happens.	I like to relax and enjoy myself.
14). When it comes to taking risks	
I like to play it safe.	I'm a high roller.
15). In conflicts with others	
I win by not fighting.	I fight to win.
16). Getting into fights	
is natural for me.	never solves an issue.

(continued)

17). When I feel like fighting I
    try to think of alternatives.   go for it.

18). Given what I know about fighting,
    it's just stupid.   he who can, fights; he who can't runs away.

19). When I'm bored
    I watch TV or read a book.   I look for excitement.

20). I like to
    drive safely, avoiding all unnecessary risks.   drive fast, right on the edge of danger.

21). So-called pick-ups should
    expect to put out.   choose their men carefully.

22). In my opinion
    some women are good for only one thing.   all women deserve the same respect as men.

23). When it comes to having sex
    I only want to have sex with someone who is in total agreement.   I never feel bad about my tactics when I
                                                         have sex.

24). I would prefer to be a
    famous scientist.   famous WWF wrestler.

25). Lesbians have a particular lifestyle
    and should be respected for it.   but really just need a good, stiff cock.

26). If someone challenges you to a fight,
    there's no choice but to fight.   it's time to talk your way out of it.

27). If you insult me,
    be prepared to back it up.   I'll try to turn the other cheek.