

Ten Years of Psychological Research on Men and Masculinity in the United States: Dominant Methodological Trends

MySha R. Whorley · Michael E. Addis

Published online: 1 December 2006
© Springer Science + Business Media, Inc. 2006

Abstract The psychological study of men and masculinity is rapidly growing. Despite this growth, few attempts have been made to critically evaluate progress in the field, and no studies to date have analyzed methodological trends in empirical research on men and masculinity. The purpose of this study was to identify the dominant trends in psychological research on men and masculinity published in the United States between 1995 and 2004. One hundred seventy-eight articles were selected and coded on 21 dimensions. Descriptive statistics revealed that quantitative, correlational, and nonobservational methods dominate research in this field. Further, the inclusion of racial/ethnic minorities was found to be quite low and researchers tended to rely on convenience samples of undergraduates. It also appears that there was little variety in the masculinity measures used. Implications and recommendations for future research are discussed.

Keywords Masculinity · Men · Methodology · Research · Review

Research in the field of men and masculinity is new and rapidly growing in the United States. In the last 6 years, several journals have been founded that are dedicated to the study of men and masculinity (Smiler, 2004). In addition, the last 10 years have shown increased representation of men and masculinity researchers in the American Psychological Association and the start of an APA division

dedicated to the psychological study of men and masculinity. The psychology of men has also become increasingly visible in the American popular media, with such publications as the *New York Times* and *Newsweek* covering stories about men and depression, men's health, and other related topics. Finally, the larger field of mental health has gained awareness of the significance of studying the psychology of men and masculinity. For example, the National Institute of Mental Health has acknowledged the importance of this field by introducing a campaign called "Real Men, Real Depression" that seeks to educate people about men's suffering. In short, the last decade has seen a dramatic increase in the amount of social scientific theory and empirical research focused on the psychology of men and masculinity.

Despite increased interest in the psychology of men and masculinity, few comprehensive reviews have been published (Betz & Fitzgerald, 1993; Good, Borst, & Wallace, 1994; Smiler, 2004), and to our knowledge there have been no reviews and critiques of methodological trends in particular. The purpose of the current article is to review methodological trends in psychological research on men and masculinity published in major US journals over the last 10 years. We chose to focus our review on research published in major US journals for several reasons. First, space constraints required narrowing the boundaries of the review. Second, and more importantly, research published in these outlets is likely to reflect the dominant methodological and conceptual trends in psychology more broadly, and these trends in turn shape the way research on men and masculinity is conceptualized and carried out. Third, dominant methodologies in any field tend to marginalize those that are less dominant and typically do so by remaining unquestioned and granted legitimacy by their status as widely accepted approaches. Thus, by revealing

Preparation of this article was supported by grant no. R34 MH073073 from the National Institute of Mental Health.

M. R. Whorley (✉) · M. E. Addis
Department of Psychology,
Clark University, 950 Main St.,
Worcester, MA 01610, USA
e-mail: mwhorley@clarku.edu

clearly those methodological trends that are most pervasive in mainstream American journals, we hoped to learn about methodologies that are *less* common but valuable for progress in the field. Finally, such dominant trends are important to analyze because they not only reflect the sorts of research questions and findings that have emerged, but they also shape and constrain the sorts of future questions that can be asked (Danziger, 1994). For example, an informal review of the literature reveals a tendency by researchers in the US to study homogeneous populations with quantitative, correlational methods and to use the same handful of masculinity measures. It also appears that much of the research is employed in the interest of demonstrating relationships between trait-like masculinity measures and other variables of interest (Addis & Mahalik, 2003; Smiler, 2004).

Although implicitly accepted methodological trends can be efficient for establishing emerging disciplines, relying too heavily on particular methods can restrict a new field's ability to advance theoretically and empirically. In order for such advancement to take place, researchers should focus not only on amassing data, but also on the larger social context within which their work is situated (Marecek, 1995). Further, in order to move forward, researchers studying men and masculinity need to understand how the methods drive which research questions are chosen and which are not. For example, trait-like measures of masculinity necessarily constrain us to asking and answering questions about stable differences between individuals rather than variability within individuals across situations. The latter questions are essential if one is interested in understanding how masculinity roles and norms can change for groups and individuals.

To date, there have been three comprehensive reviews of psychological research on men and masculinity published in the United States (Betz & Fitzgerald, 1993; Good et al., 1994; Smiler, 2004). Over 10 years ago, Betz and Fitzgerald (1993) pointed out that two main themes characterized masculinity research: a focus on the prescriptive nature of masculine socialization and a focus on the impact of that socialization on therapy with men. At that time, the field was dominated by research on gender role conflict and masculine gender role stress, thus their review focused primarily on research using those two paradigms and research on the application of those paradigms in therapy and assessment. In addition, researchers focused mostly on measuring conflict and stress related to masculinity and correlating them with negative outcomes. One year later, Good et al., (1994) noted in a comprehensive review of the field that there had been an increase in theories and models of masculinity ideology and in writings concerned with the detrimental effects (e.g., masculine role conflict) of endorsing masculinity ideology. They found

that empirical research on those same issues was not as progressive as theory, however. In their review, they remarked on the restrictive methodological designs of masculinity ideology research as well as challenges in sampling and instrument development. More recently, Smiler (2004) reviewed the empirical psychological literature on men and masculinity and its related instrumentation. He identified several approaches or movements in the study of masculinity such as the androgyny movement, the ideology movement, the strain movement, and the deconstruction movement. Smiler concluded that masculinity theory has been rather consistently dominated by an essentialist (i.e., acontextual) viewpoint. Recent theory has seen some inclusion of social and sociocultural influences, but masculinity has continued to be positioned as something residing in the individual. He argued that the field's reliance on acontextual approaches to masculinity might be one explanation for why masculinity theories are not influencing mainstream psychology.

The authors of the reviews cited above all noted that more diverse methods should be employed, but none quantitatively summarized the methods currently being used and their impact on the field. There are several reasons why researchers should be mindful of the trends in their methods. First, researchers should know if the field has relied primarily on correlational methods (i.e., methods that examine linear relationships between variables). If it is, the types of questions that can be addressed will be severely restricted to those examining simple relationships between masculinity-related variables and other variables of interest. Second, researchers should know what samples they are studying in order to keep generalizations culturally and contextually specific. For example, Good et al., (1994) reported that research using more diverse samples has tended to identify different patterns of endorsement of masculinity-related characteristics than research without diverse samples. Studying masculinity in diverse samples is important for conceptual, methodological, and empirical reasons (see Graham, 1992; Sue, 2003). Conceptually, it keeps research focused on variations in the way masculinity is culturally and socially constructed and therefore aids in avoiding implicitly essentialist models of gender. Methodologically and empirically, studying diverse samples is critical for assessing the generalizability of findings from convenience samples (e.g., largely White undergraduate students). Finally, without knowing what sorts of measures are being used, researchers cannot ensure that they are measuring constructs in creative and multiple ways. Being aware of what tools we use and how often we use them may stimulate the field to develop new, creative, psychometrically sound instruments.

The purpose of the current study was to identify dominant methodological trends in masculinity research in

the US over the last 10 years. Our specific questions were as follows:

1. At what rate are different methods being employed? Are the majority of studies quantitative and correlational, and are researchers using observational methods?
2. Are certain methods being used exclusively with certain types of samples? Is there a chronological trend in methods used?
3. At what rate are ethnic minorities being included in the research? What percentage of studies included the various minority groups, and what percentage of the sample does each minority group constitute? At what rate are women being included in the research? In addition, we were interested in the rates at which clinical samples of children and adults were being studied. This is important to assess because it has been widely suggested that masculine gender role socialization has harmful effects that may put individuals at risk for psychopathology and make them resistant to treatment. However, an informal survey of the literature suggests that clinical samples are rarely studied.
4. Is there diversity in the measures of masculinity used?

Method

Article selection

One hundred seventy-eight articles published between 1995 and 2004 were selected for inclusion in this study. To place some boundaries on our review and create a manageable analysis, we chose to focus on the last 10 years. This seemed an appropriate time for such a review because it has been roughly 10 years since the last comprehensive review of the field was conducted. In addition, one of the most widely cited texts in American research on men and masculinity, *A New Psychology of Men*, was published 10 years ago (Levant & Pollack, 1995). Because we set out to identify dominant trends in masculinity research in the US, we selected journals with the widest readership that were specifically psychologically oriented in title or reputation and most likely to publish research on men and masculinity. Therefore, all articles from *Psychology of Men and Masculinity* were included. In addition to publishing a large proportion of the research in this field, *Psychology of Men and Masculinity* also began publication during our selected timeframe. We also included select articles from other mainstream, explicitly psychological journals published in the US, including the *Journal of Counseling and Development*, *Sex Roles*, the *Journal of Counseling Psychology*, and the *Journal of Personality and Social*

Psychology because they appear to publish research on the psychology of men and masculinity regularly.

To locate articles, a *PsycINFO* search of the above journals was conducted using the following terms: men, man, masculinity, masculine, male, boys, and gender role conflict. We searched this last term specifically because gender role conflict is an extremely common construct that was developed specifically with men and masculinity in mind. Other similar constructs such as gender role stress are typically preceded by our other search terms (i.e., male or masculine) and were not used specifically as search terms. Once a preliminary list was compiled based on those search terms, the list was checked carefully for appropriateness of the articles. There were three criteria for excluding articles. First, we were careful not to include studies that focused only on sex differences because such studies are typically not focused on the psychology of men and masculinity per se. Second, we made sure that each article actually had men and/or masculinity as the subject of the research. Several studies came up in the search that were actually studies of women's experience. For example, one mistakenly labeled study of the encouragement of female athletes was found in our search (Weiss & Barber, 1995). The focus of that study was on women, not men or masculinity, and was thus ruled out. Third, we removed miscellaneous other articles that came up that were irrelevant to our research, such as book reviews or articles that were not published in the selected journals.

Article coding

The articles were coded on 21 dimensions. Seventeen of those dimensions related to the demographics of the samples: number of men in the sample; number of women in the sample; mean age of the participants in the sample; sample type (undergraduate, adult nonclinical, adult clinical, child nonclinical, child clinical); and number and percentage of European-American, African-American, Asian/Pacific Islander, Native American, Latino, and Other participants in the sample. We also examined the method section of each study and listed the specific measures of masculinity used and how many different measures were included. We drew our list of measures to code from Thompson and Pleck's (1995) review of masculinity measures as well as any other measure found in a coded study that included the words men, male, macho, masculine, or masculinity in the scale name.

The last three coded dimensions were discrete categories each related to methodology. We coded whether or not observational measures were used; a "yes" was recorded when there was any observation of participants that led to some sort of coded data whether or not it was used quantitatively (e.g., Mackey, 2001). In the predominant methodology category we coded whether a study was quantitative, qualitative, or

theoretical/conceptual. Every study was placed into one of these categories. A study was coded quantitative if there were any inferential statistics calculated (e.g., Hayes & Mahalik, 2000). Any studies where authors did not report statistics were coded qualitative. Examples of qualitative studies included those that reported interviews or content analyses (e.g., Thomas, 2003). Theoretical/conceptual studies were those that had as their purpose advancing new theories or critically reviewing existing ones (e.g., Courtenay, 2000a, 2000b). Examples of these include literature reviews and essays. We thought that these were important to include because they are indicative of the frequency with which conceptual work is being conducted in the field. The final methodological category was type of study. Again, each study was placed into one of four categories. Correlational studies were those that examined relationships between two or more variables of interest without manipulating any variables. Studies were coded experimental if variables were manipulated. Coding a study as psychometric occurred whenever the purpose of a study was to develop or validate a measure. Studies were coded "other" when they didn't fit into any of the aforementioned categories. Examples included literature reviews and any other study with no data. For articles in which several studies were described (under the headings "study 1," "study 2," and so on), each study was coded independently. This procedure yielded 167 published articles that included 198 separate studies.

Reliability

The articles were divided between two coders, the first author and an undergraduate research assistant, with a 30% overlap to assess reliability. Cohen's (1960) kappa was calculated for each of the discrete coding categories (type of sample, type of study, predominant methodology, and use of observational methods). The kappas were $K=0.90$ for type of sample, $K=0.66$ for type of study, $K=0.88$ for predominant methodology, and $K=0.46$ for use of observational methods. Fleiss (1981) considers kappas of 0.40 to 0.60 fair, 0.60 to 0.75 good, and over 0.75 excellent. Because kappa corrects for chance agreement it can be significantly lowered when particular categories are assigned at a low frequency. This was apparently the case with observational methods since they were rated infrequently ($K=0.46$, inter-rater agreement, 92.9%).

Results

Characteristics of the methods in coded studies

Figure 1 shows the frequencies of the three predominant methodologies: quantitative, qualitative, and theoretical/

conceptual. The majority of the coded studies employed quantitative methods ($N=167$, 84.3%). Frequencies follow for the four types of studies: correlational ($n=117$), experimental ($n=26$), psychometric ($n=9$), and other ($n=46$). The majority ($n=117$, 59.1%) of the coded studies used a primarily correlational method design. Descriptive analysis also shows that the overwhelming majority ($n=188$, 94.9%) of coded studies did not employ any observational methods.

To determine whether certain methods were used more often with certain populations, we selected for each type of sample (e.g., undergraduate, adult nonclinical, adult clinical, child nonclinical) and calculated chi-squares for type of study and predominant methodology. For studies with undergraduate samples, correlational methods were used significantly more often than either experimental, psychometric, and other methods, $\chi^2(3, n=89)=73.74, p<0.001$. The same pattern emerged for studies with adult nonclinical samples, $\chi^2(3, n=58)=78.55, p<0.001$ and studies using nonclinical samples of children, $\chi^2(2, n=18)=16.00, p<0.001$.

For each of the 10 years of publication covered in this study, we examined the frequencies of study type, predominant methodology, and sample type to determine whether there was a chronological trend in sampling and methodology (Figs. 1, 2, and 3, respectively). The number of correlational studies has varied considerably over the years, but with the exception of 1999, correlational studies have consistently been more abundant than other types of studies. Studies with the coding category of other have increased since 1998 but have remained fairly steady at around five or six per year. The frequency of experimental

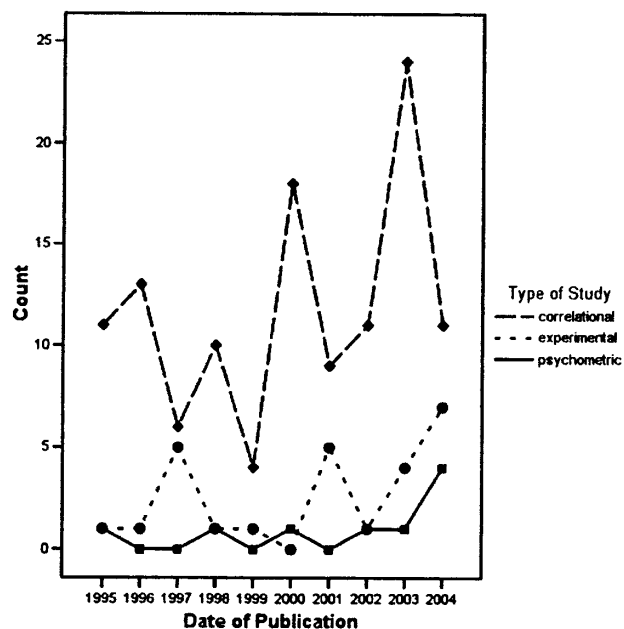


Fig. 1 Type of study frequencies by publication.

studies has also varied over the years but has been increasing steadily over the past 3 years. Psychometric studies have hovered around one per year with a possible trend upwards in the last few years.

As Fig. 2 shows, there has been a fairly consistent tendency for quantitative methods to be employed at a greater frequency than both qualitative and theoretical/conceptual methods. In the last 2 years, the frequency of using quantitative methods looks to be particularly high, while the use of qualitative and theoretical/conceptual methods has consistently been under five studies per year.

Figure 3 shows that the use of undergraduate samples has consistently been higher than the use of the other types of samples, with the exception of an equal number of adult nonclinical samples in the year 2000. The use of undergraduate samples has seen a surge in the past 2 years. Adult nonclinical samples have consistently been the second most frequently used. In our coded studies, there is no record of adult clinical samples before 1999 and the number of studies using that type of sample has remained low since then. The number of studies using child nonclinical samples has remained consistently low as well with only a slight increase in the last 5 years.

Characteristics of the samples in coded studies

Descriptive statistics were calculated to answer the research question regarding the rate at which ethnic minorities and women are being included in research. One outlier with a sample size of 58,000 was excluded from the analyses due to its extremity and potential for biasing mean estimates. The size of the samples in the remaining coded studies ($n=197$) ranged from 0 to 5,201 ($M=229.24$, $SD=454.67$). The median

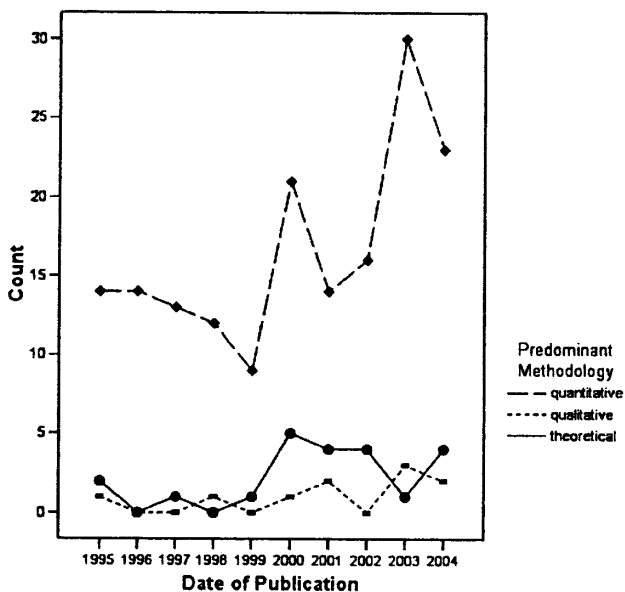


Fig. 2 Predominant methodology frequencies by publication year.

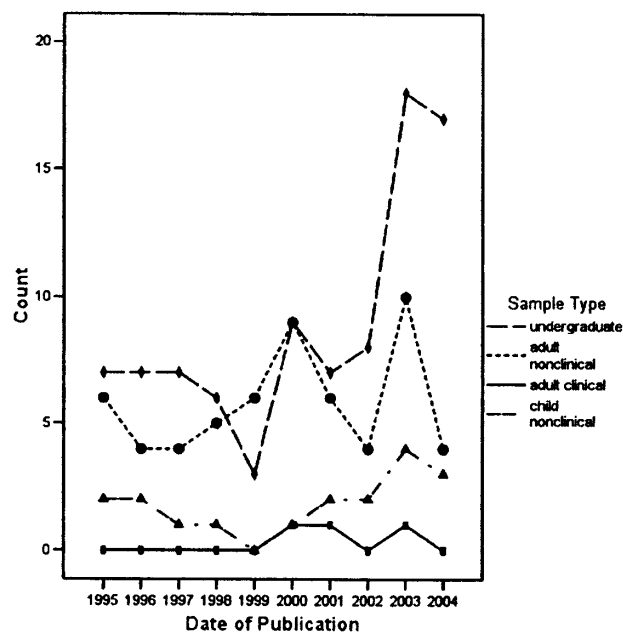


Fig. 3 Sample type frequencies by publication year.

sample size was 140. Thirty studies did not use participants either because they were theoretical/conceptual papers or because they conducted qualitative analyses of nonhuman data. For example, one study analyzed the content of children's books about fathers and thus used no participants (Fleming & Tobin, 2005). Of the remaining 167 coded studies with participants, 52 (31.1%) failed to report inclusion of racial/ethnic minorities. Out of the 115 studies that did report ethnicity, we calculated the average percentage of each minority group included in all studies and the average percentage of each minority group included in studies that reported greater than zero of that group. As Table 1 shows, there was a low rate of inclusion of ethnic minorities in research. For example, African Americans and Latinos comprised only 8.85 and 5.87%, respectively, of participants across all studies. Only 20.4% of studies reported including women. However, of those studies that did include women as participants, they made up, on average, close to 60% of the participants ($M=0.59$, $SD=0.19$).

Table 1 Minority group inclusion in samples ($N=115$).

Group	Percent studies reporting this group	Mean sample percentage	
		In all studies	In studies including this group
African American	81.74	8.85	13.75
Asian/Pacific Islander	67.83	4.92	9.22
Native American	31.30	0.57	2.30
Latino	70.43	5.87	10.58
Other	59.13	6.55	14.06

Of the 198 coded studies, 30 (15.2%) contained no sample, 89 (44.9%) had primarily undergraduate samples, 58 (29.3%) contained samples that were adult nonclinical, 3 (1.5%) had adult clinical samples, 18 (9.1%) had child nonclinical samples, and no coded studies contained clinical samples of children. The mean age of the participants in the samples was 25.54 (SD=10.26).

Diversity of masculinity measures

Descriptive statistics revealed that 108 (64.3%) of the 168 empirical studies we examined did not use any of the coded measures of masculinity. Some of these studies examined men's attitudes towards variables of interest such as intimacy and had no need for masculinity measures (see Derlega, Catanzaro, & Lewis, 2001). Other studies used measures that did not fall within our "masculinity measure" coding system such as those examining muscularity (i.e., T. G. Morrison, M. A. Morrison, & Hopkins, 2003). The modal number of masculinity measures used was 0. Of the 90 studies that did use masculinity measures, 62 (68.9%) studies used 1 measure, 20 (22.2%) studies used 2 measures, 5 (5.6%) studies used 3 measures and 3 (3.3%) studies used 4 measures of masculinity. Table 2 shows the rate of usage of the masculinity measures for which we coded. Of the studies that used a masculinity measure, 64.4% used the Gender Role Conflict Scale (GRCS; O'Neil, Helms, Gable, David, & Wrightsman, 1986).

Table 2 Frequency of masculinity measure usage.

Measure	Date of publication	Number	Percent ^a
Gender role conflict scale	1986	58	64.4
Masculine gender role stress scale	1987	16	17.8
Male role norms scale	1986	11	12.2
Personal attributes questionnaire	1978	11	12.2
Hypermasculinity inventory	1984	8	8.9
Attitudes toward women scale	1972	7	7.8
Male role norms inventory	1992	5	5.6
Others ^b	–	12	13.3

^aThese percentages were calculated based on 90 studies that reported using masculinity measures, not based on the total number of coded studies.

^bThere were seven other measures of masculinity reported. They included the Conformity to Masculine Norms Inventory (Mahalik et al., 2003), the Multicultural Masculinity Ideology Scale (Doss & Hopkins, 1998), the Auburn Differential Masculinity Inventory (Burk, Burkhart, & Sikorski, 2004), the Drive for Muscularity Scale (McCreary & Sasse, 2000), the Macho Scale (Villemez & Touhey, 1977), and the Brannon Masculinity Scale (Brannon & Juni, 1984). We grouped them because of their low frequencies.

Discussion

The purpose of this study was to examine dominant methodological trends in masculinity research published in mainstream US psychology journals over the last 10 years. We questioned whether the field was employing a diversity of methods and found that the majority of the studies in the mainstream US psychological journals we examined were quantitative and correlational. We also examined the diversity of the populations sampled and found that 48% of studies included minorities or women. The studies that did include minorities tended to include a very small proportion of each minority group. Finally, we questioned whether a diverse range of masculinity measures were being utilized in these studies and our results suggested that only a handful of masculinity measures are being used consistently to study men and masculinity.

Our review suggested that research on men and masculinity published in mainstream US journals is dominated by quantitative, correlational, nonobservational research. Over 80% of the coded studies used quantitative methodologies, almost 60% of the coded studies were primarily correlational, and an overwhelming majority (94.4%) used no observational methods. Reliance on these methods severely restricts the types of questions we can ask. For example, relying solely on quantitative methods may lead to a more general tendency of placing importance or value only on those phenomena that can be quantified. One remedy is for researchers to make greater use of qualitative methods. For example, Edley and Wetherell (1999) used discourse analysis to explore men's ideological dilemma regarding fatherhood and domestic life. Many researchers in the US and other countries have utilized a range of qualitative methods to explore various aspects of masculinity. However, these studies have rarely found their way into the major outlets that dominate mainstream US research in this area.

Dependence on correlational methods has four important consequences. First, and most importantly, correlational methods dictate that we examine simple linear relationships; even in the case of multivariate correlational techniques the question remains, "How are two or more variables linearly related?" This relatively simple analytic approach stands in stark contrast to the complexity of gender as a set of social formations and processes. Even if researchers choose to remain within a quantitative framework, there are undoubtedly aspects of the way masculinity operates that are nonlinear or systemic. We would argue that our proclivity for studying simple linear relationships has the hidden effect of limiting the relevance of findings in the psychology of men and masculinity to other subdisciplines in psychology. Second, correlational research designs do not allow us to test theories of causality and thus the

direction of influence between various masculinity norms and other psychosocial processes remains obscured. A good example of research that did allow for testing the direction of a relationship was a study by Eisler, Franchina, Moore, Honeycutt, and Rhatigan (2000). In this study, the researchers sought to test the effect of gender-relevance on men's attributions about relationship conflict. In order to do this, they showed different groups of participants different vignettes, some highly gender-relevant and others not at all gender-relevant. This design allowed them to make statements about the direction of effect between gender relevance and negative attributions during conflict. A simple correlational design examining the relationship between these variables would not have allowed them to make such statements. Third, studying simple linear relationships often leads researchers to treat contextual or demographic variables (e.g., race/ethnicity, gender, and social class) as nuisance variables instead of meaningful mediators of men's experience. Finally, correlational methods necessarily presume the existence of stable individual differences and place these at the forefront of analyses. As a result, "masculinity" most often emerges in theory and measurement as a trait-like variable, despite the common theoretical assumption that the social construction and social learning of masculinity is a historical, developmental, and fluid process (Levant & Pollack, 1995; Pleck, 1981). We found no studies, for example, that explicitly set out to examine changes in masculine norms, ideologies, or role conflicts as a function of microvariations in the social contexts of men's lives. In short, correlational analyses of assumedly stable individual differences may often be at odds theoretically with some of our most basic assumptions about the psychology of men and masculinity.

The overwhelming majority of research participants in the studies coded were nonclinical, European-American, male undergraduates. Clinical samples were used in only 1.5% of the studies we coded. Only 20.4% of the studies included women in their samples. Similarly, of all studies reporting minorities (men and women), the average percentage of a given minority group included was never above 15%. It could be argued that the comparatively low numbers of ethnic minority participants in research is consistent with US census figures and is therefore not problematic. However, such an argument continues to presume that it is less important to study ethnic minorities simply because they are statistically in the minority. Moreover, when minority participants *remain in the minority within a particular research sample*, less emphasis is typically placed on understanding particular ethnic or racial groups in their own right. Instead, they are viewed in comparison to the majority of the sample that is most often White.

What is especially surprising is that over 30% of the coded empirical studies did not report any racial/ethnic

breakdown at all. Of studies that did report a breakdown of race and ethnicity, 52% reported that no minorities were included in their samples. This is concerning because there is ample evidence that masculinities vary along racial, ethnic, and social class dimensions (e.g., Connell, 1995). It appears researchers publishing in these journals are at least attempting to incorporate minority populations into studies. However, researchers might benefit from conducting research focused specifically on these populations per se in order to more fully understand how masculinity influences the lives of a wide range of men and women.

One limitation of the current study is that we did not code for sexual orientation. This was primarily because few studies that we coded actually reported this information. In retrospect, it would have been worth quantifying whether authors reported the sexual orientation breakdown in the sample, and future studies should explore the rate of inclusion of sexual minorities in masculinity research. Including sexual minorities is important because particular norms may be less relevant for some men. For example, adherence to anti-homosexuality norms may or may not be important to sexual minorities.

The narrow range of populations studied has several consequences. First, neglecting to include racial/ethnic minorities in research means that we are necessarily limited in our understanding of the diversity of men's experiences. Despite the prevalence of powerful "hegemonic" masculinity norms in western cultures (Connell, 1995), there is no generic psychology of men and masculinity. Rather, men's experience with regard to a variety of important psychological processes and outcomes varies according to race, ethnicity, and class. The argument here is not necessarily for cross-racial or ethnic comparisons in research, although such studies may be beneficial depending on the particular questions at hand. What we are suggesting is that researchers publishing in high visibility journals in the US need to do a much better job integrating an understanding of diversity directly into our analysis of the psychology of men and masculinity (Liu, 2005).

Second, the lack of research on diverse populations probably decreases the likelihood of attracting researchers interested in studying these populations. For example, although the relative lack of research on Men of Color may motivate some researchers to begin work in this area, for others it may create significant obstacles such as the absence of a body of literature to build on and a small or nonexistent community of colleagues with similar interests. One way out of this catch-22 is for established researchers in the field of men and masculinity to make a stronger effort to study diverse populations of men.

The tendency to underrepresent clinical populations has three important consequences. First, the findings from nonclinical populations of men may not be replicable in

clinical populations. Much of the research on the psychology of men and masculinity is conducted with an eye toward bettering men's, women's, and children's lives by understanding and ultimately reducing the harmful effects of restrictive masculinity norms (e.g., Levant & Pollack, 1995). Particularly with research on men and various aspects of psychopathology, it is critical to replicate findings from "healthy" undergraduates with men from clinical populations; the significant public health implications of such research demand that it be replicated.

Second, it is possible that the attitudes of nonclinical men in regard to such issues as help-seeking are different than those of men in clinical populations. For example, many men in the nonclinical population have never sought help for a mental health problem. By virtue of never having sought help, some of these men may believe that there is a lack of a sufficient reason to seek psychological help for problems or that severe problems can be handled independently. Alternatively, men who have a history of seeking help for anxiety or depression may believe that seeking help is vital for maintaining mental health. In other words, having interacted with the mental health care system may change attitudes toward future interactions.

A third consequence is that the phenomena that arise for study may be different in nonclinical versus clinical populations. In clinical populations, for example, it is of interest to researchers how men feel about different types of therapy. The same question may have no relevance to nonclinical populations of men. Similarly, studying alcohol consumption in nonclinical college students may obscure the interaction between anxiety or depression and substance use in clinical populations. Thus, it is crucial to study diverse populations and keep both questions and conclusions population-specific.

One consequence of primarily studying young populations of college students is that we may be ignoring the possibility that what it means to be a man, or the role of masculinity in one's life, changes over time. The college years are a period of transition and examination of one's identity. Masculinity may be one facet of personality that emerges for examination during this time and it may emerge as problematic. Generalizing findings about masculinity from studies of undergraduates may not be appropriate to middle-aged or older men. Fuller (2001) found, for example, that Peruvian men all seemed to share the same general definition of masculinity but that the emphasis on particular qualities seemed to vary based on stage in life. Similarly, Gradman (1994) discussed the changes in men's masculine identity during the transition from work to retirement and found that many men successfully adjusted to life without their career. Thus, problematic aspects of masculinity that are salient for individuals in late adolescence or early adulthood may be less so for older individuals.

The underrepresentation of women is particularly problematic. Researchers in the field have long maintained that masculinity is distinct from biological sex (i.e., being male). Thus in theory, although masculinity norms may be particularly salient for men, they are also capable of affecting women (Zamarripa, Wampold, & Gregory, 2003). For example, masculinity norms regarding hard work and independence can have the same range of negative physical and emotional consequences for women in the workplace as for men. Some of these consequences include increased vulnerability to cardiovascular disease (Good, Sherrod, & Dillon, 2000) as well as problems with anxiety and depression (Courtenay, 2000a, 2000b). Studying the role of masculinity norms and ideologies in women's experience can shed a clear light on how particular norms (e.g., restrictive emotionality, competitiveness) operate without slipping into an essentialist view of gender (Bohan, 1997). However, when we study masculinity only in men we can easily fall prey to an implicit essentialism by failing to distinguish the two; thoroughly studying masculinity means understanding how it operates in the lives of both men and women.

We found little diversity in masculinity measures used across the sample of studies. Almost 70% of the studies that used masculinity measures used only one measure, and the Gender Role Conflict Scale (GRCS; O'Neil et al., 1986) was the one used in almost 65% of those studies. In fact, the GRCS was used 3.5 times more frequently than the next most common measure. Many of the measures widely used in the field of men and masculinity have established psychometric properties, which may explain the frequency with which they continued to be employed. Nonetheless, it is important to remember that no one measure captures everything of interest about the social construction or social learning of masculinities. Questionnaires themselves restrict the range of what we can understand about men's and women's experiences of gender and reliance on one or two of the same measures for sake of convenience runs the risk of intellectual atrophy in the field. Finally, there is the perennial problem in psychology of allowing our measures to dictate the questions we ask in our research.

Our review indicated that methodologies have not changed much over the last 10 years. Although there has been more research published in the field of men and masculinity, there has been little change in how much one methodology is employed relative to another. This was a disappointing finding because previous reviews over 10 years ago made a strong call for increasing the diversity of methods used in studying the psychology of men and masculinity (Betz & Fitzgerald, 1993; Good et al., 1994). For example, there has been a slight increase in the number of experimental studies over the past 3 years, but correlational research continues to be employed at a greater

frequency. The big picture does suggest that the increase in publication in our field is bringing with it an increase in nonquantitative methods. We hope that this trend will continue so that as research in our field gains momentum, the methodologies we employ will become increasingly diverse.

Conclusions, Limitations, and Future Directions

Although the field has been growing rapidly over the last 10 years, methodological critiques of research on men and masculinity have been lacking. We found in the current study that men and masculinity researchers in the US who published in widely read psychology journals were not taking advantage of a range of methodologies including sampling of diverse populations, use of experimental and observational methods, and multiple measurement. It is important to note that with the publication of the journal *Psychology of Men and Masculinity*, the number of empirical articles published on this topic has increased dramatically. As a result, our characterization of where the field of men and masculinity is headed in the US is heavily influenced by this journal. Experimental and qualitative research is rarely utilized, and there is an underrepresentation of ethnic minorities and women. Researchers typically used few measures of masculinity and the tendency is to select from a small group of comparatively popular measures. One obvious limitation in the current study is that we only coded articles in mainstream, explicitly psychological journals published in the US. We acknowledge that there is excellent work being conducted on men and masculinity issues by researchers both inside and outside of the US that would benefit from a thorough review. For example, Wetherell and Edley's (1999) discursive work should be given attention and explored for use in answering questions about the meaning and construction of masculinity. In addition, work is being done on gender identity formation and socialization in specific populations in Latin America (Escobar, 1998; Gutman, 1996; Ramírez, 1993; Valdés & Olavarria, 1997; see Vigoya, 2001 for a review).

The current study was also limited in that we examined only the methods and not the content of the studies we coded. Similarly, we did not code for different types of qualitative work. For example, some qualitative studies conducted interviews with participants while others explored the content of books or explored children's interactions with fathers. It is possible that within the broad categorization of qualitative work a variety of methods for answering research questions are being utilized that may have been missed in this review.

The current research is a starting point from which men and masculinity researchers in the US can begin to examine

their methods. This is by no means an in-depth, all-inclusive review. The field might benefit from such a review, however, and we encourage other researchers to look at specific examples of research in the field and thoroughly explore how some of the possible drawbacks noted in this review might play out in specific studies. What is perhaps most curious to note is that many of the recommendations that reviewers have made previously have not been taken up by men and masculinity researchers in the US to a large degree. These recommendations fall into roughly three categories: recommendations that we diversify our samples, recommendations that we employ more longitudinal/developmental methods, and recommendations that we continue to develop diverse measures. A prospective remedy to the first of those recommendations is to increase our efforts to understand the cultural issues that may impede research with minorities or community populations. Sensitivity to the practical issues (health and financial restraints) associated with conducting research with these populations is also essential (Areán & Gallagher-Thompson, 1996). Realistically, we should continue to integrate our concerns with those of other subfields (e.g., developmental psychology, clinical psychology) in order to attract researchers with the expertise to address some of these issues. Additionally, working to integrate our interests with the concerns of other subfields will help to bring the diversity in methods associated with those fields to bear on our research. These researchers would also likely bring diverse perspectives to instrumentation that could serve to increase the inclusion of measures that may be more relevant to older men and Men of Color. Hopefully our findings will serve as a starting point for researchers to examine the influence of their methods on their research questions and to introduce diversity of participants, methods, and measures into their research.

Acknowledgment The authors would like to thank Nathaniel Alapat for his help in coding the articles for this study.

References

- Addis, M. E., & Mahalik, J. R. (2003). Men, masculinity, and the contexts of help-seeking. *American Psychologist*, *58*, 5–14.
- Areán, P. A., & Gallagher-Thompson, D. (1996). Issues and recommendations for the recruitment and retention of older ethnic minority adults into clinical research. *Journal of Consulting and Clinical Psychology*, *64*, 875–880.
- Betz, N. E., & Fitzgerald, L. F. (1993). Individuality and diversity: Theory and research in counseling psychology. *Annual Review of Psychology*, *44*, 343–381.
- Bohan, J. S. (1997). Regarding gender: Essentialism, constructionism, and feminist psychology. In M. M. Gergen & S. N. Davis (Eds.), *Toward a new psychology of gender: A reader* (pp. 31–48). New York: Routledge.

- Brannon, R., & Juni, S. (1984). A scale for measuring attitudes toward masculinity. *JSAS Catalog of Selected Documents in Psychology*, 14, 6 (Ms. 2012).
- Burk, L. R., Burkhart, B. R., & Sikorski, J. F. (2004). Construction and preliminary validation of the Auburn differential masculinity inventory. *Psychology of Men and Masculinity*, 5, 4–17.
- Cohen, J. A. (1960). Coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20, 37–46.
- Connell, R. W. (1995). *Masculinities*. Berkeley, CA: University of California Press.
- Courtenay, W. H. (2000a). Constructions of masculinity and their influence on men's well-being: A theory of gender and health. *Social Science and Medicine*, 50, 1385–1401.
- Courtenay, W. H. (2000b). Engendering health: A social constructionist examination of men's health beliefs and behaviors. *Psychology of Men and Masculinity*, 1, 4–15.
- Danziger, K. (1994). *Constructing the subject: Historical origins of psychological research* (2nd ed.). New York: Cambridge University Press.
- Derlega, V. J., Catanzaro, D., & Lewis, R. J. (2001). Perceptions about tactile intimacy in same-sex and opposite sex pairs based on research participants' sexual orientation. *Psychology of Men and Masculinity*, 2, 124–132.
- Doss, B. D., & Hopkins, J. R. (1998). The multicultural masculinity ideology scale: Validation from three cultural perspectives. *Sex Roles*, 38, 719–741.
- Edley, N., & Wetherell, M. (1999). Imagined futures: Young men's talk about fatherhood and domestic life. *British Journal of Social Psychology*, 38, 181–194.
- Eisler, R. M., Franchina, J. J., Moore, T. M., Honeycutt, H. G., & Rhatigan, D. L. (2000). Masculine gender role stress and intimate abuse: Effects of gender relevance of conflict situations on men's attributions and affective responses. *Psychology of Men and Masculinity*, 1, 30–36.
- Escobar, L. A. (1998). Los hombres y sus historias: Reestructuración y masculinidad en México. *La Ventana*, 8, 122–173.
- Fleiss, J. L. (1981). *Statistical methods for rates and proportions*. New York: Wiley.
- Fleming, L. M., & Tobin, D. J. (2005). Popular child-rearing books: Where is daddy? *Men and Masculinity*, 6, 18–24.
- Fuller, N. (2001). The social constitution of gender identity among Peruvian men. *Men and Masculinities*, 3, 316–331.
- Good, G. E., Borst, T. S., & Wallace, D. L. (1994). Masculinity research: A review and critique. *Applied and Preventive Psychology*, 3, 3–14.
- Good, G. E., Sherrod, N. B., & Dillon, M. G. (2000). Masculine gender role stressors and men's health. In R. M. Eisler & M. Hersen (Eds.), *Handbook of gender culture and health* (pp. 63–81). Mahwah, NJ: Erlbaum.
- Gradman, T. J. (1994). Masculine identity from work to retirement. In E. H. Thompson, Jr. (Ed.), *Older men's lives* (pp. 104–121). Thousand Oaks, CA: Sage.
- Graham, S. (1992). "Most of the subjects were White and middle class": Trends in published research on African Americans in selected APA journals, 1970–1989. *American Psychologist*, 47, 629–639.
- Gutman, M. C. (1996). *The meanings of macho: Being a man in Mexico City*. Berkeley, CA: University of California Press.
- Hayes, J. A., & Mahalik, J. R. (2000). Gender role conflict and psychological distress in male counseling center clients. *Psychology of Men and Masculinity*, 1, 116–125.
- Levant, R. F., & Pollack, W. S. (1995). *A new psychology of men*. New York: Basic Books.
- Liu, W. M. (2005). The study of men and masculinity as an important multicultural competency consideration. *Journal of Clinical Psychology*, 61, 685–697.
- Mackey, W. C. (2001). Support for the existence of an independent man-to-child affiliative bond: Fatherhood as a biocultural invention. *Psychology of Men and Masculinity*, 2, 51–66.
- Mahalik, J. R., Locke, B. D., Ludlow, L. H., Diemer, M. A., Scott, R. P. J., Gottfried, M., et al. (2003). Development of the conformity to masculine norms inventory. *Psychology of Men and Masculinity*, 4, 3–25.
- Marecek, J. (1995). Gender, politics, and psychology's ways of knowing. *American Psychologist*, 50, 162–163.
- McCreary, D. R., & Sasse, D. K. (2000). An exploration of the drive for muscularity in adolescent boys and girls. *Journal of American College Health*, 48, 297–304.
- Morrison, T. G., Morrison, M. A., & Hopkins, C. (2003). Striving for bodily perfection? An exploration of the drive for muscularity in Canadian men. *Psychology of Men and Masculinity*, 4, 111–120.
- O'Neil, J. M., Helms, B., Gable, R., David, L., & Wrightsman, L. (1986). Gender role conflict scale: College men's fear of femininity. *Sex Roles*, 14, 335–350.
- Pleck, J. H. (1981). *The myth of masculinity*. Cambridge, MA: MIT Press.
- Ramírez, R. (1993). *Dime capitan: Reflexiones sobre la masculinidad* (What it means to be a man: Reflections on Puerto Rican masculinity). Rio Piedras, Puerto Rico: Ediciones Huracán.
- Smiler, A. (2004). Thirty years after the discovery of gender: Psychological concepts and measures of masculinity. *Sex Roles*, 50, 15–26.
- Sue, S. (2003). Science, ethnicity, and bias: Where have we gone wrong? In A. E. Kazdin (Ed.), *Methodological issues & strategies in clinical research* (pp. 173–188). Washington, DC: American Psychological Association.
- Thomas, S. P. (2003). Men's anger: A phenomenological exploration of its meaning in a middle class sample of American men. *Psychology of Men and Masculinity*, 4, 163–175.
- Thompson, E. H., & Pleck, J. H. (1995). Masculinity ideologies: A review of research instrumentation on men and masculinities. In R. F. Levant & W. S. Pollack (Eds.), *A new psychology of men* (pp. 129–163). New York: Basic Books.
- Valdés, T., & Olavarria, J. (Eds.) (1997). *Masculinidad/es poder y crisis*. Santiago, Chile: ISIS Internacional/FLASCO.
- Vigoya, M. V. (2001). Contemporary Latin American perspectives on masculinity. *Men and Masculinities*, 3, 237–260.
- Villemez, W. J., & Touhey, J. C. (1977). A measure of individual differences in sex stereotyping and sex discrimination: The Macho Scale. *Psychological Reports*, 41, 411–415.
- Weiss, M. R., & Barber, H. (1995). Socialization influences on collegiate male athletes: A tale of two decades. *Sex Roles*, 33, 129–140.
- Wetherell, M., & Edley, N. (1999). Negotiating hegemonic masculinity: Imaginary positions and psycho-discursive practices. *Feminism & Psychology*, 9, 335–356.
- Zamarripa, M. X., Wampold, B. E., & Gregory, E. (2003). Male gender role conflict, depression, and anxiety: Clarification and generalizability to women. *Journal of Counseling Psychology*, 50, 333–338.