Title

Measuring Men's Gender Norms and Gender Role Conflict/Stress in a High HIV-Prevalence South African Setting

Journal

AIDS and Behavior, 20(8)

ISSN

1090-7165

Authors

Gottert, Ann
Barrington, Clare
Pettifor, Audrey
et al.

Publication Date

2016-08-01

DOI

10.1007/s10461-016-1374-1

Peer reviewed

HHS Public Access

Author manuscript

AIDS Behav. Author manuscript; available in PMC 2020 June 02.

Published in final edited form as:

AIDS Behav. 2016 August ; 20(8): 1785–1795. doi:10.1007/s10461-016-1374-1.

Measuring men's gender norms and gender role conflict/stress in a high HIV-prevalence South African setting

Ann Gottert[a,b,*], **Clare Barrington**[a], **Audrey Pettifor**[b,c], **Heath Luz McNaughton-Reyes**[a], **Suzanne Maman**[a], **Catherine MacPhail**[c,d,e], **Kathleen Kahn**[c,f], **Amanda Selin**[g], **Rhian Twine**[c], **Sheri A. Lippman**[h]

[a]Department of Health Behavior Gillings School of Global Public Health University of North Carolina at Chapel Hill, Chapel Hill, NC

[b]Department of Epidemiology Gillings School of Global Public Health University of North Carolina at Chapel Hill, Chapel Hill, NC

[c]MRC/Wits Rural Public Health and Health Transitions Research Unit (Agincourt), School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa

[d]Wits Reproductive Health and HIV Institute (WRHI) University of the Witwatersrand, Johannesburg, South Africa

[e]School of Health University of New England Armidale NSW Australia

[f]Umeå Centre for Global Health Research, Division of Epidemiology and Global Health, Department of Public Health and Clinical Medicine, Umeå University, Umeå, Sweden

[g]Carolina Population Center University of North Carolina at Chapel Hill, Chapel Hill, NC

[h]Department of Medicine University of California at San Francisco Center for AIDS Prevention Studies (CAPS), San Francisco, CA

Abstract

Gender norms and gender role conflict/stress may influence HIV risk behaviors among men; however scales measuring these constructs need further development and evaluation in African settings. We conducted exploratory and confirmatory factor analyses to evaluate the Gender Equitable Men's Scale (GEMS) and the Gender Role Conflict/Stress (GRC/S) scale among 581 men in rural northeast South Africa. The final 17-item GEMS was unidimensional, with adequate model fit and reliability (alpha=0.79). Factor loadings were low (0.2–0.3) for items related to violence and sexual relationships. The final 24-item GRC/S scale was multidimensional with four factors: *Success, power, competition*; *Subordination to women*; *Restrictive emotionality*; and *Sexual prowess*. The scale had adequate model fit and good reliability (alpha=0.83). While GEMS is a good measure of inequitable gender norms, new or revised scale items may need to be explored in the South African context. Adding the GRC/S scale to capture men's strain related to gender roles could provide important insights into men's risk behaviors.

[*]**Corresponding author:** Ann Gottert, PhD, MPH, 1153 Belfair Way, Chapel Hill, NC 27517, agottert@live.unc.edu, Tel: (202) 904-0234, Fax: (919) 966-2921.

Keywords

Gender norms; Gender Equitable Men's Scale; Gender Role Conflict/Stress scale; HIV; masculinities; gender role; sexual behavior; violence; South Africa

Introduction

There has been a global call for research to assess beliefs, cultural norms, and experiences related to gender roles to inform programs to prevent HIV, gender-based violence, and other adverse health outcomes [1, 2]. In particular, research and programming focusing on the connections between men's beliefs and experiences related to gender roles, relationship power dynamics between men and women, and men's sexual risk and violence perpetration behaviors is gaining momentum [1–3]. Such efforts are particularly pertinent in communities with high HIV prevalence such as in Mpumalanga province in South Africa, where over one-fifth of adults are living with HIV and most risk behavior is enacted within intimate relationships between men and women [4, 5].

One challenge to such research and programming is operationalizing and quantitatively measuring psychosocial constructs related to gender roles. Such measures should be carefully derived from theory, validated in the specific study population or context, and, for new or adapted measures, built on past evidence–based measures to the extent possible [6]. In this study, we evaluate two scales measuring related constructs: gender norms and gender role conflict/stress (GRC/S). Gender norms are beliefs and expectations about appropriate roles and behavior for men and women, and are commonly conceptualized as ranging from equitable to inequitable [2]. GRC/S refers to men's experiences of and emotional responses to these expected roles, and can be defined as a psychological state in which restrictive definitions of masculinity limit individual wellbeing and overall human potential [7].

Among the many available measures of gender norms/ideology (for a review see Smiler & Epstein (2010)[8]), in the past decade, the Gender Equitable Men's Scale (GEMS) has become the most common measure used in HIV and violence prevention research and program evaluations in developing country settings [1, 9, 10]. However GEMS has rarely been subject to rigorous factor analysis and a better understanding of the scale's factor structure and performance is needed. Gender role conflict/gender role stress has been identified as a key predictor of men's psychological problems and risk behaviors in masculinities and health research [7, 11–14]. To date this construct has received little attention in HIV prevention research, particularly in the African setting. Therefore in the current study we evaluated the factor structure and performance of both the GEMS scale and a new scale we called the Gender Role Conflict/Stress scale, which we created for the South African context by combining elements of two other scales. Below, we describe the gender norms and gender role conflict/stress constructs and measures in more detail.

Pulerwitz and Barker (2008), working with young men in Brazil, developed GEMS as an alternative to available scales developed in the United States related to masculine ideology [9]. GEMS included normative statements about both men's and women's roles and behavior that reflect developing-country realities and relate to health outcomes of

programmatic interest to the developers, primarily sexual, reproductive, and violence behaviors [9]. Since its development, GEMS has been used in dozens of research studies and program evaluations worldwide, including in South Africa, and has consistently achieved good internal consistency reliability, with Cronbach's alphas ranging from the mid 0.70s to mid 0.80s, as well as being significantly associated with behaviors of interest among both men and women [1, 2, 10, 15–17]. However, to our knowledge there are only two factor analyses of GEMS described in the literature, one from the original scale development paper from Brazil, and the other from a study in Ghana and Tanzania [9, 15]. An in-depth understanding of the measure is particularly important in the context of the parent study for our analyses in South Africa, in which GEMS is the main trial endpoint/outcome [18]. This is a community randomized controlled trial to assess the effectiveness of a community mobilization intervention to change gender norms and reduce HIV risk behaviors, particularly among men [18].

Alongside the gender norms construct, the study team was interested in deepening our understanding of men's psychological strain related to expected gender roles, and how this strain may influence men's HIV risk behaviors. We began to explore concepts referred to as masculine gender role strain, men's gender role conflict, and masculine gender role stress.

Beginning in the late 1970s and early 1980s, a number of masculinities theorists, researchers and psychologists began to express concern about the harm that restrictive masculine expectations can cause to men's psychological wellbeing, health and health related-behavior. Pleck developed the Gender Role Strain Paradigm as a set of theoretical propositions describing how cultural standards of masculinity, as implemented in gender role socialization, have potentially negative effects on men [19]. He theorized that men's experience of trying and failing to live up to masculine norms often leads to anxiety and depression and causes many men to try to overcompensate in ways that are harmful to themselves and others [19, 20]. Such notions resonate with recent qualitative and ethnographic research on masculinities in the South African setting, which document men's limited ability to 'prove themselves' as men in a context of poor education and high unemployment, leading to what many call a "crisis of identity" that can prevent men from implementing more gender equitable practices in their own relationships, even when they are broadly supportive of gender equality [21–25].

Drawing on Pleck's Gender Role Strain Paradigm, O'Neil developed the Gender Role Conflict Scale in the mid-1980s to capture "concrete outcomes of gender role strain that can be understood and measured" [11, 26]. He defines the construct of gender role *conflict* as a psychological state in which restrictive definitions of masculinity limit individual wellbeing and overall human potential [7]. The scale includes a series of first-person statements about the importance of or stress around fulfilling traditional masculine roles, with items worded to tap into psychological strain (e.g., "I *worry* about failing and how it affects my doing well as a man", "Making more money than a woman is a measure of *my value and personal worth*"). This scale has now been used in over 300 studies, primarily with U.S. college men, but also among men in Canada, Australia, and countries in Europe and Asia [7, 11]. The scale has consistently demonstrated good psychometric properties (validity and reliability) [11, 26] and has been found to be associated with men's anxiety, depression, intimate

partner violence (IPV) perpetration, substance use, and help-seeking attitudes, among others [7, 11–14].

Another scale that captures men's psychological strain/stress related to expected gender roles is the Masculine Gender Role Stress Scale, developed by Eisler and Skidmore (1987) [27], which asks men the extent to which masculine-gender-relevant situations (e.g. "being outperformed at work by a woman") would be stressful to them. This scale draws explicitly on Lazarus' cognitive appraisal model of stress as well as Pleck's Gender Role Strain Paradigm [27, 28]. The scale has demonstrated good psychometric properties, although less consistently than the Gender Role Conflict Scale [29], and while used in fewer studies than that scale, has been found to be associated with anger, anxiety, risk for intimate partner violence perpetration, cardiovascular reactivity in response to stress, and other adverse health outcomes [12–14, 27, 30, 31].

Because the Gender Role Conflict and Masculine Gender Role Stress Scales have not previously been applied in the African setting and also remain largely unexplored in global HIV prevention research to date, we developed a scale for the South African context called the Gender Role Conflict/Stress (GRC/S) scale. To do so, we combined elements of the Gender Role Conflict Scale and Masculine Gender Role Stress Scale [11, 27], instead of using either one in its entirety, in order to use the set of sub-dimensions and items most applicable to the South African setting and most relevant to the HIV-related outcomes of interest in this study.

We evaluated the validity and reliability of the GEMS and GRC/S scales in the South African context in four steps. First, using exploratory factor analysis, we evaluated the factor structure of the GEMS and GRC/S scales. Second, we tested the structural validity of both measures using confirmatory factor analysis. Third, we examined the reliability of the factors. Fourth, we assessed convergent validity, that is, whether men's scores on each scale were correlated with other theoretically related variables.

Methods

Sample and Study Setting

Data for this study come from the baseline survey of the study *Community Mobilization for the Prevention of HIV in Young South African Women*, a two-year cluster randomized controlled trial of an intervention to change inequitable gender norms, particularly among men [18]. A population-based household survey was conducted from March to June 2012 among men and women ages 18–35 in 22 villages in the rural Agincourt area of the Bushbuckridge sub-district in Mpumalanga Province, located near South Africa's Eastern border with Mozambique. Agincourt, like many rural areas of South Africa, is densely populated and characterized by few employment opportunities and high levels of circular or temporary migration for labor. The trial was nested within the Agincourt Health and Socio-demographic Surveillance System (Agincourt HDSS), which runs an annual census to update residence status of all members and record any vital events [32].

Recruitment and Data Collection

Participants were identified and selected through the Agincourt HDSS database, from among all male and female residents aged 18–35 living in 22 study villages enumerated in the 2011 census. For sample selection each household with 18- to 35-year-old residents was designated as either male or female, and individuals of that gender in the household were randomly assigned an order. On entering a home, the individual randomly prelisted was screened for the following more detailed eligibility criteria: person lives in the home, is 18–35 years old per confirmed date of birth, is the gender assigned to the home, and has lived in the study area for the past 12 months. If the prelisted individual did not meet these eligibility criteria, the second was screened, and so on. Only one individual was interviewed per household.

Surveys were conducted by trained male or female interviewers in the participant's household and took 1 to 2 hours to complete. Interviews were conducted in the local language Shangaan or in English, depending on the participant's preference, using computer-assisted personal interviewing (CAPI), in which the interviewer reads each question to the respondent, then enters the answer into an electronic form on a laptop computer. Questions covered socio-demographic background, gender role attitudes and experiences, sexual risk behaviors, and a range of other questions related to other personal beliefs, practices, and perceptions of the social context. The survey, including GEMS and GRC/S scale items, was translated from English into Shangaan, back-translated, and revised as necessary.

We limited our analyses to men. Among 620 eligible men, 581 men were enrolled into the study (94%); 35 refused to participate (6%) and the remaining 2 (<1%) did not enroll for other reasons. The final sample was weighted using scaled weights to account for differential sampling probabilities at the household and individual level and to represent the distribution of men aged 18–35 years in Agincourt based on the 2011 Agincourt HDSS. The study was approved by the Institutional Review Boards at the University of North Carolina-Chapel Hill and University of California-San Francisco, the Human Research Ethics Committee at the University of the Witwatersrand in South Africa, and the Mpumalanga Department of Health and Social Development Research Committee.

Scale Adaptation Process

Gender norms were measured using the GEM scale [9], which includes a series of third-person belief statements. We based our scale on an Ethiopian adaptation of GEMS that had achieved high internal consistency and included 24 items representing inequitable gender norms, 18 of which came from the original GEMS [33]. We changed the wording of a few items in consultation with the local research team to increase appropriateness for the local social context. We also added six "reverse-coded" items from the original GEMS in an effort to include gender-equitable alongside gender-inequitable norms, for a total of 30 items. The content of items addressed both men's and women's roles and behavior related to four domains: sexual relationships, violence, reproductive health and disease prevention, and domestic chores and daily life, in line with the original theoretical work [9]. Response categories included "Do not agree at all," "Somewhat agree," and "Agree a lot," consistent

with the three response categories commonly used for GEMS in other studies [1, 9, 15]. We coded all GEMS items such that a higher score represented more inequitable norms.

Gender role conflict/stress was measured using the Gender Role Conflict/Stress (GRC/S) scale, which we developed in this study for the South African context by combining sub-dimensions and items from the Gender Role Conflict Scale [11, 26] and the Masculine Gender Role Stress scale [27]. As described above, these two scales share a common theoretical foundation in the Gender Role Strain Paradigm, and studies have shown that the two full scales are correlated at around 0.5 [12, 34]. We chose the Gender Role Conflict Scale format as the basis of our scale and included two of the four original domains (*Success, power, competition* and *Restrictive emotionality*). Based on input from local study team members, we discarded two domains deemed less relevant to the local context (*Restrictive affectionate behavior between men* and *Conflicts between work and family relations*). *Restrictive affectionate behavior between men* was specifically eliminated due to recent studies that suggest that items that include references to homosexuality may tap into a different latent construct than those related to masculine ideology [15, 35]. *Conflicts between work and family relations* was deemed less relevant because of the local context of high unemployment and low marriage/cohabiting rates. We also included two domains from the Masculine Gender Role Stress Scale (*Subordination to women* and *Physical inadequacy*) [27]. Another sub-dimension from that scale, Performance Failure (Work and Sex), provided additional content for a number of items in our *Success, power, and competition* sub-dimension.

We reviewed a draft of the resulting GRC/S scale with members of the local study team with years of experience in working with men, who confirmed that overall the scale had salient content (i.e., face) validity and that the domains were appropriate. The team recommended dropping seven items, adding one item, and editing the wording of nine other items. This process resulted in a final 28-item scale. Response categories included “Do not agree at all,” “Somewhat agree,” and “Agree a lot,” unlike the original Gender Role Conflict Scale, which includes five or six response categories ranging from “strongly disagree” to “strongly agree” [11]. The study team chose three categories based on experience at the site that respondents prefer fewer response categories and because they also make the most sense in the local language, but acknowledge that using only three response categories may introduce truncated variance and/or limit the performance of items in the scale. All GRC/S questions were worded and coded such that higher scores represent more strain.

Analysis

We carried out a split-sample exploratory and confirmatory factor analysis (EFA/CFA) using Classical Test Theory [6]. We randomly split the sample of men in half, conducted EFA on the first half to better understand dimensionality and identify the most plausible factor structure and CFA on the second half to test the structural validity of the structure selected based on the EFA [36, 37]. All factor analyses were performed using Mplus software version 7.11 [38]. We used the MLR estimator in Mplus, which uses maximum likelihood with a robust standard error that accounts for the cluster sampling design by village. Two participants had all GRC/S scale items missing and were dropped from that analysis. Other

missing data was minimal (<1% per item) and was handled using full information maximum likelihood in Mplus [38].

Factor structure.—To determine the factor structure through the EFA, we followed the approach and criteria recommended by DeVellis [6], adding other techniques to provide additional insight when needed. Two techniques were used to determine the number of factors for extraction. First, we used a scree plot such that the primary bend in the plot indicated the number of factors to extract. Second, we used Kabacoff's (2003) parallel analysis procedure, which minimizes bias due to random variance in the data by running simulations on multiple randomly sampled datasets to determine the maximum number of factors to retain [6, 39]. Based on the scree plot and parallel analyses we specified a plausible range of number of factors and, due to expected high correlations between the factors, used oblique (geomin) rotation methods to produce interpretable factor loadings. To determine the factor structure to be retained for testing using CFA we used the following criteria: (1) interpretability (i.e., the extent to which items within each factor seemed to be tapping into a common theme), (2) significance (i.e., factor loadings that were significant at $p<0.05$ across most factor solutions were retained), and (3) adequacy of reliability of each factor (in the same manner as described below). Finally, we assigned a name to each factor based on its item composition.

Structural validity.—We assessed whether the factor structure of the latent variable is valid and measures what the latent variable is intended to measure by applying the factor structure suggested by the EFA to the CFA following procedures recommended by Bollen [36]. We began by inspecting loadings of each item on its factor and retained items that had significant loadings ($p<0.05$). We then assessed the adequacy of model fit based on commonly recommended cut-off criteria. Because the chi square statistic tends to suggest poor model fit when the sample size is relatively large [40], we used the indicator of good fit suggested by Segars and Grover of chi square being within three times the degrees of freedom [41]. We also assessed the Root Mean Square Error of Approximation (RMSEA, with a cut-off value <0.06 indicating good fit), the Comparative Fit Index (CFI) and the Tucker-Lewis Index (TLI) (both with cut-offs >0.95), and the Standardized Root Mean Squared Residual (SRMR, cut-off < 0.08) [42]. We then reviewed modification indices, added plausible correlated errors (e.g., when two items had very similar wording or sentence structure), and re-fit the model. In addition, we looked at the degree of correlation between factors and assessed whether a higher-order factor was present, following Brown's (2012) recommended procedures [43]. We assessed the adequacy of final model fit based on the same cutoff criteria described above.

Reliability.—To assess reliability, we calculated Cronbach's coefficient alpha, a widely used measure of internal consistency reliability. We also calculated Raykov's ρ (rho), a measure of reliability similar to alpha but more suitable for categorical response categories, less influenced by number of scale items, and accounting for dimensionality in reliability estimates for multidimensional scales [44]. We deemed an alpha and rho above 0.7 to indicate adequate reliability [6].

Convergent validity.—Finally, we assessed whether GEMS and GRC/S were correlated with other theoretically related variables. We first assessed the extent to which GEMS and GRC/S (including sub-dimensions) were correlated with each other, expecting a moderate degree of correlation. We also assessed whether each was correlated with three other items included in the survey that assessed men's broader support for gender equality.

Results

Among the 581 men in the final sample, the mean age was 22.4 (range 18–35) (Table I). Most participants (85%) had never been married. Sixty percent had some high school education, 26% had completed high school, and 2% had attended University or Technikon (tertiary level practical/trade qualifications). Thirty-one percent of men had earned any income in the past 3 months.

GEMS

For the EFA for GEMS (n=291), evaluation of the scree plot suggested 1 factor, however the parallel analysis suggested 4 factors. From an initial inspection of factor loadings it became apparent that the 6 items worded to represent "equitable" gender norms all had negative loadings after undergoing reverse-coding, compared to positive values for other items in the unidimensional and most other factor solutions, suggesting they should be dropped [6]. In addition, one item ("It disgusts me when I see a man acting like a woman") had non-significant loadings across multiple factor solutions, suggesting it is not measuring the same construct as the other items. Therefore, we dropped these 7 items from the scale and re-ran the EFA. After assessing interpretability among the various factor structures, we decided that a four-factor solution resulted in the most interpretable item groupings. These fell along the lines of content areas often used to group GEMS scale items in surveys and publications: *Sexual relationships, Violence, Reproductive health and disease prevention*, and *Domestic chores and daily life*. However, we found reliability of the factors to be unacceptably low for all factors except *Domestic chores and daily life,* ranging from 0.43–0.54 for alpha and 0.41–0.55 for rho (0.78/0.66 for *Domestic chores and daily life*). Because of these low reliabilities we chose to test a unidimensional model in the CFA.

In the CFA for GEMS (n=290) we retained all but the 7 items noted previously when testing the unidimensional model. GEMS CFA results are presented in Table II. Many factor loadings were low (0.2–0.4), especially for items related to violence and sexual relationships; 6 additional items failed to load significantly on the latent factor and were therefore dropped. All factor loadings for final items were significant at $p<0.05$, with $p=0.000$ for all but three items. After reviewing modification indices for the uncorrelated model, we added nine plausible correlated errors. Model fit statistics for the uncorrelated and correlated (final) models are presented in Table IV. Overall, the final correlated 17-item model had adequate fit based on most cutoff criteria. Although as anticipated the chi-square value was highly significant, its value was within three times the degrees of freedom [41]. In addition, it would be preferable for CFI and TLI values to be above 0.95 rather than 0.90. SRMR values indicted good fit. Reliability of the final scale was good, at 0.79 for alpha and 0.71 for rho.

GRC/S Scale

The EFA for the GRC/S scale (n=289) suggested four factors, similar to those originally hypothesized. Evaluation of the scree plot was ambiguous because there was not a clear bend in the plot. However, the parallel analysis suggested four factors. From an initial inspection of factor loadings, we identified one item ("Making money is part of my idea of being a successful man") that was not performing well, likely because 91% of men had "agreed a lot" with the statement, limiting variation. In addition, similar to GEMS, one item related to homosexuality consistently loaded poorly ("Affection with other men makes men tense"). Therefore, we dropped these 2 items from the scale and re-ran the EFA. After assessing interpretability among the various factor structures, we decided that a four-factor solution had the most interpretable item groupings. The factors were labeled as: *Success, power, competition*; *Subordination to women*; *Restrictive emotionality*; and *Sexual prowess*. These groupings were similar to the groupings we had expected based on previous analyses of the two scales used to create this measure, with two exceptions: the "physical inadequacy" items—instead of forming a separate factor, loaded on *Success, power, competition*—and items related to sexual performance/prowess, which we believed would tap into *Success, power, competition,* instead formed a separate factor in themselves. Reliabilities assessed at this stage were near or above 0.7 for each factor, therefore we proceeded to test these factors in the CFA.

In the CFA for the GRC/S scale (n=290) we tested the four-factor model found in the EFA, retaining all but the 2 items noted previously. These CFA results are presented in Table III. Two items were dropped because they had non-significant factor loadings. All factor loadings for final items were significant at $p<0.05$, with $p=0.000$ for all but two items. After reviewing modification indices for the uncorrelated model, we added 12 correlated errors. The correlated model had adequate fit based on most cut-off criteria (Table IV), with the same findings in relation to cut-off criteria as described for GEMS above.

The four GRC/S scale factors were moderately to highly correlated (from 0.38 to 0.71), indicating that they are sufficiently distinct from each other but that these factors may be part of a higher-order, multi-dimensional construct [43]. Factor loadings of the four first-order factors on the second-order factor ranged from 0.56 to 0.87 (Table III). Incorporating this higher-order latent variable met criteria for plausibility in that it is theoretically justifiable, the factors are correlated, and the higher-order factor model had adequate fit, as shown in Table IV [36, 43, 45].

Reliability of the final multidimensional GRC/S scale was 0.83 for alpha and 0.83 for rho. For each factor, reliabilities were as follows (alpha/rho): 0.80/0.73 for *Success, power, competition*; 0.65/0.69 for *Subordination to women*; 0.65/0.72 for *Restrictive emotionality*; and 0.68/0.73 for *Sexual prowess.*

Finally, to assess convergent validity of the GEMS and GRC/S scales we examined their association with variables they should theoretically be related to. First, we found that GEMS and the GRC/S scale are themselves moderately correlated at 0.48 ($p<0.0001$). GEMS is also correlated with each GRC/S scale sub-dimension ($r=0.15–0.43$; all $p<0.001$). GEMS and the GRC/S scale are also correlated with three items included in our survey that assessed

broader beliefs about gender equality, at r=0.14–0.40 (all p<0.001): “Rights for women mean that men lose out;” “Gender equality, meaning that men and women are equal, has come far enough already;” and “It is a good thing that women have more rights than ever before” (reverse-coded).

Discussion

We sought to evaluate the validity and reliability of measures of gender norms and gender role conflict/stress in the rural South African context, where cultural adherence to gender norms and the experience of strain associated with these norms play a role in the HIV and violence epidemics. Though GEMS is commonly used in program evaluation, it has rarely been subject to a rigorous factor analysis. This is the first critical look at the scale’s performance in South Africa and, to our knowledge, is its first application in rural Mpumalanga province and among the Shangaan ethnic group that predominates in that province and adjacent Mozambique. We developed the GRC/S scale to tap into gender role conflict and stress in the South African context, by combining two scales validated in Western contexts. This is the first attempt to explore the psychometric properties of a scale to measure this construct in Africa, filling a gap in gender-based research in African populations, specifically in sub-Saharan Africa, with a widespread, generalized HIV epidemic [46].

GEMS Factor Analysis Findings

The unidimensional GEMS had adequate fit, although only three of five fit indicators met cut-off criteria. The scale also demonstrated good reliability. The two published factor analyses of GEMS, of which we included many of the same or similar items, also suggested a unidimensional scale for items related to inequitable gender norms, although dimensionality was not explored in detail [9, 15]. GEMS items cover a wide range of domains, from violence to sex to appropriate household roles; therefore it was not surprising to us that the EFA suggested multiple factors. However, the four groupings can also be conceptualized as content areas in which gender norms manifest rather than having clear theoretical distinctions, which may explain why factor reliabilities were unacceptably low.

GEMS factor loadings were relatively low for items related to violence, and many of these items were eventually dropped. It is possible that condoning violence against women may not stem from the same psychological orientation as endorsing gendered roles in sexual relationships, household decision-making, or division of labor. It could also be that items endorsing violence against women elicited more socially desirable responses than other GEMS items, perhaps due to predominantly female interview staff as well as recent national discourse, policy and media coverage about this issue in South Africa, including support for gender equality under the Constitution [47]. Although most of the violence-related items we dropped appear to have performed adequately in other studies, we believe that there is a need for further theoretical work to explore the relationships of the content areas to the larger construct of gender norms and exploring new scale items related to violence. Improved measurement of beliefs about violence against women in the South African context is particularly salient given the severity and extent of gender-based violence there [48, 49].

Factor loadings for items related to sexual relationships also tended to be low (0.2–0.5), which could suggest a need to revise and update these items, at least for the South African context. In particular, given the centrality of GEMS to research on HIV/sexually transmitted infections (STI) transmission, there may be a need for more items related to men adhering to norms regarding multiple and concurrent partnerships, as well as items related to sex work and women's ability to refuse sex.

GRC/S Scale Factor Analysis Findings

In the factor analysis for the GRC/S scale, we found that a multidimensional four-factor model had good structural validity and adequate model fit, although like GEMS only three of five fit indicators met cut-off criteria. The multidimensional scale had good reliability with an alpha and rho of 0.83. Three of the four sub-dimensions also had adequate reliability (i.e., ≥0.70) based on Raykov's rho, which as mentioned is less sensitive than alpha to the relatively low number of items [44]. To improve model fit, some adjustments to scale items should be made for future use. A few of the *Success, power, competition* items had limited variability and should be revised to elicit a broader range of responses. For example, revising the dropped item "Making money is part of my idea of being a successful man" to read "A man is only successful if he makes money" could elicit more varied responses and help separate men experiencing extreme conflict/stress from those with less. Fine-tuning the scale, for example by revising items, developing new items and conducting cognitive interviewing, should further improve validity and reliability.

It is interesting to note that *Sexual prowess* emerged as a separate factor among this sample of South African men, a departure from U.S. scales for which sexual performance is integrated into overall conceptualizations of success. Perhaps in response to limited work opportunities, South African men have defined sexual prowess as a separate realm in which to seek achievement as men. Luyt (2005) similarly found that "sexuality" emerged as a separate factor when assessing the factor structure of the MANI II scale measuring masculine ideology in South Africa [35]. Unfortunately, there were no items in our scale related to stress around having and showing off multiple sexual partners as an indicator of sexual prowess; we recommend that future versions of this scale explore such items.

Because due to time restrictions we did not engage in formative research or cognitive interviewing before survey implementation, the new GRC/S scale may not cover the full range of men's experiences related to gender role strain. We believe that additional work to describe the construct in the African context is critical. This should include qualitative research and an exploration of the various dimensions in related U.S. research that were not explored in our data. Two original Gender Role Conflict Scale dimensions were not explored. Specifically, the dimension *Conflicts between work and family relations* may manifest differently in South Africa, where high unemployment and low marriage/cohabiting rates are more common than the United States. However, this tension between work and family still merits exploration in Mpumalanga, particularly among men who migrate to find work in cities or the mines. In addition, although as mentioned, recent research suggests that measures relating to attitudes towards homosexuality tend to load on a separate latent factor from other aspects of masculinity [15, 35], a sub-dimension like

Restrictive affectionate behavior between men may warrant further exploration as it may be an important source of strain that in turn could affect risk behaviors. Additional research is also needed to evaluate how well the GRC/S scale performs in other settings, including confirmation of the factor structure and scale reliability.

These findings should be interpreted with several study limitations in mind. First, as mentioned previously, we did not conduct formative research to develop the GRC/S scale as is ideal when creating a scale or adapting it to a new context. Second, social desirability bias could be present; men could have responded to scale items in ways they thought were most socially acceptable rather than representing their true opinions [50]. It would be advantageous for future surveys that include the GEMS and/or GRC/S scales to include a measure of social desirability. Future measurement efforts should also explore whether and how interviewer gender may bias participant responses. Third, we did not assess test-retest reliability to establish consistency of scale responses and scores over time; this should be a priority for future research. Fourth, our sample only included non-migrating men in a rural location with high out-migration for work (about 30% of men in the study area move for work in any given year [32]). Further, most men were on the younger end of the 18–35 age range, which limits the generalizability of our findings to other age groups.

Conclusions

Having valid and reliable measures of theoretical constructs related to gender roles is essential to research on men's sexual and violence behaviors and evaluation of programs seeking to modify behaviors that place men and their partners at risk. Our findings add to growing literature on the measurement of gender norms. We found that GEMS is a good measure of inequitable gender norms, but that new or revised scale items may need to be explored in the South African context. Complementing the GEM scale's more cognitive appraisal of gender roles, the GRC/S scale is a promising multidimensional scale that captures men's experience of and emotional response to expected roles for them as men. The GEMS and GRC/S scale were moderately but not highly correlated with each other, indicating that they are measuring related but distinct constructs and can be used together. We believe that future research using both of these scales will further illuminate men's risk behaviors and the effectiveness of gender transformative interventions gaining favor in communities worldwide.

Acknowledgements

We would like to thank the study participants for sharing their time and perspectives with us. We would also like to thank the staff of the MRC/Wits Rural Public Health and Health Transitions Research Unit (Agincourt) for their support of the parent study, including data collection. Tamu Daniel, Sheree Schwartz and Rushina Cholera also assisted with data collection. Dr. Cathy Zimmer at the UNC ODUM Institute for Research in Social Science provided valuable advice on the data analysis. This study was supported by a Ruth L. Kirschstein National Research Service Institutional STD/HIV Pre-Doctoral Training Award (National Institutes of Health DHHS/NIH/NIAID: 5 T32 AI 07001–35) through the University of North Carolina at Chapel Hill Institute for Global Health and Infectious Diseases. The National Institute of Mental Health provided funding for the parent study for this research (1RO1MH087118, A. Pettifor, PI; 1R21MH090887–01, S. Lippman, PI).

References

1. Barker G, Contreras JM, Heilman B, Singh AK, Verma RK, Nascimento M. Evolving Men: Initial Results from the International Men and Gender Equality Survey (IMAGES). Washington, D.C.: International Center for Research on Women (ICRW) and Instituto Promundo, 2011.
2. World Health Organization. Engaging men and boys in changing gender-based inequity in health: evidence from programme interventions. World Health Organization, 2007 9241595493.
3. Rivers K, Aggleton P. Men and the HIV Epidemic. New York: United Nations Development Programme, 1999.
4. Gómez-Olivé FX, Angotti N, Houle B, Klipstein-Grobusch K, Kabudula C, Menken J, et al. Prevalence of HIV among those 15 and older in rural South Africa. AIDS care. 2013;25(9):1122–8. [PubMed: 23311396]
5. Shisana O, Rehle T, Simbayi L, Zuma K, Jooste S, Zungu N, et al. South African national HIV prevalence, incidence and behaviour survey, 2012. Cape Town: HSRC Press, 2014.
6. DeVellis RF. Scale development: Theory and applications: Sage Publications, Inc; 2011.
7. O'Neil JM. Men's gender role conflict: Psychological costs, consequences, and an agenda for change: American Psychological Association; 2015.
8. Smiler AP, Epstein M. Measuring gender: Options and issues Handbook of gender research in psychology: Springer; 2010 p. 133–57.
9. Pulerwitz J, Barker G. Measuring attitudes toward gender norms among young men in Brazil. Men and Masculinities. 2008;10(3):322.
10. Pulerwitz J Promoting more gender-equitable norms and behaviors among young men as an HIV/AIDS prevention strategy. Horizons Program and Instituto Promundo: Population Council; 2006.
11. O'Neil JM. Summarizing 25 Years of Research on Men's Gender Role Conflict Using the Gender Role Conflict Scale. The Counseling Psychologist. 2008;36(3):358–445.
12. Jakupcak M, Lisak D, Roemer L. The role of masculine ideology and masculine gender role stress in men's perpetration of relationship violence. Psychology of Men & Masculinity. 2002;3(2):97.
13. Gallagher KE, Parrott DJ. What Accounts for Men's Hostile Attitudes Toward Women? The Influence of Hegemonic Male Role Norms and Masculine Gender Role Stress. Violence Against Women. 2011;17(5):568. [PubMed: 21531691]
14. Eisler RM, Franchina JJ, Moore TM, Honeycutt HG, Rhatigan DL. Masculine gender role stress and intimate abuse: Effects of gender relevance of conflict situations on men's attributions and affective responses. Psychology of Men & Masculinity. 2000;1(1):30.
15. Shattuck D, Burke H, Ramirez C, Succop S, Costenbader B, Attafuah JD, et al. Using the Inequitable Gender Norms scale and associated HIV risk behaviors among men at high risk for HIV in Ghana and Tanzania. Men and Masculinities. 2013;16(5):540–59.
16. Dworkin SL, Treves-Kagan S, Lippman SA. Gender-Transformative Interventions to Reduce HIV Risks and Violence with Heterosexually-Active Men: A Review of the Global Evidence. AIDS and Behavior. 2013;17(9):2845–63. [PubMed: 23934267]
17. Jewkes R, Sikweyiya Y, Morrell R, Dunkle K. Gender inequitable masculinity and sexual entitlement in rape perpetration South Africa: findings of a cross-sectional study. PLoS One. 2011;6(12):e29590.
18. Pettifor A, Lippman SA, Selin AM, Peacock D, Gottert A, Maman S, et al. A cluster randomized-controlled trial of a community mobilization intervention to change gender norms and reduce HIV risk in rural South Africa: study design and intervention. BMC public health. 2015;15(1):752. [PubMed: 26245910]
19. Pleck J The gender role strain paradigm: An update In: Levant R, Pollack W, editors. A New Psychology of Men. New York, NY: BasicBooks; 1995 p. 11–32.
20. Brooks G, Silverstein L. Understanding the dark side of masculinity: An interactive systems model In: Levant R, Pollack W, editors. A New Psychology of Men. New York, NY: BasicBooks; 1995 p. 280–333.
21. Hunter M Masculinities, multiple-sexual-partners, and AIDS: the making and unmaking of Isoka in KwaZulu-Natal. Transformation: critical perspectives on Southern Africa. 2004;54(1):123–53.

22. Dworkin SL, Colvin C, Hatcher A, Peacock D. Men's Perceptions of Women's Rights and Changing Gender Relations in South Africa Lessons for Working With Men and Boys in HIV and Antiviolence Programs. Gender & Society. 2012;26(1):97–120.
23. Morrell R The times of change: Men and masculinity in South Africa In: Morrell R, editor. Changing men in southern Africa. Pietermaritzburg, South Africa: University of Natal Press; 2001 p. 3–37.
24. Shefer T, Crawford M, Strebel A, Simbayi LC, Dwadwa-Henda N, Cloete A, et al. Gender, power and resistance to change among two communities in the Western Cape, South Africa. Feminism & Psychology. 2008;18(2):157–82.
25. Sideris T "You have to change and you don't know how!": Contesting what it means to be a man in a rural area of South Africa. African Studies. 2004;63(1):29–49.
26. O'Neil JM, Helms BJ, Gable RK, David L, Wrightsman LS. Gender-role conflict scale: College men's fear of femininity. Sex Roles. 1986;14(5):335–50.
27. Eisler RM, Skidmore JR. Masculine gender role stress. Behavior Modification. 1987;11(2):123–36. [PubMed: 3508379]
28. Lazarus RS, Folkman S. Stress, Appraisal, and Coping. New York, NY: Springer Publishing Company, Inc; 1984.
29. Thompson EH, Pleck JH, Ferrera DL. Men and masculinities: Scales for masculinity ideology and masculinity-related constructs. Sex Roles. 1992;27(11):573–607. [PubMed: 12322226]
30. Eisler RM. The relationship between masculine gender role stress and men's health risk: The validation of a construct In: Levant R, Pollack W, editors. A New Psychology of Men. New York, NY: BasicBooks; 1995 p. 207–25.
31. Eisler RM, Skidmore JR, Ward CH. Masculine gender-role stress: Predictor of anger, anxiety, and health-risk behaviors. Journal of Personality Assessment. 1988;52(1):133–41. [PubMed: 3361409]
32. Kahn K, Collinson MA, Gómez-Olivé FX, Mokoena O, Twine R, Mee P, et al. Profile: Agincourt health and socio-demographic surveillance system. International journal of epidemiology. 2012;41(4):988–1001. [PubMed: 22933647]
33. Pulerwitz J, Hughes L, Mehta M, Kidanu A, Verani F, Tewolde S. Changing Gender Norms and Reducing Intimate Partner Violence: Results From a Quasi-Experimental Intervention Study With Young Men in Ethiopia. American Journal of Public Health. 2014(0):e1–e6.
34. Fischer AR, Good GE. Men and psychotherapy: An investigation of alexithymia, intimacy, and masculine gender roles. Psychotherapy: Theory, Research, Practice, Training. 1997;34(2):160.
35. Luyt R The Male Attitude Norms Inventory-II A Measure of Masculinity Ideology in South Africa. Men and Masculinities. 2005;8(2):208–29.
36. Bollen KA. Structural Equations with Latent Variables. New York, NY: Wiley; 1989.
37. Fabrigar LR, Wegener DT, MacCallum RC, Strahan EJ. Evaluating the use of exploratory factor analysis in psychological research. Psychological methods. 1999;4(3):272.
38. Muthén L, Muthén B. Mplus. Statistical analysis with latent variables. 2007;3.
39. Kabacoff R, editor. Determining the dimensionality of data: a SAS macro for parallel analysis. Proceedings of the 28th Annual Meeting of SAS Users Group International Seattle, WA; 2003.
40. Bentler PM, Bonett DG. Significance tests and goodness of fit in the analysis of covariance structures. Psychological bulletin. 1980;88(3):588.
41. Segars AH, Grover V. Re-examining perceived ease of use and usefulness: A confirmatory factor analysis. MIS quarterly. 1993;17(4):517–25.
42. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling: A Multidisciplinary Journal. 1999;6(1):1–55.
43. Brown TA. Confirmatory factor analysis for applied research: Guilford Press; 2012.
44. Raykov T, Marcoulides GA. Introduction to psychometric theory: Taylor & Francis; 2010.
45. Johnson RE, Rosen CC, Chang C-HD, Djurdjevic E, Taing MU. Recommendations for improving the construct clarity of higher-order multidimensional constructs. Human Resource Management Review. 2012;22(2):62–72.
46. UNAIDS. UNAIDS Report on the Global AIDS Epidemic 2013. Geneva: 2013.

47. Constitutional Assembly, Constitution of the Republic of South Africa. Cape Town, (1996).
48. Gass JD, Stein DJ, Williams DR, Seedat S. Intimate partner violence, health behaviours, and chronic physical illness among South African women. SAMJ: South African Medical Journal. 2010;100(9):582–5. [PubMed: 20822646]
49. Jewkes RK, Dunkle K, Nduna M, Shai N. Intimate partner violence, relationship power inequity, and incidence of HIV infection in young women in South Africa: a cohort study. The Lancet. 2010;376(9734):41–8.
50. Podsakoff PM, MacKenzie SB, Lee J-Y, Podsakoff NP. Common method biases in behavioral research: a critical review of the literature and recommended remedies. Journal of applied psychology. 2003;88(5):879. [PubMed: 14516251]

Table I.

Weighted Sample Characteristics (n = 581)

	Mean (range) or %
Age (mean)	22.4 (18–35)
Marital status	
Never married	85.2%
Divorced/separated/widowed	5.3%
Married (legal or traditional)	9.5%
Highest education level completed	
No school/some primary	3.1%
Completed primary	8.4%
Some high school	60.2%
Completed high school	26.4%
University/technikon	2.0%
Earned any income in the past 3 months	31.3%

Table II.

Confirmatory Factor Analysis Results for the Final GEM scale (n=290)

ITEM	Factor loading
A woman should tolerate violence to keep her family together.	.33
If someone insults a man he should defend his reputation with force if he has to.	.19
A man using violence against his wife is a private matter that shouldn't be discussed outside the couple.	.26
It is the man who decides what type of sex to have.	.39
Men are always ready to have sex.	.27
Men need sex more than women do.	.25
You don't talk about sex, you just do it.	.27
A woman who has sex before she marries does not deserve respect.	.25
Women who carry condoms on them are easy.	.45
It is a woman's responsibility to avoid getting pregnant.	.32
Only when a woman has a child is she a real woman.	.56
A real man produces a male child.	.57
Changing diapers, giving a bath, and feeding kids are the mother's responsibility.	.59
A woman's role is taking care of her home and family.	.69
The husband should decide to buy the major household items.	.72
A man should have the final word about decisions in his home.	.70
A woman should obey her husband in all things.	.54
Dropped items[a]	
There are times when a woman deserves to be beaten.	
It is alright for a man to beat his wife if she is unfaithful.	
A man can hit his wife if she won't have sex with him.	
A man needs other women even if things with his wife/partner are fine.	
It disgusts me when I see a man acting like a woman.	
A woman should not initiate sex.	
A man should be outraged if his wife/partner asks him to use a condom.	

[a]Dropped items shown here represent inequitable norms only (6 items representing equitable norms were also dropped).

All analyses incorporated sampling weights and accounted for the cluster sampling design.

All factor loadings were significant at $p<0.05$; $p=0.000$ for all but three items.

Table III.

Confirmatory Factor Analysis Results for the Final GRC/S Scale (n=290)

FACTOR AND ITEM	Factor loading[a]
Success, power, competition	**.73**
I worry about failing and how it affects my doing well as a man.	.32
I am often concerned about how others evaluate my ability to provide for my family.	.27
I strive to be more successful than others.	.55
I sometimes define my personal value by my ability to make money or find work.	.54
Feeling that I am in good physical condition is important to me as man.	.57
Being physically stronger than other men is important to me.	.68
I always strive to win in sports competitions.	.72
Having a girlfriend or wife is part of my idea of being a successful man.	.47
Subordination to women	**.87**
Making more money than a woman is a measure of my value and personal worth.	.38
Being outperformed at work by a woman would make me uncomfortable.	.57
I would be concerned if my friends knew I live with a woman and did any housework.	.50
I do not like to let a woman take control of the situation.	.43
I would be concerned if my friends knew I stayed at home to take care of children while my wife goes to work.	.51
Having a female boss would be difficult for me.	.53
Restrictive emotionality	**.61**
I have difficulty telling others I care about them.	.63
Talking about my feelings during or after sex is difficult for me.	.58
I often have trouble finding words to describe how I am feeling.	.62
I do not like to show my emotions to other people.	.67
Having someone see me cry would be difficult for me.	.52
Sexual prowess	**.56**
Being able to perform sexually is important to me as a man.	.78
I feel that I always need to be ready for sex with my partner, even if I am tired.	.44
I worry about being unable to become sexually aroused when I want.	.60
It is important to me to know I can drink as much or more than others.	.44
Having sex is part of being a successful man.	.49
Dropped items	
Making money is part of my idea of being a successful man.	
I often feel like I need to be in charge of those around me.	
I like to feel superior to other people.	
Affection with other men makes me tense.	

[a]Loadings for the factors (in bold) are loadings of those factors on the higher-order construct.

All analyses incorporated sampling weights and accounted for the cluster sampling design.

All factor loadings were significant at $p<0.05$; $p=0.000$ for all but two items.

Table IV.

Model Fit Statistics for Base and Correlated GEMS and GRC/S Scale Models in CFAs (n=290)

	χ^2	df	p	$\chi^2\Delta^a$	RMSEA	RMSEA 90% CI	CFI	TLI	SRMR
Value indicating good model fit:			*>0.05*		*<0.05*		*>0.95*	*>0.95*	*<0.08*
GEMS base model (17 items)	226.04	119	0.0000	--	0.056	[0.044,0.067]	0.81	0.79	0.065
GEMS final correlated model (17 items)	154.94	110	0.0032	−71.1 *	0.038	[0.022,0.051]	0.92	0.90	0.053
GRC/S scale base model (4 factors, 24 items)	492.55	246	0.0000	--	0.059	[0.051,0.066]	0.76	0.73	0.081
GRC/S scale final correlated model (4 factors, 24 items)	330.21	234	0.0000	−162.3 *	0.038	[0.028,0.047]	0.91	0.89	0.070
GRC/S scale correlated model with higher-order factor	337.66	236	0.0000	+7.5	0.039	[0.029,0.047]	0.90	0.88	0.070

CFI=Comparative Fit Index; CI=Confidence interval; RMSEA=Root Mean Square Error of Approximation; TLI=Tucker–Lewis Index; SRMR= Standardized Root Mean Squared Residual.

[a] $\chi^2\Delta$ is the chi-square difference statistic for comparing the current model to the previous model in the table.

* $p < .01$