Psychometric properties of the Gender Role Conflict Scale for Adolescents among Irish boys.

Article in Psychology of Men & Masculinity - March 2014
DOI: 10.1037/a0036018

CITATIONS: 4
READS: 251

3 authors, including:
Ronan Conway
University College Dublin
11 PUBLICATIONS 45 CITATIONS
Todd G Morrison
University of Saskatchewan
130 PUBLICATIONS 2,407 CITATIONS

Some of the authors of this publication are also working on these related projects:

- psychology View project
- Prevalence, Characteristics and Impacts of Macro- and Microaggressions Directed Toward LGBTT Persons in Canada View project

All content following this page was uploaded by Todd G Morrison on 29 January 2015.
The user has requested enhancement of the downloaded file.
Psychometric Properties of the Gender Role Conflict Scale for Adolescents Among Irish Boys

Cormac O’Beaglaoich and Ronan Conway
National University of Ireland

Todd G. Morrison
University of Saskatchewan

The psychometric properties of the Gender Role Conflict Scale for Adolescents (GRSC-A; Blazina, Pisecco, & O’Neil, 2005) were investigated using 2 samples of Irish males (Sample A: N = 317; Sample B: N = 199). Confirmatory factor analyses indicated that neither a 1-factor model nor a 4-factor model met established fit criteria. Across both samples, scores on the Restricted Emotionality (RE) and Conflict between Work, School, and Family (CWSF) subscales correlated consistently with self-esteem, and scores on the RE, CWSF and Need for Success and Achievement (NSA) subscales correlated with depression. Twelve items were problematic (i.e., standardized residuals and modification indices identified these items as contributing significantly to model misfit) suggesting that revision of the scale is warranted. Limitations and strengths associated with the current research are discussed and directions for future research are outlined.

Keywords: gender role conflict, masculinity, psychometrics, adolescence, Ireland
authors took into account the common “methodological sin” (Smith, McCarthy, & Anderson, 2000, p. 105) whereby the same or similar items are correlated twice (i.e., questions on the GRCS-A and their counterparts on the GRCS).

More recently, Jacobson, Marrocco, Kleinman, and Gould (2011) found that RE correlated positively with depressive symptoms and suicide ideation/intent. During a two-phase suicide screening project, the RE factor of the GRCS-A was administered to a large sample of high-school students in the United States (N = 2,189; 58.3% male; 13- to 18-years-old). Using a median split procedure, those placed in the high RE group (scores ranging from 20 to 54) were 11 times more likely than those in the low RE group (scores ranging from 9 to 19) to be depressed (i.e., based on a Beck Depression Inventory cutoff score of 16 or higher); three times more likely to report suicide ideation (after controlling for depressive symptoms, gender, age, ethnicity, and type of school attended); and more than twice as likely to report attempting suicide (again, after controlling for the aforementioned variables).

Wester, Kuo, and Vogel (2006) investigated whether, among Chinese Canadian adolescent males (N = 179; 13- to 19-years-old), GRC would contribute to psychological distress through the adoption of different coping strategies. They found support for aspects of their proposed mediation model. That is, the adoption of engagement coping strategies (i.e., confronting stressors through direct action, planning and positive thinking), in comparison with avoidance coping strategies (i.e., avoiding or accepting a situation as it is), decreased the levels of psychological distress associated with higher RE scores. Conversely, participants with higher RE scores, who adopted an avoidance coping strategy, reported greater psychological distress.

It is important to note that the GRCS-A has been adapted for use in countries outside Canada and the United States. Kim, Choi, Kim, and Park (2009) translated the 29-item GRCS-A into Korean (K-GRCS-A) and distributed the resultant measure to a sample of Korean adolescents (N = 374; 15- to 19-years-old). Cronbach’s alpha coefficients for the four subscales ranged from .67 to .80, and scores on the K-GRCS-A correlated positively with measures of depression (r = .29) and test anxiety (r = .48) and negatively with self-esteem (r = -.26). Further research using the K-GRCS-A found that self-esteem mediated the relationship between gender role conflict and depression; that is, individuals experiencing greater levels of conflict evidenced lower levels of self-esteem, which, in turn, was associated with greater levels of depression (Choi, Kim, Hwang, & Heppner, 2010).

The psychosocial view of gender acknowledges that cultures have different norms and different expectations about masculinity. In other words, masculinity “does not exist within a person but, rather, within an interaction between a person’s experience and the norms of [his or her] culture” (Kahn, 2009, p. 211). Masculinity ideology, which is a central concept within the gender role conflict paradigm, refers “to beliefs about the importance of men adhering to culturally defined standards for male behavior” (Pleck, 1995, p. 19), and implies that within each culture there are varied expectations for males. Consequently, the ways in which GRCS manifests itself may differ cross-culturally. O’Neil (2010) acknowledged cross-cultural variability in GRC and stated that using “American gender role concepts exclusively without grasping the cultural history of another country oversimplifies and clouds the real differences between two cultures” (p. 384). Further, O’Neil explained that theories emanating from any country may “have limited ability to explain attitudes and psychological processes in any other country” (p. 368). This statement supported several researchers’ contentions that cross-cultural analyses of GRC are essential (O’Neil, 2008b; Wester, 2008).

Hearn et al. (2002) noted that research in Ireland has been “slow to incorporate the study of men and masculinities into gender studies, and men as gendered subjects have remained largely outside of the gaze of critical [inquiry]” (p. 393). According to the authors, in comparison to research conducted in North America and the United Kingdom, empirical work on men and masculinity within an Irish context has “barely begun” (p. 394). Traditional Irish masculinity has been characterized as being “essentially rural, based heavily around family, marriage and celibacy” (Ferguson, 2001, p. 120); a view maintained by influential societal structures such as the state, the Gaelic Athletics Association (GAA) and the Catholic Church. Magennis and Mullen (2011) highlighted rapid changes within Irish society in recent years, which include the weakening of the Catholic Church, through the liberalization of laws on divorce and homosexuality, and the ideology of the Irish Free State. Further, Ireland underwent radical economic transformations from the 1990s to the present day, transitioning from a period of economic growth (i.e., the Celtic Tiger) to a recession in 2007. The concept of a “monolithic” masculinity also has been challenged by: (a) a GAA player (Donal Óg Cusak) acknowledging that he is gay; (b) the 2009 GAA footballer of the year, Paul Galvin, becoming a fashion columnist for one of Ireland’s leading newspapers; and (c) an openly gay man (Senator David Norris) running for the office of president of Ireland. Given Ireland’s rapidly changing society, it appears that “the time is ripe to reevaluate the representation of masculinity in the Irish context” (Magennis & Mullen, 2011, p. 1).

Psychometric testing is an incremental process necessitating the application of novel tests of validation and the use of different samples to gauge, for example, the reproducibility of a measure’s dimensional structure (Byrne, 2006). Although there exists acknowledgment that GRC and, by extension, the way in which it is measured, may differ as a function of cultural values (e.g., O’Neil, 2010), to the authors’ knowledge, the psychometric characteristics of the GRCS-A have yet to be investigated outside the parameters of North America. Therefore, the purposes of this study were to: (a) assess whether the factor structure of the GRCS-A could be replicated using two independent samples of Irish adolescents; (b) investigate the scale score reliability of the measure; and (c) test its validity, focusing specifically on the association between the GRCS-A and a multidimensional indicator of adolescent masculinity (e.g., Meanings of Adolescent Masculinity Scale [MAMS]; Oransky & Fisher, 2009) as well as self-esteem and depression. It should be noted that the authors are unaware of any published research testing these relationships among Irish youth.

Method

Participants

Sample A. Sample A consisted of 325 participants from a secondary school along the west coast of Ireland. In total, eight cases were deleted due to eight or more items of the GRCS-A being omitted (see Norwalk, Vandiver, White, & Englar-Carlson,
Asian background, and three participants didn’t specify their ethnicity. Approximately 91.4% (n = 325) and the analytic sample (n = 317).

The participants whose data were analyzed varied in age from 12 to 19 years (M = 15, SD = 1.61). They were first (n = 68; 22.1%), second (n = 81; 26.3%), third (n = 49; 15.9%), fourth (n = 36; 11.7%), fifth (n = 33; 10.7%), and sixth year students (n = 41; 12.9%), who attended a coeducational secondary school along the west coast of Ireland.1 Nine participants didn’t specify their age. Approximately 93.4% (n = 295) identified as Irish2, 6.6% (n = 21) identified as any other White background, and one participant didn’t specify his ethnicity. Approximately 91.4% (n = 286) reported being exclusively heterosexual, 4.2% (n = 13) as more heterosexual than gay, 1.9% (n = 6) as bisexual, 1% (n = 3) as more gay than heterosexual, and 1.6% (n = 5) as exclusively gay. Four participants did not report their sexual orientation.

Sample B. Sample B consisted of 207 boys from a secondary school also situated along the west coast of Ireland. In total, eight cases were deleted due to eight or more items of the GRCS-A being missed (Norwalk et al., 2011). Little’s missing completely at random test (MCAR) was statistically nonsignificant, χ²(1102) = 1240.43, p < .05, suggesting the data were missing either at random (MAR) or not completely at random (NMAR). Thus, the expectation maximization (EM) algorithm for imputing missing data was employed (see Fish, McGuire, Hogan, Morrison, & Stewart, 2013). Comparisons across age, socioeconomic status (SES) and GRCS subscales indicated there was no difference between the initial sample (n = 207) and the one retained for data analysis (n = 199; M age = 14.41, SD = 1.2).3 First (n = 60; 30.5%), second (n = 49; 24.9%), third (n = 61; 31%), fourth (n = 11; 5.6%), and fifth (n = 16; 8.1%) year students were included, all of whom attended a coeducational secondary school along the west coast of Ireland. Two participants did not specify their age. A majority (n = 155, 79.1%) identified as Irish, 15.8% (n = 31) identified as any other White background, 3.6% (n = 7) identified as African, 1% (n = 2) identified as Chinese, and 0.5% (n = 1) identified as any other Asian background, and three participants didn’t specify their ethnicity. Approximately 95.8% (n = 182) identified as exclusively heterosexual, 2.1% (n = 4) as more heterosexual than gay, 0.5% (n = 1) as bisexual, 1.1% (n = 2) as more gay than heterosexual, and 0.5% (n = 1) identified as exclusively gay. Nine participants did not report their sexual orientation.

Sample A—Validation measures.

The Adolescent Masculinity Ideology in Relationships Scale (AMIRS; Chu et al., 2005). The 12-item AMIRS is a unifactorial measure of adolescent boys’ internalization of masculine norms, attitudes and behaviors displayed within interpersonal relationships (e.g., “I think it’s important for a guy to go after what he wants even if it means hurting other people’s feelings”). This scale employs a 4-point Likert-type response format (i.e., disagree a lot, disagree, agree, agree a lot). Total scores can range from 12 to 48, with higher scores denoting stronger endorsement of masculine norms as they pertain to relationships. Five of the items are reverse scored. In support of the measure’s concurrent validity, scores on the AMIRS correlated positively with scores on the Male Role Attitudes Scale (r = .58; Blazina et al., 2007). The convergent validity of the AMIRS was suggested by its negative association with self-esteem scores (r = −.32; Chu et al., 2005), and positive association with the GRCS-A (r = .45; Blazina et al., 2007). In the current study, Cronbach’s alpha was .77 (95% CIs [.73, .80]).

Gender Role Conflict Scale for Adolescents (GRCS-A; Blazina et al., 2005). As noted earlier, the GRCS-A is a 29-item self-report scale focusing on: RE (nine items), RAM (seven items), CWSF (seven items), and NSA (six items). For all items, a 4-point Likert-type response format is used (i.e., strongly disagree, disagree, agree, strongly agree). Total scores can range from 29 to 116, with higher scores indicating greater gender role conflict.

Meanings of Adolescent Masculinity Scale (MAMS; Oransky & Fisher, 2009). The MAMS is a multidimensional scale that aims to assess the degree to which adolescent boys “endorse or reject traditional male roles” (p. 59). The scale contains 27 items, which fall into one of four subscales: Constant Effort (CE; seven items; i.e., boys must maintain a confident, tough and strong image in order to perform masculinity); Heterosexism (H; eight items; i.e., boys need to be stoic and refrain from sharing their feelings with others in order to be masculine); Heterosexism (H; eight items; i.e., one must not show any attitudes and behaviors associated with femininity or homosexuality); and Social Teasing (ST; five items; i.e., to be masculine, boys must be able to tease, and stand up to teasing from, other boys). For all items, a 4-point Likert-type response format is used (i.e., strongly disagree, disagree, agree, strongly agree). Total scores can range from 27 to 108, with higher scores indicating greater endorsement of masculine norms. Individual subscale scores also can be computed. Oransky and Fisher (2009) reported the following scale score reliability coefficients: CE, α = .79; ER, α = .80; H, α = .80; and ST, α = .61. In support of the measure’s convergent validity, the authors reported various statistically significant associations between scores on the MAMS’ subscales and measures such as peer popularity (CE, r = .50); intimate exchange among friends (ER, r = .44); anti-femininity (H, r = .66); and beliefs about aggression (ST, r = .35). In the current study, Cronbach’s alpha coefficients were: ER (α = .72; 95% CIs [.67, .76]); H (α = .82; 95% CIs [.79, .85]); ST (α = .73; 95% CIs [.68, .78]); and CE (α = .79; 95% CIs [.75, .82]). For the total scale, Cronbach’s alpha was .91 (95% CIs [.90, .93]).

Rosenberg Self-Esteem Scale (RSES; Rosenberg, 1965). The RSES is a widely adopted 10-item unifactorial trait measure of global self-esteem (e.g., “I take a positive attitude toward myself”). Five items are reverse scored, with higher total scores reflecting greater levels of self-esteem (scores can range from 1 to 40). The RSES uses a 4-point Likert-type response format.

---

1 First year students are typically aged 12 years old. Second year students are typically aged 13- to 14-years-old.

2 Model fit was tested for participants who identified as “Irish” and “heterosexual.” As there was no difference in model fit scores (e.g., fit indices for a four factor model for Sample A were: χ²/df = 2.06; TLI = .83; NNFI = .75; SRMR = .07; RMSEA = .06), all participants were included in the final analysis.

3 A statistically significant difference in age was observed between Samples A and B, t (514) = 4.45, p < .001, d = .51.
(i.e., not at all, somewhat, moderately so, very much so). In a comprehensive assessment of the RSES’ psychometric properties, Schmitt and Allik (2005) distributed this measure to approximately 17,000 respondents located in 53 nations. The authors concluded that the “internal reliability and factor structure of the RSES is psychometrically sound” (p. 639). Validation evidence also was obtained (i.e., scores on this measure of self-esteem correlated in anticipated directions with personality variables such as extraversion and neuroticism). The Republic of Ireland was not included in this cross-cultural study. However, using data from the cross-national Irish Health Behavior in School Children (HBSC) survey, Mullan and NicGabhann (2002) found that, in accordance with past research, female respondents obtained lower scores on the RSES, and self-esteem appeared to decrease with age (i.e., those between the ages of 10 and 12 reported higher scores than those between the ages of 15 and 17). In the current study, Cronbach’s alpha was .82 (95% CIs [.78, .85]).

Sample B—Validation measures.

Center for Epidemiologic Studies Depression scale (CES-D; Radloff, 1977). The CES-D is a 20-item screening measure of depression employed with nonclinical samples of adults, children, and adolescents. It uses a 4-point Likert response format (Rarely or none of the time [less than 1 day]; some or a little of the time [1–2 days]; occasionally or a moderate amount of the time [3–4 days]; most or all of the time [5–7 days]). Scores can range from 0 to 60, with a score of 16 points or more indicating depression. If a participant omits more than four questions, it is not recommended that his or her score be calculated. With respect to the measure’s psychometric soundness, O’Farrell, Flanagan, Bedford, James, and Howell (2005) found that, in keeping with established group differences, the female adolescents in their study obtained higher scores on the CES-D than did the male adolescents. As well, those evidencing lower levels of self-esteem, as measured by the Rosenberg Self-esteem Scale, were more likely to have a CES-D score equal to (or greater than) 24, which the authors used as the cut-off denoting significant levels of depressive symptomology. In the current study, Cronbach’s alpha was .86 (95% CIs [.83, .89]).

Gender Role Conflict Scale for Adolescents (GRCS-A; Blazina et al., 2005). A description of this measure is provided in the materials distributed to Sample A.

Rosenberg Self-Esteem Scale (RSES; Rosenberg, 1965). With this sample, Cronbach’s alpha was .82 (95% CIs [.78, .85]).

Procedure

Ethical approval was obtained through the Research Ethics Committee at the National University of Ireland, Galway. Schools were contacted by letter, which detailed the protocol involved in the research study. Two secondary schools (Samples A and B) in Western Ireland agreed to take part and, following approval from the boards of management of each school, the principal asked the guidance counselors to oversee the data collection procedure. Consent forms were distributed to boys during class time and prospective respondents were asked to bring the information sheets and consent forms to their parent(s)/guardian(s). Participants whose parents did not want their child to take part in the research were asked to bring the signed consent form to the guidance counselor. Approximately 7 days later, the guidance counselor went to each class to remind participants to bring the consent forms home. Two weeks later, questionnaires were distributed to all students. Participants took approximately 15 to 20 minutes to complete the survey.

Results

Preliminary Analysis

Data from Samples A and B were analyzed using SPSS version 20.0. Table 1 provides the means, standard deviations, reliability estimates, and correlations for the GRCS-A and its subscales. There were no statistically significant differences between the means of Samples A and B for the CWSF and NSA subscales. However, statistically significant differences be-

<table>
<thead>
<tr>
<th>Subscale</th>
<th>RE</th>
<th>RAM</th>
<th>NSA</th>
<th>CWSF</th>
<th>GRCS-A</th>
<th>M A (SD)</th>
<th>M B (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RE</td>
<td></td>
<td>.58*</td>
<td>.34*</td>
<td>.33*</td>
<td>.78*</td>
<td>30.8 (8.7)</td>
<td>28.3 (8.6)*</td>
</tr>
<tr>
<td>RAM</td>
<td>.55*</td>
<td></td>
<td>.35*</td>
<td>.32*</td>
<td>.75*</td>
<td>25.3 (7.6)</td>
<td>21.8 (7.0)*</td>
</tr>
<tr>
<td>NSA</td>
<td>.25*</td>
<td>.19*</td>
<td></td>
<td>.50*</td>
<td>.70*</td>
<td>22.9 (6.2)</td>
<td>22.1 (6.7)</td>
</tr>
<tr>
<td>CWSF</td>
<td>.36*</td>
<td>.22*</td>
<td>.46*</td>
<td></td>
<td>.73*</td>
<td>22.7 (7.3)</td>
<td>23.1 (7.4)</td>
</tr>
<tr>
<td>GRCS-A</td>
<td>.80*</td>
<td>.72*</td>
<td>.62*</td>
<td>.71*</td>
<td></td>
<td>102.2 (21.6)</td>
<td>96.5 (21.5)*</td>
</tr>
<tr>
<td>Alpha scores</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample A</td>
<td>.82</td>
<td>.82</td>
<td>.77</td>
<td>.78</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>95% CI</td>
<td>[.79, .85]</td>
<td>[.71, .80]</td>
<td>[.73, .81]</td>
<td>[.74, .81]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample B</td>
<td>.85</td>
<td>.73</td>
<td>.81</td>
<td>.80</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>95% CI</td>
<td>[.81, .88]</td>
<td>[.67, .80]</td>
<td>[.76, .85]</td>
<td>[.75, .84]</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. RE = Restrictive Emotionality; RAM = Restrictive Affection between Men; NSA = Need for Success and Achievement; CWSF = Conflict between Work and Family Relations. M A = Mean for Sample A; M B = Mean for Sample B. Correlations above the diagonal are for Sample A and below the diagonal are for Sample B. *p < .01.
tween the two samples were observed for RE and RAM, with scores being higher in Sample A. In addition, the total GRCS-A was significantly higher for Sample A than Sample B. Inter-subscale correlations ranged from .32 to .58 in Sample A and from .19 to .55 in Sample B. Cronbach’s alpha coefficients for each subscale ranged from .77 to .82 in Sample A and .73 to .85 in Sample B.

**Confirmatory Factor Analysis**

Using AMOS 20 for Windows, a series of confirmatory factor analyses (CFAs) were conducted on the GRCS-A for Samples A and B. A competing models approach was employed to test two nested models: a one-factor first-order model and a four-factor first-order model.

Several criteria were used to assess the goodness of fit of the hypothesized models to the observed data. The first criterion was the chi-square statistic. Due to its sensitivity to sample size (Cole, 1987), additional criteria were employed: (a) a chi-square to degrees of freedom ratio < 2 (Byrne, 2006); (b) three incremental indexes (comparative fit index [CFI] > .95, Byrne, 2006; Hu & Bentler, 1995; Tucker-Lewis Index [TLI] > .95, Hu & Bentler, 1995; and non-normed fit index [NNFI] > .95, Hu & Bentler, 1995); and (c) the standardized root-mean-square residual [SRMR] < .05, Hu & Bentler, 1995) and root mean square error of approximation [RMSEA] < .05, Byrne, 2006).

**Testing Competing Models**

**Sample A.** The one-factor model was rejected due to its failure to meet recommended model-fit criteria (see Table 2). Although the four-factor model provided better fit to the data, cut-off values for indicators of fit were not satisfied (see Table 2). Examination of the standardized residuals and modification indices identified various items that contributed significantly to model misfit (e.g., Items 7 and 11, which reflect NSA, cross-loaded onto the other three factors; CWSF, RE, and RAM). Following the removal of Items 3, 7, 9, 10, 11, and 13, the resultant model was tested, with improvements in fit observed:

\[ \chi^2/df = 1.77, CFI = .90, TLI = .89, RMSEA = .05; 90\% CI [.04, .06]. \]

However, when this “new” model was examined with Sample B, suboptimal fit indices emerged: \[ \chi^2/df = 2.145, CFI = .87, TLI = .83, RMSEA = .07; 90\% CI [.05, .09]. \]

A three-factor model also was tested (i.e., the NSA was deleted as only one of its items (specifically, Item 11) appears to map onto conflict as it is traditionally understood (see O’Beaglaoich et al., 2013) but, again, fit criteria were not satisfied (see Table 2).

**Sample B.** Consistent with the results of Sample A, the one-factor model was rejected. Also, as noted with Sample A, the four-factor model provided better but not optimal fit (see Table 2). The standardized residuals and modification indices suggested that Items 1, 3, 11, 18, 26, and 28 be removed improving the fit; however, the model fit statistics did not improve when these items were tested with Sample A.

**Validation Indices**

As shown in Table 3, the GRCS-A and its subscales correlated with a number of the validation measures used in this study. However, results were mixed in terms of associations between the GRCS-A and measures of masculinity (i.e., the MAMS and the AMIRS). No discernible pattern was evident with respect to the correlations between the GRCS-A and MAMS subscales. In contrast, statistically significant positive associations were observed between the AMIRS and each of the GRCS-A subscales (except for the NSA subscale). In both samples, the GRCS-A correlated negatively with self-esteem, as did two of its subscales (CWSF and RE). Specifically, boys that reported greater restrictive emotionality and conflict between school and lifestyle pressures also reported lower self-esteem. As predicted, respondents’ scores on the GRCS-A, and three of its four subscales (CWSF, NSA, and RE), correlated positively with their level of depression.

**Discussion**

The purpose of this study was to investigate the dimensionality and the validity of the GRCS-A with two samples of Irish
adolescent males. Confirmatory factor analyses indicated that a
one-factor model, three-factor model, and four-factor model
failed to meet established fit criteria (i.e., conventional indices
for CFI, NNFI, TLI, RMSEA, and SRMR). Across both samples, 12 items4
were identified as being problematic. Item 3, for instance, cross-loaded on three factors (RE, RAM, and NSA) as
did Item 11 (the NSA factor), which cross-loaded on RAM, CWSF, and RE. Items specific to each sample (Sample A: Items 5, 7, 9, 10, 13, and 25; Sample B: items 1, 18, 26, and 28) cross-loaded on two or more factors suggesting that alterations to the GRCS-A are required. Revisions and alterations could
emerge for the Irish participants. In a recent qualitative study,
obtained from the “Homophobia” scale of the original GRCS (O’Neill et al., 1986), correlated positively with the H subscale of the MAMS; however, neither RE nor CWSF subscale scores were associated with scores on the
MAMS. Surprisingly, the RE subscale of the GRCS-A, which measures restrictive emotionality, and the ER subscale of the
MAMS, which assesses emotional restriction, did not intercor-
late. The absence of an association suggests that endorsement of
cultural standards of masculinity relating to emotional expression do not necessarily imply conflict on this dimension, which, in turn, challenges the assumption that masculine ideology measures
should be used to validate indicators of gender role conflict.6
Further, examination of the ER items (i.e., “Guys should not talk
about their worries with each other” and “It is weird for a guy to
talk about his feelings with other guys”) and RE items (i.e., “I have
difficulty telling others I care about them” and “Expressing my
feelings makes me feel open to attack by others”), suggests the
former is measuring general attitudes about how expressive boys
are in the context of other males while the latter is measuring self-reported difficulties with expressing emotions.

The results of this study indicate that for Irish boys there was no
significant correlation between the GRCS-A and the endorsement
of masculine standards on dimensions measured by the MAMS
(i.e., CE, ER, and ST) However, in contrast to these findings was
the significant positive correlation between the GRCS-A and the
AMIRS (r = .25). The correlations between the GRCS-A and the
MAMS raise theoretical questions concerning the suitability of
using masculine ideology measures to determine the concurrent
validity of GR scale. With this type of validity, which falls
why that would be the case, one interviewee noted: “They’re all
about loving men; who’s going to take that seriously?”

Similar to past research, the GRCS-A (and three of its subscales)
correlated positively with the AMIRS. However, a statistically
significant correlation was not observed for the RAM subscale.
With respect to the MAMS, no consistent pattern of correlations emerged between this measure of masculinity and the GRCS-A.
The RAM subscale, which originally derived from the “Homopho-
bia” scale of the original GRCS (O’Neill et al., 1986), correlated
dependent in nature. The absence of an association suggests that endorsement of
males identifies that a particular item measuring “conflict” (O’Beaglaoich et al., 2013),
emerged in all tested models as problematic; thus, its deletion
apparent concern was that any model found to be satisfactory with Sample A failed to replicate with Sample B or vice versa suggesting that the
models meeting established fit criteria were highly idiosyncratic
and nonreproducible.
The wording of scale items offers one possible explanation for
why the anticipated four factor model of the GRCS-A did not
emerge for the Irish participants. In a recent qualitative study,
O’Beaglaoich et al. (2013) found general support for GR theory
among samples of Irish male adolescents. However, when asked
to scrutinize the content of the GRCS-A, participants expressed con-
cerns about the wording of nine items. To illustrate, Item 1
(“Verbally expressing my love to another man is hard for me”) was
regarded as inappropriate for boys, and better suited for girls and
gay males, because of its use of the word “love.” (Indeed, one
participant wrote beside this item: “No, because I don’t find any
man attractive.”) The phrase “hard for me” also was identified as
problematic and often triggered laughter among discussants be-
cause boys associated “hard” with “getting an erection.”
O’Beaglaoich et al. (2013) note that some participants believed the
inclusion of these sorts of items would result in boys not answering
the questionnaire in a thoughtful or reflective manner. When asked

<table>
<thead>
<tr>
<th>Subscale</th>
<th>ER</th>
<th>CE</th>
<th>H</th>
<th>ST</th>
<th>AMIRS</th>
<th>RSES</th>
<th>CESD(B)</th>
<th>RSES(B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RE</td>
<td>.07</td>
<td>.03</td>
<td>.11</td>
<td>.06</td>
<td>.28**</td>
<td>-.21**</td>
<td>.31**</td>
<td>-.33**</td>
</tr>
<tr>
<td>RAM</td>
<td>.19*</td>
<td>.00</td>
<td>.29**</td>
<td>.02</td>
<td>.07</td>
<td>-.05</td>
<td>.13</td>
<td>-.11</td>
</tr>
<tr>
<td>NSA</td>
<td>-.15*</td>
<td>-.01</td>
<td>-.05</td>
<td>-.06</td>
<td>.20**</td>
<td>-.02</td>
<td>.24**</td>
<td>.05</td>
</tr>
<tr>
<td>CWSF</td>
<td>-.02</td>
<td>.08</td>
<td>.01</td>
<td>.03</td>
<td>.22**</td>
<td>-.22**</td>
<td>.36**</td>
<td>-.22**</td>
</tr>
<tr>
<td>GRCS-A</td>
<td>.08</td>
<td>.02</td>
<td>.18**</td>
<td>.01</td>
<td>.25</td>
<td>-.17**</td>
<td>.33**</td>
<td>-.19</td>
</tr>
</tbody>
</table>

*Correlations statistically significant at p < .05. **Correlations statistically significant at p < .01.

Note. MAMS = Meaning for Masculinity Ideology Scale consisting of: ER = Emotional Restriction; CE = Constant Effort; H = Homophobia; and ST = Social Teasing; AMIRS = Adolescent Masculinity in a Relationship Scale; RSES = Rosenberg Self-Esteem; CES-D = Center for Epidemiologic Studies Depression; B = Sample B.

Table 3
Correlations Between Gender Role Conflict and Indices of Masculine Ideology, Self-Esteem, and Depression

4 Items 1, 3, 5, 7, 9, 10, 11, 13, 18, 25, 26, and 28 were problematic.
5 Details about these various permutations are available upon request.
6 O’Neill (2008a) proposes evidence of convergent validity of the GRCS highlighting that it has been significantly correlated with masculinity-related constructs in the range of .32 and .49. He states that “these significant correlations suggest that the GRCS is related to these masculinity scales, but the low to moderate correlations suggest that the GRCS measures a different construct” (p. 372).
under the rubric of criterion-related validation, one examines the relationship between a “gold standard” measure and another indicator of the same construct (see Carmines & Zeller, 1979). In the current study, the maximal correlation between the GRCS-A and the four subscales of the MAMS was .18 (i.e., 3% shared variance); thus, these measures appear to be assessing constructs that are, at most, marginally interrelated. Little research has been conducted with adolescent males in Ireland and, to the authors’ knowledge, this study is the first one to employ both the AMIRS and the MAMS; thus, conclusions based on these results are tentative until further research has been conducted. Future studies should examine their respective factor structure when completed by Irish youth. Preliminary estimates suggest that the AMIRS (.77) and each subscale of the MAMS (rs = .72-.79) possess satisfactory score reliability when completed by adolescent boys residing in Ireland; however, researchers should investigate the factor structure of these measures to justify their continued use among members of this cultural group.

The absence of a strong relationship between scores on the GRCS-A and indicators of masculinity is not surprising, given O’Neil’s (2008a) contention that the GRC is experienced in numerous situational contexts. First, males may have a gender role transition or face difficult developmental tasks over the life span. Second, they may deviate from or violate gender role norms of masculinity ideology (MI). Third, they may try to meet or fail to meet gender role norms of MI. Fourth, they may note discrepancies between their real self-concepts and their ideal self-concepts, based on gender role stereotypes and MI. Fifth, men may personally devalue, restrict, and/or violate themselves for failing to meet MI norms. Sixth, they may experience personal devaluations, restrictions, and/or violations from others for conforming to or deviating from MI. Seventh, and finally, men may personally devalue, restrict, and/or violate others because of their deviation from or conformity to MI norms (O’Neil, 2008a). In light of these various contexts, it is unclear why positive correlations between scores on the GRCS and traditional indicators of masculinity should invariably emerge. For example, if a male adolescent reports low masculinity ideology scores (i.e., he does not endorse masculine standards), it is still possible for him to experience GRC.

Predicted associations between adolescents’ gender role conflict and their levels of self-esteem and depression emerged for some of the GRCS-A subscales. Specifically, scores on the RE and CWSF correlated consistently with self-esteem across both samples. These correlations suggest that Irish boys having greater difficulties expressing themselves and/or experiencing greater pressures about school and social/family life obligations, report lower self-esteem. Given the associative nature of this study, it is similarly possible that Irish adolescent boys reporting lower self-esteem are, in turn, less able to cope with school, social, and family life obligations. Scores on the RE, CWSF, and NSA subscales also correlated with self-reported levels of depression suggesting that greater levels of gender-based conflict may be associated with decrements in mental well-being.

It is worth noting that the RAM subscale did not correlate with depression or self-esteem (across both samples). As this is the first study to investigate the GRCS-A using Irish adolescents as participants and also the first to test the relationship between the GRCS-A and self-esteem, it is difficult to decipher whether these results are idiosyncratic or a reflection of cross-cultural variation.

To the authors’ knowledge, this is the first study to attempt to confirm the factor structure of the GRCS-A using adolescent participants outside of North America. It should be noted, however, that the results of the current study are similar to those obtained in previous CFA studies using the GRCS (adult version). To date, no CFA has conclusively supported the factor structure of the GRCS. The observed fit indices in this study mirror those noted in earlier research (see Good et al., 1995; Moradi, Tokar, Schaub, Jome, & Serna, 2000; Norwalk et al., 2011; Rogers, Abbey-Hines, & Rando, 1997). Therefore, it is possible that the GRCS-A has inherited an essentially “misfitting” model from its adult counterpart. In the development of the GRCS-A, the researchers superimposed the factor structure used with adults without consulting adolescents, resulting in the elision of potential age-specific GRC patterns. This practice should be contrasted with more recently developed adolescent masculinity measures (i.e., the AMIRS and the MAMS), in which items were generated based on focus groups and individual interviews with adolescent males.

Standards of measurement have progressed in the past 26 years with the emergence of SEM cut-off points; standards that were not available during the period when the GRCS was developed (Norwalk et al., 2011). Examinations of the psychometric properties of the GRCS (adult) by means of CFAs have consistently fallen short of established fit criteria. Further, as items on the GRCS have remained unchanged for over 25 years, some researchers have suggested that this scale requires modification (Norwalk et al., 2011). Traditionally, the use of American concepts of masculinity and conflicts has been deemed suitable for other cultures (i.e., the Korean version of the GRCS-A: Kim et al., 2009); however, our findings suggest that it is important to test the psychometric properties of a measure prior to its wide-scale importation.

As with all research, the current study possesses limitations that warrant discussion. Given that factor analysis is based on asymptotic (i.e., large-sample) theory (Cole, 1987), improper and (often) spurious model solutions and nonconvergence have been found when sample sizes are small (Cole, 1987). Due to its sensitivity to sample size and the relatively small Ns used in the current study (307 and 199 respondents for Samples A and B, respectively), questions relating to the suitability of the factor structure as a function of age were left unexplored (i.e., it is unclear whether the four-factor model is appropriate for younger adolescents, but not older ones, etc.). Also, for each of the schools involved in this study, an estimate of the number of students who opted out of participating was not recorded. Thus, the degree to which these findings are (potentially) compromised by self-section bias is unclear.

References


short-form development. *Psychological Assessment, 12*, 102–111. doi:
10.1037/1040-3590.12.1.102


Wester, S. R. (2008). Thinking complexly about men, gender role conflict,
doi:10.1177/001100007310971

Chinese-Canadian adolescents, male gender role conflict, and psycho-
logical distress. *Psychology of Men & Masculinity, 7*, 87–100. doi:
10.1037/1524-9220.7.2.83

research on men and masculinity in the United States: Dominant meth-
9120-1

Received February 5, 2013
Revision received November 25, 2013
Accepted November 26, 2013

This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.