

GENDER CONFORMITY NORMS: A SYSTEMATIC REVIEW OF VALIDITY EVIDENCE

Julia Caciano da Silva
Universidade Salgado de Oliveira, Niterói-RJ, Brasil

Felipe Valentini
Universidade São Francisco, Campinas-SP, Brasil

Clarissa Pinto Pizarro de Freitas
Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro-RJ, Brasil

Abstract

The aim of this study was to conduct a systematic review to verify the validity evidence of the internal structure of the Conformity to Feminine Norms Inventory and the Conformity to Masculine Norms Inventory. The results of this review allow us to verify the applicability of the underlying structure, as well as the relationship between the items and the construct assessed in the instruments. The search was performed in PUBMED, PsycINFO and SCOPUS. The analysis of the psychometric properties reported in the publications was based on the recommendations of competent institutions in the area. Nineteen publications were included, which contained a sample composed predominantly of heterosexual college students. The publications found contain data that corroborate the validity and reliability of the instruments, however, inconsistencies are found around the most appropriate theoretical model: correlated factors or bifactor.

Keywords: gender norms; femininity; masculinity; psychometrics.

Conformidade com normas de gênero: Revisão sistemática das evidências de validade

Resumo

O objetivo deste estudo foi realizar uma revisão sistemática para verificar as evidências de validade da estrutura interna do Inventário de Conformidade com as Normas Femininas e do Inventário de Conformidade com as Normas Masculinas. Os resultados desta revisão permitem verificar a aplicabilidade da estrutura subjacente, bem como a relação entre os itens e o construto avaliado nos instrumentos. A busca foi efetuada nas bases PUBMED, PsycINFO e SCOPUS. A análise das propriedades psicométricas relatadas nas publicações pautou-se nas recomendações de instituições competentes na área. Foram incluídas 19 publicações, as quais continham uma amostra composta predominantemente por jovens universitários heterossexuais. As publicações encontradas contêm dados que corroboram a validade e a fidedignidade dos instrumentos, mas são encontradas lacunas em torno do modelo teórico mais adequado: fatores correlacionados ou bifactor.

Palavras-chave: normas de gênero; feminilidade; masculinidade; psicometria.

Conformidad con las normas de género: Revisión sistemática de las evidencias de validez

Resumen

El objetivo de este estudio fue realizar una revisión sistemática para verificar la evidencia de validez de la estructura interna del Inventario de Conformidad con las Normas Femeninas e del Inventario de Conformidad con las Normas Masculinas. Los resultados de esta revisión nos permiten verificar la aplicabilidad de la estructura subyacente, así como la relación entre los ítems y el constructo evaluado en los instrumentos. La búsqueda se realizó en las bases de datos PUBMED, PsycINFO y SCOPUS. El análisis de las propiedades psicométricas reportadas en las publicaciones se basó en las recomendaciones de instituciones competentes en el área. Se incluyeron diecinueve publicaciones, que contenían una muestra compuesta principalmente por estudiantes universitarios heterossexuales. Las publicaciones encontradas contienen datos que corroboran la validez y confiabilidad de los instrumentos, pero las brechas se encuentran alrededor del modelo teórico más apropiado: factores correlacionados o bifactor.

Palabras clave: normas de género, feminidad, masculinidad, psicometría.

The terms sex and gender can be treated as equal or opposite. In this study sex refers to biological body and gender refers to the social interpretation of this biological body (Helgeson, 2017). Although distinct, sex and gender are connected (Fernández, 2010). Within psychological science, gender can be evaluated from different perspectives. The Conformity to Feminine Norms Inventory (CFNI, Mahalik et al., 2005) and the Conformity to Masculine Norms Inventory (CMNI, Mahalik et al., 2003) are instruments that evaluate gender roles from a multidimensional perspective. The instruments are based on a *continuum* degree of agreement-disagreement. Thus, items measure the attitudes, beliefs, and behaviours associated with traditional and non-traditional feminine and masculine norms (Mahalik et al., 2003, 2005; Parent & Moradi, 2011a).

In view of the social construction of gender (Scott, 1995), it is essential to identify how the structure of CFNI and CMNI is presented in different languages and cultures. As much as differences between the existing versions of the instruments are expected, it is necessary to review how the evaluated items and constructs behave, as well as their psychometric properties.

The CFNI is composed of 84 items divided into eight female norms that involve: friendly relationships with others, involvement with children, pursuit of a thin body, keep sexual intimacy within one committed relationship, refrain from calling attention to one's qualities or abilities, invest self in romantic relationship, maintain the home and invest resources to maintaining and improving physical appearance. The CMNI is composed of 94 items divided into eleven male norms that involve: emotional restriction and suppression, drive to win, propensity for physical confrontation, desire for multiple and noncommitted sexual relations and emotional distance from sex partners, aversion to asking for assistance, propensity to high-risk behaviours, general desire to have personal control over situations, perceived control over women at both personal and social levels, viewing work as a major focus of life, being pleased with being though as important, and aversion to the prospect of being gay or being perceived as gay from some else (Mahalik et al., 2003, 2005; Parent & Moradi, 2009).

Femininity and masculinity are understood as the degree of agreement or disagreement with gender

norms for women and men. This interpretation allows the expression of femininity and masculinity to also be defined in opposition to the aforementioned social norms (Mahalik et al., 2003, 2005).

Due to the large number of publications that use both CFNI and CMNI for gender measurement, a systematic review of studies reporting validity evidence based on internal structure was performed. Validity, according to Standards for Educational and Psychological Testing (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2014), refers to the degree to which evidence and theories support interpretations of test scores for their intended use. The Standards report five sources of validity evidence: test content; response processes; internal structure; relations to other variables; and consequences of testing. Validity evidence based on internal structure gathers information on the degree to which relations between test items and test components agree with the construct on which instrument scores are proposed and interpretations are based (AERA, APA, & NCME, 2014). Thus, we verify how the model proposed by Mahalik et al. (2003, 2005) occurs in several countries and cultures, we check versions of the instruments, and we examined the psychometric indicators used in such studies.

Considering that CFNI and CMNI assess the level of conformity to gender-role norms, reviewing the validity evidence based on internal structure allows us to identify whether the factors presented are corroborated in different countries and cultures, to verify whether psychometric indicators are satisfactory, and to analyse possible changes in structure of the instruments recommended by the researchers. Finally, we sought to present similarities and discrepancies between the selected publications, suggesting future studies involving psychometric analysis of the instruments.

Method

Systematic review is a research method that recovers previous studies to verify the main advances and the possible gaps existing in the literature, within the theme it proposes to analyse (Sampaio & Mancini,

2007). Based on the descriptions presented by Grant & Booth (2009), a critical literature review and a comprehensive search process were associated. Thus, the extent and quality of psychometric studies involving the instruments are pointed out, as well as the similarities and absences in the characteristics of the studies: type of sample, tested models, language, reliability coefficients, and other aspects.

Consulted Databases

The following databases were consulted: PUBMED, PsycINFO, and SCOPUS. PUBMED is a database of publications focused on the medical and health fields. PsycINFO is the database of the American Psychological Association (APA), reference in studies in Psychology. SCOPUS is a broad database of publications from various scientific fields.

Definition of Search Terms

Two descriptors were used for each search, with the aid of the boolean indexer “OR” (“OU”, “O”). The descriptors correspond to the title of the instruments in English, Spanish and Portuguese: Conformity to Feminine Norms Inventory / Conformity to Masculine Norms Inventory; *Inventario de Conformidad con las Normas Femeninas/Inventario de Conformidad con las Normas Masculinas; Inventário de Conformidade com as Normas Femininas/ Inventário de Conformidade com as Normas Masculinas*. These descriptors were considered due the acronyms of the instruments (CFNI or CMNI) could restrict the number of studies found or refer to different instruments.

Inclusion Criteria

Inclusion criteria for the studies were: (1) to be peer-reviewed article; (2) describe an empirical research; (3) to involve validity evidence based on internal structure of at least one of the instruments; (4) be written in English, Spanish or Brazilian Portuguese.

Exclusion Criteria

To refine the search, studies were excluded if they reported: (1) book chapters; (2) literature reviews; (3) only other types of CFNI and CMNI validity evidences or only other instruments' validity evidences.

Collection Procedures

In the advanced search tool, the descriptors were added in separate fields and the boolean operator “OR” was selected. In the bases PUBMED and SCOPUS the search for the areas “title” or “summary” or “keywords” was added, whereas in PsycINFO the search for the area “summary” was added. Then, filters were used for the year of publication, selecting the period 2003 to 2018, and, when available, for the languages “English”, “Spanish” or “Portuguese”.

The search in the databases returned a total of 172 studies. After the added filters and selection criteria, remained 140 studies. Then, 30 duplicate studies were discarded. After reading the abstracts of 110 selected studies, 91 studies were excluded that did not involve validity evidence based on internal structure of the instruments. At the end of the research, 19 studies were used in the systematic review. Figure 1 presents the flowchart with the steps of the selection of the studies.

Data Analysis Procedures

Given Standards recommendations (2014), the publications were evaluated against the validity evidence based on internal structure collected. Rios and Wells (2014) argue that there are three basic aspects of the internal structure: dimensionality, measurement invariance and reliability. Dimensionality checks whether correlations between items support the test scores in which inferences are developed. The measurement invariance demonstrates whether the characteristics of the items are comparable between different groups. Reliability indicates the consistency of the test scores in repeated administrations. Thus, the publications will be presented considering the mentioned aspects. All procedures informed were performed from February 2018 to January 2019.

Results

All studies analysed were cross-sectional, in which data collection is performed in a single moment. The total number of participants ranged from 229 to 920. The average age of participants ranged from 19.00

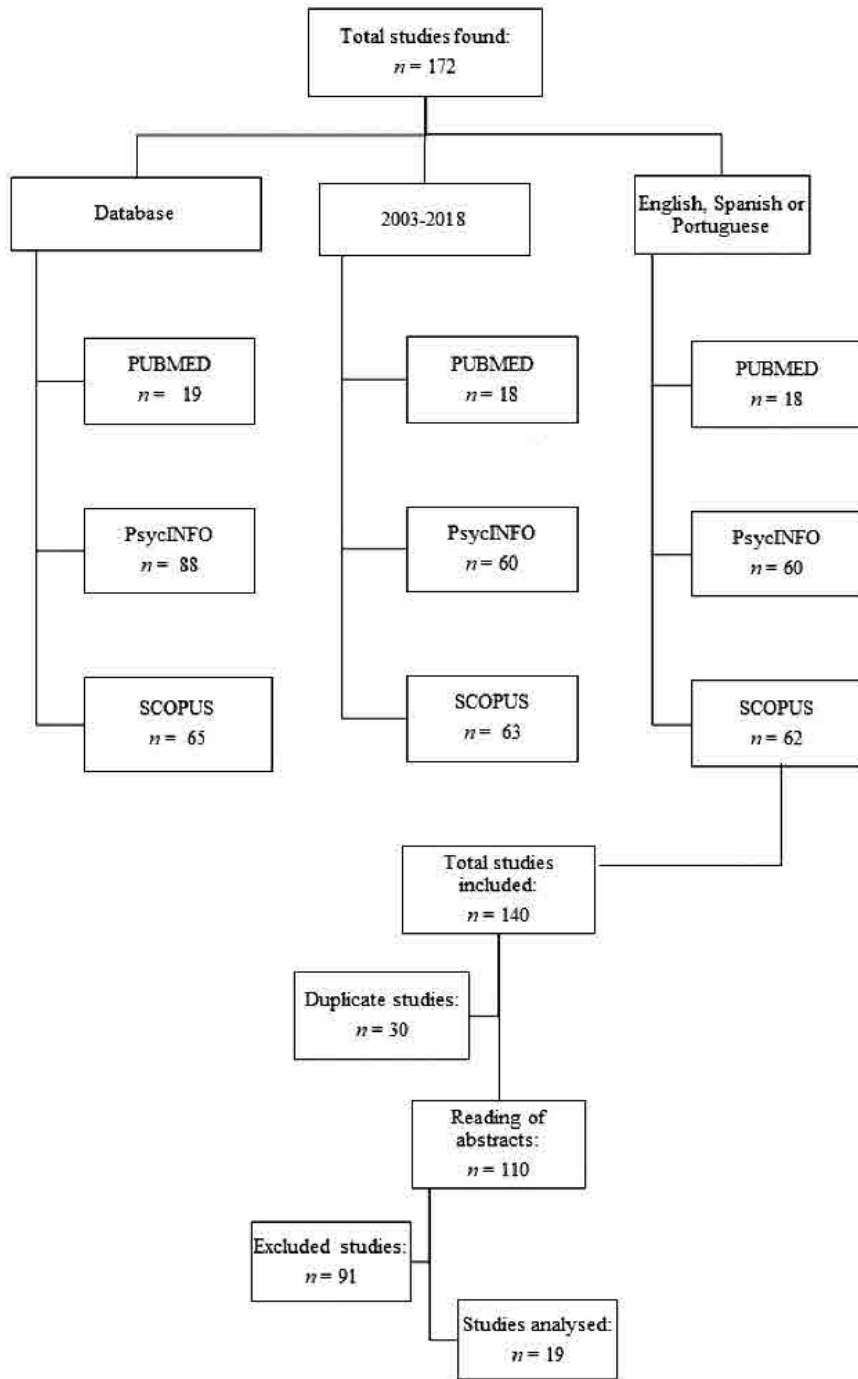


Figure 1. Flowchart with the steps of the selection of the systematic review studies.

years (SD = 1.65) to 46.74 years (SD = 16.86). In nine studies the sample consisted exclusively of university students. The studies found belong to researchers from countries of the American continent (United States =

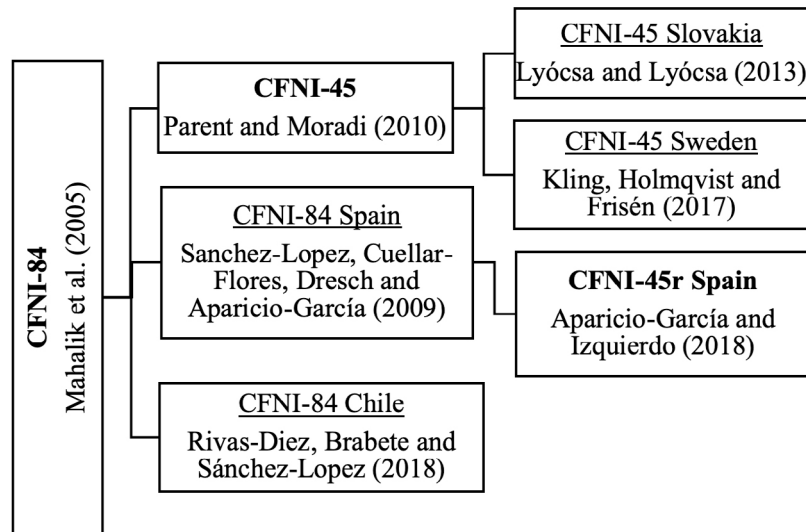
10), European (Slovakia = 1, Spain = 5, United Kingdom = 1, Sweden = 1) and Asian (China = 1).

The main objectives of the studies can be divided into: (1) describe the construction of complete or

reduced versions; (2) evaluate the factorial structure of CFNI and CMNI; (3) investigate how instrument factors are expressed in different groups. From the original publications by Mahalik et al. (2003, 2005) were found in the literature review six versions involving CFNI and six versions involving CMNI. The Figures 2 and 3 present a flowchart based on cross-cultural adaptations and new versions of instruments.

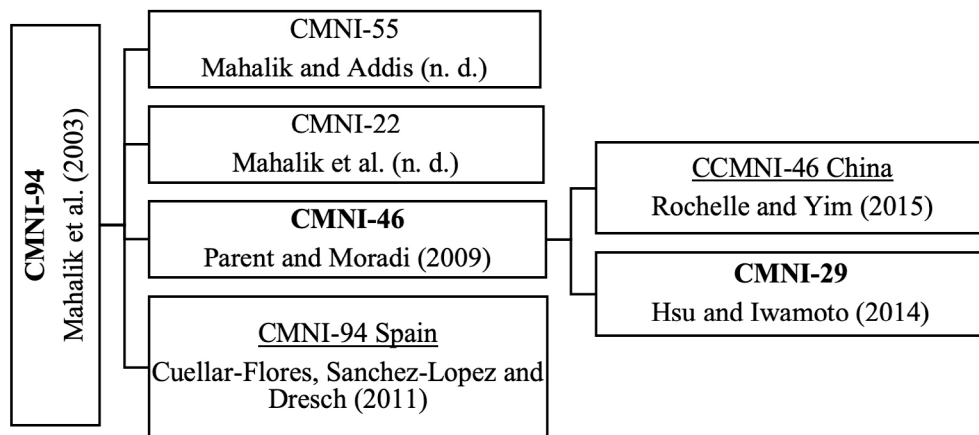
Dimensionality CFNI

Regarding the development of CFNI (Mahalik et al., 2005), the factor analysis performed with American women identified eight factors. The explained variance was 34.5%. In the CFNI adaptation to the Spain, the findings of Sánchez-López et al. (2009) supported the eight proposed factors, which explained 39.3% of the data variance. In another study with Spanish



Note. Bold refers to the original full or reduced version study. Underline refers to the study of cross-cultural adaptation.

Figure 2. CFNI versions found in the scientific literature



Note. Bold refers to the original full or reduced version study. Underline refers to the study of cross-cultural adaptation.

Figure 3. CMNI versions found in the scientific literature

participants, Sánchez-López and Cuéllar-Flores (2011) confirmed the original structure that explained 37.9% of the data variance. In Chile, the adaptation study by Rivas-Diez, Brabete and Sánchez-López (2018) found that the eight-factor structure explained 32.4% of the data variance.

The confirmatory factor analysis of CFNI performed by Parent and Moradi (2010) demonstrated a distinct structure to the one proposed by Mahalik et al. (2005). Data collected from Canadian women pointed out that it was more appropriate a nine-factor model. The results showed that the items gathered in the Nice in Relationships factor measure distinct behaviours (niceness and the importance of having and maintaining friendships) and should be treated as distinct factors. Furthermore, when considering the factor loadings of the items, the authors proposed a reduced version of the instrument, the Conformity to Feminine Norms Inventory-45 (CFNI-45), consisting of 45 items and nine factors (Parent & Moradi, 2010). In another study conducted by Parent and Moradi (2011b) with American women, CFNI-45 presented the following model fit indices: Comparative Fit Index (CFI) = 0.90, Root Mean Square Error of Approximation (RMSEA) = 0.04, and Standardized Root Mean Squared Residual (SRMR) = 0.06.

The adaptation of the CFNI-45 to the Slovak culture performed by Lyócsa and Lyócsa (2013) presented the following adjustment indices to the nine factor model: RMSEA = 0.04, CFI = 0.94 and Tucker-Lewis Index (TLI) = 0.95. However, due to lower standardized loadings, internal consistency reliability measures, and cross loadings observed, the authors argue that Sweet and Nice and Relational factors should be reconsidered. Thus, they present that only seven factors were well defined with the studied sample: Thinness, Care for Children, Sexual Fidelity, Domestic, Romantic Relationship, Invest in Appearance and Modesty (Lyócsa & Lyócsa, 2013).

When assessing CFNI-45 dimensionality, Aparicio-García and Alvarado-Izquierdo (2018) argue that the bifactor model presented the best fit to the data. Moreover, Sweet and Nice factor contributes more to the model as a general factor than as a specific (Aparicio-García & Alvarado-Izquierdo, 2018). This condition disagrees with the proposal presented by the authors of

the reduced version, who state that the nine correlated factors model is the most appropriate for the instrument (Parent & Moradi, 2011b).

Dimensionality CMNI

Concerning the development of CMNI (Mahalik et al., 2003), the factor analysis performed identified eleven factors. The explained variance was 44%. The adaptation of CMNI to Spain by Cuéllar-Flores, Sánchez-López, and Dresch (2011) confirmed the model presented by Mahalik et al. (2003), and the eleven factors explained 42.2% of the variance.

The factorial structure of two reduced versions of CMNI were studied by Owen (2011). Such versions have the eleven factors of the original instrument and consist of the two (CMNI-22) or five items (CMNI-55) of each factor that have the highest factor loading in the original study by Mahalik et al. (2003). With a sample of Americans, the CMNI-55 presented adequate adjustments, even though the CFI did not reach the cutoff point (0.88). However, the CMNI-22 obtained inadequate adjustments, indicating an inconsistent model that does not represent data satisfactorily (Owen, 2011).

The CMNI confirmatory factor analysis with Canadian men (Parent & Moradi, 2009) indicated that only nine of the eleven original factors were consistent, as well as items with weak factor loading. After that, the Conformity to Masculine Norms Inventory-46 (CMNI-46, Parent & Moradi, 2009), a shorter inventory of 46 items and nine factors was developed. The factorial structure of the CMNI-46 was further evaluated by Parent and Moradi (2011a) and the model fit indices were: CFI = 0.92, RMSEA = 0.04 CI: 0.04, 0.05, SRMR = 0.06.

When examining the factorial structure, loads, and psychometric properties of three measures used to assess masculinity, Alt, Lewis, Liu, Vilain, and Sánchez (2014) found that the model which presented the best fit was the CMNI-46 with nine correlated factors in a sample composed by homosexual and bisexual men. In another study, Hsu and Iwamoto (2014) proposed a reduced version with 29 items (Conformity to Masculine Norms Inventory-29, CMNI-29). The model fit indices to the CMNI-29 were: CFI = 0.93, RMSEA = 0.04 CI: 0.04, 0.04, SRMR = 0.04.

In China, the confirmatory factor analysis of CMNI-46 performed by Rochelle and Yim (2015) corroborates the multidimensional model presented by Parent and Moradi (2009). When the fit indices were verified, the CFI was inadequate (0.86).

Hammer, Heath and Vogel (2018) performed tests of different CMNI-46 models with a sample of college and general population men. The tests showed that the bifactor model was more adequate in the sample general population men, while the correlated factors model was more appropriate in the university men. Given this disagreement, the authors performed further analysis of auxiliary bifactor measures. These measurements supported the conceptualization of the CMNI-46 as a multidimensional instrument. Moreover, the authors emphasize that other versions of the instrument may present different dimensionality, which would impact the reliability of assigning a total score.

Measure Invariance CFNI

The CFNI assesses perceived norms within a specific culture. Thus, it is important to verify if the instrument holds the internal structural across different subgroups. Parent and Moradi (2011b) demonstrate the partial metric invariance of CFNI-45 between white (dominant group) and non-white participants (non-dominant group). The results indicate that the factor analysis and item loadings of the CFNI-45 maintained a similar structure regardless of the group to which the subject belonged, except for the items: “I get ready in the morning without looking in the mirror very much” (Invest in Appearance), “My life plans of not relying on my romantic relationship” (Romantic Relationship) and “I don’t feel guilty if I lose contact with a friend” (Nice in Relationships).

Measure Invariance CMNI

The CMNI-46 was applied to men and women, and the results indicated the partial metric invariance of the instrument in relation to the respondent’s gender. Although it has a similar structure, gender differences were found in the five items: “I would only have sex if I was in a committed relationship” (Playboy), “Being thought of a gay is not a bad thing” (Disdain for

Homosexuals), “I hate asking for help” (Self-Reliance), “Winning is not my first priority” (Winning) and “I hate it when people ask me to talk about my feelings” (Emotional Control) (Parent & Smiler, 2013).

The partial metric invariance of CMNI-46 was verified between the groups of white American and Asian American men, indicating that the understanding of what is masculine can be distinct from the considered ethnicity (Hsu & Iwamoto, 2014). Winning and Power Over Women factors presented different associations with the other instrument factors among the evaluated participants.

Reliability CFNI

The internal consistency of CFNI in the original study by Mahalik et al. (2005) was $\alpha = 0.88$ for the total CFNI score, ranging from $\alpha = 0.77$ (Romantic Relationship) to $\alpha = 0.92$ (Care for Children). The temporal stability of the instrument in this study was $r = 0.94$ for the total CFNI score over a period of two to three weeks.

In Spain, the internal consistency in the study by Sánchez-López et al. (2009) was $\alpha = 0.87$ for the total CFNI score, ranging from $\alpha = 0.64$ (Romantic Relationship) to $\alpha = 0.86$ (Care for Children). In another study, Sánchez-López and Cuéllar-Flores (2011) reported that the internal consistency for the total CFNI score was $\alpha = 0.86$, ranging from $\alpha = 0.72$ (Nice in Relationships and Romantic Relationship) to $\alpha = 0.91$ (Care for Children). In Chile (Rivas-Diez, Brabete & Sánchez-López, 2018), the internal consistency was $\alpha = 0.83$ for the total score, ranging from $\alpha = 0.52$ (Romantic Relationship) to $\alpha = 0.89$ (Care for Children).

The internal consistency of CFNI-45 in the original study by Parent and Moradi (2010) was $\alpha = 0.79$, ranging from $\alpha = 0.68$ (Sweet and Nice) to $\alpha = 0.89$ (Care for Children). In Slovakia (Lyócsa & Lyócsa, 2013), the internal consistency of CFNI-45 ranged from $\alpha = 0.48$ (Relational) to $\alpha = 0.90$ (Thinness). Finally, in Sweden (Kling, Holmqvist Gattario, & Frisén, 2017), the internal consistency of CFNI-45 ranged from $\alpha = 0.57$ (Sweet and Nice) to $\alpha = 0.90$ (Thinness). The coefficient for the total score has not been reported in Slovak and Swedish studies above.

Reliability CMNI

The internal consistency of CMNI in the original study by Mahalik et al. (2003) was $\alpha = 0.94$ for the total score, ranging from $\alpha = 0.72$ (Pursuit of Status) to $\alpha = 0.91$ (Emotional Control). The temporal stability of the total CMNI score was $r = 0.95$ over a period of two to three weeks (Mahalik et al., 2003). In Spain, the study by Cuéllar-Flores et al. (2011) demonstrated that the internal consistency of the CMNI was $\alpha = 0.89$ for the total score, ranging from $\alpha = 0.64$ (Pursuit of Status) to $\alpha = 0.81$ (Disdain for Homosexuals).

The internal consistency of CMNI-46 in the original study by Parent and Moradi (2009) was $\alpha = 0.88$ for the total score, ranging from $\alpha = 0.77$ (Primacy of Work) to $\alpha = 0.91$ (Disdain for Homosexuals). In another study by Parent and Moradi (2011a) the internal consistency of the instrument was $\alpha = 0.85$ for the total score, ranging from $\alpha = 0.78$ (Winning) to $\alpha = 0.89$ (Emotional Control and Disdain for Homosexuals).

In the study by Alt et al. (2014), the internal consistency was $\alpha = 0.86$ for the total score, ranging from $\alpha = 0.77$ (Power Over Women) to $\alpha = 0.90$ (Emotional Control). In China (Rochelle & Yim, 2015), the internal consistency was $\alpha = 0.79$, ranging from $\alpha = 0.60$ (Winning) to $\alpha = 0.80$ (Disdain for Homosexuals).

With a sample of college and general population men (Hammer et al., 2018), the internal consistency of CMNI-46 was $\alpha = 0.91$ (college men) and $\alpha = 0.87$ (general population men). The coefficients ranged from $\alpha = 0.82$ (Primacy of Work and Playboy) to $\alpha = 0.93$ (Disdain for Homosexuals) in the sample of college men and ranged from $\alpha = 0.80$ (Power Over Women and Risk-Taking) to $\alpha = 0.90$ (Emotional Control and Disdain for Homosexuals) in the sample of general population men.

Discussion

In this study, a systematic review was conducted to investigate the validity evidence based on internal structure of CFNI and CMNI. The findings provide psychometric support regarding the evidences of internal structure validity, confirm the proposed

multidimensionality by the instruments and demonstrates the model application in several cultures.

Concerning dimensionality, confirmatory factor analyses were performed with the instruments. To the CFNI, the structure composed of 84 items distributed in eight factors was corroborated with Spanish (Sánchez-López et al., 2009) and Chilean (Rivas-Diez et al., 2018) samples, but no model fit analysis was verified in these studies. The structure composed of 45 items distributed in nine factors was corroborated with Slovak (Lyócsa & Lyócsa, 2013), Spanish (Aparicio-García & Alvarado-Izquierdo, 2018) and Swedish (Kling et al., 2017) samples.

Despite this, the instrument did not obtain adequate CFI indices in some studies. The studies by Aparicio-García and Alvarado-Izquierdo (2018), and Parent and Moradi (2010, 2011b) found that the original CFNI-45 structure with nine correlated factors obtained CFI indices equal to 0.93, 0.77 and 0.90, respectively. Hu and Bentler (1999) recommend a minimum cut off point of 0.95 for the CFI index, in order to decrease type II error rates. Specifically, for the CFNI-45 the goodness of fit beneath the cut off must due the item complexity: some items must present non modelled cross loadings, as well as residual covariance. Increase the model complexity could boost the fit index, however it would likely reduce the generalizability.

Confirmatory factor analyses are also found for the model presented in the original (94 items) and reduced (46 items) versions of the CMNI. With a sample of Spanish men, the structure of eleven correlated factors was corroborated in the study by Cuéllar-Flores et al. (2011). Regarding the CMNI-46, the structure of nine correlated factors found satisfactory adjustments (Alt et al., 2014; Parent & Moradi, 2011a; Rochelle & Yim, 2015), except for the CFI indices, condition also reported in the female version.

Only three studies were found to verify the measure invariance of the instruments. The CMNI-46 has two studies reporting partial metric invariance in relation to respondent gender and ethnicity (Hsu & Iwamoto, 2014; Parent & Smiler, 2013). In the case of CFNI-45 just one study reporting partial metric invariance in relation to white and non-white people (Parent & Moradi, 2011b). The measure invariance is fundamental for an instrument to be considered fair.

The Standards (AERA et al., 2014) recommend studies involving this type of validity evidence and emphasize that psychometric justice is related to the measure invariance among the various subgroups in a sample.

CFNI and CMNI internal consistency indices were evaluated in 17 studies using Cronbach's alpha coefficient (Cronbach, 1951). Although this indicator is one of the most widely used, is not limitless and can be influenced by the number of instrument items or sample size (Nunnally & Bernstein, 1994; Rios & Wells, 2014). There are other indicators of internal consistency available, which could enrich the study proposals and provide further evidence of the validity of the internal structure of the instruments evaluated.

The lowest internal consistency indices of CFNI were found in Modesty, Nice in Relationships, Relational (CFNI-45), Sweet and Nice (CFNI-45), Romantic Relationship and Domestic factors. The highest internal consistency indices were found in Care for Children and Thinness factors.

Modesty factor is questioned among researchers. Parent and Moradi (2010) report on care in the return of the results of this factor in the clinic, since both modesty and docility are desirable aspects for women in some societies to disempower them. Lyócsa and Lyócsa (2013) reconsider the use of the factor as a female norm, because they found that it obtained low correlations, sometimes negative with the other factors, besides contributing little to the total score in the models tested.

About the CMNI, the Pursuit of Status and Dominance factors had the lowest internal consistency indices. These factors were in the original version of the instrument developed by Mahalik et al. (2003) and were not included in the CMNI-46 (Parent & Moradi, 2009). The highest internal consistency indices were found in the Emotional Control and Disdain for Homosexuals factors.

The reduced versions of CFNI and CMNI presented by Parent and Moradi (2009, 2010), in which the instruments contain, respectively, 45 and 46 items linked to nine factors contains the largest number of publications involving their psychometric properties. Only one study reporting factorial structure and internal consistency analyses of the reduced 22-item and 55-item CMNI versions was found. The lack of publications

about the psychometric properties of these versions precludes proper use and deeper analysis of the model.

Most validity evidence based on internal structure was collected from samples of predominantly white and heterosexual college participants. Only one study analysed (Alt et al., 2014) was composed of gay men. As previously noted, it is necessary to check how the underlying structure of the instruments behaves under different conditions. The use of samples from men and women from the general population, from other ethnicities, with different socioeconomic levels and/or sexual minorities will enrich the analyses involving validity evidence of these instruments. The instruments do not have any study showing validity evidence in Portuguese-speaking countries. Therefore, it is recommended that future studies seek to collect validity evidence in these countries as well as in Latin American countries.

Other evidence of instrument validity is presented in some studies, such as convergent and discriminant validity (Alt et al., 2014; Mahalik et al., 2003, 2005; Parent & Moradi, 2011a, 2011b). However, as it escapes the objectives of this systematic review, these types of validity evidence have not been reported, the reader may refer to such studies.

Limitations and Recommendations

Data obtained from the systematic review revealed that the CFNI and CMNI provide satisfactory validity evidence based on internal structure. The analysis made in the review of the studies consisted of a qualitative approach, as it was not possible to access the database of publications, this could generate a richer evaluation (meta-analysis). It is noteworthy that only peer-reviewed articles investigating the internal structure of the instruments were evaluated. Thus, some of the studies that were disregarded in this review, although they could discuss aspects involving this type of validity, were not evaluated.

The study by Hammer et al. (2018) suggested that the calculation of a general CMNI-46 score is inappropriate. In contrast, Aparicio-García and Alvarado-Izquierdo (2018) reported that CFNI-45, upon a bifactor model, allows the attribution of a general score. There is inconsistency in the underlying structure of the instruments, as both were developed to

assess female and male norms from a multidimensional perspective. Thus, it is important that researchers test different structural models to verify the best indicators with the studied sample.

References

- AERA, APA, & NCME. (2014). *Standards for Educational and Psychological Testing*. Washington, DC: American Educational Research Association.
- Alt, M., Lewis, A. M., Liu, W. M., Vilain, E., & Sánchez, F. J. (2014). On the Validity of Popular Masculinity Rating Scales with Gay Men. *Archives of Sexual Behavior*, 43(8), 1547-1557. doi: 10.1007/s10508-014-0363-6
- Aparicio-García, M. E., & Alvarado-Izquierdo, J. M. (2018). Is there a “Conformity to Feminine Norms” Construct? A Bifactor Analysis of two Short Versions of Conformity to Feminine Norms Inventory. *Current Psychology*, 1, 1-11. doi: 10.1007/s12144-018-9815-9
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297-334. doi: 10.1007/BF02310555
- Cuéllar-Flores, I., Sánchez-López, M. D. P., & Dresch, V. (2011). El inventario de conformidad con las normas de género masculinas (CMNI) en la población española. *Anales de Psicología*, 27(1), 170-178. Retrieved from https://www.researchgate.net/publication/237025460_El_Inventario_de_Conformidad_con_las_Normas_de_Genero_Masculinas_CMNI_en_la_poblacion_espanola
- Fernández, J. (2010). El sexo y el género: Dos dominios científicos diferentes que debieran ser clarificados. *Psicothema*, 22(2), 256-262. Retrieved from <http://www.psicothema.com/pdf/3724.pdf>
- Grant, M. J., & Booth, A. (2009). A typology of reviews: an analysis of 14 review types and associated methodologies. *Health Information and Libraries Journal*, 26, 91-108. doi: 10.1111/j.1471-1842.2009.00848.x
- Hammer, J. H., Heath, P. J., & Vogel, D. L. (2018). Fate of the Total Score: Dimensionality of the Conformity to Masculine Norms Inventory-46 (CMNI-46). *Psychology of Men & Masculinity*, Advance online publication. doi: 10.1037/men0000147
- Helgeson, V. S. (2017). *Psychology of Gender*. (5th ed.). New York: Routledge.
- Hsu, K., & Iwamoto, D. K. (2014). Testing for measurement invariance in the conformity to masculine norms-46 across white and Asian American college men: Development and validity of the CMNI-29. *Psychology of Men and Masculinity*, 15(4), 397-406. doi: 10.1037/a0034548
- Hu, L. T., & Bentler, P. (1999). Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria Versus New Alternatives. *Structural Equation Modeling*, 6(1), 1-55. doi: 10.1080/10705519909540118
- Kling, J., Holmqvist Gattario, K., & Frisé, A. (2017). Swedish women’s perceptions of and conformity to feminine norms. *Scandinavian Journal of Psychology*, 238-248. doi: 10.1111/sjop.12361
- Lyócsa, I., & Lyócsa, S. (2013). Confirmatory factor analysis of the Abbreviated Conformity to Feminine Norms Inventory. *Social Work Research*, 37(4), 414-422. doi: 10.1093/swr/svt034
- Mahalik, J. R., Locke, B. D., Ludlow, L. H., Diemer, M. A., Gottfried, M., Scott, R. P. J., & Freitas, G. (2003). Development of the Conformity to Masculine Norms Inventory. *Psychology of Men & Masculinity*, 4(1), 3-25. doi: 10.1037/1524-9220.4.1.3
- Mahalik, J. R., Morray, E. B., Coonerty-Femiano, A., Ludlow, L. H., Slattery, S. M., & Smiler, A. P. (2005). Development of the Conformity to Feminine Norms Inventory. *Sex Roles*, 52(7-8), 417-435. doi: 10.1007/s11199-005-3709-7
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric Theory* (3rd ed.). New York: McGraw-Hill.
- Owen, J. (2011). Assessing the Factor Structures of the 55- and 22-Item Versions of the Conformity to Masculine Norms Inventory. *American Journal of Men’s Health*, 5(2), 118-128. doi: 10.1177/1557988310363817
- Parent, M. C., & Moradi, B. (2009). Confirmatory factor analysis of the Conformity to Masculine Norms Inventory and development of the Conformity to Masculine Norms Inventory-46. *Psychology of*

- Men & Masculinity*, 10(3), 175-189. doi: 10.1037/a0015481
- Parent, M. C., & Moradi, B. (2010). Confirmatory factor analysis of the Conformity to Feminine Norms Inventory and development of an abbreviated version: the CFNI- 45. *Psychology of Women Quarterly*, 34, 97-109. doi: 10.1037/a0015481
- Parent, M. C., & Moradi, B. (2011a). An abbreviated tool for assessing conformity to masculine norms: Psychometric properties of the Conformity to Masculine Norms Inventory-46. *Psychology of Men & Masculinity*, 12(4), 339-353. doi: 10.1037/a0021904
- Parent, M. C., & Moradi, B. (2011b). An abbreviated tool for assessing feminine norm conformity: Psychometric properties of the Conformity to Feminine Norms Inventory-45. *Psychological Assessment*, 23(4), 958-969. doi: 10.1037/a0024082
- Parent, M. C., & Smiler, A. P. (2013). Metric invariance of the Conformity to Masculine Norms Inventory-46 among women and men. *Psychology of Men & Masculinity*, 14(3), 324-328. doi: 10.1037/a0027642
- Rios, J., & Wells, C. (2014). Validity evidence based on internal structure. *Psicothema*, 26(1), 108-116. doi: 10.7334/psicothema2013.260
- Rivas-Diez, R., Brabete, A. C., & Sánchez-López, M. D. P. (2018). Evaluación de la Variable Género: CFNI en Mujeres Chilenas. *Revista Iberoamericana de Diagnóstico y Evaluación - e Avaliação Psicológica*, 3(48), 117-126. doi: 10.21865/RIDEP48.3.10
- Rochelle, T. L., & Yim, K. H. (2015). Assessing the Factor Structure of the Chinese Conformity to Masculine Norms Inventory. *The Journal of Psychology*, 149(1), 29-41. doi: 10.1080/00223980.2013.837023
- Sampaio, R. F., & Mancini, M. C. (2007). Estudos de revisão sistemática: um guia para síntese criteriosa da evidência científica. *Revista Brasileira de Fisioterapia*, 11(1), 83-89. Retrieved from http://www.scielo.br/scielo.php?script=sci_arttext&pid=S14133555200700010_0013
- Sánchez-López, M. D. P., Cuéllar-Flores, I., Dresch, V., & Aparicio-García, M. E. (2009). Conformity to Feminine Gender Norms in the Spanish Population. *Social Behavior and Personality*, 37(9), 1171-1186. doi: 10.2224/sbp.2009.37.9.1171
- Sánchez-López, M. D. P., & Cuéllar-Flores, I. (2011). Comparison of Feminine Gender Norms among Spanish and American College Women. *Psychology*, 02(04), 300-306. doi: 10.4236/psych.2011.24047
- Scott, J. (1995). Género: uma categoria útil de análise histórica. *Educação & Realidade*, 20(2), 71-99. Retrieved from <https://seer.ufrgs.br/educacaoerealidade/article/view/71721/40667>

Recebido em: 11/11/2019
 Revisado em: 15/02/2020
 Aprovado em: 28/03/2020

Acknowledgments:

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brazil (CAPES) - Finance Code 001.

Valentini has a research grant from CNPq (PQ 2).

About the authors:

Julia Caciano da Silva

Doutoranda do Programa de Pós-Graduação Stricto Sensu em Psicologia da Universidade Salgado de Oliveira, Niterói-RJ, Brasil.

ORCID: <https://orcid.org/0000-0003-4445-1113>

E-mail: juliacaciano@gmail.com

Felipe Valentini

Doutor em Psicologia, Docente do Programa de Pós-Graduação Stricto Sensu em Psicologia da Universidade São Francisco, Campinas-SP, Brasil.

ORCID: <https://orcid.org/0000-0002-0198-0958>

E-mail: valentini.felipe@gmail.com

Clarissa Pinto Pizarro de Freitas

Doutora em Psicologia, Docente do Programa de Pós-Graduação Stricto Sensu em Psicologia Clínica da Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro-RJ, Brasil.

ORCID: <https://orcid.org/0000-0002-2274-8728>

E-mail: freitas.cpp@gmail.com

Contact:

Julia Caciano da Silva

Estrada Frei Orlando 150A casa 02, Piratininga

Niterói-RJ, Brasil

CEP: 24350-200

Telefone: (+55 21) 99713-3877