

What Is the Measure of a Man? A Systematic Scoping Review of 50 Years of Masculinities Scales

Men and Masculinities

2026, Vol. 0(0) 1–29

© The Author(s) 2026



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/1097184X261459614

journals.sagepub.com/home/jmm

Peter Fisher¹ , Brian Heilman¹, Ryon C. McDermott²,
Jasleen Chhabra^{3,4,*}, Ruben Benakovic^{3,4,*},
Kieran O’Gorman^{3,4,*} , Simon Rice^{3,4,5,6}, and Zac Seidler^{3,4,5}

Abstract

The dominant conceptual narrative in research on and measurement of masculinities has emphasized that greater adherence to traditional masculine norms is associated with a range of harmful outcomes. The primacy of this deficit-focused may have shaped the development of measures of masculinities in ways that underrepresent adaptive, relational, or health-promoting dimensions of masculine identity, even as it remains an open empirical question whether such dimensions exist in ways that can be meaningfully defined and measured. This systematic review uses a masculinities theory lens to offer a conceptual and methodological analysis of 65 measures of masculinity published between 1974 and 2024 with a focus on five key issues: the primacy of deficit-based approaches; limited attention to relationality or social context; lack of cultural attunement; sample homogeneity; and inconsistent psychometric standards. Results indicate that 78 percent of measures adopted a more deficit-based approach, while only 22 percent incorporated strengths-based elements. Most measures

¹Movember Institute of Men’s Health, Santa Monica, CA, USA

²Department of Professional Studies, University of South Alabama, Mobile, AL, USA

³Centre for Youth Mental Health, The University of Melbourne, Melbourne, VIC, Australia

⁴Orygen, Melbourne, VIC, Australia

⁵Movember Institute of Men’s Health, Melbourne, VIC, Australia

⁶Centre for Mental Health and Community Wellbeing, Melbourne School of Population and Global Health, University of Melbourne, Melbourne, VIC, Australia

*Authors contributed equally.

Corresponding author:

Peter Fisher, Movember Institute of Men’s Health, 1616 17th St., Santa Monica, CA 90404, USA.

Email: peter.fisher@movember.com

demonstrated limited relational focus and cultural attunement, and 66 percent relied on samples lacking demographic diversity. Psychometric findings were mixed: internal validity and reliability were generally adequate, but only 9 percent of measures demonstrated strong evidence for external validity. These findings highlight the need for a more balanced measurement landscape that integrates strengths-based approaches, improves contextual precision and relational framing, addresses cultural considerations using diverse samples, and strengthens evidence for generalizability. Future directions for developing the next generation of masculinities measures are discussed.

Keywords

masculinity scales, masculinities, measurement, strengths-based, systematic review

Introduction

Measures of adherence to masculine norms have evolved considerably over the past five decades. In the 1970s, the Bem Sex Role Inventory (BSRI; [Bem 1974](#)) and Personal Attributes Questionnaire (PAQ; [Spence et al. 1974](#)) marked an important departure from earlier measures which cast masculinity and femininity as poles on a single spectrum, instead allowing individuals to score separately on masculine and feminine dimensions. In the 1990s and early 2000s, influential measures such as the Male Role Norms Inventory (MRNI; [Levant et al. 1992](#)) and the Conformity to Masculine Norms Inventory (CMNI; [Mahalik et al. 2003](#)) were developed to assess adherence to culturally dominant masculinity norms. These tools emerged from Pleck's theory of gender-role strain ([1981, 1995](#)), which describes the difficulties of attempting to conform to a gender role that includes both dysfunctional and contradictory aspects, and aimed to link traits thought to be commonly associated with masculinity (such as emotional restriction, dominance, and self-reliance) to health and psychological outcomes.

Measures of masculinities have continued to advance, with some scholars seeking greater contextual specificity, for instance exploring young men's masculinity ideology in the context of peer relationships ([Chu et al. 2005](#); [Reigeluth and Addis 2016](#)) or measuring men's recollections of their fathers' expectations of masculine conformity ([Levant et al. 2018](#)). Others have opted for qualitative and mixed methods approaches, allowing respondents to define masculinity for themselves ([Kachel et al. 2016](#); [Wong et al. 2011](#)), or applied existing measures to new geographic contexts ([Equimundo 2022](#)).

As [Smiler \(2004\)](#) notes, measures of masculinities have evolved alongside shifting theories of gender. Early work explored masculinity as a set of personality traits; later frameworks emphasized adherence to masculine norms and then gender role strain. Across these shifts, certain trends have nonetheless remained the same. Among the most recognizable hallmarks of this literature is an approach whereby higher scores on

masculinities measures are hypothesized to be associated with worse – not better – outcomes (O’Neil 2008, 2015; Smiler 2004; Wong et al. 2017). This orientation is well justified. It reflects a foundational understanding of masculinity not merely as a set of neutral traits but as a system of power that privileges some men while harming others as well as women and people of other gender identities. This approach has also produced consistent and valuable empirical findings. Scores on the MRNI and CMNI, as well as other instruments taking a similar approach, have been associated with increased depressive symptoms, reduced help-seeking, substance misuse, emotional restriction, and risky health behaviors (Gerdes et al. 2018; O’Neil 2008; Levant et al. 2011; Levant et al. 2006). These and similar measures have also been applied in diverse domains such as interpersonal violence (e.g., Gage and Lease 2021), sexual aggression (e.g., Cole et al. 2020), reduced mask-wearing during the COVID-19 pandemic (e.g., Levant et al. 2022; Mahalik et al. 2022), and other such outcomes.

The knowledge base resulting from this work, establishing a wide range of risks and harms associated with masculine norm adherence, is valuable. However, focusing so centrally on deficits – negative outcomes and behaviors as the primary expectation – has also opened the field to critique. A recent content analysis of studies published in *Psychology of Men and Masculinities* over an 18-year period (including but not limited to studies using masculinities measures), for instance, found that only 15 percent had a positive focus (Cole et al. 2021). If such a negative or “deficit-based” focus becomes the only lens through which links between masculinity and health are examined, critics argue, the field risks systematically overlooking positive, health-promoting, or strengths-based dimensions of men’s lives (Englar-Carlson and Kiselica 2013).

It remains an open empirical question whether protective, prosocial, and/or health-promoting dimensions of masculine identity exist in ways that can be consistently defined, measured, and linked to positive outcomes. Some scholars argue that efforts to define “strengths-based” or “healthy” masculinity risk reinforcing the very structures they aim to reform, and advocate instead for deconstructing masculinity altogether as a system of power and exclusion (Beasley 2015; Connell 1995; Waling 2019). Others contend that masculinity should be actively reconstructed in ways that support well-being, equality, and connection (Kiselica and Englar-Carlson 2010). The Positive Psychology Positive Masculinity Paradigm (PPPM), for example, has advanced research in clinical practice with boys and men based on tapping into strengths tied to male identity (Kiselica and Englar-Carlson 2010).

Still others have emphasized the need for more pragmatic, context-sensitive approaches that draw on both reconstructive and deconstructive strategies depending on the setting and goals (see Addis and Mahalik 2003). More recent work continues this line of thinking by exploring how strengths-based framings may offer one productive tool among many for engaging men in ways that are resonant, values-aligned, and critically aware (Bridges and Pascoe 2014; Hoffmann and Addis 2024). It is also possible that strengths-based approaches offer more resonant pathways for health campaigns, public messaging, and interventions than approaches rooted exclusively in narratives of harm, which can put some men on the defensive (Wells et al. 2024).

Another issue which applies to some, if not all, measures of masculinities is a lack of clarity around what a respondent's score reflects (Thompson and Bennett 2015). Devoid of any clear relational framing or instructions to the respondent, a prototypical item stating "a real man should [norm statement]" with Likert-scale response options, can obscure what agreement actually signals. In their systematic review on the topic, Thompson and Bennett (2015) highlight this persistent ambiguity: "Researchers could interpret respondents' scores on the same masculinity ideology measure as evidence of people's agreement that the cultural norms exist, as evidence that the norms are desirable, and/or as evidence that the norms are personal values." (p. 443). Without greater precision about men's relationship to the norms being measured - and particularly, without attention to the relationality of those views, such as the social contexts of personal values, peer influences, or broader cultural messaging - research risks misinterpreting the underlying drivers of men's gendered behavior. Additionally, observations that most men seem to disagree with items assessing traditional masculine norms, as evidenced by item means below the scale midpoint, raise further concerns about the interpretation of such measures (e.g., Levant 2015; McDermott et al. 2017).

A lack of attunement to cultural norms in masculinities measures may also risk reinforcing narrow, static understandings of masculinities that do not translate meaningfully across communities. Measures very often reference hegemonic, White, Western, and heterosexual understandings of manhood (Levant and Richmond 2016). Many tools were developed with little attention to variation across race, ethnicity, class, geography, or lived experiences of non-dominant groups (Griffith et al. 2012). Although examples such as the Black Men's Experiences Scale and the Inventory of Subjective Masculine Experiences have sought to address these gaps through culturally specific adaptation or parallel scale design (Bowleg et al. 2016; Wong et al. 2020), the field still lacks tools designed to measure masculinities in broad, culturally-attuned ways.

Closely related to the question of cultural attunement is the issue of sample diversity. As highlighted by Gerdes et al. (2018) and Griffith et al. (2012), masculinity research has historically relied heavily on White, U.S.-based undergraduate samples. This threatens the generalizability of findings, especially given that experiences of masculinity often shift across life stages from adolescence to young adulthood to midlife; are shaped differently for men coming of age during periods of economic or political upheaval versus relative stability; and vary across cultural contexts (Meadows and Davidson 2007; Whorley and Addis 2006; Wong et al. 2020).

Finally, attention to the psychometric properties of masculinities measures is essential to ensuring that research findings are both valid and reliable. Although not a field-specific concern, masculinities measures risk offering potentially misguided insights into men's lived experiences without careful attention to these foundations - especially the generalizability and transportability of otherwise rigorously validated measures.

Aims and Objectives

Previous reviews (e.g., Thompson and Bennett 2015; Thompson and Pleck 1995) have played a foundational role in tracing the development of masculinities measures and identifying important areas for growth. Notably, Thompson and Bennett (2015) called for a “third generation” of measures that better capture men’s diverse life stages, social contexts and relationships, and cultural identities, prompting some of the active debates presented above. In the years since this review was published, both critical discourse around masculinities measurement and the number of new measures themselves have expanded considerably. The present review aims to provide a snapshot of this dynamic, multi-disciplinary field in which debates over how best to define and measure masculinities remain active and unresolved. Rather than settling those debates, the goal is to ground them in a comprehensive, theory-driven synthesis of widely used masculinities measures from the past 50 years. Following the topics outlined in this introduction, the five key areas of interest covered in the review are: (1) approach (more deficit-based vs more strengths-based); (2) relational focus; (3) cultural attunement; (4) sample diversity; and (5) psychometric properties of existing measures.

Methods

Protocol and Registration

The search process, eligibility criteria, screening procedure, coding analysis, and data synthesis were preregistered on Open Science Framework (see <https://osf.io/2cm7g>).

Data Sources and Search Strategy

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews (PRISMA - Scr) guideline was adhered to (Page et al. 2021; Tricco et al. 2018). The search was constrained to papers publicly available in English that were published in the last 50 years (1974–2024). Searches were conducted across three electronic databases in August of 2024: PsycINFO, PsycArticles, and the EMERGE gender empowerment measure repository (Raj et al. 2017). Searches were conducted on the PsycINFO/PsycArticles databases using the following query strings tabulated together using the boolean OR search function: *masculinity scale*, *masculinity measure*, *masculinities scales*, *masculinities measure*, *men’s health scale*, *men’s health measure*, *male identity measure*, *male gender role scale*, *male gender role measure*, *masculine gender role scale*, *masculine gender role measure*, *male gender norms scale*, *masculine gender norms scale*. Due to the different search functionality on the EMERGE repository, these search terms were inputted individually with no text entered in the dimension or country filters.

The ascendancy approach was applied to three reviews of masculinities measures (Thompson and Bennett 2015; Thompson et al. 1992; Thompson and Pleck 1995). The

measures and references in these reviews were examined to identify eligible measures through citation tracing.

Screening and Selection

Search results were merged with papers identified using the ascendancy approach and duplicates were removed. The lead author (PF) screened titles and abstracts using predetermined eligibility criteria and extracted data (see Table 1). The inclusion and exclusion criteria were set to capture a snapshot of the most widely used masculinities measures that researchers have employed. Rather than attempt to categorically include every single measure, these criteria aim to encompass the most common scales that researchers familiar and not familiar with the field tend to use for broad samples of men (excluding scales focused solely on specific sub-populations).

For measures with multiple derivative, abbreviated, adapted, or revised versions, only the original was included alongside the single most-cited revision (e.g., the 94-item CMNI and the CMNI-11). This allowed for comparison without skewing results by counting the same scale multiple times but also constrained the review by excluding

Table 1. Inclusion and exclusion criteria used in the screening process

	Include	Exclude
Participants	Samples of only men, only women, or mixed gender samples	Measure focuses on niche clinical or cultural population (e.g., cancer survivors or immigrant men)
Measures	Develops and validates a measure of masculine norms, masculine behavior, masculine ideology, masculinity conformity, gender-role conflict, or gender-role strain	Measure assesses masculinity as subscale within a broader measure
	Paper includes or links to the measure items, or the measure items for the scale can be identified using a secondary source	Measure only assesses sexual orientation or sexuality
	Paper reports relevant psychometric information: Internal validity, external validity, reliability	
	For derivative, revised, or abbreviated scales, only the most cited version was included along with the original paper	
Study type	Paper includes measure with at least one quantitative component Published in a peer-reviewed journal Published in English Has been cited at least 10 times	Purely qualitative measures, diagnostic tools, or assessment rubrics Dissertations and student theses Published before 1974

subsequent studies that linked measures to wider outcomes. The review was limited to scales with at least 10 citations to ensure each measure had been meaningfully used or engaged with by other researchers. See [Appendix A](#) for the full list of scales.

Extracted data included: Title, authors, publication year, number of citations, scale type (quantitative, qualitative, or mixed methods), sample size and nationality, psychometric information, and whether the target population was adults or adolescents. Extracted data were checked for accuracy by three reviewers (JS, RB, KO). Thematic content analysis was used to code studies across five conceptual categories of interest. The authors used a two-stage screening approach. First, the lead author screened papers retrieved from databases by reviewing the title and abstract. For papers that met the inclusion criteria, the full paper text and scale items were reviewed by trained researchers (JS, RB, KO). The authors did not blind any fields during the screening process.

Coding Analysis Procedure

The authors developed coding frameworks for the five areas of interest, as follows:

Approach (More Deficit-Based vs. More Strengths-Based) (0–1)

Does the scale development/items take a deficit-based approach problematizing masculinity and/or highlighting negative outcomes tied to higher scores of masculinity? Or an alternate approach highlighting at least some positive or strengths-based elements of masculinity?

0 = More deficit-based: Higher scores on the measure are generally associated with worse outcomes for men, items focus on negative aspects of masculinity, items reinforce traditional constraining and/or stereotypical views of masculinity, can include contradictory or impossible norm expectations.

Item Examples: “A real man never shows emotion”, “Men should never be responsible for domestic chores”

Methods Example: “Men with high scores on the work and emotional toughness dimensions of the scale reported poorer mental health.”

1 = More strengths-based: The measure focuses at least partially on what men do well, benefits to men; includes positive behaviors or expectations of men, highlights positive outcomes tied to masculinity, and may describe masculinity as a complex social identity.

Item Examples: “Men show strength in various ways, including displaying vulnerability”, “The important people in my life like when we show each other affection”

Methods Example: “Flexible ways of being men that allow for the expression of vulnerability, connection, and compassion ... healthy

masculinities allow boys and men to value, feel, and express the full range of their emotions”

Relational Focus (1–3)

To what extent does the scale - or the development of the scale - address the social context and incorporate a relational view of expressions of masculinities? That is, does it clearly articulate who the respondents should be considering when completing the scale items?

1 = Low: No consideration of relationality, scale items presented with unclear or contradicting social reference points; limited or unclear attention to the social context in which masculinities are expressed.

Examples: One item refers to male friend, remaining items are absent of context

2 = Moderate: Some consideration of relationality, some social reference points are included, but they may be broadly defined, some vagueness or inconsistency.

Examples: Self-reference (in the second person) for scale items, but scale instructions not explicitly presented

3 = High: Strong relational focus, social reference targets are clearly defined and articulated in methods and items; strong attention to the context in which masculinity is expressed.

Examples: “When thinking of your close male friends...”, “When you think of societal expectations of masculinity...”, “Respondents rate situations according to how stressful they feel each situation would be if it happened to them”

Cultural Attunement (1–3)

Did the authors thoughtfully discuss the cultural context for the scale, and is this reflected in the scale items? Are the cultural norms assumed (to be western) and not explicitly mentioned or considered in the scale development?

1 = Low: No mention or consideration of cultural norms; cultural assumptions (often Western) are implicit and unacknowledged.

Example: No mention of culture or cultural norms

2 = Moderate: Some consideration of cultural norms; the authors mention the impact or influence of culture on the relevant construct but may not extend this consideration to the item development process.

Example: Societal basis of gender roles discussed in introduction, future work with other cultures in discussion

3 = High: Full consideration of cultural norms and context; The measure and items are situated within cultural norms and show deliberate consideration of how the target population relates to them.

Example: Authors describe how sample is situated within larger cultural norms at the national level, describes how development of items was influenced by culture, describes boundary conditions

Sample Diversity (0–1)

Is the sample population adequately described, and does it represent diverse demographic characteristics (e.g., race, social class, age)? Do they go beyond a sample from a single US university?

0 = No: Sample is not well-documented or diverse, may consist of primarily white undergraduates from Western settings.

Example: Sample of 150 USA undergraduates from a single university, mean age of 20 years, greater than 75 percent of sample identify as White

1 = Yes: Sample is well-documented and diverse in terms of age, race, and social class.

Example: Sample of 500 participants, 250 undergraduates from two US universities and 250 community members with a mean age of 52, 40 percent identify as White, large variation in education and income

Psychometric Properties

Evidence for internal validity was coded as high when multiple comparison scales were used to establish convergent and divergent validity, moderate when some validity was established but not thoroughly, and low if only one type of validity (e.g., content, criterion, or construct) was established or mentioned. Reliability cutoffs were set according to established ranges in the field. Evidence for external validity was coded using the M-STOUT (Mechanisms, Settings, Treatments, Outcomes, Units, Time) framework for evaluating external validity (Findley et al. 2021), with an emphasis on whether the scales prevent no evidence (low), some evidence (moderate) or thorough evidence (high) for generalizability and transportability.

Evidence of Internal Validity (1–3)

Does the paper provide sufficient evidence of the internal validity of the masculinity measure (e.g., construct, convergent, discriminant validity)?

1 = Low: Only one type of validity (e.g., content, criterion, or construct) is established or mentioned.

2 = Moderate: Some evidence of validity is demonstrated, but not thoroughly (e.g., only one comparison scale used to establish discriminant validity).

3 = High: Evidence of validity is thoroughly demonstrated; multiple comparison scales are used to support the measure's convergent and discriminant validity.

Evidence of External Validity (1–3)

To what extent does the measure apply to a broader population or other target populations, using the M-STOUT framework as a reference (Findley et al. 2021)?

1 = Low: The measure is detached from real-world applications, with little explanation of broader generalizability or transportability beyond the study sample.

2 = Moderate: High-level conceptual explanation of generalizability is provided, but evidence is limited or unclear, transportability evidence is limited.

3 = High: The measure shows clear applicability to populations beyond the sample; generalizability and transportability are clearly described and are well supported.

Reliability (1–3)

To what extent does the measure yield consistent results across multiple iterations and/or timepoints?

1 = Low: Cronbach's alpha is ≤ 0.60 , Omega ≤ 0.70 , or no clear reliability data provided.

2 = Moderate: Alpha is between 0.60 and 0.79, Omega between 0.70 and 0.79

3 = High: Alpha or Omega > 0.80

Coder Reliability

Coders coded a selection of 10 papers individually and then compared codes. Fliess' Kappa values were calculated for each of the coding categories to test for interrater reliability. There was substantial or greater agreement (above 0.60) for two of the categories, moderate agreement (0.40–0.60) for two of the categories, and fair or slight agreement for the three remaining categories (see [Appendix B](#)). To strengthen the reliability of the coding framework, particularly for categories with moderate or below agreement, a reconciliation process was used to code to consensus via collaborative meetings where the coding criteria and relevant information from the paper were discussed in a group. The remaining 55 papers were coded by individual coders after the coding instructions were agreed upon and understood, and disparate codes were reconciled. Throughout the coding process, the lead author moderated weekly meetings where the coders checked in and shared questions, clarifications, and updates to the codebook as a group to ensure uniform extraction. For the coding instructions and codebook, see [Appendix C](#).

Results

The search generated 4,784 references. Of these, 4,683 were screened out and 101 potentially relevant abstracts were assessed. The screening process resulted in a final sample of 65 papers, with 59 original measures and 6 adapted or revised measures. See [Figure 1](#) for the PRISMA flow diagram of the selection process.

The mean publication year of the 65 papers included in the analysis was 2004 (range: 1974–2021; median = 2007), the median citation count was 124. Scores for all scales across the five coding dimensions of approach, relational focus, cultural attunement, sample diversity, and psychometric properties are presented in [Appendix D](#). See [Tables 2](#) and [3](#) below for summary statistics for each coding category.

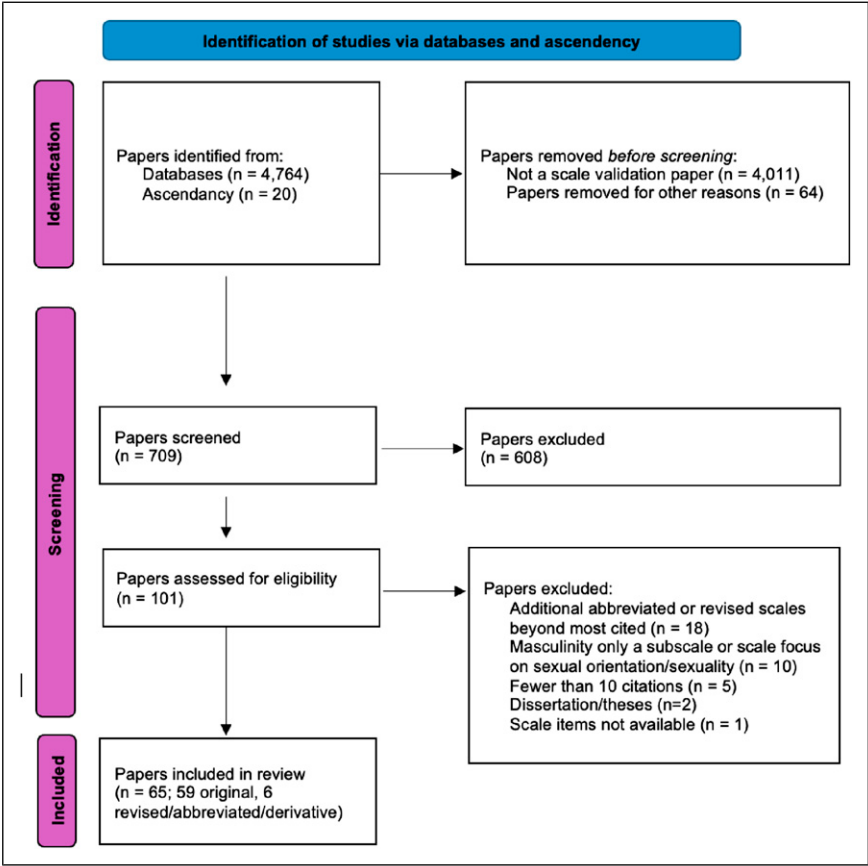


Figure 1. PRISMA diagram detailing literature search and selection process.

Table 2. Descriptive statistics by continuous coding category

Category		Low % (n)	Moderate	High
Relational focus		43 percent (28/65)	19 percent (12/65)	35 percent (22/65)
Cultural attunement		58 percent (38/65)	26 percent (17/65)	15 percent (10/65)
Psychometric Properties	Evidence of internal validity	23 percent (15/65)	43 percent (28/65)	34 percent (22/65)
	Evidence of external validity	49 percent (32/65)	42 percent (27/65)	9 percent (6/65)
	Reliability ^a	5 percent (3/63)	36 percent (23/63)	59 percent (37/63)

^aTwo scales were excluded from the reliability analysis due to using qualitative methods.

Approach (More Deficit-Based vs. More Strengths-Based)

A large majority of the reviewed measures used a more deficit-based approach. In sum, 51 of the 65 included measures (78 percent) were coded as more deficit-based, with the remaining 14 (22 percent) coded as more strengths-based. Of the twelve scales with over 500 citations, only three (25 percent) were coded as using a more strengths-based approach. See [Table 4](#) for examples of scales coded as more strengths- or deficit-based.

It bears noting that our coding criteria were deliberately asymmetric. To be classified as more strengths-based, a scale needed to include some positive framing of masculinity, however limited. By contrast, scales were classified as more deficit-based only when their theoretical orientation and items were consistently and thoroughly organized around negative outcomes.

There was greater theoretical similarity among the more deficit-based scales—despite their far higher numbers—and greater breadth, diversity, and even disagreement of theoretical approaches among the more strengths-based scales, despite their relative scarcity. The more deficit-based measures emerged largely from interconnected theories of gender role strain ([Pleck 1981, 1995](#)) and conflict ([O’Neil et al., 1986, O’Neil 2008, 2015](#)) and/or hegemonic masculinity ([Connell 1995](#)), reflecting the predominant theories of masculinities over this period as explored above (see [Smiler 2004](#)). Scales grounded in gender role strain and conflict operationalize masculinity more through the psychological costs of conformity to restrictive norms while scales grounded in hegemonic masculinity focus more on masculinity as a

Table 3. Descriptive statistics by binary coding category

Category	Yes % (n)	No
More Strengths-based	22 percent (14/65)	78 percent (51/65)
Diverse Sample	34 percent (22/65)	66 percent (43/65)

Table 4. Examples of scale methods, participant instructions, items, and/or subscales for strengths-based and deficit-based categories

More strengths-based	More deficit-based
<i>“Based on results from this qualitative research with the study population (and the literature review), the term gender-equitable young man has been operationalized here as a man who: Seeks relationships with women based on equality, respect, and intimacy rather than sexual conquest. This includes believing that men and women have equal rights and that women have as much “right” to sexual agency as do men. Seeks to be involved in household chores and childcare, meaning that they support taking both financial and care-giving responsibility for their children and household...”</i> Subscales: Inequitable gender norms, equitable gender norms. Gender equitable men scale (Pulerwitz and Barker 2008)	<i>“MGRS is conceptualized as a man’s subjective appraisal of meeting (or not meeting) society’s expectations linked to traditional male norms ... MGRS scores in men are uniquely related to certain negative health-risk behaviors relative to women ... Findings from the literature have consistently demonstrated that MGRS scores are related to men’s perpetration of violence and aggression”</i> <i>“How stressful do you feel each situation would be if it happened to you: “Being with a woman who is more successful than you; Losing in a sports competition; Having others say that you are too emotional”</i> Subscales include: Physical inadequacy, expressing tender emotions, subordination to women, intellectual inferiority, failure at work. Abbreviated masculine gender role stress scale (Swartout et al. 2015)

hierarchical system of power that subordinates women and marginalizes non-dominant men.

Nonetheless, these theories generally draw upon and harmonize with one another. Higher scores on the most-cited scales emerging from these theoretical groundings, such as the Male Role Norms Inventory (Levant et al. 1992), the Conformity to Masculine Norms Inventory (Mahalik et al. 2003), and the Gender Role Conflict Scale (O’Neil et al. 1986), often correlate to individual outcomes such as depression, substance use, and reduced help-seeking, or broader outcomes such as violence perpetration, inequitable power in decision-making, unequal participation in domestic work, and others. A defining feature of these scales is that they employ a problematizing lens – operationalizing masculinity as a source of negative outcomes or barrier to positive outcomes.

By contrast, the scales coded as more strengths-based reflect considerable heterogeneity and even contradiction in their theoretical underpinnings. Perhaps most notable among observed contradictions is the aforementioned “reconstruct or deconstruct” dichotomy (Hoffmann and Addis 2024). Scales aiming to reconstruct masculinity either identify or elevate its positive attributes that may support well-being and prosocial behavior. For instance, the Health-Related Masculine Values scale (OliFFE et al. 2019) and the Inventory of Subjective Masculine Experiences (Wong et al. 2011) seek to capture how men derive meaning, purpose, and motivation from

aspects of their masculine identity—such as responsibility, service, or providing for family. These measures align with frameworks like the Positive Psychology/Positive Masculinity (PPPM) paradigm (Kiselica and Englar-Carlson 2010), which argues that masculine identity can in fact emphasize connection, vulnerability, and equality. Other scales coded as strengths-based function quite differently, seeking instead to deconstruct or reject masculinity in service of positive outcomes. The Sex Role Egalitarianism Scale (Beere et al. 1984) and the Gender Equitable Men Scale (Pulerwitz and Barker 2008) focus not on affirming masculine traits but on measuring rejection of traditional gender norms. These tools are “more strengths-based” only insofar as they assess movement away from harmful norms.

Additional theoretical approaches to healthy or adaptive masculinities have yet to be fully operationalized in widely used scales. There is no formal measure for the PPPM paradigm, and the relational-cultural models proposed by Di Bianca and Mahalik (2022) which emphasize connection, mutual empowerment, and emotional openness within relationships, has only one measure (Di Bianca and Mahalik 2024). Similarly, intersectional frameworks that examine how masculinities are shaped by race, class, sexuality, and geography are referenced conceptually but rarely embedded in measurement design.

Relational Focus

Few scales effectively employed a relational focus, accounting for social context in scale instructions or items and/or clearly defining social reference targets. Only 23 (35 percent) of the reviewed measures were coded as demonstrating a high relational focus, while 14 (22 percent) demonstrated a moderate focus and 28 (43 percent) did not address relationality at all. Of the twelve measures with over 500 citations, two (17 percent) were coded as demonstrating a strong relational focus, four (33 percent) were demonstrating as taking a moderate relational focus, and six (50 percent) were coded as not having a relational focus. See Table 5 for examples of scales for each coding level of relational focus.

Cultural Attunement

Cultural attunement was operationalized as the extent to which scale authors meaningfully engaged with cultural norms and theory and situated their scale within them. Ten measures (15 percent of total sample) were coded as highly considering cultural norms, with 17 measures (26 percent) coded as moderately considering cultural norms and 38 measures (58 percent) coded as not taking cultural norms into account at all. Of the 12 measures with over 500 citations, none were coded as considering cultural norms. See Table 6 for examples of scales at each coding level of cultural attunement.

Table 5. Examples of scale methods, participant instructions, items, and/or subscales for each level of relational focus

High relational focus example	Moderate relational focus example	Low relational focus example
“For each question below, please think about boys you hang out with (including friends, teammates, and/or classmates). Answer each question by circling how often each of the things below has happened to you in the past month whether in-person and/or through phone or internet communication (using the scale provided of ‘Never’ to ‘Daily’). How often in the past month did boys you hang out with give you a hard time or tease you about...your clothing style; your physical appearance or body type; not ‘hooking up’ or being sexually active enough.” Policing of masculinity scale (Reigeluth and Addis 2021)	“...we designed the Adolescent Masculinity Ideology in Relationships Scale (AMIRS) to incorporate the fact that it is often through and within the contexts of interpersonal relationships that masculine norms are introduced, reinforced, incorporated, and perpetuated in ways that become personally meaningful and directly consequential to adolescent boys... the AMIRS is intended to measure the extent to which adolescent boys internalize ... appropriate behaviors for males within the contexts of their interpersonal relationships.” “Please indicate your agreement with the following statements: If a guy tells people his worries, he will look weak; I think it’s important for a guy to go after what he wants, even if it means hurting other people’s feelings.” Adolescent masculinity ideology in relationships scale (Chu et al. 2005)	Survey administration method and participant instructions not documented. “It is wrong for a man to enter a traditionally female career; A person should be more polite to a woman than to a man” Sex role egalitarianism scale (Beere et al. 1984)

Sample Diversity

A majority of measures did not use a diverse sample in the papers included in the review; 43 (66 percent) were coded as not using diverse samples with the remaining 22 (34 percent) coded as using diverse samples. Notably, of the twelve measures with over 500 citations, only one (8 percent) was coded as having a diverse sample. See Table 7 for examples of scales with samples coded as diverse or lacking diversity.

Table 6. Examples of scale methods, participant instructions, items, and/or subscales for each level of cultural attunement

High cultural attunement scale examples	Moderate cultural attunement scale examples	Low cultural attunement scale examples*
<i>“In a general sense, exploration of South African masculinities sensitized the instrument revision to subtle cross-cultural differences in masculine display that rendered particular items inappropriate...In retrospect, this was deemed a poor means through which to operationalize the concept of heterosexual self-regulation, as cross-cultural displays of heterosexuality are varied. This militated against the use of such culturally specific (Western) notions of heterosexual behavioral practice in assessing participant endorsement of heterosexuality.”</i> Male attitudes norm inventory-II (Luyt 2005)	<i>“The resulting first subscale focused on aspects of personality (Personality subscale, PS) that are culturally believed to be more present in women than in men or vice versa. Individuals were asked to rate themselves in relation to a general population of other men and women, in reference to the culture they predominantly associated themselves with.”</i> Gender-related attributes survey – masculinity (Gruber et al. 2020)	<i>“How well does each item describe you? Aggressive; Ambitious; Competitive; Independent; Individualistic”</i> Bem sex role inventory (Bem 1974) Note: Low cultural attunement scale papers provided no indication within their manuscript that the authors had considered the effect of culture during the scale development process.

Psychometric Properties

Psychometric properties across the included measures were mixed, with relatively strong and stable evidence for internal validity and reliability but low evidence for external validity. Twenty-two measures (34 percent) were coded as demonstrating strong evidence of internal validity, 28 (43 percent) were coded as demonstrating moderate evidence of internal validity, and 15 (23 percent) were coded as demonstrating low evidence of internal validity. Reliability scores varied but were generally strong with 37 (60 percent) coded as high, 23 (35 percent) coded as moderate, and 2 (3 percent) coded as low. Only 6 measures (9 percent) were coded as demonstrating evidence of high external validity, with 27 (42 percent) coded as demonstrating moderate evidence and 32 (49 percent) coded as demonstrating low evidence. The Inventory of Subjective Masculine Experience and Implicit Masculine Self-Concept Measure were not included in the reliability analysis due to the qualitative nature of their methodologies.

Table 7. Examples of participant descriptions for scales coded as having diverse samples and samples lacking diversity

Diverse sample example	Example of sample lacking diversity
<i>“The sample from the U.S. represented diverse racial/ethnic backgrounds (47.4 percent white, 16.6 percent Black, 13.9 percent Hispanic/Latino, 19.6 percent Asian, 2.5 percent Other). The majority of the participants in the U.K. identified as white (67.8 percent). Surveyors did not assess race/ethnicity in Mexico. The most commonly reported employment status was full time employment. The majority of respondents (72.0 percent) reported some level of post-secondary training. In the U.S. and U.K., most respondents (>49.9 percent) reported being single. In Mexico, 38.0 percent of participants noted that they were dating someone, while 33.9 percent said they were single. The vast majority (>87.0 percent) identified as heterosexual.”</i> Man box Scale (Hill et al. 2020)	<i>“Two hundred and eighty-seven males and females volunteered to participate in this study ... Most participants were undergraduate students at Rutgers University, drawn from the biology, education, and engineering departments. Additional participants included undergraduates from a small private college in New Jersey and graduate students in the Department of Educational Psychology at Rutgers.”</i> Male role norms inventory (Levant et al. 1992)

Changes Observed by Decade

Chi-square tests of independence and Kruskal-Wallis tests were conducted to examine differences across the coding outcomes by publication decade (1970–2020s). Two dimensions demonstrated statistically significant variation over time: external validity ratings improved substantially across decades, $\chi^2(10) = 26.47, p = .003, H(5) = 20.87, p < .001$, with mean ratings rising from 1.00 in the 1970s to 2.50 in the 2020s (see Figure 2). Cultural attunement also showed a significant positive trend, $H(5) = 11.21, p = .047$, with mean ratings increasing from 1.00 in the 1970s to 2.25 in the 2020s, though the chi-square test did not reach significance ($p = .195$, see Figure 3). The remaining five dimensions — deficits vs. strengths approach, relationality, sample diversity, internal validity, and reliability — did not differ significantly across decades (all $ps > 0.05$, see Appendix E).

Discussion

This review found that a large majority of masculinities measures reflect a more deficit-based approach, and that few explicitly consider the impact of relationships, incorporate diverse samples, and/or attend to cultural variation. While psychometric scores for internal validity and reliability were generally strong, few measures showed evidence of generalizability or transportability. Taken together, these findings provide an overview of the current landscape of masculinities measurement, highlighting both

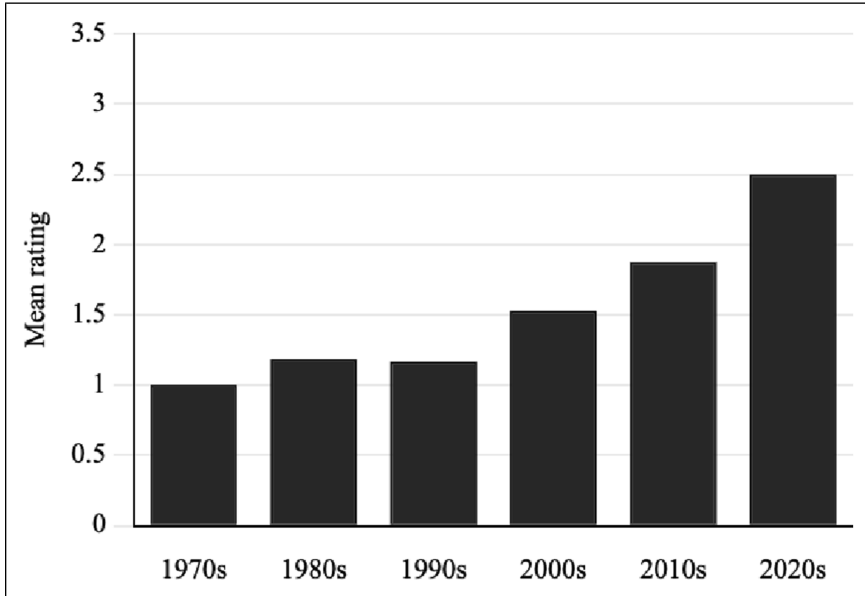


Figure 2. Mean external validity rating by publication decade (1 = low, 3 = high), $\chi^2 (10) = 26.47$, $p = .003$; Kruskal-Wallis $H (5) = 20.87$, $p < .001$.

areas of methodological consistency and conceptual frontiers where future work may expand the field’s reach, relevance, and impact.

While 78 percent of the measures reviewed took a more deficit-based approach, only 22 percent adopted a more strengths-based approach focusing at least partially on positive outcomes tied to masculinity. This ratio is comparable to the proportion of *Psychology of Men and Masculinities* articles published between 2000 and 2018 with a positive focus (15 percent) as reported by [Cole et al. \(2021\)](#). This gap affirms that the field has been far more attentive to the risks and harms associated with masculine norms than to positive alternatives. The theoretical heterogeneity observed among strengths-based measures reflects a deeper and unresolved tension in masculinities scholarship: whether the goal should be to deconstruct masculinity as a system of power and exclusion, or to reconstruct it in ways that support equity, connection, and well-being ([Beasley 2015](#); [Connell 1995](#); [Kiselica and Englar-Carlson 2010](#); [Waling 2019](#)). This debate has profound implications for measurement.

In the domain of relationality and social context, where [Thompson and Bennett \(2015\)](#) identified ambiguity in the interpretation of scores across 16 masculinities scales, the present review echoes those concerns among a much broader set of measures. The persistent lack of clarity supported by this review – particularly around whether items are capturing personal endorsement, perceived norms, cultural awareness, or some combination – represents a foundational challenge to the empirical validity of masculinities research. Without increased precision in accounting for and

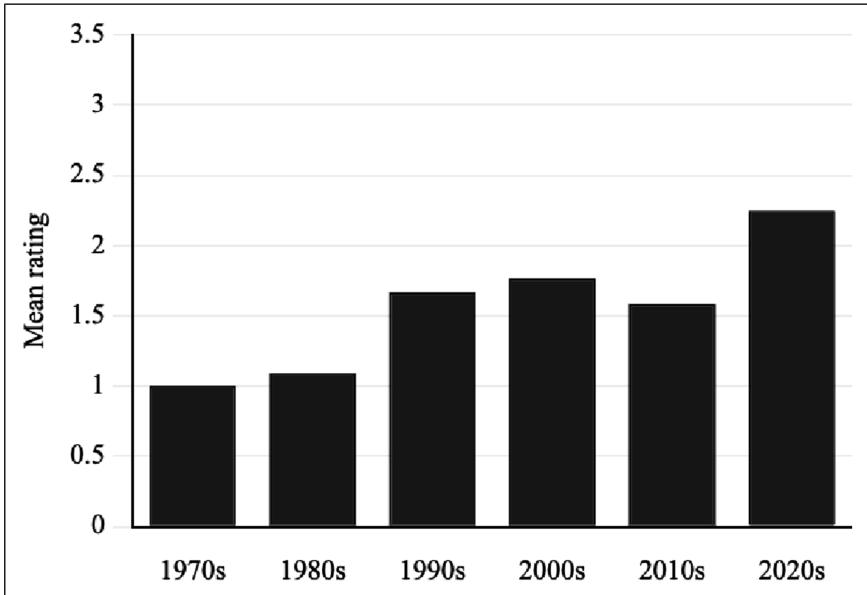


Figure 3. Mean cultural attunement rating by publication decade (1 = low, 3 = high). Kruskal-Wallis $H(5) = 11.21, p < .047$.

measuring social context and relationality, masculinities researchers will continue to face ambiguous interpretation parameters for measure scores and research participants will continue to experience confusion around what is being asked of them when reading scale instructions and items. Ambiguity around social context may leave the field vulnerable to critique or dismissal, undermining the credibility of efforts to link masculine norms – regardless of deficit- or strengths-based approach – with health, social, and behavioral outcomes.

The finding that relatively few measures demonstrated strong cultural attunement or were developed and validated using diverse samples reflects historical patterns in psychological scale development more broadly (Henrich et al. 2010). Failure to attend to cultural norms and contexts limits the applicability, reach, and reliability of data on masculinities. When measures are developed primarily with homogenous samples (most often Western, White, cisgender, and heterosexual men) they are unlikely to capture the experiences, values, or gendered realities of men from other cultural or social backgrounds. As a result, findings derived from such measures may not translate well across regions and demographic strata, potentially overlooking key sources of variation or misrepresenting masculinity in global or multicultural contexts. An important caveat to this finding is that many scholars are aware of this issue and have modified and adapted scales presented in this review to apply more directly to specific populations and cultural contexts. Some prominent examples of these scales include the Black Men’s Experiences Scale, the Masculinity Inventory Scale for Black Men,

and the African American Men's Gendered Racism Stress Inventory (Bowleg et al. 2016; Mincey et al. 2014; Schwing et al. 2013). While these efforts are important in providing tools for researchers working with specific populations, the question remains how scales designed for use with diverse populations can incorporate cultural and other identity-specific considerations into their development process.

Examining the scales in this review over time demonstrates important increases in attention to external validity and cultural attunement, paralleling other shifts in the field such as the move away from trait-based measures, while also showing how changes in other dimensions such as sample diversity and relational focus have been slower to catch on. Scales published before 2000 represent 31 percent of the sample ($n = 20$ of 65) yet account for 76.7 percent of all citations in the dataset (32,343 of 42,144 total). The oldest 20 percent of scales — just 13 measures published between 1974 and 1987 — account for 66.8 percent of all citations across the entire sample, illustrating the outsized impact older measures of masculinities have had on the field.

Future Directions

Taken together, this review of the past 50 years of masculinities measurement suggests the need for a broader and more balanced measurement landscape in the future. The next generation of masculinities measures must attend to the considerations explored in this paper and do so while also grappling with the ways conversations about gender identity, masculinities, and social norms have moved into mainstream political and media discourse, carrying new opportunities, responsibilities, and risks for researchers in the field. It is a considerable challenge - to build a body of research that is simultaneously critical, constructive, creative, and complex - but a worthy one.

Deficit-based approaches will continue to play a key role in identifying risk patterns and guiding targeted support, especially in domains such as mental health, substance use, and violence prevention. Strengths-based approaches offer a complementary option for researchers aiming to capture protective factors, adaptive strategies, and sources of resilience that may be embedded within masculine norms. Frameworks such as the Positive Psychology/Positive Masculinity (PPPM) model (Kiselica and Englar-Carlson 2010; Kiselica et al. 2016), along with intervention approaches that reframe masculine ideals to encourage prosocial behavior (Brannon et al. 2020), suggest the potential for tools that examine how gender identity can contribute positively to well-being and meaning-making among men.

McDermott et al. (2019) found that many attributes associated with masculinity — such as service, bravery, and responsibility — were widely endorsed as both desirable and more expected of men than of women, lending support to the idea that some socially prescribed masculine norms may be leveraged to support positive, pro-social outcomes. However, the study also underscores that several of these traits may reflect more moderate expressions of traditional masculinity, while others may better be understood as broadly valued human strengths not specific to any one gender. This evokes a risk asserted by Waling (2019) and others that a strengths-based approach

may over-emphasize masculinity as the primary or only mode of being for men, inadvertently reinforcing binary thinking about gender.

The development of strengths-based tools presents practical challenges alongside empirical ones. As [Brown and colleagues \(2017\)](#) note in research on thriving, capturing positive functioning requires not simply flipping the valence of traditional constructs, but actively identifying and theorizing what excellence or flourishing looks like within a given domain. Additional issues that limit many deficit-based measures – such as vague item phrasing, homogenous samples, and weak cultural grounding – can equally threaten the utility of strengths-based tools as well, if not addressed with equal rigor ([Thompson and Bennett, 2015](#)). Recent research has begun to address these issues through various approaches, including bringing a life stage lens to extant measures of masculinity (e.g., [Levant et al. 2022](#)).

Improving contextual precision and employing a relational focus will be essential for future research on masculinities, whether taking a strengths- or deficit-based approach. Without clear guidance on what a respondent's score represents, researchers risk drawing incorrect conclusions and undermining the applicability of their findings. Recent work grounded in a relational-cultural framework highlights the particular importance of men's interpersonal relationships – such as those with friends, family members, and intimate partners – as critical contexts for understanding how masculinities are shaped and experienced ([Di Bianca and Mahalik 2022, 2024](#)). In a positive sign, existing tools have begun to incorporate more explicit forms of contextual framing. The Policing of Masculinity Scale ([Reigeluth and Addis 2021](#)), for example, anchors itself in a specific peer context by instructing respondents, “For each question below, please think about the boys you hang out with (including friends, teammates, and/or classmates).” The Subjective Masculinity Scale ([Wong et al. 2013](#)) also models contextual clarity by asking participants to generate personal statements that complete the sentence “As a man...” with explicit instruction that these responses reflect their own personal thoughts, rather than social expectations. Further efforts to increase clarity in item framing and reference points will both reduce participant confusion and enable more confident and consistent interpretation of scale scores.

The field would also benefit from more systematic research on how masculinities measures function in practice in two ways: first, in terms of which types of framings are most effective in achieving improvements in health, relationships, and other prosocial behaviors; and second, in terms of how measures are adapted and applied across diverse settings. While many tools are developed for academic or clinical contexts, their findings increasingly inform public-facing campaigns and intervention efforts. Yet there is little reliable evidence on how different approaches (whether deficit-based, strengths-based, or mixed) affect men's willingness to engage, and more importantly, which framings do the most good and least harm in supporting healthier, more connected, and more equitable lives. Do deficit-based approaches provoke backlash primarily among the most vocal defenders of traditional masculinity, or do they also risk alienating more moderate or potentially receptive men? Do strengths-based messages foster genuine reflection and positive outcomes, or do they merely re-package traditional ideals and not move the needle? These are testable questions with

direct implications for measurement, messaging, and impact. At the same time, the adaptation of existing measures is rarely documented or evaluated. Researchers may revise instruments – by altering wording, shortening scales, or modifying scoring procedures – without clarity on how these changes affect psychometric integrity or interpretability. Future work should explore how masculinities measures are being implemented in applied settings and develop stronger guidance for adapting tools without compromising their validity or relevance. Addressing the identified gaps – through clearer framing, greater cultural and contextual sensitivity, and stronger evidence of real-world application – will be essential to advancing a measurement landscape that not only reflects theoretical complexity but also advances positive change.

Limitations

While this review employed a systematic approach, the quantity of measures in the masculinities research space necessitated certain constraints. The coding categories were kept at a conceptually high level in order to generate insights into overarching patterns across measures. As a result, important domain-specific richness within each category was not fully explored. For example, related to *approach* (more deficit-based vs more strengths-based), a more detailed analysis could have examined the valence of individual survey items as distinct from the overall measure orientation, or the specific domains of risk (e.g., violence, health outcomes, emotional restriction) and strength (e.g., resilience) assessed. Similarly, the *relational focus* category could be broken into researcher-centered design choices (e.g., how social context is framed in item construction) and participant-centered interpretation (e.g., how respondents understand and relate to normative prompts). Sample diversity, which was assessed holistically, could have been broken out and assessed as a number of different demographic metrics, including but not limited to age, race, sexual orientation, education level, socioeconomic status, and geographic location. The low initial agreement on the coding categories of deficit vs. strengths approach, cultural attunement, and external validity could reflect genuine conceptual ambiguity in how these constructs have been operationalized and interpreted.

This review was limited to articles published in English, which may have excluded important contributions from scholars outside English-speaking contexts. Such scholarship may include novel theoretical approaches, reflect diverse cultural perspectives, and represent important viewpoints to incorporate into future reviews.

In addition, by excluding articles with fewer than ten citations and non-peer-reviewed articles, this review omits some recent and emerging measurement scholarship, including new measures that build from the Positive Psychology/Positive Masculinity (PPPM) paradigm. Measures such as the Relationships Fostering Healthy Masculinities Scale (Di Bianca and Mahalik 2024) and the Positive Masculinity Scale (Bates 2023), reflect important advancements toward strengths-based approaches and should be considered in ongoing and future evaluations of the field.

Conclusion

In ongoing debates about whether to deconstruct, reconstruct, or integrate multiple approaches to measuring masculinities, the findings of this review empower researchers to make informed choices about how to approach developing or selecting a measure. By mapping both the field's common practices and the consequences of the primacy of deficit-based theories, this review offers a foundation for more balanced, context-sensitive, and culturally-attuned approaches to measurement. Continued progress will require not only technical refinement, but a renewed commitment to building tools that reflect the lived realities of men—and contribute meaningfully to healthier, more equitable outcomes for all.

ORCID iDs

Peter Fisher  <https://orcid.org/0000-0002-4723-8214>

Kieran O’Gorman  <https://orcid.org/0000-0003-2087-0532>

Funding

Peter Fisher, Brian Heilman, Simon Rice, and Zac Seidler are employees of Movember Foundation which funded this work.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Supplemental Material

Supplemental material for this article is available online.

References

- Addis, Michael E, and James R Mahalik. 2003. “Men, masculinity, and the contexts of help seeking.” *American Psychologist* 58 (1): 5–14. doi:[10.1037/0003-066x.58.1.5](https://doi.org/10.1037/0003-066x.58.1.5)
- Bates, David. 2023. *Development and Initial Validation of a Positive Masculinity Measure*. PhD diss. University of Cincinnati.
- Beasley, Christine. 2015. “Introduction to Special Issue of Men and Masculinities: Heterodox Hetero-Masculinities.” *Men and Masculinities* 18 (2): 135–9. doi: [10.1177/1097184X15583825](https://doi.org/10.1177/1097184X15583825).
- Beere, Carole A., Daniel W. King, Donald B. Beere, and Lynda A. King. 1984. “The Sex-Role Egalitarianism Scale: a Measure of Attitudes Toward Equality Between the Sexes.” *Sex Roles* 10 (7-8): 563–76. doi: [10.1007/BF00287265](https://doi.org/10.1007/BF00287265).
- Bem, Sandra L. 1974. “The Measurement of Psychological Androgyny.” *Journal of Consulting and Clinical Psychology* 42 (2): 155–62. doi: [10.1037/h0036215](https://doi.org/10.1037/h0036215).
- Bowleg, Lisa, Devin English, Ana Maria del Rio-Gonzalez, Gary J. Burkholder, Michelle Teti, and Jeanne M. Tschann. 2016. “Measuring the Pros and Cons of what It Means to Be a

- Black Man: Development and Validation of the Black Men's Experiences Scale (BMES)." *Psychology of Men & Masculinities* 17 (2): 177. doi: [10.1037/men0000026](https://doi.org/10.1037/men0000026).
- Brannon, Tiffany N., Peter H. Fisher, and Abigail J. Greydanus. 2020. *Selves as Solutions to Social Inequalities: Why Engaging the Full Complexity of Social Identities Is Critical to Addressing Disparities*. 1st ed. Cambridge University Press. doi: [10.1017/9781108874267](https://doi.org/10.1017/9781108874267).
- Bridges, Tristan, and Cheri J. Pascoe. 2014. "Hybrid Masculinities: New Directions in the Sociology of Men and Masculinities." *Sociology Compass* 8 (3): 246–58. doi: [10.1111/soc4.12134](https://doi.org/10.1111/soc4.12134).
- Brown, Daniel J., Rachel Arnold, David Fletcher, and Martyn Standage. 2017. "Human Thriving." *European Psychologist* 22 (3): 167–79. doi: [10.1027/1016-9040/a000294](https://doi.org/10.1027/1016-9040/a000294).
- Chu, Judy Y., Michelle V. Porche, and Deborah L. Tolman. 2005. "The Adolescent Masculinity Ideology in Relationships Scale: Development and Validation of a New Measure for Boys." *Men and Masculinities* 8 (1): 93–115. doi: [10.1177/1097184X03257453](https://doi.org/10.1177/1097184X03257453).
- Cole, Brian P., Margaret Brennan, Emily Tyler, and Willard Ryan. 2020. "Predicting Men's Acceptance of Sexual Violence Myths Through Conformity to Masculine Norms, Sexism, and 'Locker Room Talk'." *Psychology of Men & Masculinities* 21 (4): 508–17. doi: [10.1037/men0000248](https://doi.org/10.1037/men0000248).
- Cole, Brian P., Kelsey Moffitt-Carney, Tristan P. Patterson, and Willard Ryan. 2021. "Psychology of Men and Masculinities' Focus on Positive Aspects of Men's Functioning: a Content Analysis and Call to Action." *Psychology of Men & Masculinities* 22 (1): 39–47. doi: [10.1037/men0000264](https://doi.org/10.1037/men0000264).
- Connell, R. W. 1995. *Masculinities*. Polity Press.
- Di Bianca, Michael, and James R. Mahalik. 2022. "A Relational-Cultural Framework for Promoting Healthy Masculinities." *American Psychologist* 77 (3): 321–32. doi: [10.1037/amp0000929](https://doi.org/10.1037/amp0000929).
- Di Bianca, Michael, and James R. Mahalik. 2024. "Development of the Relationships Fostering Healthy Masculinities Scale." *Psychology of Men & Masculinities* 25 (1): 1–14. doi: [10.1037/men0000488](https://doi.org/10.1037/men0000488).
- Englar-Carlson, Matt, and Mark S. Kiselica. 2013. "Affirming the Strengths in Men: a Positive Masculinity Approach to Assisting Male Clients." *Journal of Counseling and Development* 91 (4): 399–409. doi: [10.1002/j.1556-6676.2013.00111.x](https://doi.org/10.1002/j.1556-6676.2013.00111.x).
- Equimundo. 2022. *The International Men and Gender Equality Survey: A Status Report on Men, Women, and Gender Equality in 15 Headlines*. Equimundo. <https://www.unfpa.org/publications/international-men-gender-equality-survey-images>
- Findley, Michael, Kyosuke Kikuta, and Michael Denly. 2021. "External Validity." *Annual Review of Political Science* 24: 365–93. doi: [10.1146/annurev-polisci-041719-102556](https://doi.org/10.1146/annurev-polisci-041719-102556).
- Gage, Alida N., and Suzanne H. Lease. 2021. "An Exploration of the Link Between Masculinity and Endorsement of IPV Myths in American Men." *Journal of Interpersonal Violence* 36 (13-14): 6145–65. doi: [10.1177/0886260518818430](https://doi.org/10.1177/0886260518818430).
- Gerdes, Zachary T., Kristen M. Alto, Sara Jadaszewski, Francesca D'Auria, and Ronald F. Levant. 2018. "A Content Analysis of Research on Masculinity Ideologies Using all Forms of the Male Role Norms Inventory (MRNI)." *Psychology of Men & Masculinities* 19 (4): 584–99. doi: [10.1037/men0000134](https://doi.org/10.1037/men0000134).

- Griffith, Derek M., Katie Gunter, and Daphne C. Watkins. 2012. "Measuring Masculinity in Research on Men of Color: Findings and Future Directions." *American Journal of Public Health* 102 (S2): S187–94. doi: [10.2105/AJPH.2012.300715](https://doi.org/10.2105/AJPH.2012.300715).
- Gruber, Freya M., Eva Distlberger, Thomas Scherndl, Tuulia M. Ortner, and Belinda Pletzer. 2020. "Psychometric Properties of the Multifaceted Gender-Related Attributes Survey (GERAS)." *European Journal of Psychological Assessment* 36 (4): 612–23. doi: [10.1027/1015-5759/a000528](https://doi.org/10.1027/1015-5759/a000528).
- Henrich, Joseph, Steven J. Heine, and Ara Norenzayan. 2010. "The Weirdest People in the World?" *Behavioral and Brain Sciences* 33 (2-3): 61–135. doi: [10.1017/S0140525X0999152X](https://doi.org/10.1017/S0140525X0999152X).
- Hill, Amber L., Elizabeth Miller, Galen E. Switzer, et al. 2020. "Harmful Masculinities Among Younger Men in Three Countries: Psychometric Study of the Man Box Scale." *Preventive Medicine* 139: 106185. doi: [10.1016/j.ypmed.2020.106185](https://doi.org/10.1016/j.ypmed.2020.106185).
- Hoffmann, Ethan, and Michael E. Addis. 2024. "To Reconstruct or Deconstruct? A Fundamental Question for the Psychology of Men and Masculinities." *Psychology of Men & Masculinities* 25 (1): 1–12. doi: [10.1037/men0000440](https://doi.org/10.1037/men0000440).
- Kachel, Sven, Melanie C. Steffens, and Claudia Niedlich. 2016. "Traditional Masculinity and Femininity: Validation of a New Scale Assessing Gender Roles." *Frontiers in Psychology* 7: 956. doi: [10.3389/fpsyg.2016.00956](https://doi.org/10.3389/fpsyg.2016.00956).
- Kiselica, Mark S., and Matt Englar-Carlson. 2010. "Identifying, Affirming, and Building Upon Male Strengths: the Positive Psychology/Positive Masculinity Model of Psychotherapy with Boys and Men." *Psychotherapy: Theory, Research, Practice, Training* 47 (3): 276–87. doi: [10.1037/a0021159](https://doi.org/10.1037/a0021159).
- Kiselica, Mark S., Sheila Benton-Wright, and Matt Englar-Carlson. 2016. "Accentuating Positive Masculinity: a New Foundation for the Psychology of Boys, Men, and Masculinity." In *APA Handbook of Men and Masculinities*, edited by Y. Joel Wong and Stephen R. Wester, 123–43. American Psychological Association. doi: [10.1037/14594-006](https://doi.org/10.1037/14594-006).
- Levant, Ronald F. 2015. "The Road Goes Ever On: an Editorial." *Psychology of Men & Masculinities* 16 (4): 349–54. doi: [10.1037/a0039695](https://doi.org/10.1037/a0039695).
- Levant, Ronald F., and Katherine Richmond. 2016. "The Gender Role Strain Paradigm and Masculinity Ideologies." In *APA Handbook of Men and Masculinities*, edited by Y. Joel Wong and Stephen R. Wester, 23–49. American Psychological Association. doi: [10.1037/14594-002](https://doi.org/10.1037/14594-002).
- Levant, Ronald F., Linda S. Hirsch, Elizabeth Celentano, et al. 1992. "The Male Role: an Investigation of Contemporary Norms." *Journal of Mental Health Counseling* 14 (3): 325–37.
- Levant, Ronald F., Glenn E. Good, Sarah W. Cook, et al. 2006. "The Normative Male Alexithymia Scale: Measurement of a Gender-Linked Syndrome." *Psychology of Men & Masculinities* 7 (4): 212–24. doi: [10.1037/1524-9220.7.4.212](https://doi.org/10.1037/1524-9220.7.4.212).
- Levant, Ronald F., David J. Wimer, and Christine Williams. 2011. "An Evaluation of the Health Behavior Inventory-20 (HBI-20) and Its Relationships to Masculinity and Attitudes Towards Seeking Psychological Help Among College Men." *Psychology of Men & Masculinities* 12 (1): 26–41. doi: [10.1037/a0021014](https://doi.org/10.1037/a0021014).

- Levant, Ronald F., Zachary T. Gerdes, Kristen M. Alto, Sara E. Jadaszewski, and Ryon C. McDermott. 2018. "Development and Evaluation of the Fathers' Expectations About Sons' Masculinity Scale (Short Form)." *Psychology of Men & Masculinities* 19 (3): 325–39. doi: [10.1037/men0000108](https://doi.org/10.1037/men0000108).
- Levant, Ronald F., Javier Martin-Fernandez, Ryon C. McDermott, and Edward H. Thompson Jr. 2022a. "Measurement Invariance by Age Cohort of Two Versions of the Male Role Norms Inventory – Short Form and Very Brief." *Men and Masculinities* 25 (3): 438–58. doi: [10.1177/1097184X211017620](https://doi.org/10.1177/1097184X211017620).
- Levant, Ronald F., Ryon C. McDermott, Shana Pryor, and Jennifer Barinas. 2022b. "Masculinity and Compliance With Centers for Disease Control and Prevention Recommended Health Practices During the COVID-19 Pandemic." *Health Psychology* 41 (2): 94–103. doi: [10.1037/hea0001119](https://doi.org/10.1037/hea0001119).
- Luyt, Russell. 2005. "The Male Attitude Norms Inventory-II: a Measure of Masculinity Ideology in South Africa." *Men and Masculinities* 8 (2): 208–29. doi: [10.1177/1097184X04264631](https://doi.org/10.1177/1097184X04264631).
- Mahalik, James R., Benjamin D. Locke, Larry H. Ludlow, et al. 2003. "Development of the Conformity to Masculine Norms Inventory." *Psychology of Men & Masculinities* 4 (1): 3–25. doi: [10.1037/1524-9220.4.1.3](https://doi.org/10.1037/1524-9220.4.1.3).
- Mahalik, James R., Michael Di Bianca, and Michael P. Harris. 2022. "Men's Attitudes Toward Mask-Wearing During COVID-19: Understanding the Complexities of Mask-Ulinity." *Journal of Health Psychology* 27 (5): 1187–1204. doi: [10.1177/1359105321990793](https://doi.org/10.1177/1359105321990793).
- McDermott, Ryon C., Ronald F. Levant, Joseph H. Hammer, Rosalie J. Hall, Daniel K. McKelvey, and Zachary Jones. 2017. "Further Examination of the Factor Structure of the Male Role Norms Inventory-Short Form (MRNI-SF): Measurement Considerations for Women, Men of Color, and Gay Men." *Journal of Counseling Psychology* 64 (6): 724–38. doi: [10.1037/cou0000225](https://doi.org/10.1037/cou0000225).
- McDermott, Ryon C., Kipp R. Pietrantonio, Brandon R. Browning, et al. 2019. "In Search of Positive Masculine Role Norms: Testing the Positive Psychology Positive Masculinity Paradigm." *Psychology of Men & Masculinities* 20 (1): 12–22. doi: [10.1037/men0000160](https://doi.org/10.1037/men0000160).
- Meadows, Robert, and Kate Davidson. 2007. "Maintaining Manliness in Later Life: Hegemonic Masculinities and Emphasized Femininities." In *Age Matters: Realigning Feminist Thinking*, edited by Toni M. Calasanti and Kathleen F. Slevin, 295–312. Routledge.
- Mincey, Krista, Moya Alfonso, Amy Hackney, and John Luque. 2014. "Being a Black Man: Development of the Masculinity Inventory Scale (MIS) for Black Men." *The Journal of Men's Studies* 22 (3): 167–79. doi: [10.1037/t39145-000](https://doi.org/10.1037/t39145-000).
- O'Neil, James M. 2008. "Summarizing 25 Years of Research on Men's Gender Role Conflict Using the Gender Role Conflict Scale: New Research Paradigms and Clinical Implications." *The Counseling Psychologist* 36 (3): 358–445. doi: [10.1177/0011000008317057](https://doi.org/10.1177/0011000008317057).
- O'Neil, James M. 2015. *Men's Gender Role Conflict: Psychological Costs, Consequences, and an Agenda for Change*. American Psychological Association. doi: [10.1037/14501-000](https://doi.org/10.1037/14501-000).
- O'Neil, James M., Barbara J. Helms, Robert K. Gable, Laurence David, and Lawrence S. Wrightsman. 1986. "Gender-Role Conflict Scale: College men's fear of femininity." *Sex Roles* 14 (5): 335–350. doi:[10.1007/BF00287583](https://doi.org/10.1007/BF00287583)

- Olliffe, John L., Simon Rice, Michelle T. Kelly, et al. 2019. "A Mixed-Methods Study of the Health-Related Masculine Values Among Young Canadian Men." *Psychology of Men & Masculinities* 20 (3): 310–23. doi: [10.1037/men0000157](https://doi.org/10.1037/men0000157).
- Page, Matthew J., Joanne E. McKenzie, Patrick M. Bossuyt, et al. 2021. "The PRISMA 2020 Statement: an Updated Guideline for Reporting Systematic Reviews." *BMJ* 372: n71. doi: [10.1136/bmj.n71](https://doi.org/10.1136/bmj.n71).
- Pleck, Joseph H. 1981. *The Myth of Masculinity*. MIT Press.
- Pleck, Joseph H. 1995. "The Gender Role Strain Paradigm: an Update." In *A New Psychology of Men*, edited by Ronald F. Levant and William S. Pollack, 11–32. Basic Books.
- Pulerwitz, Julie, and Gary Barker. 2008. "Measuring Attitudes Toward Gender Norms Among Young Men in Brazil: Development and Psychometric Evaluation of the GEM Scale." *Men and Masculinities* 10 (3): 322–38. doi: [10.1177/1097184X06298778](https://doi.org/10.1177/1097184X06298778).
- Raj, A., L. McDougal, and A. Trivedi. 2017. *EMERGE Project Report: Theoretical and Definitional Basis for Identification of Measures of Gender Equality and Empowerment*. University of California, San Diego School of Medicine, Center on Gender Equity and Health.
- Reigeluth, Christopher S., and Michael E. Addis. 2021. "Policing of Masculinity Scale (POMS) and Pressures Boys Experience to Prove and Defend Their 'Manhood'." *Psychology of Men & Masculinities* 22 (2): 306–20. doi: [10.1037/men0000318](https://doi.org/10.1037/men0000318).
- Reigeluth, Christopher S., and E. Addis Michael. 2016. "Adolescent boys' experiences with policing of masculinity: Forms, functions, and consequences." *Psychology of Men & Masculinities* 17: 74–83. doi: [10.1037/a0039342](https://doi.org/10.1037/a0039342)
- Schwing, Alison E., Y. Joel Wong, and Melanie D. Fann. 2013. "Development and Validation of the African American Men's Gendered Racism Stress Inventory." *Psychology of Men & Masculinities* 14 (1): 16–24. doi: [10.1037/a0028272](https://doi.org/10.1037/a0028272).
- Smiler, Andrew P. 2004. "Thirty Years After the Discovery of Gender: Psychological Concepts and Measures of Masculinity." *Sex Roles* 50 (1-2): 15–26. doi: [10.1023/B:SERS.0000011069.02279.4c](https://doi.org/10.1023/B:SERS.0000011069.02279.4c).
- Spence, Janet T., Robert Helmreich, and Joy Stapp. 1974. "The Personal Attributes Questionnaire: a Measure of Sex Role Stereotypes and Masculinity-Femininity." *Journal Supplement Abstract Service Catalog of Selected Documents in Psychology* 4: 43. doi: [10.1037/t02466-000](https://doi.org/10.1037/t02466-000).
- Swartout, Kevin M., Dominic J. Parrott, Amy M. Cohn, Brett T. Hagman, and Kathryn E. Gallagher. 2015. "Development of the Abbreviated Masculine Gender Role Stress Scale." *Psychological Assessment* 27 (2): 489–500. doi: [10.1037/a0038443](https://doi.org/10.1037/a0038443).
- Thompson, Edward H., and Kate M. Bennett. 2015. "Measurement of Masculinity Ideologies: a (Critical) Review." *Psychology of Men & Masculinities* 16 (2): 115–33. doi: [10.1037/a0038609](https://doi.org/10.1037/a0038609).
- Thompson, Edward H., and Joseph H. Pleck. 1995. "Masculinity Ideologies: a Review of Research Instrumentation on Men and Masculinities." In *A New Psychology of Men*, edited by Ronald F. Levant and William S. Pollack, 129–63. Basic Books.
- Thompson, Edward H., Joseph H. Pleck, and David L. Ferrera. 1992. "Men and Masculinities: Scales for Masculinity Ideology and Masculinity-Related Constructs." *Sex Roles* 27 (11-12): 573–607. doi: [10.1007/BF02651094](https://doi.org/10.1007/BF02651094).

- Tricco, Andrea C., Erin Lillie, Wasifa Zarin, et al. 2018. "PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation." *Annals of Internal Medicine* 169 (7): 467–73. doi: [10.7326/M18-0850](https://doi.org/10.7326/M18-0850).
- Waling, Andrea. 2019. "Problematising 'Toxic' and 'Healthy' Masculinity for Addressing Gender Inequalities." *Australian Feminist Studies* 34 (101): 362–75. doi: [10.1080/08164649.2019.1679021](https://doi.org/10.1080/08164649.2019.1679021).
- Wells, Lana, Sarah Fotheringham, and Laura Pascoe. 2024. "Insights on Engaging Men and Boys in Creating a More Gender Equal Future in Canada." *Violence Against Women* 30 (1): 297–322. doi: [10.1177/10778012231203621](https://doi.org/10.1177/10778012231203621).
- Whorley, Matthew R., and Michael E. Addis. 2006. "Ten Years of Psychological Research on Men and Masculinity in the United States: Dominant Methodological Trends." *Sex Roles* 55 (9-10): 649–58. doi: [10.1007/s11199-006-9120-1](https://doi.org/10.1007/s11199-006-9120-1).
- Wong, Y. Joel, Munyi Shea, Julie R. LaFollette Hickman, et al. 2011. "The Inventory of Subjective Masculinity Experiences: Development and Psychometric Properties." *The Journal of Men's Studies* 19 (3): 236–55. doi: [10.3149/jms.1903.236](https://doi.org/10.3149/jms.1903.236).
- Wong, Y. Joel, Munyi Shea, Sarah J. Hickman, Julie R. LaFollette, Nicholas Cruz, and Tamar Boghokian. 2013. "The Subjective Masculinity Stress Scale: Scale Development and Psychometric Properties." *Psychology of Men & Masculinities* 14 (2): 148–55. doi: [10.1037/a0027521](https://doi.org/10.1037/a0027521).
- Wong, Y. Joel, Moon-Ho Ringo Ho, Shu-Yi Wang, and I. S. Keino Miller. 2017. "Meta-Analyses of the Relationship Between Conformity to Masculine Norms and Mental Health-Related Outcomes." *Journal of Counseling Psychology* 64 (1): 80–93. doi: [10.1037/cou0000176](https://doi.org/10.1037/cou0000176).
- Wong, Y. Joel, Ricky M. Granderson, Nelson O. O. Zounlome, Keiko M. McCullough, Jacqueline Hyman, and Sara B. Schwabe. 2020. "The Assessment of Subjective Masculine Norms in the United States." *Psychology of Men & Masculinities* 21 (4): 545–57. doi: [10.1037/men0000254](https://doi.org/10.1037/men0000254).

Author Biographies

Peter Fisher, Ph.D. is a Research Fellow at the Movember Institute of Men's Health. His research focuses on social norms, masculinities, and men's health. Recent publications include Movember's "The Real Face of Men's Health" US report (2025) and "Where and with whom does a brief social-belonging intervention promote progress in college?" (*Science*, 2023).

Brian Heilman, M.A. is a Senior Research Fellow at the Movember Institute of Men's Health. His research examines violence prevention, masculinities, mental health, and fatherhood. Major publications include co-authoring "State of the World's Fathers" reports (2017, 2019, 2021, 2023) and "Pathways to Gender-Equitable Men: Findings from the International Men and Gender Equality Survey" (*Men and Masculinities*, 2014). He also co-authored the "International Men & Gender Equality Survey (IMAGES)" (UNFPA, 2022).

Ryon C. McDermott, Ph.D. is an Associate Professor in the Department of Counseling and Instructional Sciences at the University of South Alabama. His research focuses on traditional masculinity ideology, men's mental health, and gender role

conflict. His publications include “Further Examination of the Factor Structure of the Male Role Norms Inventory-Short Form” (*Journal of Counseling Psychology*, 2017) and “Development and Validation of a Five-Item Male Role Norms Inventory Using Bifactor Modeling” (*Psychology of Men and Masculinity*, 2019).

Jasleen Chhabra, Ph.D. is a Research Fellow in Young Men’s Mental Health at the Centre for Youth Mental Health, The University of Melbourne, and Orygen, Melbourne, Australia. Her research focuses on youth mental health, paternal perinatal mental health, and social media’s impact on young people. Recent publications include “Social Media and Youth Mental Health: Scoping Review of Platform and Policy Recommendations” (*Journal of Medical Internet Research*, 2025) and paternal perinatal mental health research published in peer-reviewed journals.

Ruben Benakovic, is a Project Manager and Research Associate at Orygen, Melbourne, Australia, and The University of Melbourne. His research examines men’s mental health, psychotherapeutic engagement, and social determinants of mental illness among men. Recent publications include co-authoring “Social Media and Youth Mental Health: Scoping Review of Platform and Policy Recommendations” (*Journal of Medical Internet Research*, 2025) and work on the Men in Mind practitioner training program and Australian men’s help-seeking pathways.

Kieran O’Gorman is a Research Assistant in the Young Men’s Mental Health team at Orygen, Centre for Youth Mental Health, The University of Melbourne, and Movember Institute of Men’s Health, Melbourne, Australia. His research focuses on men’s experiences of depression and suicidality, including male-type depression and avoidant coping. Publications include “‘The Landscape’s Changed and No One’s Giving Them a New Map’: Positive Masculinity in Boys and Young Men” (*International Journal of Men’s Social and Community Health*, 2024).

Simon Rice, Ph.D. is Global Director of the Movember Institute of Men’s Health and Honorary Professorial Fellow at the Centre for Youth Mental Health, University of Melbourne, Melbourne, Australia. His research encompasses men’s mental health, masculinities, athlete mental health, and suicide prevention. He has published over 300 peer-reviewed articles and served as Editor-in-Chief of *Australian Psychologist* (2020-2023). Major works include “The Mental Health of Elite Athletes: A Narrative Systematic Review” (*Sports Medicine*, 2016) and edited volume *Masculinities and Mental Health in Young Men* (Springer, 2024).

Zac Seidler, Ph.D. is Associate Professor at Orygen and the Centre for Youth Mental Health, The University of Melbourne, and Global Director of Research at Movember, Melbourne, Australia. His research focuses on masculinities, men’s mental health, suicide prevention, and gender-responsive mental health service design. He has published over 150 peer-reviewed articles. Key publications include “Men’s Mental Health Services: The Case for a Masculinities Model” (*Journal of Men’s Studies*, 2018) and edited volume *Masculinities and Mental Health in Young Men: From Echo Chambers to Evidence* (Springer, 2024).