This is the Accepted Manuscript version. This version is defined in the NISO recommended practice RP-8-2008 http://www.niso.org/publications/rp/


Copyright

Items in ResearchSpace are protected by copyright, with all rights reserved, unless otherwise indicated. Previously published items are made available in accordance with the copyright policy of the publisher.

This is an Accepted Manuscript of an article published in *Critical Public Health* on 5 Jul 2016, available online: http://www.tandfonline.com/doi/full/10.1080/09581596.2016.1205182

For more information, see General copyright, Publisher copyright, SHERPA/RoMEO.
Title: A Feminist Quality Appraisal Tool: Exposing gender bias and gender inequities in health research

Authors: Tessa Morgan BA, Lisa Williams PhD, Merryn Gott PhD

Institutional Affiliation: School of Nursing, Faculty of Medical and Health Sciences, The University of Auckland, Auckland, New Zealand.

Each author’s current appointment and work address:
Tessa Morgan: Research Assistant, School of Nursing, The University of Auckland
Lisa Ann Williams: Research Fellow, School of Nursing, The University of Auckland
Merryn Gott: Professor of Health Sciences, School of Nursing, The University of Auckland

Address:
School of Nursing
Faculty of Medical and Health Sciences
The University of Auckland
Private Bag 92019
Auckland, New Zealand

Contact details of corresponding author:
Tessa Morgan
School of Nursing
Faculty of Medical and Health Sciences
The University of Auckland
Private Bag 92019
Auckland, New Zealand
E-mail: t.morgan@auckland.ac.nz
Abstract

Quality appraisal tools used in systematic reviews to evaluate health literature do not adequately address issues related to gender. This oversight is significant because disparities between genders have been identified as a major health equity concern and systematic reviews are regarded as a powerful means for informing policy that could redress gender inequities. In this paper we present our Feminist Quality Appraisal Tool that offers researchers a template to undertake a comprehensive gendered analysis of studies they review. Informed by a feminist perspective, the tool addresses issues of power, gender, and inequity, thereby giving researchers the means to interrogate the scientific rigour of systematic reviews that focus on gender. Specifically, our tool outlines ways gender can be critically examined in terms of study design, data collection, analysis, discussion and recommendations. We argue that this tool has the potential to improve the provision of public health by providing solid understandings and critical reflections on the reasons why women continue to face barriers in their access to optimal health care.

Keywords: Health Inequalities, Evaluation, Feminism
Introduction

Systematic literature reviews have a critical role to play in improving health outcomes. Their usefulness lies in their ability to synthesise data around a specific research question to discern generalisability and consistency within the selected field in an organised, credible way that can be replicated by other researchers (Klassen, Jadad, & Moher, 1998; Mulrow, 1994), as well as provide a way to identify and alleviate problems of potential bias in the literature (Lohr & Carey, 1999). Because they make manageable the vast amounts of literature available, systematic reviews provide a “powerful mechanism for improving population health” (Sweet, 2007). As a result, policy makers often rely on them as evidence for decision making, (Hawker, Payne, Kerr, Hardey, & Powell, 2002) which means systematic reviews have the potential to influence the distribution of resources available to the health sector and therefore ameliorate areas of particular inequity (Sweet, 2007).

One area where health inequities are particularly prominent is in relation to gender (Denton, Prus, & Walters, 2004; Fernandez, Schiaffino, Rajmil, Badia, & Segura, 1999; Kawachi, Kennedy, Gupta, & Prothrow-Stith, 1999; Moss, 2002). Although the disparities that women and men face vary dramatically between countries and communities, women are almost unanimously devalued and marginalised within the societies in which they live (Bose, 2015). As a result women are more likely to live with co-morbidities and to experience worse health than men (Sen, George, & Östlin, 2002; World Health Organisation, 2007, 2009), in no small part due to their reduced access to the material and social conditions of life that foster health, to their differential exposure to stressful life events and to everyday stressors associated with women’s social roles (Denton et al., 2004; Kawachi et al., 1999; Nieuwenhoven & Klinge, 2010; Spitzer, 2005). Men’s health also suffers from expectations associated with ideals surrounding hegemonic masculinity, whereby they are more likely to die earlier due in part to the greater likelihood of their engagement in more physically strenuous work and leisure activities (Calasanti, 2004).

Gendered ideas surrounding health also shape the diagnosis and treatment of women and men. For example, men are far less likely to be diagnosed with depression despite exhibiting the same symptoms as women, whereas women are far less likely to be diagnosed with alcoholism, due to the gender expectations attached to these respective conditions (World Health Organisation, 2015). Such examples also reveal the relational and interconnected nature of gender expectations, which means that adverse health effects for one group have a
bearing on the other (Calasanti, 2004; Connell, 2012). Indeed, research has identified a positive correlation between the egalitarian nature of a society with the greater overall health of its population (Kawachi et al., 1999; Sen et al., 2002).

If gendered differences are due to societal factors rather than simply ‘natural’ biological processes, then there is a clear avenue for health and social policy to ameliorate inequities (Denton et al., 2004; Sen et al., 2002). In recognition of this, there is already growing demand to both systematically collate information on gender disparities in order to fill knowledge gaps, (Bieri & Sancar, 2009) as well as move toward “improving quality of evaluative activity within the social sector” (Maharey, 2003, p. 2). More recently, some governments have commissioned reports to tackle gender inequality (Northern Ireland Statistics and Research Agency, 2013) and to promote the use of a gender analysis as a method to help promote gender equality (Global Affairs Canada, 2015).

 Nonetheless, because systematic reviews remain the key way to highlight gender disparities in health (United Nations, 2010), an urgent need exists to increase their nuance and scope if they are to inform policy shifts of the type described above (Hankivsky & Christoffersen, 2008; S. M. Johnson, Karvonen, Phelps, Nader, & Sanborn, 2003). Indeed, it has been argued that more effort is required to produce “practical tools to operationalize and measure sex/gender in relation to other health determinants”(Tudiver, Boscoe, Runnels, & Doull, 2012). In particular, scholars need to develop deeper understandings of the “contexts where health interventions occur” (Hammarström, 2007; Moss, 2002; Tudiver et al., 2012). Concerns also exist regarding clarity; the lack of a unified definition of gender and sex in the literature (Calasanti & King, 2007; J. L. Johnson, Greaves, & Repta, 2009) or clear guidelines for doing gender research have been cited as reasons for researchers to refrain from undertaking a gender analysis at all (Doull et al., 2014; Macintyre, Hunt, & Sweeting, 1996). As a result, research produced varies significantly in quality (Hammarström, 2007). It was within this context that we identified a need to review existing systematic review tools which explicitly address gender and subsequently develop a new quality appraisal tool.

**Our conceptual underpinnings**

Our feminist approach does not fit neatly within a designated feminist perspective, which is to say that we draw upon and acknowledge a plethora of feminisms, such as intersectional, constructionist, liberal, and radical. We are sensitive to the historical silencing of women’s
voices both as the researcher and the researched, and we believe it is imperative this situation be redressed. We are committed in our research to capturing and valuing women’s experiences and believe this should be a central feature of all scientific research.

**Overview of existing systematic review tools addressing gender**

*Search methods*

A search strategy was developed with a subject librarian to review the literature for existing quality appraisal tools on gender. In March of 2015, and then again in December 2015, T.M. conducted a literature search of Medline, Embase, CINAHL and Sociological Abstracts databases. The following search terms were used: “systematic review*”, “appraisal*”, “assessment”, “evaluat*”, “tool*”, “framework*” “gender*”, “feminis*”, “sex”, and the search was limited to papers in English. Hand searches of reference lists were also conducted. Specific databases and journals that focus on systematic reviews such as, PRISMA, the Cochrane Systematic reviews data base and Systematic Reviews Journal were also searched.

*Findings*

The five frameworks we identified are summarised in the web appendix. Only one quality appraisal framework was identified that focused on gender/sex (Doull et al., 2014). An additional four quality appraisal tools considering questions of equity in health were also included as gender was one of the variables they were interested in capturing (Alshamsan, Majeed, Ashworth, Car, & Millett, 2010; Burstrom, Nylen, Clayton, & Whitehead, 2011; O'Neill, 2014; Welch et al., 2012). Notably, there were many quality appraisal tools that did not reference social constructionism or particular considerations of gender despite being focused on health disparities between women and men. See: the Critical Appraisal Skills Programme, (Taylor, das Nair, & Braham, 2013; Vogel, 2013); Quality Assessment Tool For Cohort Studies (Brown, Elsass, Miller, Reed, & Reneker, 2015); or Joanna Briggs Institute Extraction Tool,(Sudhakar & Admassu, 2013).

The search corroborated the findings of previous work that emphasised the disparate nature of tools used to assess evidence about gender (Einstein & Shildrick, 2009; Mackenbach, 2003). Our search identified an increased enthusiasm for looking at inequities in health, most notably by the equity ‘extension’ of already highly subscribed to tools for systematic reviews such as PRISMA and Cochrane. However, because such tools aim to cover a variety of
variables that contribute to health inequities, they provide a limited ability to engage on a
deep level with any particular variable such as gender or ethnicity. Such tool’s breadth of
analysis also requires reviewers to have a more wide-ranging knowledge of the axis of
oppression, than is perhaps feasible given the scope of any given systematic review.

In response to these identified deficiencies with existing tools, Welch et al. developed a set of
briefing notes to address sex and gender in systematic reviews (Welch et al., 2014). This
useful tool provides an outline for thinking about gender in every aspect of the research
process, thus providing a well needed addition to the systematic review literature.
Nevertheless, the tool appears to relegate qualitative studies to a secondary category by
regarding them as merely ‘additional analyses’, which contradicts a strong body of feminist
literature that has specifically emerged to emphasise the merit of qualitative research for
gaining insight and understanding women’s experiences (Hesse-Biber & Leavy, 2007;
Oakley, 1998; Olesen, 2011; Sarantakos, 1998; Westmarland, 2001). Also missing from
Welch et al. was a category that extended beyond simply outlining gaps in the research, but
rather assessed the extent to which the mooted measures for improvement and change would
actually contribute to tangible and effective solutions. Finally, it was difficult to determine
the overall quality of any given study because there was no summative grading system
provided.

**Principles underpinning the development of the Feminist Quality Appraisal Tool**

In order to address the gaps in the current review tools, we have taken an explicitly feminist
approach which involves a central focus on questioning power and gender, while also
acknowledging that the “relationship between gender and power is not straightforward”,
(Hester, Donovan, & Fahmy, 2010, p. 1364). In this way we see our tool as analogous to a
critical realist perspective that sees science and the relational ‘truths’ it espouses as part of a
“social process” itself (Bhaskar, 1989), albeit with added focus on the position of women in
society (Gringeri, Wahab, & Anderson-Nathe, 2010; Landman, 2006). This particular focus
on women is essential considering that the human sciences have historically lent themselves
to an array of social restrictions on women’s lives and bodies under the guise of ‘objectivity’
(Springer, Mager-Stellman, & Jordan-Young, 2012). Therefore taking a feminist perspective
that is sensitive to the often gendered relationship between power and knowledge will enable
reviewers, and ultimately policy-makers, to challenge gendered stereotypes that re-inscribe
disparate health outcomes. Such a perspective will be useful for highlighting the way that
gender influences every aspect of the research process (Landman, 2006; Oakley, 1998), and emphasising the need for practical implications to ameliorate gender inequities in health (Hammarström, 1999).

Furthermore, a feminist perspective underlies this framework’s definitions of “gender” and “sex”. Gender is understood to mean “the social, cultural and symbolic construction of femininity and masculinity in any given society” (Hammarström, 1999, p. 242), and sex as the “male/female differences attributed to biological characteristics, such as chromosomes, physiology and anatomy” (Nieuwenhoven & Klinge, 2010, p. 314). While we recognise the interconnectedness of gender and sex, this framework remains predicated on gender as a discrete category because we believe that inequities in public health are overwhelmingly predicated on the basis of culture not biology (West & Zimmerman, 1987).

Before presenting the explanation of our feminist quality appraisal framework, we wish to acknowledge that a reviewer need not personally identify as feminist in order to apply the template. However, a personal commitment to the equitable provision of health care, as well as a broader belief in the equality of all genders in society, is essential to successfully conducting a critical appraisal.

The development of the Feminist Quality Appraisal Tool

The Feminist Quality Appraisal Tool is presented in Table 1. A general inductive approach (Thomas, 2006) was adopted to develop the appraisal tool which enabled Tessa Morgan (TM) to draw on the extant frameworks above as the raw data for formulating the novel feminist template (Bryman, 2008). Lisa Williams (LW) and Merryn Gott (MG) subsequently reviewed the framework and in discussion with TM refined the template. Outlined below is an explanation of each category within the appraisal tool which includes: study design, data collection and analysis, how gender is contextualised in the discussion, effective recommendations for change and the overall quality of the feminist analysis. A discussion of the application of the tool to our recently published systematic review on gender and family caregiving at end of life (Morgan, Williams, Trussardi, & Gott, 2016) is also presented in order to highlight the value of the tool, as well as to show how we adapted our tool slightly in the process of application.

How to measure quality
Rather than quantifying quality in numeric terms (Alshamsan et al., 2010; Brown et al., 2015), the feminist quality appraisal tool aims to determine whether a study accomplishes a thorough, moderate or cursory feminist analysis. By appraising studies in this way, the tool emphasises the subjective nature of the appraisal process, which is ultimately contingent on a series of personal choices made by the reviewer (Kaptchuk, 2003). One commonly cited way to minimise ‘bias’ within this inherently impressionistic exercise is to have two or more reviewers appraise studies (Scholosser, 2007; Tricco et al., 2008; Uman, 2011) and then confer on their findings. Regardless of the personal bias, however, it remains important to make some kind of summative determination about quality so that researchers and policy-makers can explicitly determine the value the reviewer has placed on their included studies (Alshamsan et al., 2010). For the framework presented in this paper:

- A ‘thorough’ feminist analysis would consider gender and power in relation to the four substantial categories of the framework: conceptualisation, study design and analysis, contextualisation of findings and practical prescriptions for change.
- A ‘moderate’ feminist analysis would satisfy 2-3 of the above categories.
- A ‘cursory’ feminist analysis would satisfy 1 of the above categories.

**Conceptual underpinnings of study**

The reviewer’s first step should be to look at whether there are any obvious sign postings of the authors’ conceptual underpinnings, by which we mean the epistemologies they are influenced by; bearing in mind that these may be explicit and/or implicit. Reviewers may look at the methods section to see whether a subsection such as ‘conceptual underpinnings’ or ‘methodology’ is included. If there is no explicit discussion of the conceptualisation of the study, which is often the case, reviewers may investigate whether “topics or methods centred on women were concerned with social justice or liberation, focused on gender and gender-related issues, or critically explored or engaged issues of power and praxis” to see whether feminist aspects have been included (Gringeri et al., 2010, p. 391).

Reviewers should then look to see whether there has been a definition of gender provided. Gender may be characterised in research as a social variable, a system, a social construction, and/ or a political tool (Jarviluoma, Moisala, & Vilkko, 2004). If it simply provides a description of ‘sex’, reviewers should note the lack of conceptualisation of gender. Reviewers may also come across definitions that refer to sex/gender, whereby they should carefully note
whether this term has been used to define the intersection of the two concepts, or whether they have been conflated to mean biological sex. Reviewers should also pay close attention to the aims section to determine whether gender lies at the heart of the research project, as well as to the reference lists, and specifically to the papers used in the background and discussion sections, in order to see what theory has been used to support the central argument.

Data collection and analysis

In quantitative, qualitative and mixed-method studies it is important to think about how ideas of gender and the corresponding hierarchies of power are likely to affect both the researcher and the participant, and thus the kinds of data collected and analysed (Oakley, 1998). Consequently, reviewers should first consider the extent to which the authors have been reflexive in their data collection and analysis. They achieve this by looking for an awareness of their own subjectivities and whether they have considered them in a transparent way so that the reader understands their positionality relative to the study/participants (Gringeri et al., 2010; Hesse-Biber & Leavy, 2007; Sutherland, Ward-Griffin, McWilliam, & Stajduhar, 2015).

It is similarly important that reviewers consider how aspects of gender may be captured within aspects of the data and analysis stages. This process extends beyond using sex desegregated data, (Klinge, 2008; Nieuwenhoven & Klinge, 2010) to how participants as gendered subjects may respond to research questions, and to how surveys/interviewing conditions may further influence the nature and extent of a participant’s disclosure (Spierings, 2012). Reviewers should therefore examine how studies have considered mitigating factors, for example, the stereotype that men disclose less than women, (Fromme et al., 2005) as well as the way in which the genders of the interviewer/interviewee will impact the outcome of the interview process (Gringeri et al., 2010; Hammarström, 2007; Jarviluoma et al., 2004; Latu, Mast, & Stewart, 2015; Tang, 2002). Reviewers should also look at whether any given study design has allowed space within their model to describe the evolution of gender inequities in health care, and whether a study has included subgroup and sensitivity analyses (Higginbottom et al., 2013; Langford, 2015; Scott, 2010; Welch et al., 2014).

Gendered context in discussion section
Reviewers ought to be mindful of the ways in which the findings of any given study have been situated within their specific gendered context (Ballantyne, 1999; Macintyre et al., 1996) and particularly in relation to the overarching norms of femininity and masculinity (Lahelma, Arber, Martikainen, Rahkonen, & Silventoinen, 2001). Reviewers should look at whether there has been any discussion of socialisation or expectations and/or duties associated with being a particular gender. Given the multiplicity of gendered contexts, and the increasing plurality of many societies, reviewers ought also to be wary of over-generalisations, particularly those that portray gender norms as unchanging, (Freixas, Luque, & Reina, 2012; Lahelma et al., 2001).

Second, because gender is relational and co-constructed, it is useful for reviewers to consider how femininities and masculinities have been considered in every study, even if the particular study under review only focuses on a particular gender (Calasanti, 2004). For example, reviewers could explore whether conceptions of femininity may affect men’s reluctance to engage in caregiving activities. Gender is also relational in the sense that it intersects with other social statuses, (Hankivsky & Christoffersen, 2008; Springer et al., 2012) and as such reviewers should note, if relevant, how studies have considered intersecting factors such as ethnicity, race, age, sexuality and physical ability.

**Effective recommendations for change**

Informed by the feminist commitment to social justice and the amelioration of social inequalities, reviewers should look closely at whether the implications for future research, practice and policy that each study has provided are sophisticated enough to lay the basis for practical and equitable change (Einstein & Shildrick, 2009; Hammarström, 1999). To do so, reviewers ought consider whether prescriptions for change have been provided on both an individual level – including ways in which women and men can make changes in their own lives – and on a structural level – how government policy, as well as cultural norms, may be altered (Roen, Arai, Roberts, & Popay, 2006). In particular, reviewers should note who the recommendations are targeted to and consider how different genders have been factored into the solutions they have provided.

**The application of the Feminist Quality Appraisal Tool in a Systematic Review: an example**
We applied the feminist quality appraisal tool to our systematic review titled *Gender and family caregiving at end of life for adults over 65* (Morgan et al., 2016) and in the process refined the tool. In this section we will provide reflections on this application process to the 19 studies we identified as fitting our review’s criteria.

First, we found that very few papers provided an explicit outline of their conceptual underpinnings, and that no study in our review provided a definition of gender (Calasanti & King, 2007; Chappell, Dujela, & Smith, 2015). Interestingly, in one case we found that a study that purported to be set within a feminist paradigm was only ‘cursory’ in its feminist quality overall; the researchers were reluctant to critically assess the power structures inherent between women (Read & Wuest, 2007).

Second, as a means of avoiding considerable overlap, we amalgamated the data collection and analysis sections because they both addressed concerns about how to design methodology to capture gender. We found that there was an increasing emphasis on sex-desegregating data and the use of sub-group analyses along the lines of gender. A few studies, particularly those looking to problematize masculinity, even experimented with effective ways to best elicit male responses in spite of the tendency for men to under-report (Fromme et al., 2005). Others looked at intersections of gender and other relevant characteristics such as age and marital status (Campbell, 2010; Chappell et al., 2015).

Third, this tool highlighted the discrepant use of context within the studies. For example, one study that included gender in its title devoted only a single line to the influence of normative ideas of gender on the disparate outcomes they identified in their study (Navaie-Waliser, Spriggs, & Feldman, 2002). In contrast, other studies produced an integrated discussion of gender that stemmed from formulation of the study to its findings (Calasanti & King, 2007; Campbell, 2010).

While all studies provided at least one direct and/or practical solution to address the problems identified within their study, particularly in relation to service provision and the need for more research on a particular topic (Brazil, 2009; DiGiacomo, Lewis, Nolan, Phillips, & Davidson, 2013; Mystakidou et al., 2013), few studies were concerned with wider structural and cultural concerns (Calasanti & King, 2007; Fromme et al., 2005). One neglected to mention gender in their prescriptions (Govina et al., 2014).
Overall we found that nine of the studies identified had a thorough analysis. Seven of these were qualitative, one quantitative and one mixed-method. In particular, we found that the qualitative studies were more likely to situate their findings within a context as well as provide and reflect on their theoretical underpinnings. This is significant because quantitative studies tend to score higher in most standard quality appraisal tools, especially in terms of wider generalisability, which is specifically not a focus of this tool.

**Strengths and limitations**

While we have discussed this review in relation to the health literature, we think that this tool could be applied beyond the health sciences. We also think that this tool can be used to supplement other quality appraisal tools, as a way to promote a critical reflection about the relationship between gender, power and research, which will enhance the overall judgements of any given systematic review.

One key limitation to this framework pertains to the binary use of gender. While our review tool is designed to discuss the multiplicity of genders and gender expressions, which include non-binary genders such as trans* people, our application of the framework to the family and end of life caregiving literature inadvertently privileged cis women and men because of such literature’s exclusive focus on cis gendered people. We are confident, however, that our framework is flexible enough to incorporate non-binary genders and indeed promotes reviewers to think about and embrace gender fluidity.

A final limitation is that our literature search for critical appraisal tools, and additional research, were restricted to English language sources, which means that we may have overlooked evidence and/or review tools published in languages other than English.

**Conclusion**

The Feminist Quality Appraisal Tool offers researchers a means to analyse the ways in which gender is addressed (or not addressed) in research they are reviewing. Feminist approaches typically explore issues of power and social context. Therefore, using the tool to dig deeper into gender will help to expose ways in which ‘objective’ health research is inscribing inequities in public health into policy and practice. As a result, this tool has the potential to improve public health by providing solid understandings and critical reflections on the reasons women continue to face barriers in their access to optimal health care.
Acknowledgements

We are grateful for Gabriella Trussardi’s helpful suggestions in regards to the framework.

References


