The inclusion of quantitative techniques and diversity in the mainstream of feminist research

Niels Spierings

Radboud University, The Netherlands

Abstract

Much is written about quantitative techniques and feminist and gender studies. Despite convincing arguments in favour of utilizing these methods, they are still largely absent in the heartland of gender studies. In this article, this is related to the observation that methods are tied to epistemological positions and consequently quantitative studies are a priori associated with overgeneralization. A new perspective – the diversity continuum – is presented in order to contextualize research and make it possible to judge it relatively. This shows that quantitative research can increase the understanding of differences and similarities. Subsequently, it is discussed how quantitative research can be utilized to understand differences between people using existing basic techniques. In line with the conceptualization of diversity, this includes techniques to study intersectionality and how developments have different effects on men and women. The continuum and techniques are illustrated by an example of political interest (in the Dominican Republic).

Keywords

Diversity, feminist methods, gender studies, intersectionality, quantitative research, regression analysis, statistics

. . . sometimes these things should be judged by distance traveled rather than by current position.[1]

Introduction

In the last 15 to 20 years, several scholars have convincingly argued that quantitative techniques can make valuable contributions to feminist endeavours and gender

Corresponding author:
Niels Spierings, Radboud University, Institute for Management Research, Departments of Political Science and Economics, Thomas van Aquinostraat 5.01.78, PO Box 9108, 6500 HK Nijmegen, The Netherlands
Email: n.spierings@fm.ru.nl

studies (see Harding, 1997; McCall, 2005; Maynard, 1994; Mazur and Goertz, 2008). In many social sciences where the dominant discourse includes statistical methods, quantitative methods and gender studies are increasingly integrated. Journals such as *Feminist Economics*, *Psychology of Women Quarterly* and *Women's Health Issues* often publish articles in which quantitative techniques are used. The situation is much different in the mainstream of feminist and gender studies.[2] In this mainstream – the institutes, journals and conferences that primarily have a gender studies identity – quantitative research still seems to be the ugly duckling, or at least the lonely duckling. In February 2009 an international conference on feminist research methods accommodated 69 paper sessions, of which only one was on quantitative techniques;[3] feminist methodology books include a critique of quantitative methods at most (e.g. Harding, 1997; Ramazanoğlu and Holland, 2002; Reinharz, 1992); and of all research articles published in five of the most highly ranked multidisciplinary women's, gender and feminist studies journals in 2008, only 11% included some application or discussion of quantitative techniques (see Table 1).[4]

Despite these numbers and personal experiences in gender studies arenas, I have had many discussions with scholars who deny the absence of quantitative studies in gender studies, which would imply that the argument presented above is a straw woman. To be clear: I do recognize that there is discussion on the use of these techniques, that in interdisciplinary environments more quantitative studies might be present and that feminist researchers recognize the importance of quantitative research. My argument here is only that, despite all of this, quantitative techniques are scarcely used in the core of gender studies.

In what follows, I first argue against the claim that quantitative equals bad, and that quantitative feminist research is practically impossible. Next, drawing on McCall (2005), I argue that in mainstream gender studies, quantitative research is still *seen as* bad. McCall tries to create a common understanding by distinguishing and labelling three different ways in which intersectionality is approached, each with its own value. The way she

Table 1. Quantitative techniques in general women's/gender/feminist studies journals

Journal[a]	Year	N total[b]	Univariate or bivariate descriptives (%)	Multivariate techniques (%)[c]	Other attention for quantitative research (%)[d]	Total qnt (%)
JGS	2008	21	3 (14%)	1 (5%)	1 (5%)	3 (14%)
G&S	2008	29	4 (14%)	4 (14%)	0 (0%)	5 (17%)
EJWS	2008	16	1 (6%)	0 (0%)	0 (0%)	1 (6%)
WSIF	2008	48	4 (8%)	1 (2%)	0 (0%)	4 (8%)
SIGNS	2008	27	0 (0%)	0 (0%)	2 (7%)	2 (7%)
Total		141	12 (9%)	6 (4%)	3 (2%)	15 (11%)

[a]*Journal of Gender Studies (JGS) Gender and Society (G&S); European Journal of Women's Studies (EJWS); Women's Studies International Forum (WSIF); SIGNS: Journal of Women in Culture and Society (SIGNS)*; these journals are top ranking in the Web of Knowledge social science citation index category of 'women's studies' and do not aim at a specific discipline (such as health, geography or economics).

[b]Only research articles are included, contributions labelled as book reviews, forums, editorials, introductions and such are not taken into account.

[c]Examples are factor analysis, OLS regression analysis and logistic regression analysis.

[d]e.g. chi-square, discussion of mathematics or the quantitative/qualitative divide.

frames it, however, does not lead to a true continuum and still ties quantitative techniques to the bad 'other'. I offer a way out of the diagnosed deadlock by reconceptualizing the debate in terms of relative diversity. The utility of this approach lies in contextualizing particular studies in the existing knowledge and practices of a field. Subsequently, I discuss how existing quantitative techniques can be easily applied to understand the diversity in people's positions, illustrated by an example concerning political interest.

The lack of quantitative methods in feminist research

There are at least four possible reasons why quantitative methods are missing in the mainstream of gender studies: (1) quantitative research is bad; (2) the practical hurdles are too high; (3) quantitative research is 'not done'; and (4) it is unclear how to do good/ better quantitative gender research. Before working to a new perspective based on the last two reasons, this section briefly discusses the (historical) context of the discussion by looking at the first two reasons.

It is bad

As Gerring (2007: 11) observes, 'the purpose of a statistical sample is to reveal elements of a broader population'. In the 1970s and 1980s it was exactly this practice of generalization that was heavily criticized by feminists: assuming homogeneous societies, ignoring important differences between the position of men and women, and assuming that results concerning one half of the population (men) also hold for the other half (women) (Bernard, 1975; Harding, 1986; Jayaratna, 1983; Millman and Kanter, 1975; Oakley and Oakley, 1979). This idea is part of a larger 'war' between 'the two cultures' (qualitative and quantitative) in the social sciences (King et al., 1994: 3; Mahoney and Goertz, 2006), and steers feminist empirical research away from deriving general statements and using quantitative techniques.

However, in the current mainstream of gender studies, many take the position that it is possible (and even desirable) to also make generalizations about (groups of) persons. This resonates with (feminist) scholars who have argued convincingly that 'bad' is not synonymous to quantitative (see Andermahr et al., 1997; Harding, 1997; McCall, 2005; Sprague, 2008). The problem of bad quantitative research is not that it is 'quantitative', but that it is 'bad', implying that good quantitative research is possible as well, and some authors have indeed formulated ways to treat quantitative results carefully (e.g. McCloskey and Ziliak, 1996; Miller and Van der Meulen Rodgers, 2008). Having said this, the observation remains that quantitative techniques mainly feature in the mainstream of gender studies as an object of critique, not as a practice. As a consequence, I suggest that 'badness' is not inherent to quantitative research, but the 'badness' label still sticks.

It is practically impossible

A cluster of more practical arguments forms the second reason for the absence. Quantitative research models that include high levels of social complexity (as a good gender studies

study would) require large and rich datasets (see also Warner, 2008), highly technical models, and (top peer-reviewed) journals that accept lengthy explanations (cf. Sprague, 2008). These difficulties might be particularly problematic for gender studies scholars who focus on many different identities and their intersection (see McCall, 2005: 1787).

This argument is problematic in two ways. First, good research is equated to ideal type research. However, no matter what data or methods are used, no quantitative or qualitative study will be perfect. Knowledge will grow step by step. Second, regarding the journals, it is assumed they are ruled by 'outsiders' with their own norms (as suggested by McCall, 2005). True or not, this does not explain the observation that opened this article. As Sprague stresses, 'very high quality' gender studies journals 'with a broad readership' have been created (Sprague, 2008: 702), and these are – according to their aims and scopes – open to complex, lengthy and quantitative articles. I therefore conclude that practical impossibilities are not the problem, but that either the supply of quantitative studies is low or the review hurdle is disproportionally high. Both might be attributed to an aversion to quantitative work.

From 'bad' to 'better'

Thinking 'it is bad'

I suggested above that in the culture of mainstream gender studies quantitative research still tends to be regarded as bad. In this section, I argue that this is due to thinking in binary pairs instead of scientific progress. McCall's (2005) recent and influential attempt to build a bridge between different methods is my starting point, because it frames the problem aptly around categorization, but it does not resolve the core problem: positional thinking. McCall discusses three approaches to intersectionality in terms of categories (in other words generalization or diversity), and argues that quantitative research can contribute significantly to research on intersectionality.

The first approach – anticategoricalism – is highly critical of categorization because dividing people into groups is the fundament of power relations. Categorization entails inclusion and exclusion based on non-reality, because all categories are an artifice of language, and it denies the multiple identities each individual has (McCall, 2005: 1777–1779). Anticategoricalists set out to attack existing stereotypes and deconstruct categories. Consequently, large restrictive essentialist identities are broken down and multiple and fluid identities emphasized. The second group – intracategorical research – does not reject categorization, but complicates existing boundaries. In particular, seeing women as a single unitary group has been heavily criticized (McCall, 2005: 1778–1780). Scholars have refined this category, by constantly asking which identities are important for a particular social structure and the related power relations. People in different positions have different experiences, which should not be ignored or overgeneralized. This view focuses on diversity within a group and the role of multiple categories. The last approach – intercategoricalism – starts from the notion that in terms of power relations social groups do exist, so the research is category-based. These power relations are studied and explicated,

which 'requires the provisional use of categories' (McCall, 2005: 1785). This includes the study of how categories change and interact. In the intercategorical approach in particular all sorts of statistical techniques and models can be applied (McCall, 2005: 1784–1788).

Together these three traditions are the main building blocks of McCall's 'continuum of complexity management' and it seems that by showing how the different approaches try to manage complexity, McCall tries to create a common understanding (e.g. McCall, 2005: 1775). However, the continuum is built around three (archetypical) groups of research. This means that it is not a continuum, but a set containing three positions which resonate strongly with the epistemologies Harding (1986, 1997) distinguishes: postmodernism, standpoint feminism and feminist empiricism (see Figure 1).

McCall herself acknowledges this and also hints at its consequences: 'in the anticategorical and even intracategorical camps' 'advanced quantitative techniques and large data sets . . . [are] associate[d] negatively with the legacies of positivism, or empiricism' (McCall, 2005: 1791) (see Figure 1). However, the classification of intersectionality research strengthens this binary thinking. This is especially so because McCall mostly stresses the distinction between intercategorical research on the one hand and the intra- and anticategorical approach on the other. From the dominant position in the mainstream of gender studies (the standpoint view and postmodernism) the negative association with quantitative research then remains (see Figure 2).

In sum, McCall's classification of research in different groups enables her to show how different types of research can contribute to research on diversity, but it does not transcend the positional thinking and the negative attitudes towards quantitative research. Because research methods are tied up with epistemological positions it becomes very difficult to change these attitudes. Reconceptualizing the debate in terms of *diversity* and *judging relatively* might shed new light on the issue.
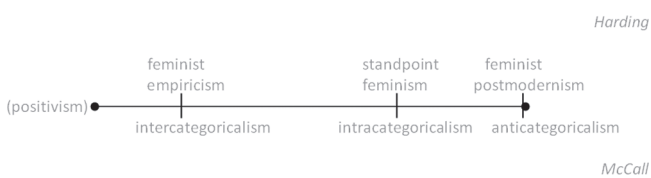

Figure 1. A continuum of complexity management and epistemological positions

Figure 2. Judging quantitative studies

Judging whether it is better

Thinking about feminist research in terms of diversity might provide us with a conceptual tool for creating a true continuum. One of the two extremes on this continuum is infinite diversity: each and every individual is unique in every aspect. The other end of the continuum is no diversity at all. Each and every individual is a specimen of one single larger group: what holds for one of them holds for all. Between these two poles, the level of diversity differs and thereby the degree of generalization differs as well.

The concept of diversity in this approach has two dimensions: the basis of diversity and the type of diversity. The basis of diversity is the dimension on which (groups of) people are distinguished: the identities that are considered relevant, such as gender, sex, class, ethnicity, race, age, sexual orientation, religion and able-bodiedness (see e.g. Lorber, 2006; Phoenix and Pattynama, 2006; Verloo, 2006; Yuval-Davis, 2006). On the continuum these are not a priori hierarchically ordered: taking one into account is not by definition more diverse than taking another into account. However, distinguishing between three or more sexes is more diverse than distinguishing between two, and distinguishing between three sexes and several ethnicities is even more diverse (ceteris paribus).

For the second dimension – the type of diversity – a similar argument holds, but there is a technical hierarchy. First there are differences in the *position* of (groups of) people based on their identities, and second, (groups of) people are differently *affected* by phenomena because of those identities. A prerequisite for understanding such an effect is establishing the position of different groups of people. Whether education has a different effect on women's and men's employment status implies knowing their employment status.

The usage of the concept of diversity as described above leads to a two-dimensional continuum (Figure 3a), on which feminist as well as androcentric research can be placed. Each study occupies a unique position – has its own degree of diversity – and hardly any study is in an extreme corner of the concept.

Placing studies on such a continuum shows how a study relates to other studies – the state of the art in the particular disciplinary field (indicated by black circle in Figure 3b).
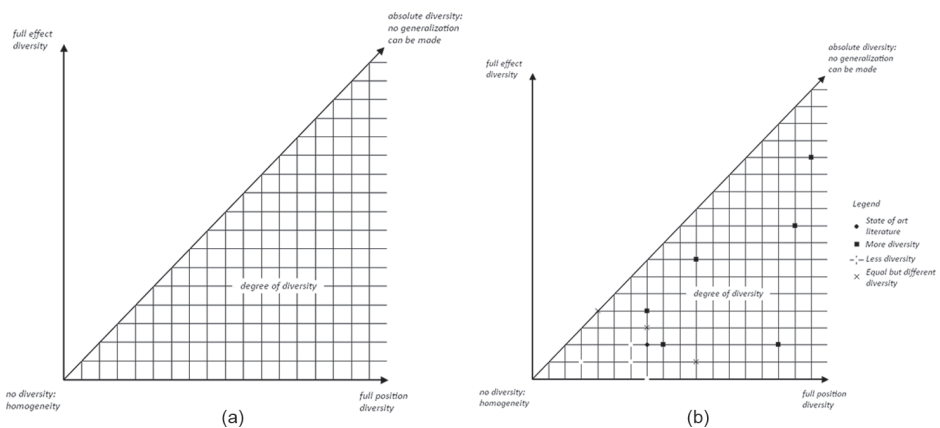

Figure 3. (a) A two-dimensional concept of diversity (b) Relative positioning on the diversity continuum

By adding one new distinction that leads to an understanding of inequality, progress is made in that field (black square in Figure 3b) if judged relatively. The focus now shifts from judging something as being 'bad' or 'good' to judging it as being 'better' or 'worse'. In that sense, differentiating between women and men in medical research might be far more progressive or feminist (cf. Hankivsky and Christoffersen, 2008: 272; Nieuwenhoven and Klinge, 2010) than studying political representation differences between straight and gay men and women in the US Black and Latino communities.

Applying the diversity continuum shows that quantitative techniques offer gender scholars a means to increase our knowledge of differences, make scientific progress and simultaneously enable scholars to relate to existing knowledge outside gender studies in a way that makes communication between the different fields possible.

Quantifying diversity

Based on the perspective taken in this article, it is certainly possible to make progress and accommodate diversity using existing statistical techniques, but we need an overview of available options. Hankivsky and Christoffersen (2008) for instance discuss a new theoretical approach that includes intersectionality in health research, but methodologically and empirically they end their contribution with the observation that 'there is little empirical work that examines intersectional differences between multiple social categories, and models and methods that are able to measure and simultaneously investigate multiple intersectionalities are only beginning to develop' (Hankivsky and Christoffersen, 2008: 278). While sharing their overall observations and line of thinking, I disagree with them on the last point: existing (basic) methods and models are already sufficient for tackling many empirical questions. For instance, the work of Sen et al. (2009), also in the field of health, illustrates this: they use quantitative techniques to analyse the effects of the intersection of class and gender on health inequalities. However, a more general overview of basic possibilities is missing.[5]

Differences in positions

This section focuses on the options for distinguishing groups, describing these groups and drawing inferences for these groups from a background of existing data and basic quantitative techniques (Table 2). Some caveats are discussed at the end.

To understand the different positions of different groups of people it is necessary to use a classification. Although large comparative datasets such as the European Social Surveys, World Values Surveys and Programme for International Student Assessment data often do not include the fine-grained sets of questions that are, for example, discussed by Browne (2008) regarding sexuality, or Völke et al. (2010) on ethnicity/race/skin colour/ancestry, current techniques and data still provide several opportunities. First, virtually all surveys provide binary information about the sex of the respondent. This information can be used to create a variable that compares women and men. Second, most surveys provide data on some other identities. These can be used to understand differences between other identities – preferably in as many

Table 2. Techniques and different diversities in quantitative methods

	Classifying	Describing	(Cor)relating and explaining
Position	• Distinguish between sexes *create a nominal sex variable* • Use the most detailed classification possible *interval > categorical > nominal AND n+1 > n categories* • Check for differences on other identities *create variables for other identity variables* • Use intersectional identities *cross the identity variables so that all subgroups are uniquely classified*	• Give descriptives per group and its subgroups *calculate the measures of centrality – mean – for all distinguished groups* • Describe not only centrality but also variability per (sub) group *include the standard deviation and minimum & maximum or interval(like) variables and the frequencies on nominal and categorical data*	• Test for structural differences in position *include identity variables in multiple regression models* • Let the different identities intersect *include interaction terms of 'crossed dummies' so that each group has its own regression coefficient* • Study whether characteristics that are identity-related are responsible for the found differences *include additional control variables that are for instance gender informed* • Include societal explanations when studying individual positions *use multilevel models that include individual and societal characteristics*
Effects			• Test whether effects differ for different (sub) groups of people, at least by sex *include interactions terms of each explanatory variable with the identity variable* • Test for 'intersectional' differences in effect *include three-way interactions or 'crossed dummies'* • Test whether societal characteristics affect different individuals differently according to their identity *include cross-level interactions*

categories as possible – and subsequently a combination of identity variables can be used to create an 'intersectional' identity matrix. If I, for example, want to know the position of a black woman, I should realize that this position is probably more than the sum of the two parts: the difference between being black and white might be different for men and women (e.g. Hancock, 2007; Wekker and Lutz, 2001). Consequently, I have to distinguish four groups.

When describing the characteristics of groups, researchers should not only provide figures on the population as a whole, but also for each of the subgroups. Furthermore, besides measures of centrality/central tendency (such as the mean, median and modus), the differences within a group can be provided by including measures of variability/dispersion (such as the standard deviation and the minimum and maximum) for interval variables. These latter measures show the heterogeneity within a group: how precise the mean is and how many people are close to it.

In order to draw inferences about the relationship between sex, gender and other identities and outcomes such as income, participation and health, intersectional positions can be included relatively easily in regression models. First, all relevant and available identity variables such as sex, gender and ethnicity can be included as a regular explanatory variable or dummy (Hardy, 1993; Pollock, 2009). The advantage of this over the techniques discussed above is that it is possible to take into account differences that correlate with these identities (see Allison, 1999). For instance, the wage gap between men and women will then be a net difference because differences caused by occupational segregation and working experience are filtered out. Second, intersectionality can be incorporated by crossing identities. This can be done by either using specific dummies in the case of ordinal or nominal variables, or by interaction terms for any type of variable (see Jaccard and Turrisi, 2003; Pollock, 2009). For instance, Almeida-Filho et al. (2004: 1347) study how the interaction of gender, ethnicity and social class influences the prevalence of depressive disorders in Bahia (Brazil) and find that among upper-middle-class respondents no notable gender differences were found, regardless of ethnicity. Among poor and working-class respondents, women were found to have a higher prevalence of depressive disorders, but not among whites.

Third, non-identity factors that are socially attached to an identity can be included in statistical models. For instance, besides sex, variables tapping into gender norms and structures – such as having children in the case of labour market participation – can be included. If after inclusion of this variable the effect of being male/female decreases, the researcher could conclude that the differences between women and men is partly due to the caretaker role women have in society. Fourth, when studying individuals, researchers can use multilevel models to see whether societal factors have an important role in causing differences (see Jones and Duncan, 1998; Painter, n.d.). For example, besides education levels, having children and marital status, societal factors such as gender equality policies and gendered labour market structures can be very relevant for women's labour market decisions and position (e.g. Spierings et al., 2010).

While these basic and available techniques can increase our understanding of diversity enormously, there are some caveats that should be considered. First, when looking at

the difference between women and men, essentialist conclusions lurk around the corner. A statistical technique only shows a correlation between factors; it says nothing about the origin of the difference or the causal mechanism at work. The tendency in, for instance, mainstream sociology to invoke biology when sex differences are found, as observed by Lorber (2006: 449), serves as a reminder that statistical packages do not resolve the nature–nurture debate. Second, drawing an inference means that the researcher generalizes the results to the population as a whole. This (theoretical) population should be clearly demarcated (see Goertz and Hewitt, 2006). Ordinarily, general conclusions can only be drawn for groups included in the sample that is analysed. Within that sample, one can test for differences in position (see above) or effect (see below). Third, creating subgroups does indeed mean deciding where to draw the line; this is both arbitrary and can create power relations. However, in line with McCall and others, I argue that the use of new and finer categories – for a particular field – is also a way to contest existing ideas and power relations.

Differences in effects

Many factors have a different impact on people with different identities. An obvious example is the effect of having children on men's and women's labour market participation. In order to study these kinds of different effects and draw inferences about them, each explanatory variable in a regression model can interact statistically with (an) identity variable(s). This will show whether the coefficients differ significantly (statistically and substantially) across sex, ethnicity, sexual orientation and so on (see Jaccard and Turrisi, 2003; Pollock, 2009). In other words, it shows whether an effect is gendered or equivalently related to other identities.

For instance, in the case of a dichotomous sex variable and a two-group ethnicity variable, the researcher can calculate interaction effects between the explanatory variable and dummies for each group of people (e.g. black women, black men, white men, white women). In the case of interval variables, three-way interaction can be used (Jaccard and Turrisi, 2003: 43–60). By using these complex structures, intersectionality is transformed from a positional approach into a causal approach: the researcher not only discusses how the positions of the different groups deviate from each other, but also tests whether and how people at different crossroads are affected differently by similar processes. The example given above concerning gender, ethnicity and social class in the prevalence of depressive disorders is basically a three-way interaction.

In multilevel models, the inclusion of interactions allows for testing the different effects of norms, policies and structures on individuals. For instance, Huisman and Smits (2009) show that the number of schools in an area influences the educational participation of girls but not boys. Multilevel modelling might then also create opportunities for understanding intersectionality in terms of the structure/agency or the institutions/individuals (see Ferree, 2009: 87).

The major drawback of including interactions is that it leads to more complex models and requires richer datasets. However, as argued above, given the state of the field, even with existing data and relatively simple models, we can make progress in our analyses of diversity. Furthermore, it is of course possible to train (new generations of) gender scholars to use more advanced techniques. Similarly, it is of course possible to collect new and

better data (e.g. Browne, 2008), among other things, to avoid the risk of 'too many variables, too few cases' (see Landman, 2008: 303).

Example: Differences in political interest

For several decades, an important research focus has been the determinants of individual interest in politics. Research shows that the political interest of women is lower than that of men, and that education increases it (see Inglehart and Norris, 2003; Koch, 1997; Verba et al., 1997). From this starting point, steps in the two diversifying directions (see Figure 3) can be made to increase our knowledge of political interest and sex differences: how do the effects of factors such as education differ by sex (studied by Morgan et al. [2008] for the Dominican Republic) and what are the differences among women (or men)? Let us start with the latter.

Morgan et al. (2008: 42) find that the political interest of men in 2007 is about 2.9 on a scale running from 0 to 6. Women score 2.3. For both groups, the median is 2. Further descriptives are not given, but the Latinobarometer 2007 data show a similar difference – the average political interest in the Dominican Republic is 2.2 (on a scale from 1 to 4),[6] with men scoring 2.3 and women 2.0 – and they permit further investigation in this study. The standard deviation is just over 1, indicating quite a wide spread, which is also shown by the frequencies of political interest (see Figure 4 for women). In addition, Figure 5 shows that Catholic women's political interest is 2.0 and that of Protestant women 2.3; a difference equalling the variation among the sexes. The descriptives for self-reported ethnicity show that the differences between mulato and white people are almost negligible, but that mestizo women have a political interest (1.94) that is about 0.2 lower on average than that of indigenous women (2.16). These results illustrate that very simple statistics (means and standard deviations) show how a category differs internally and that those differences can be just as important as differences across the sexes (Figure 5). Including this form of diversity provides a better view of the phenomenon of political interest, and it constitutes progress in terms of the diversity continuum (Figure 3).
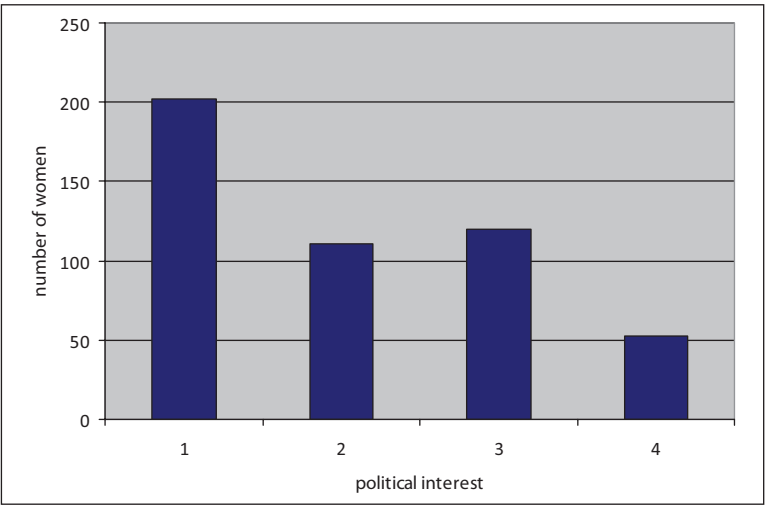

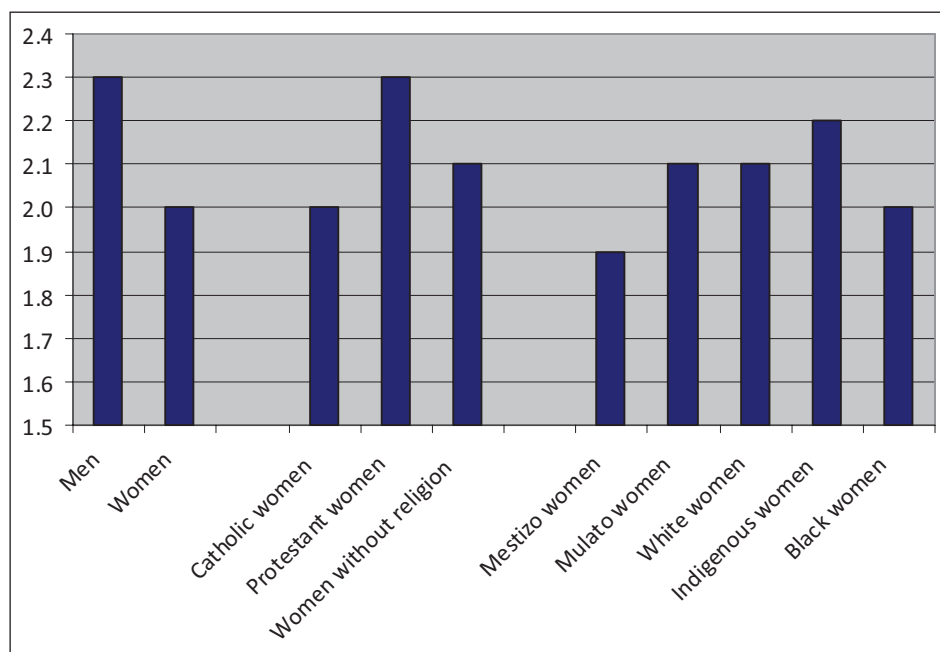**Figure 4.** Political interest among women

Figure 5. Differences in political interest

A second way to increase diversity entails modelling differences in effects as Morgan et al. have done. They include an interaction term for education and gender (see Table 3), and show that women with no education have 0.63 lower political interest than men with no education. On a seven-point scale, political interest rises by 0.33 for men once they are educated; for women this is 0.15 (i.e. 0.33 − 0.18). In other words, education has more effect on men's political interest than on women's and thus increases the political interest gap between men and women. The authors conclude that 'This sizeable difference in the effect of education suggests that while education promotes political engagement overall, there may be something about the education process . . . that encourage[s] men to take more interest in politics than women' (Morgan et al., 2008: 47–48). Their study illustrates how knowledge of the genderedness of effects on political interest is increased by a step in the upward direction on the diversity continuum (Figure 3).

Table 3. Results from Morgan et al. (2008: 45): 'OLS regression: Political interest in the Dominican Republic 2004'

Variable	B-coefficient	Standard error
Female	−0.63**	0.09
Education	0.33**	0.03
Education*Female	−0.18**	0.06

**$p < .01$.
Note: Also included were: evaluation of services, pocketbook evaluation, employed, age groups, interactions between age and gender, socioeconomic status, casada(o)/unida(o), rural.

In summary, applying existing data and techniques allows three important steps to be taken in the field of political interest and diversity. First, it was found that women are generally less interested. Second, differences among women are charted to contextualize the sex difference. Third, differences in the effects of other factors are found for different sex categories. Further progression could be found by including (more) gender(ed) characteristics instead of sex in order to explain why there is a difference, and to interact those aspects with, for example, education. Also, the interaction between identities could be included, such as whether religious differences are different for men and women.[7]

Discussion

Historically, feminist and gender studies have both nourished and been nourished by social movements. Feminist and gender studies has also tried to influence policy-makers, who seem to prefer statistics over open-ended interviews and ethnographies. This is one of the reasons why studying complexity and diversity with statistical tools is important (Kemp and Squires, 1997: 145),[8] and it can serve the goal of feminist research as formulated by Wall: 'to promote change and challenge the structures that oppress women by "documenting women's lives, experiences, and concerns [and] illuminating gender-based stereotypes and biases" ' (Wall, 2008: 221, quoting Brooks and Hesse-Biber).

The opportunities that quantitative research offers for understanding diversity will be acknowledged more easily (in the mainstream of gender studies) once the scientific enterprise is approached from a perspective of relative positions and progress. Each study can be assessed in terms of how it incorporates (possibly relevant) diversities. This can be done in terms of gender and sex, but also with regard to other identities that might be important and intersect with gender. When it is determined to what extent a study generalizes and diversifies, it is possible to assess its relative position: does the study incorporate existing knowledge concerning relevant diversities in this field, or does it show new relevant differences? Quantitative studies can be very useful for studying the degree to which diversity exists empirically.

This approach intends to transcend the current methodological debate that is linked to fierce epistemological debates characterized by static positions (cf. Harding, 1997; McCall, 2005: 1795). Constructing a debate of oppositions often obscures the issue at hand; we need a common language to engage in fruitful discussion.

Of course, for a discussion, we not only need the same terms, but also the same concepts. In this article, diversity is often mentioned as referring to substantial, significant, relevant or meaningful differences and inequalities (see McCall, 2005: 1785). Several scholars have shown convincingly that not all differences are meaningful inequalities and that not all inequalities can be treated similarly (e.g. Phoenix and Pattynama, 2006; Smith, 2009; Verloo, 2006; Yuval-Davis, 2006). What is considered to be *meaningful* diversity? There are many different ways to react to this question. First, quantitative social scientists might focus on statistical significance. Second, Warner (2008: 458), for example, argues that 'it is ultimately an empirical question as to which types of categories explain behaviour, and should be treated as such', which implies that any empirical difference is relevant, but she does not provide criteria that help us to establish when to

consider something as empirically different. Third, McCall (2005) shows that relevance may depend on the goal of a study, but it is difficult to derive specific guidelines from her discussion. Fourth, based on Young (2003), we might conclude that whether a difference is meaningful depends on the group one belongs to. Socially speaking, this implies that for a difference to be meaningful, one group's opinion suffices. My own position is a combination of a (asymmetrical) statistical approach and loosely formulated substantial criteria. If something is statistically significant, it is highly probable that the difference is meaningful, even if it is very small. Nonetheless, such a conclusion still deserves a substantiation[9] and the dispersion of cases should be taken into account (cf. McCloskey and Ziliak, 1996). However, statistically insignificant differences are not by definition meaningless, it depends on the goals set in a study. It is always up to researchers to be explicit about why a difference is meaningful in terms of strength, certainty and social relevance. There is thus at least one important meta-criterion: the judgement about meaningfulness should be substantiated by the author.

Before a judgement about differences can be offered, the differences have to be laid bare. A large part of this contribution discussed existing basic quantitative techniques, such as dummy analyses, interaction terms, measurements of variability and control variables, which all can be used to analyse differences. This does not, however, imply that only quantitative methods should be applied. It is a combination of methods that leads to the most insight, because different methods are complementary, not substitutions of each other. For instance, many critical and qualitative methods will help to formulate new hypotheses about (the shifting importance of) identities, and quantitative studies can study the general patterns regarding these (shifting) identities. Moreover, causal mechanisms can often be studied better by qualitative approaches. In sum, only mixed-methods fields can truly grasp complex phenomena; individual studies do not have to have mixed-methods designs, but scholars studying a certain phenomenon have to be familiar with what is done on that topic by scholars applying different kinds of methods.

In that respect, the starting point of this contribution was the absence of quantitative work in the mainstream of gender studies, but many of the observations discussed here also apply to incorporating knowledge about gender or feminist methods into quantitative dominated fields (as fostered by feminist scholars in the disciplines and multidisciplinary arenas). The suggestions and guidelines formulated in this contribution might inspire the gender studies community to do more quantitative research, and serve as a possible checklist or reminder for people who want to do quantitative research that pays attention to (some form of) diversity or intersectionality.

Funding

Notes

1. Davidson A (2008) *The Gargoyle*. Edinburgh: Canongate, p. 145; this book brings different worlds together.
2. 'Gender studies' and 'feminist studies' are often used interchangeably, and journals that once focused on feminist research are now using the term 'gender' for their subject matter

(e.g. *Women's Studies International Forum* and *Gender and Society*). However, there are differences. In this study, I treat feminist research as a subset of gender research: feminist research is gender research, but gender studies do not have to be feminist studies. Whatever I have written about gender studies thus also holds for feminist studies.

3.   Feminist Research Methods – An international conference, Stockholm, Sweden, 4–6 February 2009. The quantitative session was workshop 6: 'Feminism counts: Exploring the relationship between quantitative methods and feminist research'. For more information: www.kvinfo. su.se/femmet09/index.htm.

4.   Even if we consider that women's, gender and feminist studies are grounded in the social sciences and in humanities – in which text analyses are more prevalent and postmodernism more accepted (Sapiro, 1998) – this percentage is low. For instance, in about 50 of the 141 articles (35%), qualitative (in-depth, open-ended, semi-structured) interviews were used.

5.   Not only is the work of Sen et al. not a general overview of possibilities in terms of describing diversity among groups and diversities in effects, but I also believe it is not a basic approach. While they do use existing techniques in developing their method for analysing the intersections between different social inequalities, they have chosen to use a set of different statistical measurements to test for the statistical significance of differences between groups. I would argue that this is overly complicated. Regression analyses including interaction terms should not be dismissed because of the problem that coefficients are only compared to the reference group. One can easily change the reference group and test for differences between other groups. This approach can be used for both intersectional dummies as propagated by Sen et al. as well as for classic interaction terms. It provides only one type of statistical test so results are easily compared and understandable.

6.   The average in the study of Morgan et al. is 2.6. Recalculated to that scale the Latinobarometer average would be 2.4.

7.   A preliminary analysis based on the data used to differentiate among women indicates that the differences between men and women do not differ by religion (Catholicism vs Protestantism) or vice versa. This is based on an OLS regression including the variables 'sex', 'religion' and their interaction term. The interaction was by no means significant ($t = -0.624$).

8.   This might even bring back a new binding factor that can serve as a 'common cause' for feminism as politics, which seems to be lost by too much emphasis on theory and the non-existence of sameness (see Soper, 1997: 289).

9.   To be sure, a statistically significant relationship can be meaningless if it is really small (the question what constitutes 'really small' is ignored here for the sake of argument); however, this is highly unlikely because the statistical significance is strongly related to the size of the found difference. 'Only' when the 'n' of a study is very large (for now I also ignore the issue of what constitutes 'very large') do small differences become statistically significant, but even then that difference is often meaningful. In those cases it is particularly important to compare the impact of that particular factor with the impact of other factors to contextualize the results.

References

Allison P (1999) *Multiple Regression: A Primer*. Thousand Oaks, CA: Pine Forge Press.

Almeida-Filho N, Lesse I, Magalhães L et al. (2004) Social inequality and depressive disorders in Bahia, Brazil: Interactions of gender ethnicity, and social class. *Social Science and Medicine* 59(7): 1339–1353.

Andermahr S, Lovell T and Wolkowitz C (1997) *A Concise Glossary of Feminist Theory.* London: Arnold.

Bernard J (1975) *Women, Wives and Mothers: Values and Options.* Chicago, IL: Aldine Publishing.

Browne K (2008) Selling my queer soul or queerying quantitative research? *Sociological Research Online* 13.

Davidson A (2008) *The Gargoyle*. Edinburgh: Canongate.

Ferree MM (2009) Inequality, intersectionality and the politics of discourse: Framing feminist alliances. In: Lombardo E, Meier P and Verloo M (eds) *The Discursive Politics of Gender Equality: Stretching, Bending and Policymaking*. London: Routledge, pp. 86–104.

Gerring J (2007) *Case Study Research: Principle and Practices.* Cambridge: Cambridge University Press.

Goertz G and Hewitt JJ (2006) Concepts and choosing populations. In: Goertz G (ed.) *Social Science Concepts: A User's Guide*. Princeton, NJ: Princeton University Press, pp. 159–176.

Hancock AM (2007) When multiplication doesn't equal quick addition: Examining intersectionality as a research paradigm. *Perspectives on Politics* 5(1): 63–79.

Hankivsky O and Christoffersen A (2008) Intersectionality and the determinants of health: A Canadian perspective. *Critical Public Health* 18(3): 271–283.

Harding S (1986) *The Science Question in Feminism.* New York: Cornell University Press.

Harding S (1997) Is there a feminist method? In: Kemp S and Squires J (eds) *Feminisms.* Oxford: Oxford University Press, pp. 160–170.

Hardy MA (1993) *Regression with Dummy Variables*. Thousand Oaks, CA and London: Sage.

Huisman J and Smits J (2009) Effects of household- and district-level factors on primary school enrolment in 30 developing countries. *World Development* 37(1): 179–193.

Inglehart R and Norris P (2003) *Rising Tide: Gender Equality and Cultural Change Around the World.* Cambridge: Cambridge University Press.

Jaccard J and Turrisi R (2003) *Interaction Effects in Multiple Regression*, 2nd edn. Thousand Oaks, CA and London: Sage.

Jayaratne TE (1983) The value of quantitative methodology for feminist research. In: Bowles G and Klein RD (eds) *Theories of Women's Studies.* Boston, MA: Routledge and Kegan Paul.

Jones K and Duncan C (1998) Modelling context and heterogeneity: Applying multilevel models. In: Scarbrough E and Tanenbaum E (eds) *Research Strategies in the Social Sciences*. Oxford: Oxford University Press, pp. 95–123.

Kemp S and Squires J (1997) *Feminisms.* Oxford: Oxford University Press.

King G, Keohane RO and Verba S (1994) *Designing Social Inquiry: Scientific Inference in Qualitative Research.* Princeton, NJ: Princeton University Press.

Koch J (1997) Candidate gender and women's psychological engagement in politics. *American Politics Research* 25(1): 118–133.

Landman T (2008) *Issues and Methods in Comparative Politics: An Introduction*, 3rd edn. London: Routledge.

Lorber J (2006) Shifting paradigms and challenging categories. *Social Problems* 53(4): 448–453.

McCall L (2005) The complexity of intersectionality. *SIGNS: Journal of Women in Culture and Society* 30(3): 1771–1800.

McCloskey DN and Ziliak ST (1996) The standard error of regressions. *Journal of Economic Literature* 34(1): 97–114.

Mahoney J and Goertz G (2006) A tale of two cultures: Contrasting quantitative and qualitative research. *Political Analysis* 14(3): 227–249.

Maynard M (1994) Methods, practice and epistemology: The debate about feminism and research. In: Maynard M and Purvis J (eds) *Researching Women's Lives from a Feminist Perspective*. London: Taylor and Francis, pp. 10–26.

Mazur AG and Goertz G (2008) Introduction. In: Goertz G and Mazur AG (eds) *Politics, Gender, and Concepts: Theory and Methodology.* Cambridge: Cambridge University Press, pp. 1–13.

Miller JE and Van der Meulen Rodgers J (2008) Economic importance and statistical significance: Guidelines for communicating empirical research. *Feminist Economics* 14(2): 117–149.

Millman M and Kanter RM (1975) *Another Voice: Feminist Perspectives on Social Life and Social Science*. New York: Anchor Press.

Morgan J, Espinal R and Hartlyn J (2008) Gender politics in the Dominican Republic: Advances for women, ambivalence from men. *Politics and Gender* 4(1): 35–63.

Nieuwenhoven L and Klinge I (2010) Scientific excellence in applying sex- and gender-sensitive methods in biomedical and health research. *Journal of Women's Health* 19(2): 313–321.

Oakley A and Oakley R (1979) Sexism in official statistics. In: Irvine J, Miles I and Evans J (eds) *Demystifying Social Statistics.* London: Pluto.

Painter J (n.d.) Designing multilevel models using SPSS 11.5 mixed model. Available at: www.unc.edu/~painter/SPSSMixed/SPSSMixedModels.PDF.

Phoenix A and Pattynama P (2006) Editorial: Intersectionality. *European Journal of Women's Studies* 13(3): 187–192.

Pollock PH (2009) *An SPSS Companion to Political Analysis*, 3rd edn. Washington, DC: CQ Press.

Ramazanoğlu C and Holland P (2002) *Feminist Methodology: Challenges and Choices.* London: Sage.

Reinharz S (1992) *Feminist Methods in Social Research.* New York: Oxford University Press.

Sapiro V (1998) Feminist studies and political science and vice versa. In: Phillips A (ed.) *Feminism and Politics*. Oxford: Oxford University, pp. 67–89.

Sen G, Iyer A and Mukherjee C (2009) A methodology to analyse the intersections of social inequalities in health. *Journal of Human Development and Capabilities* 10(3): 397–415.

Smith DE (2009) Categories are not enough. *Gender and Society* 23(1): 76–80.

Soper K (1997) Feminism, humanism, postmodernism. In: Kemp S and Squires J (eds) *Feminisms*. Oxford: Oxford University Press, pp. 286–292.

Spierings N, Smits J and Verloo M (2010) Micro and macro-level determinants of women's employment in six MENA countries. *Journal of Marriage and Family* 72(5): 1391–1407.

Sprague J (2008) Sociology: The good, the bad, and the public. *Gender and Society* 22(6): 697–704.

Verba S, Burns N and Schlozman K (1997) Knowing and caring about politics: Gender and political engagement. *Journal of Politics* 59(4): 1051–1072.

Verloo M (2006) Multiple inequalities, intersectionality and the European Union. *European Journal of Women's Studies* 13(3): 211–229.

Völke K, Van der Haar M and Yanow D (2010) Trouble taxonomies: 'Types of women', 'continent of birth', and 'race'. Working Paper, ECPR Joint Sessions, Workshop 15, Category-making and public policy, Münster, 22–27 March.

Wall S (2008) Of heads and hearts: Women in doctoral education at a Canadian university. *Women's Studies International Forum* 31(3): 219–228.

Warner LR (2008) A best practices guide to intersectional approaches in psychological research. *Sex Roles* 59(6): 454–463.

Wekker G and Lutz H (2001) Een Hoogvlakte met koude winden. De geschiedenis van het gender- en etniciteitsdenken in Nederland. In: Botman M, Jouwe N and Wekker G (eds) *Caleidoscopische Visies – de zwarte, migranten- en vluchtelingen vrouwenbeweging in Nederland*. Amsterdam: KIT Publishers, pp. 25–50.

Young RJC (2003) *Postcolonialism: A Very Short Introduction.* Oxford: Oxford University Press.

Yuval-Davis N (2006) Intersectionality and feminist politics. *European Journal of Women's Studies* 13(3): 193–209.