

RESEARCH ARTICLE

Beyond traditional understanding of gender measurement: the gender (re)presentation approach

Russell Luyt*

Department of Psychology, Anglia Ruskin University, East Road, Cambridge CB1 1PT, UK

(Received 7 June 2012; final version received 23 June 2013)

This paper considers different approaches to measuring gender. It critically reviews gender role theorising and describes how this has informed two approaches to measuring gender as an individual phenomenon: gender orientation (the assessment of individual traits) and gender ideology (assessing individual endorsement, and internalisation, of social norms). It is argued here that social constructionist perspectives offer a viable alternative to gender role theory and that these inform an alternative approach to measuring gender as a social phenomenon: gender (re)presentation. This approach assesses group level endorsement of dominant gender representations. Endorsement is not seen to reflect individual traits or internalised social norms. Rather, it is understood as a social practice, made meaningful through shared understanding of dominant gender representation. This approach is introduced through a critique of the traditional concept of attitudes and a reformulation thereof. The practical measurement implications and benefit of this reformulation are outlined.

Keywords: gender ideology; gender orientation; gender representation; gender role; social constructionism

Introduction

Attempts to measure gender have paralleled theoretical shifts in its understanding. This body of work is arguably most often associated with psychology due to its long-standing and well-established interest in the quantitative assessment of individual differences (Coaley 2009). This paper reviews gender theorising and demonstrates how psychological perspectives underpinning gender role theory have informed two approaches to measuring gender as an individual phenomenon: gender orientation and gender ideology (Thompson and Pleck 1995). Description of gender role theory as well as these approaches to measuring gender is not new. Yet, an explicit link between theoretical and measurement literatures has yet to be drawn. This discussion is then extended by introducing a novel approach to measuring gender as a social phenomenon: gender (re)presentation. This is informed by theoretical shifts associated with social constructionism. The practical measurement implications and benefit of this approach are also outlined. Its central claim, that gender should only be measured at a social level, signifies a radical departure from established psychological thinking. It offers the potential that gender measurement, customarily the methodological preserve of psychological research, and its benefits, such as providing large-group normative data, may appeal to those working

*Email: russell.luyt@anglia.ac.uk

outside the confines of traditional psychology. Here gender is more readily understood as a social, as opposed to an individual, level phenomenon.

Many gender-related measures are available. These operationalise an array of constructs. Beere (1990a, 1990b), for example, categorises 18 types of gender-related measures. The most recognisable include those assessing 'gender roles' (such as the Bem Sex-Role Inventory [Bem 1974]); 'stereotypes' (such as the Beliefs About Women Scale [Belk and Snell 1986]); 'attitudes toward gender roles' (such as the Ambivalence toward Men Inventory [Glick and Fiske 1999]). This paper critically considers measures operationalising the constructs of 'gender roles' and 'attitudes towards gender roles'. The prominence of these constructs in literature justifies their critical examination. This is illustrated well through a keyword search of the terms 'gender role' or 'sex role' using the literature database PsycINFO. These terms appeared in 4292 peer-reviewed journal articles at the time of writing. This paper therefore does not seek to offer an all-encompassing overview of gender-related measures and associated theory. There are better texts for this purpose (e.g. Beere 1990a, 1990b; Smiler and Epstein 2010). Rather it focuses upon and interrogates a concept – gender roles – that has featured significantly in the psychological measurement of gender.

Theoretical rationale of measuring gender as an individual phenomenon

Until relatively recently, the majority of research concerning gender has been informed by gender role theory (Connell 1992). Gender role theories are often referred to in the singular, as is seen in the term 'gender role theory'. These suggest gender roles exist as a group of attributes considered appropriate for one sex rather than the other (Constantinople 1979). They describe processes of sex typing through which an individual acquires these attributes (Mischel 1966) and develops a related gender role identity (Garnets and Pleck 1979).

However, the term 'gender role theory' fails to reflect the often substantial differences between theories. These concern both the processes by which an individual acquires sex-typed characteristics as well as meta-theoretical assumptions concerning gender (Roopnarine and Mounts 1987). Difference among theories risks rendering the term 'gender role theory' a misnomer. A more subtle categorisation of theories is possible. In the account that follows, older perspectives of gender role acquisition are described as belonging to the gender role identity paradigm, as originally suggested by Pleck (1976), whilst more recent perspectives of gender role acquisition are said to represent the gender role beliefs paradigm (originally described by Pleck 1981 as the gender role strain paradigm).

The gender role identity paradigm includes some of the most well-known psychological explanations of gender (Pleck 1987). However, few authors have attempted to provide a systematic overview of this. Its review is difficult. Theories belonging to it have been modified, and others added, so as to accommodate problems in earlier theoretical formulations (Pleck 1981). Nevertheless, three theoretical perspectives stand out in their contribution: classical psychoanalytic (e.g. Freud 1905/2000); (cognitive) social learning (e.g. Bandura 1977); and developmental (e.g. Kohlberg 1966). To a greater or lesser extent, each of these perspectives suggests that individuals have an inner psychological need to affirm their biological sex through the acquisition of sex-typed characteristics. It is believed a female who acquires male sex-typed characteristics will develop an inappropriate gender role identity and vice versa. This is considered dysfunctional and is thought to result in negative psychological and social consequences (Pleck 1976, 1981, 1987). It is seen clearly in, for example, classical psychoanalytic theory

where failure to resolve the Oedipal and Electra complexes is believed to be the cause of neuroses (Roopnarine and Mounds 1987).

The gender role beliefs paradigm may be distinguished from the gender role identity paradigm. This distinction captures the major theoretical shifts that have taken place in gender role theorising (Pleck 1995). These shifts were in large part due to critique concerning the assumption that biological sex determines which gender role identity is 'appropriate' and functional (Pleck 1981). Information processing perspectives (e.g. Bem 1981a, 1981b, Martin and Halverson 1981) illustrate these shifts. Theoretical perspectives that share the assumptions of the gender role beliefs paradigm argue that males and females experience external social pressure to achieve gender-related social norms. Individuals evaluate themselves, and others, against these norms during gender role acquisition. They are motivated to do so in order to avoid negative social consequences (Pleck 1981). This is well illustrated by, for example, Bem's (1981a, 1981b) gender role schema theory where prevailing societal definitions of gender determine the content of the gender schema which in turn helps guide individual perception and organises self-concept.

The theoretical assumptions underlying the gender role identity and beliefs paradigm are clearly distinguishable in the measurement of gender as an individual phenomenon. Constantinople (1973) is credited as providing an early and seminal critique of existing measures (Williams and Best 1990a). However, it was only later that Thompson and Pleck (1986) suggested that limitations primarily resulted from an inconsistent understanding and measurement of gender as either individual traits or social norms. Others have suggested similar distinctions (e.g. Williams and Best 1990a, 1990b).

Trait approaches are largely grounded in the gender role identity paradigm. It is suggested that masculinity and femininity exist as a configuration of fixed individual attributes. Self-concept ratings on paper-and-pencil questionnaires or scales are understood to reflect real differences between men and women. Studies predominantly seek to account for the acquisition and consequences of sex-typed characteristics through behavioural and personality correlates (Thompson and Pleck 1995). Accordingly, masculinity and femininity are defined as 'relatively enduring traits which are more or less rooted in anatomy, physiology' or 'early experience, and which generally serve to distinguish males from females in appearance, attitudes, and behaviour' (Constantinople 1973, p. 390).

Alternatively, normative approaches suggest gender is socioculturally defined. These are typified by explanations offered by the gender role beliefs paradigm. The fact that gender role theory informs both trait and normative approaches may appear confusing. The assumption embedded in the gender role beliefs paradigm, that gender is determined through social norms but ultimately internalised by the individual as relatively enduring traits, results in this confusion. Greater emphasis is placed on the contextual specificity of gender. This challenges the unexamined assumption that an individual's gender remains relatively fixed. However, the assumption that gender-related social norms are internalised is not questioned. Thompson et al. (1992) argue that until recently literature has not distinguished clearly enough between trait and normative approaches. This is reflected in the theoretical confusion of many measures that include, to varying degrees, the assumptions of both.

In sum, it is argued that although the development of gender-related measures has largely proceeded in an a-theoretical fashion (Hoffman 2001), the assumptions of gender role theory are implicit within their design. So too are positivist epistemological assumptions where measures seek 'to produce "factual" knowledge about an objectively

present, and so observable and measurable, external world', and, one might add, through behavioural inference, the internal world (Wilkinson 2001, p. 18).

A more subtle theoretical distinction can be made between the gender role identity and beliefs paradigm. The former has informed the measurement of gender identity as an individual trait, whilst the latter has suggested the measurement of gender attitudes as individually endorsed and internalised social norms. These different measurement approaches have been defined as gender orientation and ideology approaches, respectively. Thompson and Pleck (1995) report that the distinction has met with some resistance. Nevertheless, it holds advantages in that it informs appropriate research-specific use of instruments and aids in the development of new theoretically sound measures. In addition, evidence exists that instruments of gender orientation and ideology measure independent constructs and have differential correlates (Thompson et al. 1992). These two concepts are discussed below. This provides an overview of the fundamental and changing theoretical assumptions guiding gender measurement over the last 70 years. It also provides a critical foundation upon which to consider how social constructionism may inform a gender (re)presentation approach to measurement.

The 'gender orientation' approach to measurement

The early measurement of masculinity–femininity (M–F) (such as the Attitude Interest Analysis Test [Terman and Miles 1936]) was based on four key assumptions, overlapping with many of those espoused by the gender role identity paradigm, and arguing that gender fixed: a fixed individual attribute; not directly observable through overt behaviour; innate to individuals and as such a determinant of their mental health; and existing on a bipolar¹ and unidimensional² continuum defined by sociocultural stereotypes of masculinity and femininity (Morawski 1987). Thus, individuals were assessed along a continuum, ranging from masculine to feminine at each extreme, in order to determine their sex typing. Gender was considered something individuals had, either psychologically or biologically, and this was judged appropriate depending upon whether it matched their biological sex. As such, non-normative gender traits (such as a woman who appeared 'masculine') were considered maladaptive (Windle 1987). Hoffman (2001) suggests that this conceptualisation served as the basis for M–F measurement over subsequent decades.

It was only in the mid-1970s that a well-developed critique of traditional gender theory emerged – contributing towards the gender role identity paradigm and related measurement (Williams and Best 1990a). The critique revolved around three major issues. First, it was claimed that gender did not exist as a bipolar construct. Critics argued that masculinity and femininity should not be considered to be mutually exclusive. That is to say 'what it is to be a man' is not related in linear opposition to 'what it is to be a woman'. Rather, the constructs were said to operate independently of each other – although often in systematic opposition (Morawski 1987). Empirical evidence surfaced that undercut the notion of bipolarity in gender measurement. For example, findings indicate a positive rather than a negative relationship between the correlates of masculinity and femininity (Constantinople 1973). Second, the supposed unidimensionality of the M–F construct was said to be overly simplistic. This assumption was most obvious in its measurement as a single score (Morawski 1987). Evidence has been presented to substantiate its multidimensionality through a plethora of correlational as well as factor analytic studies.³ Accordingly, it is argued that M–F may best be assessed through a number of subscores (Constantinople 1973). Finally, scepticism surrounded both the empirical and theoretical foundation of existing M–F measurement. The most common

means of measurement development was to administer a large pool of items to participants. Sex-based differences in response served as a means of item selection (Morawski 1987); for example, the Minnesota Multiphasic Personality Inventory (MMPI) Masculinity–Femininity Scale (Mf) (Hathaway and McKinley 1943). This method⁴ was clearly rooted in the widely held view that masculinity and femininity reflected a basic biological dimorphism. This view became increasingly untenable along with the realisation that gender was not so easily categorised (Williams 1987). As such, critics argued that the empirical foundation of M–F measurement often relied on a false equation between psychological gender and biological sex (Constantinople 1973).

Theoretical critique focused specifically on the rigidity of gender conceptualisation and its subsequent measurement. Psychological gender was understood as encompassing a set of stable and universal traits. These were seemingly uninfluenced by individual agency as well as social factors. That is to say, this understanding overlooked the notion of agentic self-concept in which individuals actively contribute towards the development of their own unique sense of gender self. Similarly, the role played by sociocultural factors in defining gender categories differently over context and time was ignored. Thus, whilst the use of standard M–F measures in assessing supposedly fixed gender traits may have seemed reasonable (Hoffman 2001), recognition that all measures find production within a unique sociocultural moment was absent. Their use persisted despite obvious variation in item content from one measure to another (Constantinople 1973).

The failure to appreciate the potential impact of individual agency and social factors in determining psychological gender held negative consequences. Constantinople (1973) notes that in its insistence that the M–F construct included a set of fixed individual traits, deemed ‘healthy’ dependent upon an individual’s biological sex, gender measurement was guilty of casting the non-normative as abnormal.

Theoretical critique therefore encouraged the reassessment of existing measures. Researchers took the criticism that M–F did not exist as a bipolar construct seriously, as is evidenced in the Bem Sex-Role Inventory (Bem 1974) where individuals were identified as one of four types: ‘masculine’, ‘feminine’, ‘androgynous’ or ‘undifferentiated’. Recognition that masculinity and femininity were independent constructs held important implications on both a methodological and a theoretical level. Methodologically, this indicated that gender could no longer be measured along a continuum in which a single score determined an individual’s sex typing. Theoretically, it implied that healthy men and women could possess cross-gender traits, and even went so far as to suggest the developmentally restrictive nature of traditional gender roles. As such, psychological gender was no longer understood as ‘appropriately’ related to biological sex. The concept of androgyny captured this new awareness. Crucially, however, the issue of construct multidimensionality remained unexplored (Windle 1987).

Yet, by the mid-1980s theorising increasingly attacked what was perceived as an overly rigid portrayal of gender experience (Morawski 1987). Existing gender role explanations were said to neglect its contextual specificity through their emphasis on fixed individual traits. In disregarding the pivotal role played by sociocultural factors, these explanations were argued to adopt an apolitical outlook, which effectively ignored the impact power relations play in both compliance with, and endorsement of, social norms (Thompson et al. 1992). Morawski (1987) notes, however, that despite these criticisms the question of social power continues to remain largely absent from trait approaches. Unlike social constructionism, which adopts an alternative metatheoretical position as discussed below, these approaches do not consider power adequately within their analysis. They may therefore not easily be open to change. It is worth considering Constantinople’s (1973)

early observation that weaknesses in gender orientation measurement may reside not only in the method of measurement but also in the utility of the construct they assess – gender as a fixed individual attribute.

It is also necessary to acknowledge the inherent cultural bias of gender orientation measurement (Williams and Best 1990a). This arises due to the use of gender role stereotypes in guiding item development. Numerous authors have questioned whether these successfully capture what it is to be masculine and feminine (Morawski 1987). At the very least, continued dependence on the use of gender role stereotypes, underlines the sociocultural specificity of these measures.

The ‘gender ideology’ approach to measurement

From the mid-1980s onwards, trait approaches to measurement met with increasing criticism. The normative approach was offered as a feasible alternative. Gender is described as a sociocultural product rather than a relatively fixed individual attribute. An individual’s endorsement of traditional gender norms is thought to vary across context and time. This accounts somewhat for inconsistency and contradiction in individual attitudes towards gender (Levant et al. 1992, 1996). The notion that gender exists as a sociocultural product holds far-reaching implications for quantitative measurement, not least of which, the suitability of techniques that arguably objectify, and hence potentially reify gender, as well as universalise related attitudes, ideas, beliefs and experience.

The term gender ideology is used to describe individual attitudes towards gender-related social norms (Levant 1996). Measures of gender ideology are described as assessing individual ‘endorsement and internalisation of cultural belief systems about masculinity (femininity) and male (female) gender, rooted in the structural relationship between the two sexes’ (Pleck 1995, p. 19). These definitions reflect the underlying assumptions of the gender role beliefs paradigm where gender is determined through social norms but ultimately internalised by individuals.

Measures of gender ideology consider the constructs of masculinity and femininity as multidimensional and independent. Multidimensionality suggests the need to assess support for dominant norms rather than a single masculinity or femininity script. Individual endorsement of these is seen to vary over context and time, as well as across age, culture, sexual orientation, social class and ‘race’ (Levant and Majors 1997, Levant and Fischer 1998). In addition, the assumed independence of masculinity and femininity requires that attitudes towards these constructs are assessed apart. Accordingly, individuals are assessed along a continuum in order to determine their endorsement of independent and multiple masculine and/or feminine norms.

However, a lack of systematic theoretical focus has resulted in a number of weaknesses in many instruments. It is argued, for instance, that measures of gender ideology ought to incorporate three key characteristics: an emphasis on multiple masculinities or femininities (construct multidimensionality); a clear distinction between masculinity, femininity and gender ideologies (construct independence); and theoretically appropriate content in measuring ideology (Thompson et al. 1992, Thompson and Pleck 1995).

Understanding that masculinity and femininity are multidimensional constructs provides a more sophisticated means with which to explore sociocultural variation in gender conceptualisation. It accounts for differential endorsement of gender norms, at an individual or group level, across dimensions. For example, empirical studies using the Male Role Norms Inventory (Levant et al. 1992) have found variable support for masculine norms along gender (Levant et al. 1996, Levant and Majors 1997, Levant et al.

1998), 'race' (Levant and Majors 1997) and cultural lines (Levant et al. 1996). Findings even suggest urban–rural variation (Levant et al. 1998). The complexity of masculinity and femininity ideology is easily understood once interactions between these mediating variables are considered.

It is also recognised that attitudes towards masculinity, femininity and gender do not constitute an internally consistent set of beliefs. Men's and women's experiences differ due to dissimilar life opportunities (Thompson et al. 1992). Stress is therefore placed on absolute gender characteristics. These typify either masculinity or femininity, but do not necessarily differentiate between the two, as relative characteristics would (Levant et al. 1992). This implies that attitudes towards men will not be systematically related to those held of women or gender relations in general. Whilst these constructs are likely to be empirically correlated, their independence is demonstrated through their differing correlation to theoretically distinct constructs, or other meaningful variables. Measures of masculinity or femininity ideology that fail to appreciate this independence often include gender comparative items. This is argued to incorrectly assess gender ideology, that is to say attitudes towards gender relations (Pleck 1981, Thompson et al. 1992).

Finally, it is argued that measures of gender ideology need to include appropriate content. Thompson and Pleck (1995) suggest that third-person statements are most beneficial in evaluating ideology, whereas the use of the first person may better assess gender orientation. Apparent inconsistent use of the third person in gender ideology measures is evident in Tolman and Porche's (2000) Adolescent Femininity Ideology Scale which is written in the first person, whilst Chu et al.'s (2005) Adolescent Masculinity Ideology in Relationships Scale is written in the third person. A distinction has also been made between descriptive stereotypes (that is, 'what men or women are like') and prescriptive norms ('what men or women should be like'). Prescriptive statements are believed to convey the normative assumptions underlying gender ideology. Descriptive statements assess the notion of gender orientation in their focus on individual characteristics (Levant et al. 1992, Thompson et al. 1992). Neither of the above propositions has been assessed empirically and as such offer potential areas of enquiry. Thompson and Pleck (1995) also suggest the use of the plural, as opposed to the singular, helpfully underlines an awareness of construct multidimensionality (femininities or masculinities).

In sum, measures of gender ideology have successfully addressed many criticisms of earlier gender measurement, in particular the lack of importance afforded to construct independence and multidimensionality. That is to say, they have developed the notion of construct independence, beyond that which was achieved by later orientation measures, by suggesting that all forms of gender comparison should be avoided. This is most commonly seen in their use of absolute gender characteristics. Additionally, ideological measures have been the first to achieve multidimensionality even though early measurement critique questioned the dominant conceptualisation of gender as a unidimensional construct. A commitment to multidimensional assessment has enabled these instruments to account somewhat for the contextual specificity of gender. This is seen in inconsistent and often contradictory endorsement of traditional norms at both an individual and group level. Finally, gender ideology has adopted greater critical awareness surrounding the limitations of gender measurement. Specifically, this has involved the careful definition of masculinity and femininity and has relied far more on a theoretically guided process in instrument development. The Male Role Norms Inventory-Revised (MRNI-R) (Levant et al. 2007) provides a recent example. This measure underwent initial validation by means of a student sample (n = 170) in the USA. It is meant to assess the endorsement of

traditional masculinity ideology through appraisal of a total ('MRNI-R Total Scale') and seven subscale scores (i.e. 'Avoidance of Femininity', 'Fear and Hatred of Homosexuals', 'Extreme Self-Reliance', 'Aggression', 'Dominance', 'Non-relational Sexuality' and 'Restrictive Emotionality'). The discussion above has outlined two distinct approaches to measuring gender as an individual phenomenon: gender orientation and gender ideology. These are grounded in different measurement approaches, distinguished as assessing either individual traits or social norms, and informed by the two broad paradigms underpinning gender role theory. This paper argues that a third distinct approach to measuring gender as a social phenomenon – the gender (re)presentation approach – might beneficially be adopted. This approach, its underlying theoretical rationale of social constructionism, its measurement implications and benefits are described below.

Theoretical rationale of measuring gender as a social phenomenon

Social constructionism, which is grounded in broader postmodernist metatheory, has increasingly informed gender theorising. Epistemologically, it asserts "that facts" are always dependent on the particular forms of language and the particular language communities which have created and maintained them' and that therefore we 'cannot "know" the external world' or, through behavioural inference, the internal world, 'because all knowledge is mediated by ... the specificities of language' (Wilkinson 2001, p. 24).

Pleck (1995, p. 22) argues that the gender beliefs paradigm is 'in a broad sense, a social constructionist perspective that simply predated the term'. Yet, this reflects misunderstanding. Social constructionism recognises gender as relatively flexibly (re)produced within situated interaction, mediated by language and other shared symbolic systems, rather than an internalised individual attribute.

Rubin's (1975) valuable distinction between 'sex' and 'gender', which differentiates between the socioculturally defined 'feminine' and 'masculine' and biologically defined 'female' and 'male', acts as a useful starting point for discussion. Numerous authors have considered the relationship between these two concepts (Deaux 1985, Unger and Crawford 1993, Diamond 2000, Pryzgodna and Chrisler 2000) and the extent to which they are compatible with social constructionist theorising. It is claimed that these categories are not able to account for observed variability in gender practices over context and time. Neither gender (Williams 1987) nor sex (Fausto-Sterling 1993) necessarily appears as dichotomous. West and Zimmerman (1987) suggest the sex/gender distinction ought to be expanded in order to distinguish between the notions of 'sex', 'sex category' and 'gender': sex is defined on the basis of socioculturally agreed upon biological/physiological characteristics. An individual's normative allocation to a sex category – such as male or female – depends upon their possessing suitable such characteristics. In daily life, individuals are not allocated to sex categories on the basis of these agreed upon characteristics, but rather through an appraisal of their gendered social practices. These social practices may be understood as embodying an individual's gender. They are guided by dominant gender representations that suggest what practices are appropriate for members of each sex category. Individuals reinforce their membership to specific categories through adopting 'appropriate' social practices.

Thus, social constructionists recognise that sex categories act as primary organising principles in society but challenge the notion that these are an objective reality. 'Sex, like gender, draws meaning from shifting cultural understandings and ever-changing social practices' (Marecek et al. 2004, p. 207). 'Sex' is nothing more than socioculturally agreed

upon biological/physiological characteristics. Likewise, individuals do not 'have gender'. It does not exist as an individual attribute, characteristic or traits – however acquired. Rather, it is considered a product of situated interaction. Individuals are therefore seen to 'do gender' (West and Zimmerman 1987). They constantly lay claim to specific sex categorisation through social practices. At the same time, they also (re)produce meanings concerning what it is to be a woman or a man in society (Bohan 1993).

Social constructionism stresses the importance of understanding gender as a situated social practice. This accounts for its observed variability within and across contexts. Prescriptions and proscriptions concerning gender differ from situation to situation. In order to 'do gender' successfully, individuals are required to adapt their practices in accordance with these subtle situational demands (Bohan 1993). Individuals always remain accountable for their practices in that these are open to social appraisal. They are aware of this fact and consider how their practices may be judged through salient sociocultural standards before undertaking them. Constant accountability ensures that 'doing gender', whether socially acceptable or not, is unavoidable (West and Zimmerman 1987). Individuals who fail to do gender appropriately are likely to experience negative consequences (Bohan 1993), whereas appropriate gender practices confer a sense of social competence (West and Zimmerman 1987) as well as possible material reward.

'Doing gender' largely reinforces dominant representations concerning what constitutes appropriate gender practice. In doing so, it also legitimates sociocultural structures that are based on these meanings (West and Zimmerman 1987). This stresses the close relationship between systems of knowledge and power. Gender is constructed in such a way so as to perpetuate women's subordination in society (Flax 1987).

Empirical evidence supports the social constructionist argument of gender, demonstrating variation in its sociocultural (e.g. Luyt 2012a) and situated practice (e.g. Stokoe 1998). Thus, as Beall (1993) argues, cross-cultural research provides especially firm evidence. But so too does more micro-level analysis of human interaction. This theoretical perspective is able to account for variability in gender practices through its emphasis on context. It recognises that gender is in no way essential or enduring. Men and women may supposedly choose to adopt any form of practice. Crucially, however, in everyday life, this choice is restricted through ideological and structural constraints. Social constructionism nevertheless underlines the radical extent to which individual as well as social change may be possible (Bohan 1993).

Some social constructionists raise concerns regarding the use of quantitative measurement for the purpose of gender research. These criticisms may be categorised as emanating from either a weak or a strong/strict constructionism. They are particularly damning of normative measures of gender.

Weak constructionism argues phenomena acquire different meanings depending upon their interpretation within specific sociocultural contexts (Wilkinson 2001). These phenomena are believed to exist independently outside of interpretation. They therefore act to limit potential meaning as well as human action. Social constructions are also not merely considered fleeting or transient. This implies that sweeping change in meaning is unlikely to occur over the short term. Traditional research objectives such as theory testing and empirical generalisation may therefore remain worthwhile (Stryker 1995).

Critics from this perspective suggest quantitative measurement objectifies, and hence potentially reifies, gender, as well as universalising related attitudes, ideas, beliefs and experience. Yet, this paper, which supports quantitative measurement as informed by social constructionism, argues that any use of linguistic categories unavoidably objectifies reality and makes possible its reification (Berger and Luckmann 1966). This is the case in

both quantitative and qualitative analyses. Normative measures are also argued to oppose the assumption of universality through their emphasis on multidimensionality. This is partially able to account for contradiction and complexity in attitudes towards traditional gender norms. Yet, Thompson and Pleck (1995, p. 135) note that some still claim that ‘these scales still assume one monolithic male role, albeit with component dimensions’. However, there are two further ways through which to counter such universalism. Researchers should be prepared to evaluate evidence of an instrument’s content validity as well as construct validity among the population in which it is to be disseminated and re-evaluate this evidence across time. This is rarely undertaken (McHugh and Frieze 1997). When it is, in a cross-cultural context, this may result in separate measures being developed. Although such an eventuality ‘complicates or totally eliminates the possibility of cross-group comparisons’ (Floyd and Widaman 1995, p. 295), it still provides a worthwhile quantitative means with which to describe cross-cultural endorsement of traditional gender norms.

Strong or strict constructionism suggests a far more radical ontological position. Meanings assigned to phenomena are believed to be constructed. However, so too are the phenomena themselves, which are ‘brought into existence precisely by the discourse that presumes to “describe” them’ (Bohan 2002, p. 75). It is argued that a reality independent of interpretation either does not exist or cannot be known. As such, there is no objective way or method through which to judge the ‘truth’ between competing claims concerning the world (Stryker 1995, Wilkinson 2001). A strong/strict constructionist critique is evident in the argument presented by, for example, discursive psychology. The concept of attitudes and its measurement is the focus of especially stinging criticism (Potter 1996, 1998, Wiggins and Potter 2003). This paper claims that traditional assumptions underlying the concept may be reformulated, and in so doing make it theoretically and methodologically congruent with strict constructionist perspectives. One such possible reformulation is briefly described below. Table 1 summarises the three approaches to measuring gender and their theoretical assumptions.

Table 1. Approaches to measuring gender and their theoretical assumptions.

		Approaches to measuring gender		
		Gender orientation	Gender ideology	Gender representation
Theoretical assumptions	Epistemological perspective	Positivism	Positivism	Constructionism
	Theoretical perspective	Gender role theory (gender role identity paradigm)	Gender role theory (gender role beliefs paradigm)	Social constructionism
	Gender assessed as an individual trait	... internalised individual attitudes towards social norms	... as a subject position relative to dominant gender representations
		Gender reducible to the ...		

... individual ...
individual ... social

The 'gender (re)presentation' approach to measurement

The concept of attitudes is central to the normative approach to measuring gender as an individual phenomenon. It has also featured prominently in social psychological debate for most of its history (Potter and Wetherell 2006). For these reasons, the gender (re)presentation approach to measurement will be introduced through its critique and reformulation.

Despite its prevalence, it is difficult to locate an all-encompassing definition of the concept as its meaning has shifted over time (Potter and Wetherell 2006). Schwartz and Bohner (2001) note that early definitions of attitudes stressed that they remained stable and were closely related to behaviour. More recent definitions have emphasised their evaluative nature. These argue an attitude is expressed when an individual situates an object of thought along dimensions of evaluative judgement. Or, in simpler terms, an attitude is expressed when a person evaluates a perceived phenomenon either positively or negatively to varying degrees. The expression of an attitude is considered to represent an internal, pre-formed and stable mental state (Billig 1998a, 1998b). That is to say, attitudes cannot be observed directly, but only through the assessment of self-reports or behaviour (Schwartz and Bohner 2001). This explanation is compatible with the notion of gender ideology. As is noted above, measures of gender ideology are commonly described as assessing individual 'endorsement and internalisation of cultural belief systems about ... gender' (Pleck 1995, p. 19).

Discursive psychologists reject the traditional conceptualisation of attitudes. In particular it is argued that their stability is overstated. They point towards the fact that individuals may not express an evaluation in exactly the same way across, as well as within, situations (Potter and Wetherell 2006). Such variability in attitudes proves problematic for traditional perspectives seeking to demonstrate the existence of stable underlying mental constructs (Potter 1998). Schwartz and Bohner (2001) note that self-reported attitudes appear somewhat context dependent. Traditional perspectives account for variability in two main ways: stable attitudes are seen to exist, where varying responses are merely seen to reflect measurement error due to the influence of changing situational variables; or, relatively stable attitudes exist, but individuals access either a range of memory structures or multiple attitudes about an object, when responding to questions.

It is nonetheless claimed that these explanations do not sufficiently account for observed variability as is seen in the analysis of social interaction. Evaluations are not seen to reflect some internal, pre-formed and stable mental state, but are rather believed to be situation dependent. It has been suggested that attitudes be described as evaluative judgements that are made solely on the basis of situationally accessible information (Schwartz and Bohner 2001). Discursive psychology extends this line of argument. Individuals are argued to draw on available discursive resources in order to make evaluations. Furthermore, when individuals make evaluations they are seen to be doing something, or more precisely performing an action. This implies individuals make attitude claims for specific purposes within any situation. Attitudes are therefore best re-conceptualised as evaluative practices serving particular functions (Potter 1998, Potter and Wetherell 2006).

It is commonly argued that this re-conceptualisation makes the use of quantitative measures inappropriate. It is only supposedly in applying assumptions characteristic of social cognition work, which stress the importance of stable internal internal states as opposed to meaning making through social interaction, that warrant their use (Potter 1998). In particular, Potter and Wetherell (2006) note the core assumption in traditional theory; that attitudes are clearly separate from the 'object of thought'. A questionnaire

item serves as the object of thought in attitude measurement. Participants are asked to reflect upon the item in order to express their attitude. Assessment of attitudes held by different individuals is made possible because objects of thought are believed to hold the same meaning for all. Most discursive psychologists disagree. They argue that attitude measurement is meaningless because 'the object is formulated and constructed in discourse in the course of doing evaluation' (Potter and Wetherell 2006, p. 207), or to paraphrase, the object only acquires meaning through the specific interaction in which the individual is asked to make the evaluation.

Discursive psychology suggests evaluative practices are most suitably explored through an analysis of social interaction. Potter (1998, p. 242) underlines the perspective. He:

... treats action as fundamental or, as Schegloff (1995) puts it, omnirelevant. So if you want to understand evaluations you need to consider carefully what people are doing with them in their 'home' environments, rather than in the more arcane contexts of filling in attitude scales.

The extract reveals the extent to which the critique concerning the use of attitude measurement is largely based on a preference for bottom-up as opposed to top-down approaches to discourse analysis (Edley and Wetherell 1997). The bottom-up approach is informed by the work of theorists such as Sacks (1964–1965/1992) and Garfinkel (1967). Fine-grained analysis of textual features is undertaken in order to describe the action-orientation of social interaction (Edley and Wetherell 1997). Austin (1962) was the first to identify 'talk as action'. He argued that language carries not only meaning but also force. That is to say, individuals are able to do and achieve things through the use of language. This suggests analysis should focus on what individuals are accomplishing through social interaction (Potter and Wetherell 2001). These arguments are evident in Potter's (1998) emphasis on the action-orientation of text ('action as fundamental') and his explicit mention of a well-known conversation analyst ('Schegloff 1995').

From a bottom-up perspective, which stresses the need for fine-grained or micro-level analysis of textual features in order to describe the action-orientation of social interaction, it would clearly be unsuitable to make use of scales or questionnaires. However, from a top-down perspective that explores how individuals are constituted or positioned through discourse, these research materials may prove useful. This approach draws heavily on the work of theorists such as Foucault (1978) and Marx (1867–1894/1981). It focuses on broad concepts such as ideology and power in order to explore how individuals are constituted or positioned through discourse (Wooffitt 2005).

Potter and Wetherell's (2006, p. 207, emphasis added) original observation that 'the object is formulated and constructed in discourse in the course of doing evaluation' reflects a bottom-up approach to analysis. Here discourse is characterised as a practice. Questionnaire items are not thought to hold a similar meaning for individuals because they are seen to be caught within ceaseless discursive debate. It therefore becomes important to understand the function of evaluative practices through features of the proximate context (participant understanding surrounding the type of conversation to which they are contributing; the actions made possible through sequences of talk; and the roles participants are assigned or assume).

Those in favour of a top-down approach to analysis may have alternatively observed that the object is formulated and constructed by discourse in the course of doing evaluation. Here discourse is characterised as an entity. Questionnaire items are thought to hold a similar meaning for most members of a sociocultural group because they are believed to share broad systems of meaning. It therefore becomes important to understand

the function of evaluative practices through features of the distal context (such as participant's age, ethnicity, 'race' and social class; the research site; and the sociocultural and ecological milieu in which it is embedded). Quantitative measures are suited to this task. They allow us to ask a number of useful inter- and intra-group questions, for example, are traditional gender norms conceptualised differently across sociocultural groups and are they endorsed to the same extent across divisions such as age and social class within these groups?

In seeking to answer these questions, the gender (re)presentation approach holds a number of assumptions concerning features of the research context, which distinguishes it from other perspectives. Both traditional research and discursive psychology view measurement research environments as problematic.

In traditional research, these environments are deemed problematic in that they introduce potential bias. This is seen, for example, in the concept of social desirability which is defined as 'the tendency for a person to respond in a way that seems socially appealing, regardless of his or her true characteristics' (Furr and Bacharach 2008, p. 246). It is argued that we are better able to assess an individual's 'true characteristics' through controlling for such bias.

In discursive psychology, these environments are considered problematic due to their claimed artificiality. As noted above, Potter (1998, p. 242) views these as 'arcane contexts' that are removed from the everyday realities experienced by people in their "home" environments'. Individual evaluations in these contexts may serve particular functions, but these are not especially useful given their alleged remoteness from ordinary life.

The gender (re)presentation approach, by contrast, does not view measurement research environments as problematic but rather as constituting an important and meaningful context for data elicitation. They are not seen as introducing bias that clouds our apprehension of an individual's 'true characteristics' – as suggested by traditional research. Indeed, as a social constructionist approach, 'true' individual attributes, characteristics or traits, however acquired, are not believed to exist. These are not understood to exist in individual heads but are rather a product of social interaction. Nor is the measurement research environment considered an artificial setting that is removed from 'naturally' occurring social interaction and hence superfluous – as suggested by discursive psychology. The research environments in which questionnaires or scales are completed are considered unique, but no more so than any other. This includes our so-called 'home environments' that do not exist 'naturally', inherently or independently of the social interaction that constitutes them. All are equally meaningful (Edley and Litosseliti 2010).

It is important therefore to describe some of the potential meanings that frame measurement research environments. The gender (re)presentation approach argues that the way these research environments are understood by participants is indispensable for the meaningful elicitation and interpretation of data. Measures of gender (re)presentation seek to assess endorsement of dominant gender representations. It is therefore important that participants complete such measures within a research environment characterised by the normative. As described below, measures of gender (re)presentation need to exhibit five key characteristics, including appropriate content that is specifically tailored to the assessment of dominant gender representation. Through the use of, for example, third-person statements and prescriptive norms, these measures actively encourage participants to orient towards these dominant discourses and position themselves in relation to them. In this sense, a well-crafted measure of gender (re)presentation needs to reflect the same discourses of power that structure everyday social practices. So, when individuals complete these measures, they are 'doing gender' as informed by a normative

environment. They draw upon features of the distal context, such as their perceived group membership, and apply available systems of meaning in order to interpret dominant gender representations, and so respond in a socially meaningful way.

In sum, most discursive psychologists argue that the concept of attitudes is flawed and should be discarded. Potter (1998, p. 241) typifies this stance:

... it would be easy to give the impression that there are pre-existing objects – attitudes – which are understood in one way, or one set of ways, in mainstream social cognition work, and then understood in another way in discursive social psychology. However, this would understate the degree to which the attitude notion is constituted out of social psychological theory, and the extent to which the notion is dissolved in discursive social psychology into a range of other considerations.

This stance is surprising. Recent debate among social psychologists in the UK, in particular, has generated calls for productive engagement between adherents of traditional versus critical perspectives [The Psychologist (2005), vol. 18, provides particularly interesting reading in this regard]. The suggestion that the concept of attitudes should be ‘dissolved’ in the face of discursive critique is not conducive to engagement but rather serves to polarise debate. It should be recognised that the concept has inhabited a central position in social psychological theory. This is unlikely to change. As such, it would seem more sensible to reformulate it from a discursive perspective. Contrary to Potter’s (1998) assertion, a reformulation of this kind is possible. It will allow discursive psychologists, among other social constructionists, to make use of quantitative measures when applying a top-down approach to discourse analysis. This would be in keeping with Abell and Walton’s (2010, p. 686) observation that:

... discursive approaches to social psychology are a very broad church and the members of its congregation demonstrate in their empirical investigations dazzling inventiveness in their combination of approaches, methods, epistemological, and ontological positions. Further such a state of affairs is entirely consistent with the view that what matters is empirical utility rather than methodological purity.

This proves especially helpful when exploring gender at a social or group, as opposed to an individual, level. Through the use of large-group normative data, researchers may gain understanding of dominant gender representations and theorise the functions any variation in their endorsement serve within and between sociocultural groups. Such variations may, in particular, point towards possible conflicts and challenges in gender ideology within any given society.

This paper argues that measures of gender (re)presentation should include five key characteristics: an emphasis on multiple masculinities, femininities and other gender practices (construct multidimensionality); a clear distinction between masculinity, femininity and other gender-related concepts (construct independence); a focus on social or group level, as opposed to individual, phenomena; suitable evidence of measurement validity; and theoretically appropriate content.

Construct multidimensionality is emphasised by measures of gender (re)presentation as is the case in measures of gender ideology. This facilitates a more nuanced and complex account of gender attitudes by capturing often inconsistent and contradictory endorsement of dominant gender representations where, for example, men of a particular sociocultural group may support traditional notions of masculine success but at the same time distance themselves from notions of traditional masculine toughness. It allows useful inter- and intra-group comparisons. As noted above, we may be interested in asking whether traditional gender representations are conceptualised differently across sociocultural groups as well as whether they are endorsed in a similar fashion across divisions such as age and social class

within these groups. We might, for instance, discover that working-class men endorse ideas of traditional masculine toughness to a greater extent than middle-class men of the same sociocultural group, who, by contrast, endorse ideas of traditional masculine success more strongly. Lastly, it facilitates our being able to trace changes in the variable endorsement of dominant gender representations over time (McHugh and Frieze 1997). We may find, for example, working-class men increasingly distance themselves from notions of traditional masculine toughness. We could then ultimately theorise the functions any apparent variations serve within and between sociocultural groups. For instance, we might argue men of different social classes are most likely to endorse gender representations that they are able to practice most easily, as a means with which to position themselves as 'real' men. Multidimensionality thus provides a means with which to explore social, political and historical variation in gender conceptualisation, and in doing so highlights difference in gender attitudes, ideas, beliefs and experience.

Construct independence is of central importance in measures of gender (re)presentation as well as gender ideology. Levant et al.'s (1992) observation that measures should be developed in order to assess absolute, as opposed to relative, gender characteristics is useful. Absolute gender characteristics typify constructs such as masculinity and femininity but do not differentiate between them. This recognises the independence of these constructs at both an instrument and an item level. That is to say, separate measures should be developed for constructs such as masculinity versus femininity representation, and items should avoid gender comparison.

Social constructionists argue that gender is something we 'do' rather than something we 'have'. Exploration of attitudes as an individual level phenomenon is, from this perspective, inappropriate. They are not considered to reflect some internal, pre-formed and stable mental state. Rather, they are understood as evaluative practices based upon broad shared systems of meaning, which serve particular functions. Measures of gender (re)presentation are therefore applied in order to understand gender as a social or group level phenomenon. We are interested in how gender is constructed in specific sociocultural contexts by discourse and the subject positions made possible as a result. Dynamism is recognised where the interplay of agency and structure allow attitudes to appear both stable and changing, unidimensional and multidimensional, but only meaningful when dominant gender representation is understood in the same way by a sociocultural group.

Variability in gender understanding across context and time demands that we carefully obtain evidence in support of an instrument's measurement validity before making use of it. Gibbons et al. (1997) note most cross-national research makes use of an existing measure, for which evidence in support of its reliability and validity in a specific cultural context exists (most frequently the USA), and applies this to a new cultural context. All too often, such measures are applied among populations for which there is no such evidence. At best, this implies a false universalism or an imposed *etic* in gender experience, one that is for the most part 'white', western and middle class. At worst, it reflects nothing more than idle research practice. As noted above, researchers should be prepared to evaluate evidence of an instrument's content as well as construct validity among the population in which it is to be disseminated. A mixed-method approach, drawing upon the strengths of both qualitative and quantitative techniques, should ideally be adopted in order to do so (Luyt 2012b). In cross-cultural research, for example, the equivalence of measures should not be assumed. Measurement validation procedures should always be undertaken. Qualitative methods, such as focus groups, may be useful in assessing evidence in support of a measure's content validity. The cross-cultural meaningfulness of items can be determined. In particular, the constructs they operationalise, as well as the settings (such as

higher education) and tasks (such as asking someone on a date) they describe, ought to be considered carefully (Gibbons et al. 1997). This may indicate the useful addition or removal of culture-specific constructs (Gibbons et al. 1997) or that separate emic measures should be developed for use within different sociocultural contexts. Where justified, we should be tolerant of the so-called ‘scale proliferation’ (McHugh and Frieze 1997, p. 2). It is recognised that developing equivalent measures may neither be possible nor necessarily desirable due to differences in gender understanding across such contexts. This may complicate cross-cultural comparison. Yet, a far richer, albeit more complex, set of findings may result. For example, replicatory or confirmatory factor analysis might indicate cross-cultural equivalence of some but not all underlying constructs. In such cases, comparisons would be limited to constructs demonstrating sufficient equivalence. Gibbons et al. (1997) also suggest that analysis of individual items, as opposed to total scores, may be helpful in such circumstances.

Finally, measures of gender (re)presentation need to include appropriate content that is specifically tailored to the assessment of gender representations. As in the case of measures of gender ideology, this includes the use of third-person statements; prescriptive norms (‘what women or men should be like’ as opposed to ‘what women or men are like’); and the use of the plural as a means to highlight construct multidimensionality (femininities or masculinities as opposed to femininity or masculinity) (Thompson and Pleck 1995). The importance of theory in guiding measurement development is clearly evident. Table 2 summarises the three approaches to measuring gender and their measurement assumptions.

This paper suggests social constructionism may inform a novel approach to measuring gender, beyond that already offered by the gender orientation and gender ideology approach; this known as the gender (re)presentation approach. Measures of gender (re)presentation share an emphasis on the independence and multidimensionality of gender with measures of gender ideology. Yet, they may alternatively be said to index the extent to which groups endorse dominant gender representations, which serve to legitimate

Table 2. Approaches to measuring gender and their measurement assumptions.

		Approaches to measuring gender		
		Gender orientation	Gender ideology	Gender representation
Measurement assumptions	Construct dependence	Interdependent/independent	Independent	Independent
	Construct dimensionality	Unidimensional	Multidimensional	Multidimensional
	Cross-cultural measurement equivalence	Possible and desirable	Possible and desirable	Often impossible as well as undesirable
	Item characteristics	Absolute/relative	Absolute	Absolute
		First person Descriptive stereotype	Third person Prescriptive norm	Third person Prescriptive norm
Level of analysis	Individual/group	Individual/group	Group	

and (re)produce unequal gender relations, and in so doing make specific gender subject positions available. Gender is not understood as a set of internalised social norms. It is alternatively understood as a social practice made meaningful through shared representations held by members of the same social group. Interest therefore lies in its measurement at a group or social, rather than individual, level. This serves to compliment fine-grained or micro-level situated analysis where individuals are seen to strategically adopt varying gender practices and claim subject positions in the course of unfolding social interaction.

It may be argued that this paper offers an alternative interpretative lens through which to view existing measures rather than an entirely new approach to gender measurement. Yet, an argument of this kind would fail to appreciate the important relationship between theory and method as outlined above. Measurement assumptions change only gradually across different approaches to measuring gender (see Table 2). But these are informed by underlying theoretical assumptions that contrast more starkly (see Table 1). This is especially so when comparing the gender (re)presentation approach against existing approaches. Its theoretically informed claim, that gender should only be measured at a group or social level, signifies a radical departure from traditional psychological thinking. It offers the potential that gender measurement may appeal to those working outside the confines of traditional psychology where gender is more readily understood as a social, as opposed to an individual, level phenomenon. This brings, in particular, the benefit of large-group normative data. The challenge now exists to develop adequate measures of gender (re)presentation. It is hoped this paper serves as a foundation.

Acknowledgements

Research funds were made available through the Commonwealth Commission and Universities UK.

Notes

1. A unipolar continuum contains numerical scores that extend in one direction and are interpreted in terms of a single end point. A bipolar continuum contains numerical scores that extend in two directions and are interpreted in terms of opposite end points (Lavrakas 2008).
2. A unidimensional continuum assesses a construct through interpreting numerical scores on a single measure. A multidimensional continuum assesses a construct through interpreting numerical scores on two or more measures (Rust and Golombok 2009).
3. Whilst different M–F measures share variance, indicating that they have something in common, they are not entirely comparable. It would be fair to expect relatively strong correlations to emerge between M–F measures should they be assessing the same construct – particularly given their high individual reliability (Constantinople 1973). Yet, these correlations remain modest at best. These findings are strengthened by results indicating that correlations between measures are considerably lower than their individual reliabilities (Williams and Best 1990a). Factor analytic studies of M–F measures, which seek to determine their main underlying conceptual structure through simplifying data, also question the unidimensionality of the M–F construct (e.g. Ratliff and Conley 1981).
4. This method of measurement development is often referred to as criterion keying. It involves selecting items that correlate to a relevant criterion variable. It is a purely empirical approach as there is no underlying conceptual or theoretical rationale to measurement development. It is only interested in identifying items that can discriminate between pre-defined groups. Items demonstrating a weak relationship to the relevant criterion variable are dropped. These measures may contain items that contain little meaningful relationship to the criterion (Coaley 2009).

Notes on contributor

Russell Luyt is a senior lecturer in the Department of Psychology at Anglia Ruskin University. He is particularly interested in identity processes as well as debates concerning research philosophy and qualitative methodology. His work is grounded within a critical approach and has, to date, focused primarily on the social psychology of gender. This has included, for example, the critique of traditional masculinity, femininity and gender measurement; media representations; masculinities and aggression; as well as the intersection of gender and other social categories. He is currently extending these lines of research and developing his interest in radicalisation among marginalised communities from a gender perspective.

References

- Abell, J. and Walton, C., 2010. Imagine: towards an integrated and applied social psychology. *British journal of social psychology*, 49, 685–690.
- Austin, J., 1962. *How to do things with words*. Cambridge, MA: Harvard University Press.
- Bandura, A., 1977. *Social learning theory*. New York: General Learning Press.
- Beall, A.E., 1993. A social constructionist view of gender. In: A.E. Beall and R.J. Sternberg, eds. *The psychology of gender*. New York: Guilford Press, 127–147.
- Beere, C.A., 1990a. *Gender roles: a handbook of tests and measures*. Westport, CT: Greenwood.
- Beere, C.A., 1990b. *Sex and gender issues: a handbook of tests and measures*. Westport, CT: Greenwood.
- Belk, S.S. and Snell, W.E., 1986. The Beliefs About Women Scale (BAWS): scale development and validation. *Social and behavioral sciences documents*, 16, 10.
- Bem, S.L., 1974. The measurement of psychological androgyny. *Journal of clinical and consulting psychology*, 42, 155–162.
- Bem, S.L., 1981a. Gender schema theory: a cognitive account of sex typing. *Psychological review*, 88, 354–364.
- Bem, S.L., 1981b. The BSRI and gender schema theory: a reply to Spence and Helmreich. *Psychological review*, 88, 369–371.
- Berger, P.L. and Luckmann, T., 1966. *The social construction of reality: a treatise in the sociology of knowledge*. New York: Anchor.
- Billig, M., 1998a. Rhetoric and the unconscious. *Argumentation*, 12, 199–216.
- Billig, M., 1998b. Rhetorical and discursive analysis: how families talk about the royal family. In: N. Hayes, ed. *Doing qualitative analysis in psychology*. Hove: Psychology Press, 39–54.
- Bohan, J.S., 1993. Regarding gender: essentialism, constructionism, and feminist psychology. *Psychology of women quarterly*, 17, 5–21.
- Bohan, J.S., 2002. Sex differences and/in the self: classic themes, feminist variations, postmodern challenges. *Psychology of women quarterly*, 26, 74–88.
- Chu, J.Y., Porche, M.V. and Tolman, D.L., 2005. The Adolescent Masculinity Ideology in Relationships Scale: development and validation of a new measure for boys. *Men & masculinities*, 8, 93–115.
- Coaley, K., 2009. *An introduction to psychological assessment and psychometrics*. London: Sage.
- Connell, R.W., 1992. A very straight gay: masculinity, homosexual experience, and the dynamics of gender. *American sociological review*, 57, 735–751.
- Constantinople, A., 1973. Masculinity-femininity: an exception to a famous dictum? *Psychological bulletin*, 80, 389–407.
- Constantinople, A., 1979. Sex-role acquisition: in search of the elephant. *Sex roles*, 5, 121–133.
- Deaux, K., 1985. Sex and gender. *Annual review of psychology*, 36, 49–81.
- Diamond, M., 2000. Sex and gender: same or different? *Feminism and psychology*, 10, 46–54.
- Edley, N. and Litosseliti, L., 2010. Contemplating interviews and focus groups. In: L. Litosseliti, ed. *Research methods in linguistics*. London: Continuum, 155–179.
- Edley, N. and Wetherell, M., 1997. Jockeying for position: the construction of masculine identities. *Discourse & society*, 8, 203–217.
- Fausto-Sterling, A., 1993. The five sexes: why male and female are not enough. *The sciences*, 33, 20–24.
- Flax, J., 1987. Postmodernism and gender relations in feminist theory. *Signs*, 12, 621–643.
- Floyd, F.J. and Widaman, K.E., 1995. Factor analysis in the development and refinement of clinical assessment instruments. *Psychological assessment*, 7, 286–299.

- Foucault, M., 1978. *The history of sexuality*. Vol. 1. R. Hurley, Trans. Harmondsworth: Penguin.
- Freud, S., 1905/2000. *Three essays on the theory of sexuality*. New York: Basic Books.
- Furr, R.M. and Bacharach, V.R., 2008. Response biases. In: R.M. Furr and V.R. Bacharach, eds. *Psychometrics: an introduction*. London: Sage, 239–264.
- Garfinkel, H., 1967. *Studies in ethnomethodology*. Englewood Cliffs, NJ: Prentice Hall.
- Garnets, L. and Pleck, J.H., 1979. Sex role identity, androgyny, and sex role transcendence: a sex role strain analysis. *Psychology of women quarterly*, 3, 270–283.
- Gibbons, J.L., Hamby, B.A. and Dennis, W.D., 1997. Researching gender-role ideologies internationally and cross-culturally. *Psychology of women quarterly*, 21, 151–170.
- Glick, P. and Fiske, S.T., 1999. The Ambivalence toward Men Inventory: differentiating hostile and benevolent beliefs about men. *Psychology of women quarterly*, 23, 519–536.
- Hathaway, S. and McKinley, J., 1943. *The Minnesota Multiphasic Personality Inventory*. Minneapolis: University of Minnesota Press.
- Hoffman, R.M., 2001. The measurement of masculinity and femininity: historical perspective and implications for counselling. *Journal of counseling and development*, 79, 472–485.
- Kohlberg, L., 1966. A cognitive-developmental analysis of children's sex-role concepts and attitudes. In: E.E. Maccoby, ed. *The development of sex differences*. London: Tavistock, 82–173.
- Lavrakas, P.J., 2008. *Encyclopedia of survey research methods*. London: Sage.
- Levant, R.F., 1996. The new psychology of men. *Professional psychology*, 27, 259–265.
- Levant, R.F. and Fischer, J., 1998. The Male Role Norms Inventory. In: C.M. Davis, et al., eds. *Handbook of sexuality-related measures*. Thousand Oaks, CA: Sage, 469–472.
- Levant, R.F. et al., 1992. The male role: an investigation of contemporary norms. *Journal of mental health counselling*, 14, 325–337.
- Levant, R.F. and Majors, R.G., 1997. Masculinity ideology among African American and European American college women and men. *Journal of gender, culture and health*, 2, 33–43.
- Levant, R.F., Majors, R.G. and Kelley, M.L., 1998. Masculinity ideology among young African American and European American women and men in different regions of the United States. *Cultural diversity and mental health*, 4, 227–236.
- Levant, R.F. et al., 2007. Initial validation of the Male Role Norms Inventory-Revised (MRNI-R). *The journal of men's studies*, 15, 83–100.
- Levant, R.F., Wu, R. and Fischer, J., 1996. Masculinity ideology: a comparison between U.S. and Chinese young men and women. *Journal of gender, culture and health*, 1, 207–220.
- Luyt, R., 2012a. Constructing hegemonic masculinities in South Africa: the discourse and rhetoric of heteronormativity. *Gender and language*, 6, 47–77.
- Luyt, R., 2012b. A framework for mixing qualitative and quantitative methods in measurement development, validation and revision: a case study. *Journal of mixed method research*, 6, 294–316.
- Marecek, J., Crawford, M. and Popp, D., 2004. On the construction of gender, sex, and sexualities. In: A.H. Eagly, A.E. Beall and R.J. Sternberg, eds. *The psychology of gender*. London: Guildford Press, 192–216.
- Martin, C.L. and Halverson, C.F., 1981. A schematic processing model of sex typing and stereotyping in children. *Child development*, 52, 1119–1134.
- Marx, K., 1867–1894/1981. *Capital: a critique of political economy*. Vols. 1–3. Harmondsworth: Penguin.
- McHugh, M.C. and Frieze, I.H., 1997. The measurement of gender-role attitudes: a review and commentary. *Psychology of women quarterly*, 21, 1–16.
- Mischel, W., 1966. A social-learning view of sex differences in behaviour. In: E.E. Maccoby, ed. *The development of sex differences*. London: Tavistock, 56–81.
- Morawski, J.G., 1987. The troubled quest for masculinity, femininity, and androgyny. In: P. Shaver and C. Hendrick, eds. *Sex and gender*. Newbury Park, CA: Sage, 44–69.
- Pleck, J.H., 1976. The male sex role: definitions, problems, and sources of change. *Journal of social issues*, 32, 155–164.
- Pleck, J.H., 1981. *The myth of masculinity*. Cambridge, MA: MIT Press.
- Pleck, J.H., 1987. The theory of male sex-role identity: its rise and fall, 1936 to the present. In: H. Brod, ed. *The making of masculinities*. Boston, MA: Allen & Unwin, 21–38.
- Pleck, J.H., 1995. The gender role strain paradigm: an update. In: R.F. Levant and W.S. Pollack, eds. *A new psychology of men*. New York: Basic Books, 11–32.

- Potter, J., 1996. Attitudes, social representations, and discursive psychology. In: M. Wetherell, ed. *Identities, groups and social issues*. London: Sage, 119–173.
- Potter, J., 1998. Discursive social psychology: from attitudes to evaluative practices. *European review of social psychology*, 9, 233–266.
- Potter, J. and Wetherell, M., 2001. Discourse analysis. In: A. Smith, R. Harré and L.V. Langenhove, eds. *Rethinking methods in psychology*. London: Sage, 80–92.
- Potter, J. and Wetherell, M., 2006. Unfolding discourse analysis. In: M. Wetherell, S. Taylor and S.J. Yates, eds. *Discourse theory and practice: a reader*. London: Sage, 198–209.
- Pryzgodna, J. and Chrisler, J.C., 2000. Definitions of gender and sex: the subtleties of meaning. *Sex roles*, 43, 553–569.
- Ratliff, E.S. and Conley, J., 1981. The structure of masculinity–femininity: multidimensionality and gender differences. *Social behavior and personality*, 9, 41–47.
- Roopnarine, J.L. and Mounts, N.S., 1987. Current theoretical issues in sex roles and sex typing. In: D.B. Carter, ed. *Current conceptions of sex roles and sex typing: theory and research*. New York: Praeger, 7–31.
- Rubin, G., 1975. The traffic in women: notes on the political economy of sex. In: R. Reiter, ed. *Towards an anthropology of women*. New York: Monthly Review Press, 157–210.
- Rust, J. and Golombok, S., 2009. *Modern psychometrics: the science of psychological assessment*. London: Routledge.
- Sacks, H., 1964–1965/1992. *Lectures on conversation*. Oxford: Basil Blackwell.
- Schwartz, N. and Bohner, G., 2001. The construction of attitudes. In: A. Tesser and N. Schwarz, eds. *Blackwell handbook of social psychology: intraindividual processes*. Oxford: Blackwell, 436–457.
- Smiler, A.P. and Epstein, M., 2010. Measuring gender: options and issues. In: J.C. Chrisler and D.R. McCreary, eds. *Handbook of gender research in psychology: Vol. 1. Gender research in general and experimental psychology*. New York: Springer, 133–157.
- Stokoe, E.H., 1998. Talking about gender: the conversational construction of gender categories in academic discourse. *Discourse & society*, 9, 217–240.
- Stryker, S., 1995. Review essay: progress? Well, yes and no: the states of sociological social psychology. *Social psychology quarterly*, 58, 326–331.
- Thompson, E.H. and Pleck, J.H., 1986. The structure of the male role norms. *American behavioural scientist*, 29, 531–543.
- Thompson, E.H. and Pleck, J.H., 1995. Masculinity ideologies: a review of research instrumentation on men and masculinities. In: R.F. Levant and W.S. Pollack, eds. *A new psychology of men*. New York: Basic Books, 129–163.
- Thompson, E.H., Pleck, J.H. and Ferrera, D.L., 1992. Men and masculinities: scales for masculinity ideology and masculinity-related constructs. *Sex roles*, 27, 573–607.
- Tolman, D.L. and Porche, M.V., 2000. The Adolescent Femininity Ideology Scale: development and validation of a new measure for girls. *Psychology of women quarterly*, 24, 365–376.
- Unger, R.K. and Crawford, M., 1993. Commentary: sex and gender: the troubled relationship between terms and concepts. *Psychological science*, 4, 122–124.
- West, C. and Zimmerman, D.H., 1987. Doing gender. *Gender & society*, 1, 125–151.
- Wiggins, S. and Potter, J., 2003. Attitudes and evaluative practices: category vs. item and subjective vs. objective constructions in everyday food assessments. *British journal of social psychology*, 42, 513–531.
- Wilkinson, S., 2001. Theoretical perspectives on women and gender. In: R.K. Unger, ed. *Handbook of the psychology of women and gender*. Hoboken, NJ: Wiley, 17–28.
- Williams, J.E. and Best, D.L., 1990a. *Sex and psyche: gender and self viewed cross-culturally*. Newbury Park, CA: Sage.
- Williams, J.E. and Best, D.L., 1990b. *Measuring sex stereotypes: a multination study*. Newbury Park, CA: Sage.
- Williams, W.L., 1987. Women, men, and others. *American behavioral scientist*, 31, 135–141.
- Windle, M.S., 1987. Measurement issues in sex roles and sex typing. In: D.B. Carter, ed. *Current conceptions of sex roles and sex typing: theory and research*. New York: Praeger, 33–45.
- Wooffitt, R., 2005. *Conversation analysis and discourse analysis: a comparative and critical introduction*. London: Sage.