

**Exploring the meaning of gender:
Evaluating and revising the Bem Sex-Role
Inventory (BSRI) for a Swedish research
context (BSRI-SE)**

Roland S. Persson

INSIKT 1999:1

Vetenskapliga rapporter från HLK

Abstract

This research evaluates the Bem Sex-Role Inventory (BSRI) for use in a Swedish setting, and in the process of so doing also compares the distribution of American and Swedish gender roles as elicited by the inventory. A sample of 118 individuals (48 males and 70 females) was used in order to arrive at norms better suited to a Swedish context than those provided by Bem (1974). Reliability and factor analyses were performed and a revised version of the inventory, with acceptable psychometric properties for the Swedish sample, was arrived at and termed BSRI-SE. Comparisons between the original American and the Swedish sample show that differences in the way that femininity, masculinity, androgyny and undifferentiated gender are distributed are not statistically significant. Raw score data for the BSRI-SE, as well as the scoring sheet, are provided for continued standardisation of the inventory. Methods for classification and further research are discussed.

Field of study: Social psychology

Keywords: Gender, Androgyny, Masculinity, Femininity, Psychological Measurement, Cross-cultural comparisons, Identity, Differentiation, Personality.

Index

Introduction,	4
The Bem Sex-Role Inventory (BSRI),	8
The BSRI in a Swedish context,	17
The nature of the Swedish sample,	17
Reliability of the inventory scales,	20
Procedure for classification and standardisation,	27
The median split method,	28
The M-F Difference score,	29
Are Swedes masculine, feminine or androgynous?	32
Sex-typing in the sample subgroups,	35
Concluding remarks,	36
References,	40
Appendices	
I Test sheet for the BSRI-SE	
II Raw scores of the Swedish sample	

Introduction

Few issues in recent years have been subject to so much debate, research, exploitation and politics as have gender roles, their meaning and social impact. While researchers agree that sex is biologically determined they do not agree on when, why and how psychological differentiation between gender roles occur. To argue that something is typically male or female is by no means theoretically straight-forward (Sternberg, 1993), and there is a great deal of variety across cultures of what is regarded as typically male or female behaviour. Although there are some universals with regard to gender, societies differ in the value attached to maintenance of differences between the sexes. Generally, males are more likely to be physically aggressive and express dominance over women rather than vice versa. Women are more likely in a global perspective to conform, defer, comply and submit to authority, especially if the authority is male (Segall, Dasen, Berry & Poortinga, 1990). Men and women's self-perceptions are less stereotypical in more economically and socially developed countries, and it appears that as sexual-equality ideals spread, for example through feminist ideology, behavioural differences between the sexes diminish and prevailing stereotypes are increasingly being blurred (Williams & Best, 1989).

Gender is a multi-faceted phenomenon. Most would think of gender in terms of feminine and masculine, thus expressing that gender is confined to a dichotomy. But there is in fact no agreement amongst behavioural scientists on how many genders there are! Is it

feasible to have as many genders as there are sexes, namely two? Or are there also “nuances” of both, which should perhaps be regarded as separate genders? Note that by gender is meant the psychological array of behaviours, that to a large extent are socially determined, and that we have a tendency to identify as typical of either sex. In some cases there are biological causes that could possibly qualify the conceptualisation of more genders than the traditional two. Sex chromosome disorders like Turner’s Syndrome (individuals born with only one X chromosome), Klinefelter’s Syndrome (individuals born with three X chromosomes or two X chromosomes and one Y chromosome) or the XYY Syndrome are all genetic disorders, which have distinctive influences on an individual’s secondary sexual characteristics like body hair, breasts, body build and so on. With these disorders follow some differentiation in cognitive functioning (cf. Willerman & Cohen, 1990 for an overview). In addition, there are a number of different sexual behaviours that challenge the traditional division of gender restricted to the dichotomy of male and female. Transsexualism, for example, describes incongruity between gender identity and anatomical sex, the aetiology of which could possibly be due to endocrine factors rather than in various ways social learning (Hoenig & Kenna, 1974). Similarly, male transvestism could be understood as the co-existence of two selves: male and female (Larsson, 1997). Homosexuality is also in a sense a challenge to traditional gender roles (Bailey, 1996; Ross, 1985; Williams, 1996), particularly since recent research is increasingly showing that homosexuality may be genetically and/or hormonally determined (Ellis, 1996; LeVay, 1993). A majority of researchers tend to stress

social aspects instrumental to developing sexual behaviour (Howells, 1987). Money and Erhardt (1972), for example, reviewed a large number of studies on sex hormone anomalies and concluded that the most important factor in developing sex role and gender identity is the assigned sex of rearing. However, such research may well be in error if understanding homosexuality as merely a social construction of no or little evolutionary significance, as elegantly argued by McKnight (1997). Genetics and human physiology seem to play important, but hitherto often wilfully ignored, roles in the development of a homosexual identity. Murphy (1990) provocatively points out that “the incentive to discover the origins of homosexuality seems to belong to those who find homosexuality a pathological, sinful, immoral or criminal condition. At least on the basis of these views there is reason to try and understand the origins of homosexual behaviour if only to prevent and eliminate it. It is ordinarily some deficit which prompts medicine and the rest to reach for a causal explanation of behaviour” (p. 134).

Thus, on the basis of culture, and invariably at some level, also biological sex—which at times by no means is always straight forward—it may well be possible to construe gender roles in terms of a masculine male, a feminine male, a masculine woman, a feminine woman or any nuance of these.

Rudberg and Bjerrum-Nielsen (1994) point out, that “gender is still relevant [in modern society]. There is nothing in either the family’s gender socialisation or the history of modernity that suggests that gender should be ‘suspended’ as a psychological or social category; rather it seems that gender will survive in a new form” (p. 49-50). The

understanding of gender is a field of knowledge in flux not only because it intrigues social scientists due to its complexity, but also because the results of research on possible causes and effects in this context have strong political repercussions. It is historically true in the Western world that women, often at great risk, have struggled in order to obtain the same rights that by unquestioned tradition have been the privilege of men only. However, during the process by which such rights and privileges are increasingly recognised for both sexes, the biased function of rationalisation should also be recognised.

While it is politically essential to argue that men and women have equal rights in a democracy it is simultaneously paramount for science to investigate the nature of differentiation. But where the political agenda is set differences between gender roles would appear not to be a favoured object of discussion or debate—at least in a context of equal opportunities on the job market. Rather issues are focussed which may justify a certain political development. In such a situation it is far more opportune to focus how men and women are similar rather than different, since anything different from what is considered typically masculine not infrequently is interpreted as inferior. Although many stereotypes about gender differentiation are false—like for example the notion prevalent amongst men and school teachers, that women generally are believed to be less intelligent than men (Broverman, Vogel *et al.*, 1972)—it is my impression that real differences often are underplayed in favour of making women more like men (cf. Lipman-Blumen, 1989). Baumeister (1988), for example, argues that differences between sexes should not be studied at all, whereas McHugh, Koeske and Frieze (1986) suggest that gender differences

should only be reported under limited circumstances

This research targets the social differentiation of gender believing it to be essential. First, it sets out to evaluate the Bem Sex-Role Inventory (Bem, 1974) in a Swedish context, revising it to make it more akin to Swedish culture rather than American culture, and to provide norms by which the revised inventory may be compared. Similar efforts have been done in several Hispanic settings (e.g. Kaschak & Sharrat, 1983; Kranau, Green & Valencia-Weber, 1982). Some studies have used the BSRI for cross-cultural comparisons in Israel (Maloney, Wilkof & Dambrot, 1981) and Australia (Rowland, 1977), but both of these used the BSRI original norms on the slightly dubious assumption that the understanding of gender identity and behaviour is roughly the same as in the United States.

An instrument such as the BSRI, however, has value for research purposes, since gender consideration often is an important variable in Social Science research. Few instruments are developed or translated and standardised for a Swedish population. For example, the publishing branch of the Swedish Psychological Society in their 1997 catalogue lists 111 different tests and test batteries, which should be compared to The American Psychological Association's estimate that some 20.000 tests are developed annually (APA, 1993).

Second, and a result of the process of evaluating the BSRI in a Swedish context, is an exploration into the meaning of Swedish gender roles. How does the Swedish sample compare to the American sample, which served as the basis for developing the American norms by which BSRI scores are compared? Are Swedes more or less masculine,

feminine, androgynous than Americans?

The Bem Sex-Role Inventory (BSRI)

The BSRI was developed by Stanford-psychologist Sandra Bem (1974; 1975; 1981) in an effort to accommodate the fact that there are more nuances to being masculine or feminine than the theoretical constructs at the time would allow to be expressed (e.g. The California Psychological Inventory by Gough, 1957). This sex-role dichotomy, Bem (1974) argues, “has served to obscure two very plausible hypotheses: first, that many individuals might be ‘androgynous’; that is, they may be *both* masculine and feminine ... depending on the situational appropriateness of these various behaviours; and conversely, that strongly sex-typed individuals might be seriously limited in the range of behaviours available to them as they move from situation to situation” (p. 155).

In Bem’s research effort masculinity and femininity are regarded as two independent dimensions rather than two ends of a single dimension. This conceptualisation of gender allows any individual to indicate whether he or she is high on both dimensions (and thus androgynous), low on both dimensions (and therefore undifferentiated) or high on one and low on the other (either feminine or masculine). Three scales were originally developed in order to arrive at an instrument, which would allow these distinctions to be made in reference to stereotypical gender roles, namely Femininity, Masculinity and Social Desirability.

An item pool was selected by Bem and her students on the

criterion that descriptions of masculinity and femininity should describe something positive and stereotypically male or female. A personality characteristic qualified as typically masculine if independently judged by both females and males to be significantly more desirable for a man than a woman (Table 1). Similarly, a characteristic qualified as typically feminine if judged by both males and females to be significantly more desirable for a woman than for a man (Table 2).

The items relating to social desirability were culled from characteristics, which seemed to be neither typically feminine nor masculine but rather applicable to both sexes and thus neutral in terms of sex-appropriateness. Of these half were positive in tone and half were negative. Items were judged neutral if they were independently judged by both males and females to be no more desirable for one sex than for the other (Table 3). Note, however, that this scale mainly served an initial purpose in constructing the instrument “to insure that the inventory would not simply be tapping a general tendency to endorse socially desirable traits” (Bem, 1974, p. 156).

Table 1. *Items of the BSRI Masculinity Scale (Bem, 1974; 1981). Translation into Swedish included as well as their variable labels by which they were processed.*

English original labels	Swedish translation	Variable
Self-reliant	Full av självförtroende	M1SJLVFR
Defends own beliefs	Försvarande mina övertygelser	M2FORSV
Independent	Oberoende	M3OBERO
Athletic	Sportslig	M4SPORT
Assertive	Bestämd	M5BESTMD
Strong personality	En stark personlighet	M6STPER
Forceful	Kraftfull	M7KRFTFL
Analytical	Analytisk	M8ANALYT
Has leadership abilities	En ledartyp	M9LEDARE
Willing to take risks	Villig att ta risker	M10RISK
Makes decisions easily	Lätt för att ta beslut	M11BSLUT
Self-sufficient	Själv tillräcklig	M12SJLVT
Dominant	Dominant	M13DOMIN
Masculine	Maskulin	M14MASK
Willing to take a stand	Villig att fatta beslut och stå för dem	M15BSLST
Aggressive	Aggressiv	M16AGGRS
Act as a leader	Att agera ledare	M17LEDAR
Individualistic	Individualist	M18INDIV
Competitive	Ha en tävlingsmentalitet	M19TAVL
Ambitious	Ambitiös	M20AMBI

Table 2. *Items of the BSRI Femininity Scale (Bem, 1974; 1981). Translation into Swedish included as well as their variable labels by which they were processed.*

English original	Swedish translation	Variable labels
Yielding	Undfallande	F1UNDFAL
Cheerful	Munter	F2MUNTR
Shy	Blyg	F3BLYG
Affectionate	Tillgiven	F4TILLGV
Flatterable	Mottaglig för smicker	F5SMICKR
Loyal	Lojal	F6LOJAL
Feminine	Feminin	F7FEMINI
Sympathetic	Sympatisk	F8SYMPAT
Sensitive to the needs of others	Känslig för andras behov	F9KNSLBH
Understanding	Förstående	F10FORST
Compassionate	Medlidsam	F11MDLID
Eager to sooth hurt feelings	Angelägen att lindra sårade känslor	F12LINDR
Soft spoken	Mild och stillsam	F13MILD
Warm	Varm person	F14VARM
Tender	Ömsint	F15OMSIN
Gullible	Lättlurad	F16LTLUR
Childlike	Barnslig	F17BARNS
Does not use harsh harsh language	Använder inte svordomar	F18SVOR
Loves children	Barnkär	F19BARKR
Gentle	Varsam	F20VARSM

Table 3. *Items of the BSRI Social Desirability Scale (Bem, 1974; 1981). Translation into Swedish included as well as their variable labels by which they were processed.*

English original	Swedish translation	Variable labels
Helpful	Hjälpsam	SD1HJALP
Moody	Lynnig	SD2LYNN
Conscientious	Samvetsgrann	SD3SMVTG
Theatrical	Teatralisk	SD4TEATR
Happy	Glad	SD5GLAD
Unpredictable	Oförutsägbar	SD6OFRSB
Reliable	Pålitlig	SD7PLIT
Jealous	Svartsjuk	SD8SVART
Truthful	Sanningsenlig	SD9SANN
Secretive	Hemlighetsfull	SD10HEML
Sincere	Uppriktig	SD11UPPR
Conceited	Inbilsk	SD12INB
Likeable	Trevlig	SD13TREV
Solemn	Allvarlig	SD14ALLV
Friendly	Vänlig	SD15VANL
Inefficient	Ineffektiv	SD16INEF
Adaptable	Anpassningsbar	SD17ANPA
Unsystematic	Osystematisk	SD18OSYS
Tactful	Finkänslig	SD19FINK
Conventional	Konventionell	SD20KONV

During construction judges used a seven-point Likert-scale, ranging from one (not at all desirable) to seven (extremely desirable) when rating items. Note that the inventory outlines *American* stereotypes of gender roles, which are not necessarily transferable to other nations and cultures. Bem is aware of the fact that the inventory outlines feminine and masculine traits desirable to an American

population. It was emphasised during the construction of the instrument that participants should voice what they believed to be an opinion *general* to American society rather than uniquely their own.

The psychometric properties of Bem's original instrument are fair. Internal consistency was evaluated by allowing separate samples of men and women provide scores for both the feminine and masculine scales respectively: Both scales as scored by women show that the F-scale ($\alpha = .75$) as well as the M-scale ($\alpha = .78$) have an acceptable degree of internal consistency. Similarly, both scales as scored by the male sample show an acceptable degree of consistency as well: F-scale ($\alpha = .87$) and M-scale ($\alpha = .86$).

Test-retest reliability is also acceptable, estimated in the same way with men and women providing separate scores. Test-retest reliability ranges from $r = .76$ to $r = .94$ (i. e. from the female sample the F-scale yields $r = .82$ and M-scale $r = .94$, whereas for the male sample F-scale yields $r = .89$ and M-scale $r = .76$).

The neutral Social Desirability Scale serves the purpose of safeguarding the conceptualisation of masculinity and femininity as separate constructs, but serves no purpose when using the completed and validated inventory. By allowing females and males separately to judge both the F-scale and the M-scale, Bem showed that sex-appropriate behaviour—which is deemed to be socially desirable—as judged by either sex, was significantly higher when females judged feminine items and when males judged masculine items in comparison to when males judged feminine items and women judged masculine item. There was, however, no such significant difference when both sexes judged the neutral Social Desirability Scale. This prompted Bem

to conclude that Femininity and Masculinity are indeed separate constructs. In addition, she goes on to show that both the participating men and women are in significant agreement as to what qualifies as “sex-appropriate” behaviour and “sex-inappropriate” behaviour. In other words, when the men felt that a certain characteristic is sex-inappropriate for a man or a woman the female participants would tend to agree. And conversely, when women felt that a certain characteristic is sex-inappropriate for a man or a woman the participating men would largely concur.

Androgyny, the measure of which is the ultimate objective of reconceptualising masculinity and femininity, reflects the relative levels of masculinity and femininity that any individual includes in their self-description. Androgyny is thus expressed as an index that shows the extent to which a person is sex-typed on the basis of how they score on the M-scale and F-scale together. The greater the value of the Androgyny Index (AI) the more the person is sex-typed or sex-reversed. The smaller the value of AI the more an individual could be considered androgynous.

Bem has proposed three ways of calculating AI. Originally AI was defined as Student’s t - ratio for the difference between masculine and feminine self-endorsement. In other words AI is the difference between an individual’s Masculinity and Femininity scores normalised with respect to the standard deviations of the scores from both scales and expressed as the Student’s t - ratio. In this line of reasoning individuals are classified as sex-typed, masculine or feminine if the androgyny t - ratio reaches statistical significance ($|t| \geq 2.2025$, $df = 38$, $p < .05$). An individual is classified as androgynous if the absolute

value of the t - ratio is less or equal to one. Bem points out that these cut-off points are somewhat arbitrary and may be adjusted if need be.

Bem offers a simpler way of assessing the AI. If multiplying the difference score between the F-scale and the M-scale by a conversion factor of 2.322 one arrives at an index which correlates nigh perfectly with the t - ratio calculation ($r = .98$). This conversion factor was derived from the original American normative sample ($n = 917$), but Bem does not convey by what means. Note that this “short-cut” was used by, for example, Hassler (1991) in her experimental studies of spatial abilities, musical composition ability and their relationship to the Bem’s construct of psychological androgyny.

The original three-way classification by Bem into gender-role as masculine, feminine or androgynous, however, was criticised by Spence, Helmreich and Stapp (1975). They argued that classification based on Student’s t - ratio obscures one important distinction made possible by understanding femininity and masculinity as separate constructs, namely the fact that it is also possible to score high on both scales or low on both scales rather than preferring one to the other. Bem (1977) investigated the possibility to consider a four-way classification, and found that the critique of merely a three-way classification was warranted. At a later stage in the development of the inventory, therefore, as the BSRI became commercially available, the procedures for classification slightly changed. Bem (1981) now recommends that classification be done by means of a median split classifying individuals into four categories of gender role orientation: Undifferentiated, Androgynous, Feminine and Masculine. The step-by-step procedure by which to apply the median split for classification

will be discussed below and in relation to the Swedish revision of the BSRI.

The normative samples used by Bem in developing the instrument consists of approximately 900 individuals made up of the following groups: Caucasian undergraduates ($n = 32$), Afro-American undergraduates ($n = 63$), Hispanic undergraduates ($n = 35$), Adolescents aged 14 - 17 ($n = 29$), Age group 20 - 30 ($n = 108$), Age group 31 - 65 ($n = 55$), Older adults ages >62 ($n = 60$) and Psychiatric in-patients ($n = 55$). The averaged norms for these groups are given in Table 4 below.

The distribution of categorisations in the normative sample according to Bem's (1974) original classification scheme, and as based on Student's averaged t - ratios, shows that amongst males 7 % are sex-typed as feminine, 7 % as near-feminine, 39 % as androgynous, 18 % as near-masculine and 29 % as masculine. Amongst females on the other hand, 37 % are sex-typed as feminine, 14 % as near-feminine, 32 % as androgynous, 9 % as near-masculine and 8 % as masculine.

Table 4. Raw score means, medians and standard deviations for the *F*-scale, *M*-scale and the *F*-minus-*M* difference score of the normative sample. Note that the category sexes combined has been statistically weighted in order to correct the unequal numbers of men and women making up the sample (Bem, 1981)

	Sexes combined	Females (<i>n</i> = 340)	Males (<i>n</i> = 476)	<i>t</i>
Femininity				
Mean	4.82	5.05	4.59	11.95*
<i>MD</i>	4.90	5.10	4.60	
<i>SD</i>	.59	.53	.55	
Masculinity				
Mean	4.95	4.79	5.12	7.03*
<i>MD</i>	4.95	4.80	5.10	
<i>SD</i>	.68	.66	.65	
F-minus-M				
Mean	-.01	6.30	-6.33	13.09*
<i>MD</i>	.97	6.83	-6.50	
<i>SD</i>	14.94	13.35	13.37	

* $p < .001$

Table 5. Distribution of subjects in the combined normative samples into sex-role categories according to the four-way classification using the median split method (Bem, 1974)

Sex-role category	Females (%)	Males (%)
Feminine	39	12
Masculine	12	42
Androgynous	30	20
Undifferentiated	18	27

Distribution of categories in the same sample but as based on median splits are accounted for in Table 5.

It is possible to evaluate how the two classification schemes compare if the values of near-feminine and near-masculine are combined and regarded—for the purpose of comparison—as more or less the same as being undifferentiated. In the male sample the t - ratio classification corresponds only moderately to the median split classification ($r = 0.45$), whereas in the female sample there is little difference between two schemes ($r = 0.949$).

The BSRI in a Swedish context

The nature of the Swedish sample

The Swedish sample ($n = 118$) is relatively small in comparison to the normative sample used by Bem. The objective of the present research, however, is mainly to investigate how the BSRI fares in a Swedish context rather than to strictly standardise the instrument anew in a fully fledged large-scale effort. On the other hand, Swedish norms will be produced and the present data may indicate the psychometric value of the instrument as such. At the very least this evaluation may certainly provide a base of data, which can be used to probe psychometric properties further. For this purpose the raw mean scores are appended to this report (Appendix II). A breakdown of the Swedish is provided in Table 6. Note that there is an over-representation of females in the sample (41 % males and 59 % females). This will need to be considered as the normative scores are produced. The same over-

representation problem is evident in almost exactly the same proportions in Sandra Bem's normative sample. However, in Bem's sample over-representation is reversed (42 % are female and 58 % are male).

The original American BSRI from 1974 was translated into Swedish and the administered answering sheet included all three scales (see Tables 1 - 3 above for a comparison between the original American items and their translation into Swedish). A majority of the respondents filled in the translated during class and always because course content qualified or included a discussion of gender-roles or socialisation. The inventory was always completed *prior* to any discussion of gender-roles, however, and respondents were only provided with the information needed complete the task.

Scale items on the scoring sheets were compiled in such a manner that order effects were avoided. It was not possible for participants to perceive that in fact three scales rather than one were being rated. The order of the items was set in groups of three in the following order: Masculine, Feminine and Neutral, which is also the order used by Bem. Note that Swedish participants were encouraged to rate themselves, which is contrary to Bem's instructions to the American sample. Bem asked that participants should express what they felt was typically American in a general sense rather than using the BSRI to provide a personal profile.

Table 6. *The different subgroups of the Swedish test sample. Note that the percentages pertaining to each sex is the relative frequency of the subgroup, whereas the relative frequency of n relates to N.*

Subgroup	n	M	F	Age			
				Mean	Min	Max	SD
Pre-school student teachers Relative frequency (%)	24	4 (20)	20 (17)	23.3 (83)	19	40	5.01
Communication Studies students Relative frequency (%)	49	11 (42)	38 (22)	24.7 (78)	20	44	5.21
Amateur football players Relative frequency (%)	12	12 (10)	-- (100)	39.4 --	25	54	9.1
Pre-school teachers Relative frequency (%)	16	4 (14)	12 (25)	48.9 (75)	24	61	9.55
Comprehensive student teachers Relative frequency (%)	12	12 (10)	-- (100)	36.0 --	21	51	10.45
Various university staff Relative frequency (%)	5	5 (4)	-- (100)	43.0 --	18	60	16.07
All subgroups (N) Relative frequency (%)	118	48 (100)	70 (41)	31.2 (59)	18	6	112.00

All the Swedish participants completed the answering sheets impeccably. At all occasions students were provided with immediate feedback as everyone had completed the task. That is, after completion the students themselves were allowed to score their sheets after being informed which items belonged to which scale. Scoring was straightforward and students were able to estimate themselves whether they were high - high, low - low or high - low on either the M-

scale or F-scale. The neutral Social Desirability Scale was ignored. Participants only had access to their own score so as to keep the ethical integrity of the testing situation. No further conclusions were drawn by students except for their relative standing in comparison to the American-elicited items.

However, the administration tended to trigger interesting discussions on gender and the face validity of the instrument. In some cases, particularly with regard to the Communication Studies students, the inventory caused some merriment in the class. They argued that items were “out-dated” and not at all appropriate for outlining masculine and feminine stereotypical behaviour in Sweden.

In evaluating the BSRI in the following estimations are made on the basis of three types of scores: judgements by the men of the sample, judgements made by the women of the sample and both men and women combined. This is also the strategy chosen by Bem for producing the original norms for BSRI.

Reliability of the inventory scales

Cronbach's Coefficient α was used as a measure of internal consistency also for the translated version of BSRI (Table 7). It would appear that reliability from this perspective largely concurs with that of Bem's normative sample with one exception. The Social Desirability Scale as judged by the Swedish sample shows poor consistency ($\alpha = .56$ for the combined sample), whereas the same scale in Bem's original 1973 sample is acceptable ($\alpha = .70$ for the combined Stanford sample). However, the SD-scale is nevertheless less reliable than the M-scale and the F-scale also in the Bem normative sample. After criticism from

Walkup and Abbot (1978), Bem proposed that the Social Desirability Scale should no longer play a part in the inventory. The items are still included in the inventory but serve only the purpose of being “fillers”; providing a general environment. They are rated by respondents but not considered by the research or clinician making use of the inventory. The Social Desirability Scale has been removed completely in the Swedish revision of BSRI.

The fact that the SD-scale is of questionable reliability in the Swedish setting strongly suggests that it should not be used to qualify or verify the independence of androgyny as a construct, which was originally done by Bem in the normative sample. She showed that androgyny was nearly uncorrelated with androgyny as expressed by Student’s *t* - ratio (that is $r = .03$ for males and $r = -.10$ for females).

Table 7. *Coefficient alpha for both the Swedish sample and Bem’s 1978 Stanford sample*

Scale	Males	Females	Combined
Masculinity			
Swedish sample	.80	.82	.82
Stanford sample	.86	.87	.86*
Femininity			
Swedish sample	.84	.74	.81
Stanford sample	.78	.78	.80*
Social Desirability			
Swedish sample	.63	.47	0.56
Stanford sample	n/a	n/a	.70*

* *These values are from Bem’s 1973 Stanford sample*

Either an indigenously derived scale of Social Desirability for a Swedish context be constructed to replicate Bem's construct procedure, or the averaged F-scale and M-scale means are correlated with each other to estimate the degree to which they correspond. A high degree of correspondence would suggest that Masculinity and Femininity as separate constructs is not a successful one, whereas a low degree of correspondence would suggest the opposite. Bem offers this reliability check also in the commercially published version of the inventory and thus shows that the two remains uncorrelated (Females $r = .00$ and Males $r = -.05$).

However, this is not the case for the Swedish sample. Whilst correspondence between the scales as judged by the female participants is low ($r = .16$) the correspondence as judged by the male participants is moderate ($r = .41$). The same procedure applied to the total Swedish sample, including both men and women, also signifies a degree of correspondence although weak ($r = .30$).

These results suggest that the use of BSRI in a Swedish setting in merely a translated form is not straightforward in spite of acceptable alpha values for both scales. There appears to be some overlap between the Masculinity Scale and the Femininity Scale, which are likely due to cultural differences. This conceptual discrepancy between the samples requires closer scrutiny.

Performing an inter-item total correlation analysis shows that a number of variables are either weakly correlated to the scales or that men and women strongly disagree on certain items as being sex-appropriate (Table 8). For example, it appears that being athletic (M4SPORT), analytical (M8ANALYT), competitive (M19TAVL) or

ambitious (M20AMBI) is considered by both men and women not to be typically masculine. Similarly all participants, irrespective of sex, seem to agree that yielding (F1UNDFAL),

Table 8. *Inter-item total correlation as based on female and male participant scores separately and combined. To decipher item labels see Table 1 & 2 above. Items which either correlate weakly to the scale or sample to which they apply, or items regarding which men and women appear to disagree strongly, are boldened and enlarged.*

Scale items (M)	M	F	Comb	Scale items (F)	M	F	Comb
M1SJLVFR	.51	.50	.52	F1UNDFAL	.17	-.01	.06
M2FORSV	.50	.54	.53	F2MUNTR	.23	.17	.25
M3OBERO	.13	.61	.40	F3BLYG	.22	.21	.16
M4SPORT	.01	-.10	-.03	F4TILLGV	.64	.48	.59
M5BESTMD	.65	.52	.56	F5SMICKR	.31	.31	.25
M6STPER	.63	.34	.43	F6LOJAL	.45	.32	.39
M7KRFTFL	.53	.54	.56	F7FEMINI	.36	.51	.49
M8ANALYT	.15	.32	.27	F8SYMPAT	.57	.40	.50
M9LEDARE	.39	.58	.59	F9KNSLBH	.64	.39	.56
M10RISK	.33	.53	.46	F10FORST	.48	.38	.48
M11BSLUT	.42	.39	.44	F11MDLID.61	.53	.58	
M12SJLVT	.53	.31	.35	F12LINDR	.55	.42	.52
M13DOMIN	.58	.56	.55	F13MILD	.41	.46	.41
M14MASK	.39	.34	.38	F14VARM	.51	.54	.57
M15BSLST	.39	.58	.53	F15OMSIN.73	.61	.70	
M16AGGRS	.37	.09	.21	F16LTLUR	.40	-.05	.17
M17LEDAR	.39	.48	.48	F17BARNS	.25	.07	.15
M18INDIV	.14	.55	.34	F18SVOR	.12	.22	.20
M19TAVL	.21	.25	.29	F19BARKR	.41	.23	.38
M20AMBI	.32	.23	.21	F20VARSM	.72	.42	.59

cheerful (F2MUNTR), shy (F3BLYG), flatterable (F5SMICKR) and childlike (F17BARNS) do not describe characteristics that are typically feminine, nor is use or no use of harsh language (F18SVOR) a variable

appropriate to describe femininity. Observe, however, that men disagrees with women on the status of being independent (M30BERO), aggressive (M16AGGRS) and individualistic (M18INDIV). The women suggest that independence and individualism are typical male characteristics whereas men do not. On the other hand, men feel that aggression to some degree is typically male while women appear to disagree. In a like manner men and women do not agree on whether women are typically gullible (F16LTLUR). Men say they are while women object. Also, men seem to think, to a higher degree than do women, that it is typically feminine to love children (F19BARKR).

In order to come to terms with the response patterns of the Swedish sample—deviant in comparison to the American normative sample—and increase the integrity of the two scales, seven items were removed from the Masculinity Scale (M3, M4, M8, M16, M18, M19 and M20) and eight items from the Femininity Scale (F1, F2, F3, F5, F16, F17, F18, F19). A new reliability analysis was performed on the resulting shortened version of the translated inventory.

This procedure increased internal consistency as expressed by Coefficient Alpha for both scales and with regard to females, males and men and women combined (Table 9). Inter-item correlations have also become more robust (Table 10). To further confirm the conceptual basis of the BSRI in a Swedish setting a factor analysis (principal components with varimax rotation) was performed on the combined and revised scales. A first run revealed a solution of seven factors explaining 66.5 % of total variance. However, in this solution two factors dominate and together explain 41.4 % of total variance. The variance of the remaining five factors ranges from 4.0 to 7.3 %. In a

second run the principal component analysis was restricted to a two-factor solution only, which would seem to corroborate the existence of two separate constructs (Table 11) explained by approximately equal proportions of the total variance (22.0 % and 19.3 % respectively). However, there is still some overlap between the two scales. Masculinity (M14MASK) loads moderately on both factors and the same is true of Soft-spoken (F13MILD).

Table 9. *Coefficient alpha for both the unrevised and the revised Swedish version of BSRI*

Sample	Masculinity			Femininity		
	Unrevised	Revised	<i>t</i>	Unrevised	Revised	<i>t</i>
Males	.80	.85	33.0*	.84	.87	57.0*
Females	.82	.84	83.0*	.74	.83	17.4
Combined	.82	.84	83.0*	.81	.86	33.4*

* $p < .01$

Table 10. *Inter-item total correlation as based on the corrected Femininity and Masculinity Scales. To decipher item labels see Table 1 & 2 above*

Scaleitems (M)				Scale items (F)			
M	F	Combined		M	F	Combined	
M1SJLVFR	.43	.54	.51	F4TILLGV	.59	.58	.60
M2FORSV	.51	.46	.48	F6LOJAL	.43	.34	.39
M5BESTMD	.66	.52	.57	F7FEMINI	.39	.44	.49
M6STPER	.64	.40	.47	F8SYMPAT	.58	.46	.52
M7KRFTFL	.52	.56	.57	F9KNSLBH	.65	.56	.65
M9LEDARE	.59	.63	.64	F10FORST	.58	.49	.57
M10RISK	.40	.47	.44	F11MDLID	.61	.57	.59
M11BSLUT	.47	.45	.49	F12LINDR	.65	.46	.58
M12SJLVT	.48	.30	.32	F13MILD	.32	.37	.32
M13DOMIN	.58	.58	.56	F14VARM	.55	.58	.61
M14MASK	.42	.30	.37	F15OMSIN	.77	.71	.76
M15BSLST	.43	.54	.51	F20VARSM	.73	.47	.61
M17LEDAR	.43	.53	.52				

Table 11. Factor analysis (principal components, varimax rotation) of the revised M and F-scales combined. Note that the cut-off point is set to $>.35$. Loadings included on this criterion are boldened and enlarged in the table.

Items	Factors		Items	Factors	
	I	II		I	II
M1SJLVFR	.08	.58	F4TILLGV	.70	.02
M2FORSV	.14	.58	F6LOJAL	.48	.19
M5BESTMD	.11	.66	F7FEMINI	.62	-.17
M6STPER	.21	.60	F8SYMPAT	.61	.12
M7KRFTFL	.01	.68	F9KNSLBH	.74	.00
M9LEDARE	-.08	.75	F10FORST	.65	.03
M10RISK	.04	.53	F11MDLID	.67	-.08
M11BSLUT	-.22	.55	F12LINDR	.68	-.06
M12SJLVT	.14	.41	F13MILD	.37	-.39
M13DOMIN	-.19	.66	F14VARM	.71	.19
M14MASK	-.40	.44	F15OMSIN	.82	.03
M15BSLST	-.06	.56	F20VARSM	.67	-.10
M17LEDAR	.00	.64			

Cut-off point for factor loadings were set to $>.35$, which is in accordance with Bem's (1981) factor analysis of the original BSRI in an effort to arrive at a shorter version of the inventory. This cut-off point is more stringent than recommended for the current number of variables and sample size. Child (1990), for example, recommends a cut-off point of $>.20$ at $p < .05$ and $>.26$ at $p < .01$ when $N = 100$ and the analysis involves no more than 30 variables (see also Burt, 1952).

In summary, it seems not advisable to use the BSRI in its original American rendition in a Swedish context, since it contains items describing gender characteristics that appear not similarly understood in Sweden. However, in view of the fact the integrity of the inventory seems to hold in the revised version, this could be used in a Swedish research setting where it is desirable to compare gender to other variables. In so doing, however, one must bear in mind that the Swedish revised version (which will be termed BSRI-SE in the following) only accounts for 41.4 % of total variance, which suggests there is more to the understanding of gender roles in Sweden than is expressed by the BSRI. Ideally a new item pool should be derived from a larger and more representative Swedish sample than has been used for this study, and a new and entirely indigenous set of scales be constructed.

Procedure for classification and standardisation

The classification procedure suggested by Bem (1981) to replace the dismissed Student's *t* - ratio as a basis for deciding to what category a respondent might belong is to some extent straight-forward. The median split method is simple and quick given there are norms available by which to check arrived-at scores. The estimation of the strength of sex-typing, however, is more complex and the BSRI manual is somewhat paradoxical and far from straight forward in explaining how the four-way classification procedure relates to estimating the degree to which a person may be regarded as typically masculine, feminine, androgynous or undifferentiated. For the standardising

process of the SE-version of BSRI the following procedures were employed, which are based on Bem's recommendation as far as the four-way classification by means of median splits go. The use and relevance of standard scores, however, differs from what Bem suggests. This will be discussed separately below. The procedure below describes the classification process with reference to already obtained norms.

The median split method

As answering sheets have been collected, coded and fed into a suitable computer statistics package (such as the SPSS 8.0, which was used for this study), the first step is to compute each individual's arithmetic average raw score on both the Femininity Scale and the Masculinity Scale.

These average scores are then compared to the *median* scores of both scales arrived at by finding the average medians for the whole sample tested. The medians for the Swedish sample are 4.92 for the Masculinity Scale and 5.19 for the Femininity Scale (see Table 12 below). These are the normative medians by which an individual's score is split from the total sample into a particular category. For example, if one respondent's averaged raw scores on the F-scale and the M-scale is 4.69 and 4.75 respectively. He or she would be classified as undifferentiated because 4.69 is below the normative median for the F-scale (i.e. below 4.92) and 4.75 is also below the normative median for the M-scale (i.e. below 5.19). Similarly, if an individual on average has scored 3.85 of the F-scale and 6.08 on the M-scale, he or she should be classified as masculine, since the F-score falls below the norm and

the M-score falls above the norm (see Figure 1). If a score falls on the norm precisely it is best regarded as falling above the norm.

Note that depending on context and purpose of research it might be advisable to create other norms specific for the sample or population involved in the research. This is also Bem's (1981) recommendation.

MASCULINITY SCORE			
Below median	Above median		
Undifferentiated (low F - low M)	Masculine (low F - high M)	Below median	
Feminine (high F - low M)	Androgynous (high F - high M)	Above median	FEMININITY SCORE

Figure 1. The BSRI four-way classification as suggested by Bem (1981).

Note that Bem recommends a slightly different procedure. She recommends that the averaged individual scores from both scales be standardised into *T*- scores (that is, first obtain *z* - scores, which is easily done with for example SPSS. Then multiply the *z* - score with 10 and add 50. The formula for converting *z* into *T* is simple: $T = z (10) + 50$. The standardised *T* - scale has a mean of 50 and a standard deviation of 10). However, Bem then suggests that the M-F Difference Score be calculated from the standardised F-score and M-score, and that this difference score in turn is standardised in the same way. This spells problems, however, and is discussed in the following. The M-F Difference Score (i.e. the result of subtracting the averaged masculine

score from the averaged feminine score) is important and should be accompany the classification. It tells the strength and direction of sex-typing.

The M-F Difference score

The median split categorises a subject *only* with reference to the norms of a certain sample. It does not consider the difference between an individual's scores on the M-scale and the F-scale. The result is inevitably that some individuals are classified as masculine or feminine in spite of the fact that their scores are fairly equal. Conversely, some may be categorised as androgynous or undifferentiated although they score very differently on masculinity and femininity. This is a problem for individuals who score near the cut-off points (i.e. the median splits). In spite of this, the median split is the method that Sandra Bem recommends for most purposes because it is a fairly straight- forward and simple procedure if norms are available.

However, M-F Difference score also has bearing on classification and could perhaps at times serve as a more precise basis for classification if appropriate ratios were decided within which scores may be classified. While Bem (1981) dismisses the Student *t* - ratio as the basis of classification replacing it by the median split, the importance of a ratio remains in the classification procedure nevertheless.

Bem proposes, that every score should be standardised—and Bem favours *T* - scores—and that the M-F Difference score be calculated and standardised as *T* - scores also. This makes little sense, however, since Bem is keen on pointing out the importance of sign (plus or minus) of

the M-F Difference score. “High scores in either direction”, Bem argues (1981), “indicate a tendency to be strongly sex-typed (or sex-reversed), positive scores indicate femininity and negative scores indicate masculinity” (p. 7). In the light of this statement observe that the whole point of applying a T - transformation is to avoid negative scores (Cohen, Swerdlik & Phillips, 1995). The problem with the BSRI Manual recommendation on how to deal with the M-F Difference scores therefore becomes twofold:

First, if the difference between the two already standardised scores is subject to a T -score transformation also the values become inordinately large. In a sense the procedure entails standardising the standardisation.

Second, the T - score transformation also removes the negative signs, which makes void the notion of a ratio in which the sign of any value provides a clue as to its direction. Such a ratio could be expressed on the basis of T - scores also, of course, but the immediate appeal of a positive or negative sign signifying direction is lost. In other words, it is *easier* to understand the significance of the M-F Difference score if this quality is retained.

For the purpose of this study, therefore, the M-F Difference score was calculated from the raw scores of the F-scale and the M-scale, and the resulting difference was transformed into z - scores rather than T - scores (Figure 2). This transformation allows for both negative and positive signs, since the mean of z - scale is 0.00 and its standard deviation 1.00. For anyone desiring to try other transformations the raw data for BSRI-SE are appended to this research report (Appendix II).

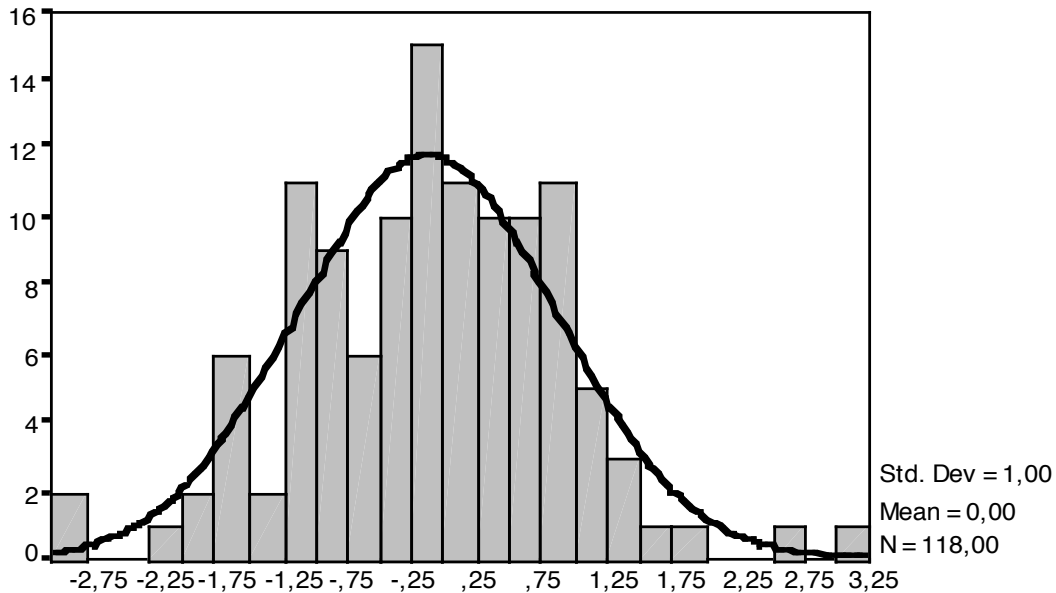


Figure 2. The distribution of z - standardised M-F Difference scores as compared to normal distribution.

The value of each standardised M-F Difference score tells not only how the individual relates to the normative sample but retaining the negative and positive signs through using z - scores also says something about the strength of sex-typing. A large negative discrepancy between the M and F scores signifies that an individual is strongly sex-typed in a masculine direction, whereas a large positive difference means that an individual is strongly sex-typed in a feminine direction. Small differences indicate that a person is either undifferentiated (i.e. low mean scores on both scales) or androgynous (i.e. high mean scores on both scales). For classifying any individual according to the BSRI or the BSRI-SE a standard score indicating the distribution and strength of sex-typing should be given together with the classification.

Are Swedes masculine, feminine, androgynous or undifferentiated?

An interesting aspect of evaluating the BSRI-SE on a Swedish sample is how Swedes fare in terms of classification distribution as compared to Bem's normative sample. The revised procedure of classifying participants on the basis of a median split—as outlined above—was followed, and the average medians of the Swedish sample were employed as norms (Table 12). Observe that these normative values have been weighted in order to compensate for the lesser number of men in the sample.¹

Differences were observed between the two samples (Table 13), and note the clerical error in the American sample: it is made up of 101 % males and 99 percent femals [sic!]. However, consider the fact that in both samples roughly half fall in either a category which is not sex-typed (undifferentiated or androgynous) or a category which is sex-typed as either masculine or feminine.

¹ For future users of the BSRI-SE it might be useful to have the weighting equation readily at hand for calculating weighted means by hand in a smaller sample: The weighted $X = \frac{\sum X_F + \sum X_M}{n_F + n_m}$ that is, the number of males multiplied by the means for males. The same procedure is performed for the female sample. The resulting sums are added and divided by the total number of the sample (i.e. males and females combined).

Table 12. *The male and female average judgments of the Masculinity and Femininity scales as based on the revised BSRI. Values for females and males combined have been weighted to compensate for the difference in numbers between the two sample groups. Boldened and enlarged values in the table serve as norms for classifying participants in the Swedish sample according to a median split.*

Sample	Masculinity			Femininity			Difference (M - F scale)		
	<i>M</i>	<i>MD</i>	<i>SD</i>	<i>M</i>	<i>MD</i>	<i>SD</i>	<i>M</i>	<i>MD</i>	<i>SD</i>
Males	5.07	5.15	1.26	4.71	4.79	1.31	.36	.36	-.05
Females	4.60	4.77	1.33	5.32	5.46	1.17	-.72	-.69	.16
Combined	4.79	4.92	1.30	5.07	5.19	1.23	-.28	-.26	.08

Table 13. *Classification distribution in samples (in percentages) for Bem's original BSRI and the revised BSRI-SE. Cp denotes cumulative percentage.*

Sex-type classes	Original BSRI sample				Revised BRSI-SE sample			
	<i>M</i>	<i>Cp</i>	<i>F</i>	<i>Cp</i>	<i>M</i>	<i>Cp</i>	<i>F</i>	<i>Cp</i>
Undifferentiated	27	27	18	18	31	31	27	27
Androgynous	20	47	30	48	21	52	20	47
Masculine	42	89	12	60	40	92	17	64
Feminine	12	101	39	99	8	100	36	100

Table 14. *Sex-typing in the different subgroups of the Swedish test sample.*

Sample subgroup	<i>n</i>	Undiff	Andr	Masc	Fem
Pre-school student teachers					
Males	4	1	3	0	1
Females	20	3	4	1	11
<i>n</i>	24	4	7	1	12
Percentage		17	29	4	50
Communication Studies students					
Males	11	4	1	5	1
Females	38	11	7	9	11
<i>n</i>	49	15	8	14	12
Percentage		31	16	29	24
Amateur football players					
Males	12	7	1	3	1
<i>n</i>	12	7	1	3	1
Percentage		59	8	25	8
Pre-school teachers					
Males	4	2	0	2	0
Females	12	4	3	2	3
<i>n</i>	16	6	3	4	3
Percentage		37	19	25	19
Comprehensive student teachers					
Males	12	0	4	7	1
<i>n</i>	12	0	4	7	1
Percentage		0	33	59	8
Various university staff					
Males	5	3	1	1	0
<i>n</i>	5	3	1	1	0
Percentage		60	20	20	0

A paired-samples t - test (pairs consisting of American males/Swedish males and American females/Swedish females) was employed to test whether differences observed were statistically significant. They were not ($t = .143, n.s.$ and $t = -.059, n.s.$ respectively), which suggests either that samples are indeed similar in terms of the distribution of gender stereotypes in the two cultures, or alternatively that the median split is not sensitive enough a method by which to classify a sample of the size used for BSRI-SE (i.e. $N = 118$).

Sex-typing in the sample sub-groups

Although the Swedish sample is small in comparison to that used by Bem, it adds to the understanding of gender-roles in Sweden to categorise participants within sample groups also (Table 14). At least it may raise a few questions as to the relationship between chosen profession or activity and professed gender stereotype. The two larger sample groups: Pre-school students teachers and Communication studies students are interesting in this respect.

Students of Communication Studies, who will graduate as Public Relation Officers with a employment opportunities in a wide variety of contexts, seem to a large extent (70 %) to be either undifferentiated or masculine in their sex-type orientation. Pre-school student teachers, on the other hand, are largely directed towards either feminine or androgynous behaviour (79 %). The latter should perhaps be compared to full-time working pre-school teachers; a group in which stereotypical gender roles are more evenly distributed: 56 % are either undifferentiated or androgynous whereas 44 % are sex-typed as

typically feminine or masculine.

Concluding remarks

It is clear from this research that the Bem Sex-Role Inventory (BSRI) should not be used in a Swedish context in its original form. It has limited face validity, as spontaneously argued by many of the respondents. It would also appear that while some scale items of the American Femininity Scale and the Masculinity Scales are indeed shared between the two cultures others are not, which weakens the statistical reliability of the inventory. Rather, the BSRI-SE should be used, which has been revised to better accommodate the cultural differences. However, it must be remembered that although the revised version has better psychometric properties, and does indeed confirm Bem's theoretical stance of regarding femininity and masculinity as separate constructs rather than extreme ends of the same dimension, the BSRI-SE only accounts for 41.4 % of total variance. There are obviously other aspects of gender in the Swedish sample, which are not accounted for in the inventory and that deserves further research.

The evaluation and revision of BSRI interestingly showed that on the basis of BSRI-SE there is no statistically significant difference between the distribution of gender role categories in the United States and in Sweden. This is different from, for example, the findings of Maloney, Wilkof and Dambrot (1981) who—using the BSRI in Israel—found that while American and Israeli women did not differ, American and Israeli men did. Israeli males were significantly less androgynous

than were American men (6 % versus 20 %).

However, the apparent similarity between American and Swedish cultures in this respect is nevertheless interesting. Sweden, unlike the United States has been described as a “feminine” country in terms of what Hofstede (1984) designates as “masculine and feminine goals” in his exceedingly extensive research into international differences in work-related values. In this

Table 15. *The Masculinity Societal Norm (Hofstede, 1984).*

Low MAS	High MAS
• People orientation	• Money and things orientation
• Quality of life and environment are important	• Performance and growth are important
• Work to live	• Live to work
• Service ideal	• Achievement ideal
• Inter-dependence ideal	• Independence ideal
• Intuition	• Decisiveness
• Sympathy for the unfortunate	• Sympathy for the successful achiever
• Levelling: don't try to be better than others	• Excelling: try to be the best
• Small and slow are beautiful	• Big and fast are beautiful
• Men need to be assertive but can also take caring roles	• Men should behave assertively and women should care
• Sex roles in society should be fluid	• Sex roles in society should be clearly differentiated
• Differences in sex roles should not mean differences in power	• Men should dominate all settings
• Unisex and androgyny ideal	• Machismoideal (ostentative manliness)

research Hofstede has calculated a Masculinity Index (MAS) for each country participating. The content on which the MAS is based is outlined below (Table 15). The United States receives an MAS Index of 62, whereas Sweden receives MAS 6, which is the lowest index value of

all the 39 countries participating in the study. Japan has the highest MAS index at 87 and Sweden is joined at the lower end by Norway (MAS 10), The Netherlands (MAS 14) and Denmark (MAS 22).

The findings of the present research seem to corroborate Hofstede's findings as far as the Swedish sample is concerned: 74 % of the participants (men and women together) are indeed classified as either undifferentiated or androgynous, thus largely conforming to the attributes outlined by Hofstede as being typical of a low MAS Index (see also, Daun, 1996). However, this is more or less the case for the American sample also: 66 % are either undifferentiated or androgynous, which is somewhat contradictory to the relatively high MAS Index of 62.

No conclusions may be drawn from these the within-group classifications done as part of the present research. The groups are too small. However, future research should for example look into whether the over-representation of women in teacher training (cf. OECD, 1993) also means an over-representation of non-masculine gender roles, which would perhaps be the common sense assumption to make. Amongst all teachers participating in the present study ($n = 52$, of which 20 are male and 32 are female) the four gender role categories are apparently evenly distributed—in spite of female over-representation: 19 % are undifferentiated, 27 % are androgynous, 23 % are masculine and 31 % are sex-typed as feminine. This would be interesting research, not only to corroborate the usefulness and reliability of the BSRI-SE further, but also to bring new light to the understanding of much of education as sexist and gender biased, which is often seen as threat to the increasing equality between the sexes (cf.

Houston, 1996; Morgan, 1996).

The division of humanity into different roles, gender roles and other, will continue. However, the meaning of gender is changing over time and across cultures with increasing speed and with the sometimes dubious aid of media and commercial promotion (e.g. Sullivan & O'Connor, 1988; Wagner & Banos, 1973). Research instruments such as the BSRI need to reflect this change (Lonner, 1990). The revision of the BSRI into BSRI-SE is such an effort.

References

- American Psychological Association (1993, January). Call for book proposals for test instruments. *APA Monitor*, 24, 12.
- Bailey, M. J. (1996). Gender identity. In R. C. Savin-Williams & K. M. Cohen (Eds.). *The lives of lesbians, gays, and bisexuals: children to adults* (pp. 71-88). New York: Harcourt Brace College Publishers.
- Baumeister, R. F. (1988). Should we stop studying sex differences altogether? *American Psychologist*, 42, 756-757.
- Bem, S. L. (1974). The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology*, 42(2), 155-162.
- Bem, S. L. (1975). Sex-role adaptability: one consequence of psychological androgyny. *Journal of Personality and Social Psychology*, 31, 634-643.
- Bem, S. L. (1977). On the utility of alternative procedures for assessing psychological androgyny. *Journal of Consulting and Clinical Psychology*, 45, 196-205.
- Bem, S. L. (1981). *Bem Sex-Role Inventory*. Palo Alto, CA: Mind Garden.
- Bjerrum-Nielsen, H. & Rudberg, M. (1994). *Psychological gender and modernity*. Oslo: Scandinavian University Press.
- Broverman, D. M., Vogel, S. R., Broverman, D. M., Clarkson, F. E. & Rosenkrantz, P. S. (1972). Sex-role stereotypes: a current appraisal. *Journal of Social Issues*, 28, 59-78.
- Burt, C. (1952). Tests of significance in factor studies. *British Journal of Psychology. Statistics Section*, 5, 109-133.
- Child, D. (1990). *The essentials of factor analysis* (2nd edition). London: Cassell.
- Cohen, R. J., Swerdlik, M. E. & Phillips, S. M. (1995). *Psychological testing and assessment: an introduction to tests and measurements* (3rd edition). London: Mayfield.
- Daun, Å. (1996). *Swedish mentality* (Translated by Jan Teeland). University Park, PA: Pennsylvania State University Press.
- Ellis, L. (1996). Theories of homosexuality. In R. C. Savin-Williams & K. M. Cohen (Eds.). *The lives of lesbians, gays, and bisexuals: children to adults* (pp. 9-34). New York: Harcourt Brace College Publishers.

- Gough, H. G. (1957). *Manual for the California Psychological Inventory*. Palo Alto, CA: Consulting Psychologists Press.
- Hassler, M. (1991). *Androgynie [Androgyny]*. Göttingen, Germany: Verlag für Psychologie Hogrefe.
- Hoening, J. & Kenna, J. C. (1974). The nosological position of transsexualism. *Archives of Sexual Behavior*, 3, 272-287.
- Hofstede, G. (1984). *Culture's consequences. International differences in work-related values*. London: Sage.
- Houston, B. (1996). Gender freedom and the subtleties of sexist education. A. Diller, B. Houston, K. P. Morgan & M. Ayim (Eds.). *The gender question in education. Theory, pedagogy and politics* (pp. 50-63). Oxford: Westview Press.
- Howells, K. (1987). Sex roles and sexual behaviour. D. J. Hargreaves & A. M. Colley (Eds.). *The psychology of sex roles* (pp.269-286). London: Harper & Row.
- Kaschak, E. & Sharratt, S. (1983). A Latin American sex role inventory. *Cross-Cultural Psychology Bulletin*, 18(1), 3-6.
- Kranau, E. J., Green, V. & Valencia-Weber, G. (1982). Acculturation and the Hispanic woman: attitudes toward women, sex-role attribution, sex-role behavior, and demographics. *Hispanic Journal of Behavioral Science*, 4, 21-40.
- Larsson, S. (1997). *Det andra jaget vid manlig transvestism. Ett jag-teoretiskt och kognitionspsykologiskt perspektiv [The second self and male transvetism. A self-theoretical and cognitive perspective]*, Doctoral dissertation. Uppsala, Sweden: Department of psychology, Uppsala University.
- LeVay, S. (1993). *The sexual brain*. London: MIT Press.
- Lipman-Blumen, J. (1989). Female leadership in formal organizations: Must the female leader go formal? In H. J. Leavitt, L. R. Pondy & D. M. Boje (Eds.). *Readings in managerial psychology* (pp. 325-344). Chicago, IL: Chicago University Press.
- Lonner, W. J. (1990). An overview of cross-cultural testing and assessment. R. W. Brislin (Ed.). *Applied cross-cultural psychology* (pp.56-76). London: Sage.
- Maloney, P., Wilkof, J. & Dambrot, F. (1981). Androgyny across two cultures: United States and Israel. *Journal of Cross-Cultural Psychology*, 21(1), 95-102.
- McHugh, M. C., Koeske, R. D. & Frieze, I. H. (1986). Issues to consider in conducting nonsexist research: a guide for researchers. *American Psychologist*, 41, 879-890.

- McKnight, J. (1997). *Straight science? Homosexuality, evolution and adaptation*. London: Routledge.
- Money, J. & Ehrhardt, A. A. (1972). *Man and woman, boy and girl*. Baltimore, MD: Johns Hopkins University Press.
- Morgan, K. P. (1996). The androgynous classroom: liberation or tyranny? A. Diller, B. Houston, K. P. Morgan & M. Ayim (Eds.). *The gender question in education. Theory, pedagogy and politics* (pp. 64-74). Oxford: Westview Press.
- Murphy, T. F. (1990). Homosexuality, a philosophical analysis. *Journal of Homosexuality*, 19, 132-137.
- OECD (1993). *Education in OECD countries. A compendium of statistical information*. Paris: OECD.
- Ross, M. W. (Ed.). (1985). *Homosexuality, masculinity and femininity*. New York: Harrington Park Press.
- Rowland, R. (1977). The Bem Sex-Role Inventory. *Australian Psychologist*, 12, 83-88.
- Segall, M. H., Dasen, P. R., Berry, J. W. & Poortinga, Y. H. (1990). *Human behavior in a global perspective. An introduction to cross-cultural psychology*. Oxford: Pergamon.
- Spence, J. T., Helmreich, R. & Stapp, J. (1975). Ratings of self and peers on sex-role attributes and their relation to self-esteem and conceptions of masculinity and femininity. *Journal of Personality and Social Psychology*, 43, 568-571.
- Sternberg, R. J. (1993). What is the relation of gender to biology and environment? An evolutionary model of how what you answer depends on just what you ask. In A. E. Beall & R. J. Sternberg (Eds.). *The psychology of gender* (pp. 1-8). London: Guilford Press.
- Sullivan, G. L. & O'Connor, P. J. (1988). Women's role portrayals in magazine advertising: 1958 - 1983. *Sex Roles*, 18, 181-188.
- Wagner, L. C. & Banos, J. B. (1973). A woman's place: a follow-up analysis of the roles portrayed by women in magazine advertisements. *Journal of Marketing Research*, 10, 213-214.
- Walkup, H. & Abbott, R.D. (1978). Cross-validation of item selection on the Bem Sex-Role Inventory. *Applied Psychological Measurement*, 2, 63-71.

Williams, W. L. (1996). Two-spirit persons: gender nonconformity among native American and native Hawaiian youths. In R. C. Savin-Williams & K. M. Cohen (Eds.). *The lives of lesbians, gays, and bisexuals: children to adults* (pp. 416-432). New York: Harcourt Brace College Publishers.

Williams, J. E. & Best, D. L. (1989). *Sex and psyche: self-concept viewed cross-culturally*. Newbury Park, CA: Sage.

BSRI-SE **The Bem Sex-Role Inventory**

Utvärdat och reviderat för en svensk kontext

Roland S. Persson, PhD

Centrum för psykologi
Högskolan för lärarutbildning och kommunikation
Högskolan i Jönköping

(Original instrument konstruerat av Sandra Bem, Stanforduniversitet, USA,
och publicerat 1981 av Mind Garden, Palo Alto, CA)

Råpoäng M	Råpoäng F	M-F poäng	Klass	Standard

BSRI-SE

Kön: Ålder: Datum: Kod:

Ta ställning till följande påståenden (1 till 25) genom att i boxen framför varje påstående skriva den siffra, som bäst motsvarar hur du upplever hur dessa utsagor stämmer på dig. Läs och bedöm *alla* utsagor! Du kan använda dig av siffrorna 1, 2, 3, 4, 5, 6 eller 7 för att beskriva hur sanna du tycker att dessa påståenden är. Siffran 1 betyder att ett påstående är aldrig eller nästan aldrig sant, medan siffran 7 betyder att ett påstående är alltid eller nästan alltid sant. Värdena däremellan beskriver olika grader av dessa två påståenden (se figuren nedan).

1 Det är aldrig eller nästan aldrig sant ← 3 → Det är alltid eller nästan alltid sant 7

- | | | | |
|------------------------------|---------------------------------|------------------------------|--|
| 1. <input type="checkbox"/> | Jag är full av självförtroende | 14. <input type="checkbox"/> | Jag är angelägen att lindra sårade känslor |
| 2. <input type="checkbox"/> | Jag är tillgiven | 15. <input type="checkbox"/> | Jag är förstående |
| 3. <input type="checkbox"/> | Jag försvarar mina övertygelser | 16. <input type="checkbox"/> | Jag har lätt för att ta beslut |
| 4. <input type="checkbox"/> | Jag är bestämd | 17. <input type="checkbox"/> | Jag är självtricklig och klarar mig själv |
| 5. <input type="checkbox"/> | Jag är mild och stillsam | 18. <input type="checkbox"/> | Jag är en lojal person |
| 6. <input type="checkbox"/> | Jag har en stark personlighet | 19. <input type="checkbox"/> | Jag är dominant |
| 7. <input type="checkbox"/> | Jag är sympatisk | 20. <input type="checkbox"/> | Jag är maskulin |
| 8. <input type="checkbox"/> | Jag är en varm person | 21. <input type="checkbox"/> | Jag är ömsint |
| 9. <input type="checkbox"/> | Jag är en kraftfull person | 22. <input type="checkbox"/> | Jag är varsam |
| 10. <input type="checkbox"/> | Jag är en ledartyp | 23. <input type="checkbox"/> | Jag är villig att fatta beslut och att stå för dem |
| 11. <input type="checkbox"/> | Jag är medlidsam | 24. <input type="checkbox"/> | Jag agerar ledare |
| 12. <input type="checkbox"/> | Jag är känslig för andras behov | 25. <input type="checkbox"/> | Jag är feminin |
| 13. <input type="checkbox"/> | Jag är villig att ta risker | | |

Tack för att du har besvarat alla påståenden!

De olika påståendena på svarsblanketten har ordnats slumpmässigt enligt följande. M anger att påståendet tillhör maskulinitetsskalan och F visar att påståendet tillhör femininitetsskalan. Siffrorna efter M respektive F anger variabelnummer och kan identifieras i ovanstående utvärdering av BSRI. M-skalan består av 13 påståenden och F-skalan av 12. I forskningsrapporten beskrivet hur revision av detta inventarium gått till väga samt hur resultat skall bedömas och räknas ut.

1.	Jag är full av självförtroende	M1
2.	Jag är tillgiven	F4
3.	Jag försvarar mina övertygelser	M2
4.	Jag är bestämd	M5
5.	Jag är mild och stillsam	F13
6.	Jag har en stark personlighet	M6
7.	Jag är sympatisk	F8
8.	Jag är en varm person	F14
9.	Jag är en kraftfull person	M7
10.	Jag är en ledartyp	M9
11.	Jag är medlidsam	F11
12.	Jag är känslig för andras behov	F9
13.	Jag är villig att ta risker	M10
14.	Jag är angelägen att lindra sårade känslor	F12
15.	Jag är förstående	F10
16.	Jag har lätt för att ta beslut	M11
17.	Jag är självtricklig och klarar mig själv	M12
18.	Jag är en lojal person	F6
19.	Jag är dominant	M13
20.	Jag är maskulin	M14
21.	Jag är ömsint	F15
22.	Jag är varsam	F20
23.	Jag är villig att fatta beslut och att stå för dem	M15
24.	Jag agerar ledare	M17
25.	Jag är feminin	F7

VIKTIGT! För närmare beskrivning av BSRI:s ursprungliga konstruktion och tolkning se Bem (1974; 1975; 1981). Användaren av detta reviderade svenska inventarium bör för ändamålet alltid införskaffa det amerikanska och kommersiellt publicerade materialet från Mind Garden, Palo Alto, Kalifornien med tanke på lagen om upphovsrätt. Detta svenska material är ett forskningsmaterial av begränsad spridning och har ett icke-kommersiellt syfte.

Appendix II

Normative raw scores for BSRI-SE

Part 1:

Subject	M-scale mean	F-scale mean	MF Diff	Diff z	Subject	M-scale mean	F-scale mean	MF Diff	Diff z
1f	4.69	4.75	-.06	.1904	34f	5.77	5.33	.44	.60383
2f	5.77	5.33	.44	.6038	35f	5.08	3.92	1.16	1.19911
3f	4.31	6.50	-2.19	-1.5706	36f	3.77	6.17	-2.40	-1.7442
4m	5.15	4.58	.57	.7113	37f	6.46	5.67	.79	.89320
5m	5.00	5.42	-.42	-.1072	38f	5.08	5.08	.00	.24005
6f	5.38	5.92	-.54	-.2064	39f	4.54	5.17	-.63	-.28082
7f	3.69	5.58	-1.89	-1.3225	40f	4.85	5.08	-.23	.04989
8m	4.46	5.42	-.96	-.5536	41f	3.85	6.08	-2.23	-1.6036
9m	5.53	5.67	-.14	.1243	42f	5.08	5.50	-.42	-.1072
10f	4.00	6.08	-2.08	-1.4796	43f	5.54	6.00	-.46	-.1402
11f	5.38	5.75	-.37	-.0658	44f	4.31	4.83	-.52	-.1898
12f	3.31	5.58	-2.27	-1.6367	45f	3.54	4.75	-1.21	-.7603
13f	4.15	4.67	-.52	-.1898	46f	4.46	5.00	-.54	-.2064
14f	4.54	5.92	-1.38	-.9009	47f	4.31	5.58	-1.27	-.8099
15f	4.01	5.33	-1.32	-.8513	48m	6.08	3.08	3.00	2.7203
16f	4.15	4.00	.15	.3640	49m	5.54	4.00	1.54	1.5132
17f	4.31	5.92	-1.61	-1.0910	50m	3.62	5.08	-1.46	-.9670
18f	5.92	5.00	.92	1.0006	51m	3.92	4.17	-.25	.0333
19f	3.54	5.33	-1.79	-1.2398	52m	5.08	4.25	.83	.9262
20f	3.92	5.42	-1.50	-1.0001	53m	6.00	4.58	1.42	1.4140
21f	3.00	6.75	-3.75	-2.8603	54f	4.38	5.58	-1.20	-.7520
22f	4.00	6.08	-2.08	-1.4796	55f	3.69	5.75	-2.06	-1.4531
23f	4.92	5.75	-.83	-.4461	56f	5.38	5.17	.21	.4136
24f	4.00	5.17	-1.17	-.7272	57f	4.08	6.25	-2.17	-1.5540
25f	4.54	4.92	-.38	-.0741	58f	5.23	6.83	-1.60	-1.0828
26m	4.54	5.25	-.71	-.3469	59f	1.69	5.25	-3.56	-2.7032
27m	5.38	4.42	.96	1.0337	60f	6.23	5.00	1.23	1.2569

28m	4.46	3.17	1.29	1.3065	61m	3.38	4.92	-1.54	-1.0331
29f	4.54	4.75	-.21	.0664	62f	4.38	3.58	.80	.9014
30f	4.46	4.33	.13	.3475	63f	5.31	4.42	.89	.9758
31m	6.46	5.92	.54	.6865	64f	5.38	4.54	.84	.9345
32f	5.23	5.42	-.19	.0829	65f	4.62	5.00	-.38	-.0741
33f	5.15	4.67	.48	.6369	66f	4.38	6.00	-1.62	-1.0993

Part 2

Subject	M-scale mean	F-scale mean	MF Diff	Diff z	Subject	M-scale mean	F-scale mean	MF Diff	Diff z
67f	4.69	5.83	-1.14	-.7024	93f	5.31	4.83	.48	.6369
68f	4.15	5.50	-1.35	-.8761	94f	4.77	4.75	.02	.2565
69f	4.15	5.50	-1.35	-.8761	95f	4.54	4.75	-.21	.0664
70f	5.38	6.25	-.87	-.4792	96f	5.00	6.17	-1.17	-.7272
71f	5.08	4.83	.25	.4467	97f	4.85	6.25	-1.40	-.9174
72f	5.15	3.92	1.23	1.2569	98f	5.23	5.42	-.19	.0829
73f	4.15	5.17	-1.02	-.6032	99f	4.08	4.08	.00	.2400
74m	4.54	4.67	-.13	.1325	100m	5.77	6.17	-.40	-.0906
75m	4.15	3.83	.32	.5046	101f	4.85	4.75	.10	.3227
76m	4.62	4.58	.04	.2731	102m	5.54	2.00	3.54	3.1668
77m	4.54	4.67	-.13	.1325	103m	5.34	3.92	1.46	1.4471
78m	6.00	5.17	.83	.9262	104m	5.08	6.00	-.92	-.5205
79m	4.69	4.33	.36	.5376	105m	6.31	5.75	.56	.7030
80m	4.77	4.58	.19	.3971	106m	6.31	5.42	.89	.9758
81m	5.00	5.83	-.83	-.4461	107m	5.46	4.42	1.04	1.0999
82m	5.69	4.00	1.69	1.6373	108m	4.92	4.33	.59	.7278
83m	5.08	4.33	.75	.8601	109m	4.54	5.67	-1.13	-.6942
84m	3.93	4.58	-.65	-.2973	110m	5.15	6.50	-1.35	-.8761
85m	6.00	5.92	.08	.3061	111m	5.00	4.83	.17	.3806
86m	4.31	4.75	-.44	-.1237	112m	6.62	4.50	2.12	1.9928
87f	4.92	5.25	-.33	-.0327	113m	5.62	4.75	.87	.9593
88f	4.85	6.00	-1.15	-.7107	114m	5.23	4.75	.48	.6369

89f	3.69	6.33	-.2.64	-1.9426	115m	4.15	4.75	-.60	-.2560
90f	4.92	4.42	.50	.6534	116m	4.54	4.33	.21	.4136
91m	5.15	3.92	1.23	1.2568	117m	4.08	3.75	.33	.51288
92m	5.08	4.33	.75	.86013	118m	5.54	6.00	-.46	-.1402