Multidimensional scaling and equivalence: is having a job the same as working?
Braun, Michael; Scott, Jacqueline

Multidimensional Scaling and Equivalence: Is *having a job* the same as *working*?

MICHAEL BRAUN AND JACQUELINE SCOTT

The question of functional equivalence in internationally comparative surveys is discussed from the viewpoint of secondary analysis. A number of data-analytical procedures – ranging from a comparison of means over establishing correlations with third variables in individual countries to Multidimensional Scaling (MDS) – are discussed and used to check problems in functional equivalence in an item battery on gender roles. The data base consists of several national samples (Italian, German, American, British and Hungarian) of the 1988 International Social Survey Programme study on Family and Changing Gender Roles.

It is concluded that (1) ex post checking of functional equivalence is useful and necessary because (2) the related problems might have less to do with translation than with ambiguities of the concepts and/or formulations used in source instruments and with differences in the social and economic realities in the respective countries, but that (3) even secondary-analytic strategies are as a rule not conclusive and should be supplemented with methodological experiments.

1. Introduction

Inadequacies in translations leading to a lack of functional equivalence of indicators, i.e., that the measures do not relate to the same underlying concepts, may be due to a variety of causes. These are related to different possible remedies, and they may appear in very different forms: from translation errors to slight variations in connotations. Some problems may arise due to a lack of carefully conducted translations. At first glance, this problem seems to be quite easy to handle by devoting more resources to this crucial step of the research process. *Ex ante*, i.e., in the context of questionnaire construction, the procedure of backtranslation seems to be appropriate. The basic idea is to translate the master questionnaire into the other languages and then to re-translate the translations. Any deviations between the original master questionnaire and the backtranslations are then discussed. Ideally, this is not only done to improve the translations themselves, but if

translation difficulties are found to be due to a general ambiguity of concepts/formulations used in the master questionnaire, also to modify the latter (Scheuch, 1968). In doing this, it would be possible to eliminate ambiguities which could be problematic even in the source language. However, it is not possible to guarantee functional equivalence by this method, only simple translation errors could possibly be eliminated by this procedure (see Harkness and Schoua-Glusberg, this volume).

Another, probably only partial remedy during the phase of questionnaire construction is to make the facet structure (Borg and Shye, 1995) of items explicit and to document the intended meaning of the questions, in order to give translators the necessary guidance and orientation (see Borg, this volume). Ultimately, however, the need to reappraise the problems of functional equivalence in the phase of data analysis remains (Scheuch, 1968; Smith, 1988). *Ex ante* procedures do not do away with the need for *ex post* procedures, i.e., detecting problems with functional equivalence during data analysis. As a rule, more than one indicator for a theoretical concept is necessary to do this.

A first step then is to check whether the ordering between different countries – with respect to the marginals or means (or any other adequate summary measure) – is similar between different indicators. If this appears not to be the case, especially if the mean differences have different signs or dramatically different magnitudes for items that are supposed to tap the same underlying aspect, the measures have to be treated with care. In this case, further inquiry is necessary, because that casts some doubt on whether the assumption of unidimensionality is fulfilled in all of the countries.

In addition, the regression of attitudes on demographic variables such as age, education or income can be used to help clarify the meaning of an attitudinal question. High correlations with age and education often give some support to an ideological interpretation, while correlation with income (net of education) make a self-interest-based understanding more likely. The question remains whether more ideological and more interest-based variables can be unambiguously singled out by this procedure. It is clear that a distinction between ideological and interest-based variables is likely to vary from

country to country or, more precisely, by the 'meaning' of the socio-demographic explicative variables. This is because it may be unclear whether what we see is generated by comparable groups giving different responses or by basically different groups in the single countries (Scheuch, 1968). Therefore, using variables like age, education or substantive variables known to measure values to inquire into the nature of an item will not always yield conclusive results. For example, in East Germany, different educational groups are closer to each other with regard to whether a pre-school child is likely to suffer if his or her mother works than are the respective groups in West Germany. (There is a 7 percentage points difference between lowest and highest educational qualifications in the East and a 15 percentage points difference in the West (cf. Braun and Bandilla, 1992). This could be due either to a different meaning of *suffering* in the two parts of Germany or to a difference with respect to the relationship between education and the attitudinal item (in the language adopted here, the latter would indicate different 'meanings' of the variable education), or both.

A wide range of more sophisticated statistical techniques (for an overview see Van de Vijver and Leung, 1997) have been proposed and used in the area of internationally comparative research, such as confirmatory factor analysis (Watkins, 1989) and psychometric approaches (Hulin, 1987; Van de Vijver and Poortinga, 1997). In this paper, in addition to simple procedures such as comparing marginals, we use Multidimensional Scaling (MDS), which has at least two advantages. First, MDS is easy to use, so that it can be applied routinely in analyzing item batteries. Second, MDS provides an instructive impression of the structure of the data. We outline details of this approach in the next section.

We apply this method to some potential problems of functional equivalence in the analysis of gender roles. The examples come from our substantive work (Alwin, Braun and Scott, 1992; Braun, Scott and Alwin (1994); Braun, Scott and Alwin, 1993). Part of the problems documented here became visible in the course of a substantive analysis of the data. Others were discovered by carrying out a backtranslation (admittedly while also knowing the source text). We also point out several ambiguities in the formulations of the

British English master questions, even if potential problems resulting from them could not be tracked in the data. Most of the issues raised here have to be tested, they should be understood as hypotheses on the effects of question wording.

2.   Data and Methods

The data come from the 1988 International Social Survey Programme (ISSP) module on Family and Changing Gender Roles (see Braun, 1994; Davis and Jowell, 1989). Eight countries participated in this study, but for our purpose, it is sufficient to concentrate on the Italian, the American and the German data, treating the British and Hungarian data as ancillary and ignoring the other countries. The 1988 questionnaire includes a battery of nine items designed to measure gender-role attitudes. These can be conceived of as tapping three different aspects: consequences of women working, gender ideology and the economic importance of work (Braun, Scott and Alwin, 1994). In the following, we concentrate on four items which present some peculiarities, especially in the Italian data. The other items are explained as needed. The items we focus on are related to the consequences and gender-ideology aspects. In the table below, italics are used to highlight the part of the items focused on.

First we consider the item *A woman and her family will all be happier if she goes out to work*. It is likely to tap several aspects: the consequences of women working for family life, gender-role ideology, and the possible contribution of a double income to the economic well-being of the family. Thus, in terms of the English text, it is obvious that this item might be highly ambiguous. Under certain conditions, ambiguity as such does not pose a big problem, provided people nevertheless understand the item in the same way, perhaps even mixing up different interpretations in a kind of summary statement (constant ambiguity mix). The problem arises if different subgroups in society or people in different countries understand the question in a different way. The problem with international comparability may be exacerbated because translations of ambiguous items are likely to emphasize some aspects more than others.

Table 1: Four language versions of four items from the 1988 ISSP module

Item 1	**Family suffers**
English source	All in all, family life suffers when the woman *has a full-time job*.
Italian	Tutto considerato la vita familiare risente negativamente se la madre *lavora a tiempo pieno*.
German	Alles in allem: Das Familienleben leidet darunter, wenn die Frau *voll berufstätig ist*.
Hungarian	A család élete megsínyli, ha a feleség *teljes munkaidöben dolgozik*.
Item 2	**Woman happier**
English source	A woman and her family will all be happier if she *goes out to work*.
Italian	Una donna e i suoi familiari sono più sereni se la donna *ha un lavoro*.
German	Wenn eine Frau *berufstätig ist*, wird sie und ihre Familie glücklicher sein.
Hungarian	A nöknek is és a családnak is job, ha a nök eljárnak dolgozni.
Item 3	**Children better**
English source	A job is all right, but what most women really want is a home and children.
Italian	Un lavoro è una buona cosa ma quello che realmente vuole la maggioranza delle donne è una casa e dei bambini.
German	Einen Beruf zu haben ist ja ganz schön, aber das, was die meisten Frauen wirklich wollen, sind ein Heim und Kinder.
Hungarian	Állásban lenni is fontos lehet, de a legtöbb nönek az az igazi vágya, hogy otthona és gyermeke legyen.
Item 4	**Housework as fulfilling**
English source	Being a housewife is just as fulfilling *as working for pay*.
Italian	Essere una casalinga è altrettanto soddisfacente quanto *avere un lavoro retribuito*.
German	Hausfrau zu sein ist genauso erfüllend wie *gegen Bezahlung zu arbeiten*.
Hungarian	A háziasszonyi teendöket jól ellátni legalább akkora teljesítmény, mint a fizetésért végzett munka.

We found that the Italian translation diverges from the English source rendering *if she goes out to work* by *se la donna ha un lavoro* (translated word-by-word: 'if the woman

has a job'). The Italian formulation may prompt, for some respondents, a comparison of employment and unemployment rather than a comparison with the voluntarily chosen homemaker role. Italian colleagues felt that in Italian 'to work' and 'to have a job' are treated and understood as equivalent expressions. However, that does not rule out the possibility that, at least for some respondents, a comparison of employment with unemployment may come to mind, thus moving the meaning of the item away from the consequences/gender-ideology aspect in the direction of the economic-function aspect. Thus, given that the original English formulation is itself ambiguous, the Italian translation only changes the ambiguity mix. There is an additional peculiarity with the Italian rendering which might or might not have an effect in the same direction. While the English formulation alludes to work outside the home *(goes out)*, some translations (both the Italian and the German versions) are more likely to include work that is done at home. This may make respondents more likely to agree to this item.

In a second item, the Italian translation deviates in a similar way from the English master questionnaire. In the item *Being a housewife is just as fulfilling as working for pay*, *working for pay* is again rendered by *avere un lavoro retribuito* (word-for-word: 'having a paid job'). Here, too, economic consequences might come to mind more easily in the Italian version.

Thus we have two pairs of items which, on the surface, appear to be very similar with regard to their intended meanings. The first pair is:

* *All in all, family life suffers when the woman has a full time job.*

* *A woman and her family will all be happier if she goes out to work.*

And the second pair is:

* *A job is all right, but what most women really want is a home and children.*

* *Being a housewife is just as fulfilling as working for pay.*

The English version uses *to work* for some items and *to have a job* for others. It is only in the second item of each pair that the Italian version deviates in the described way from the English, i.e., refers in Italian to 'to have a job' instead of 'to work'.

MDS (see Borg and Groenen, 1997) can help us to track the effects of these translation problems in the data. Technically, what MDS does is represent the intercorrelations of the items in a multidimensional space. Correlations correspond to the distances between the items which are drawn as points. The interpretation of the MDS representation focuses on the correspondence between geometrical characteristics of the configuration and the substantive characteristics of the items. The lines entered in the MDS representations are theoretically derived dividing lines between items; ideally the partitions of the space should be achievable by straight lines (only the MDS representation of the Hungarian data below does not meet that criterion).

3. Results

Do the empirical data support the assumption that the Italian item is understood in a different way from the English or the German? Let us examine the marginals first. While the Italians show quite traditional gender-role attitudes in general, on the items 2 and 4 in Tables 1 and 2, where the Italian version is different ('Woman happier' and 'Housework as fulfilling'), they turn out to be the least traditional nation (see Table 2).

Table 2: **Means of gender-role items in the US, Germany and Italy (items recoded such that high values correspond to non-traditional attitudes)**

Items	USA	Germany	Italy	North Italy	South Italy
Family suffers (1)	3.2	2.5	2.4	2.5	2.1
Woman happier (2)	2.7	2.6	3.3	3.3	3.4
Children better (3)	3.1	2.9	2.7	2.9	2.4
Housew. as fulfilling (4)	2.6	2.7	3.1	3.1	2.9

Moreover, the same trend can be seen within Italy, with Southern Italians appearing to be slightly less traditional in terms of these items than the Northern Italians, while for the other two questions Southern Italians are clearly more traditional than Northern.

The next question is: Apart from the marginals, is the cognitive representation of the items different in Italy, too? To discuss this aspect we turn to 2-dimensional Multidimensional Scaling. The graph for Northern Italy (Figure 1) shows that the items can be easily grouped preserving the theoretically postulated partitioning of the items into consequences, gender-ideology and economic-function regions (the partitioning is not produced by the program, but by the researchers reflecting theoretical considerations). The items in bold letters in the graph refer to the items affected by the deviating translation, the items in italics are those described above which do not suffer from this problem. The different partitionings are underlined. The additional items of the *consequences class* (on the left-hand side) are:

* Child suffers: *A pre-school child is likely to suffer if his or her mother works*.
* Warm relation: *A working mother can establish just as warm and secure a relationship with her children as a mother who does not work*.

The additional item of the *gender-ideology class* which appears in the middle is:

* Household W's job: *A husband's job is to earn money; a wife's job is to look after the home and the family*.

Finally, the two items of the *economic-function class* on the right-hand side are:

* Independence: *Having a job is the best way for a woman to be an independent person*.
* Double income: *Both the husband and wife should contribute to the household income.*

**Figure 1: 2-dimensional MDS for Northern Italy**
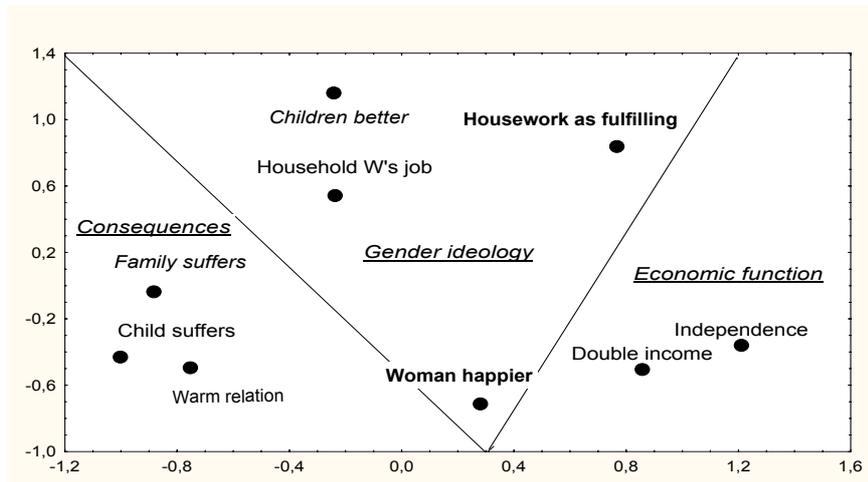
However, as Figure 2 shows, it is also possible to draw a dividing line between the consequences and gender-ideology regions on the one hand, and the economic-function region on the other, with the two items affected by the deviating translation located in the latter region. In fact, the two affected items are closer to the items of the economic-function class than to the remaining items of the consequences and gender-ideology class.

Figure 2: 2-dimensional MDS for Northern Italy, alternative partitioning
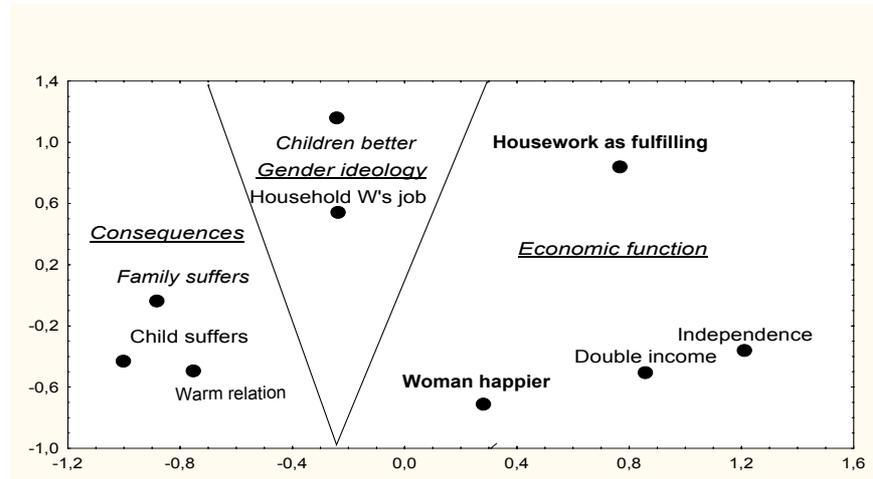
Thus, the problematic items do seem to belong more in the economic than the gender-ideology region. This leads us to ask whether this tendency is more pronounced in Southern Italy, where unemployment is higher and economic aspects might come more easily to mind than in Northern Italy. Though the pictures for Southern and Northern Italy look roughly the same, the correlations between the two items forming the above-mentioned pairs are somewhat higher in the North (for 'Woman happier' and 'Family suffers', .25 in the North and .22 in the South, and for 'Housework as fulfilling' and 'Children better', .38 in the North and .21 in the South.)

As argued above, the ambiguity of the source questionnaire items suggests that there could be a similar effect in other countries, too, with the problematic items being closer to the economic-function region. The graph for Germany (Figure 3) illustrates that the problematic items are closely related to the economic-function region, despite their 'natural home' for Germany being closer to the consequences and gender-ideology regions.

Figure 3: 2-dimensional MDS for Germany
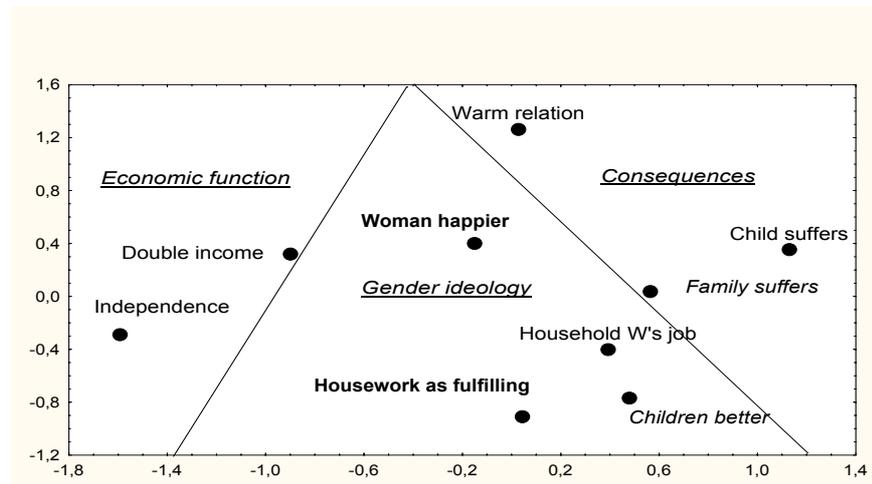
Unfortunately, the graph for the United States (Figure 4 below) does not show this as clearly. The problematic items turn out to be located somewhere in the middle, between consequences items and gender-ideology items on the one hand, and items from the economic-function class on the other. This, of course, makes our argument much less conclusive. Our *post hoc* conjecture is that, apart from language, the interpretation of the problematic items is also affected by social reality. In a society where female labor-force participation is perceived as 'normal', the gender-ideology interpretation of the problematic questions might be less salient than an economic interpretation. On the other hand, in a society where paid work is seen as contrary to the nature of women, a gender-ideology interpretation is more likely. Because attitudes towards female labor-force participation are, on the whole, more liberal in the US than in the other countries which participated in the 1988 survey, the American interpretation of the two items might be more in terms of economic function. In countries like Italy, where attitudes are in general more traditional, the gender-ideology interpretation should have been more dominant - had the translation not encouraged an economic interpretation.

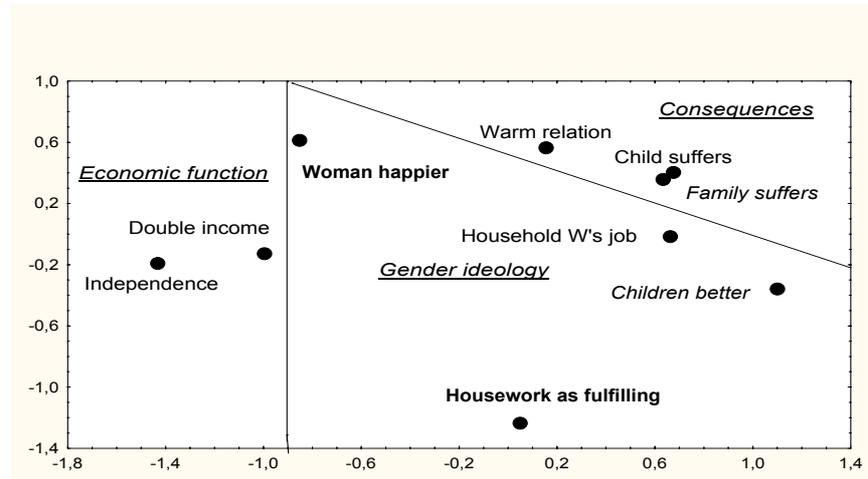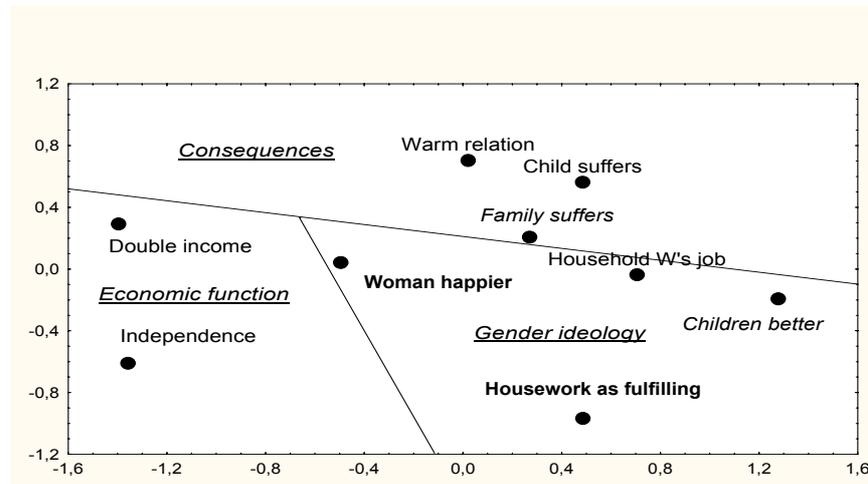**Figure 4: 2-dimensional MDS for the United States**

Figure 5: 2-dimensional MDS for Great Britain

The structure for the second English-speaking country, Britain, illustrated by Figure 5 below, lies somewhat in between those for the United States and Germany. The British show a level of acceptance for female labor-force participation similar to that of Americans. On other aspects, however, such as the perceived economic necessity of work for women and the actual level of female labor-force participation, Britain is somewhat between the United States and Germany.

As yet, we have little conclusive empirical evidence to add in support of our *post hoc* hypothesis. The Hungarian data would be an ideal candidate for further clarification, because in 1988, when the data was collected, women were obliged to work and there was virtually universal female labor-force participation. Unemployment was virtually non-existent and the attitudes of the population were extremely traditional with regard to female labor-force participation. Under these circumstances we would not expect the two problematic items to receive an economic interpretation. However, the Hungarian data is affected by a serious translation error with one of the crucial items (see below). Nevertheless, the Hungarian results are worth presenting.

Figure 6: 2-dimensional MDS for Hungary
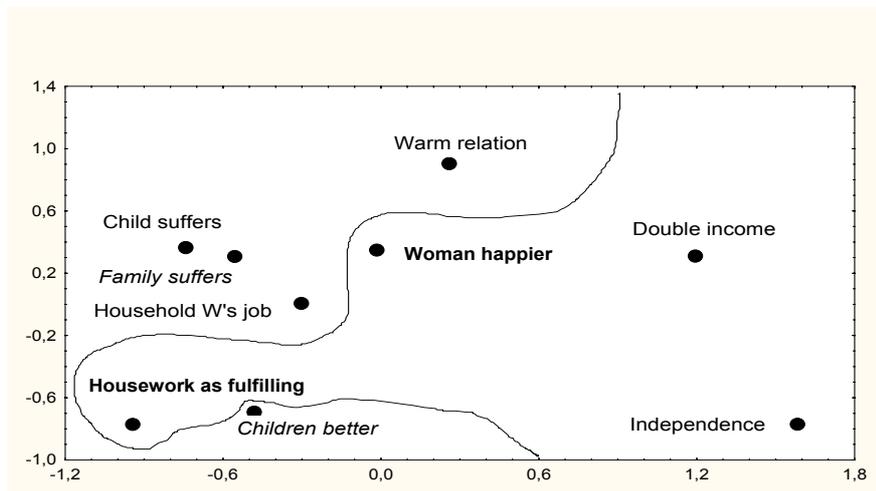
The graph for Hungary (Figure 6) fits perfectly with our expectations. No sensible dividing line could be drawn to bring the problematic items into the economic-function region (Figure 6 shows how a dividing line would have to look). As mentioned, there is a problem with the Hungarian translation: In Hungary *Being a housewife is just as fulfilling as working for pay* was translated to something like 'To do the job of a homemaker well is at least as big an achievement as paid work'. The translation focuses on *fulfill* in the sense of 'carry out a task' while the source item uses *fulfill* in emotional psychological terms. Correspondingly, for this item, it does not make much sense to compare the attitudes of the Hungarians with those of the remaining nations. Nevertheless, on the basis of an analysis of the correlations, the bias in the Hungarian translation cannot be detected (i.e., the correlations between this wrongly translated item and other items is similar to those obtained from other countries). The most likely reason for this is that attitudes measured by the original source questionnaire question and the wrong translation may be highly correlated: Respondents who think that looking after the family and house is as big an achievement as paid work should also be more likely to think that this is just as fulfilling as working for pay. What should remain different, however, are the marginals – and they are. To conclude, while the marginals for the Hungarians are incomparable with the other countries, the concept measured is very likely to be the same. Therefore, we contend that the Hungarian case lends at least some support to our interpretation.

4.   General Discussion

We have demonstrated the importance of checking for problems of functional equivalence at the data-analysis stage. We advocate the use of MDS at different stages in data analysis. It should be used either if the marginals yield unexpected results or to check whether problems in translation are evident in the data. However, it could also be routinely used before the substantive analysis of any battery of items because functional equivalence could be violated, even if the marginals do not look suspicious and translation problems are not detected.

It will be clear that some of the usual remedies for problems of functional equivalence might have been of little help. In the present case, giving more attention to translation and using backtranslation would not have guaranteed functional equivalence, because here we are not dealing with simple translation errors which could have been eliminated by this procedure. Moreover, the English formulations are, themselves, vague or ambiguous. As a result, even in English-speaking countries, they may be interpreted in different ways by different respondents. Perhaps the most important aspect, however, is the way social reality and language interact. Identical questions might be understood differently against different backgrounds. A reading of *to have a job* as referring to a contrast between a voluntarily chosen role as a worker and unemployment requires that unemployment is seen as a possible outcome. The degree to which such notions exist in different cultures varies considerably and it is very difficult to anticipate the effects of such variations when drafting a questionnaire and adapting it for different cultures. Thus, exploring problems of functional equivalence is a necessary part of data analysis; otherwise substantive interpretations may well be misleading. Such investigations will lead, incrementally, to improving the validity of the data and, as a by-product, investigators can gain new insights into the methodology of intercultural research. However, there is no guarantee that secondary-analytic techniques will bring conclusive results. The present discussion illustrates only too well that alternative interpretations may remain. Therefore, we advocate that split-ballot methodological experiments are also used so as to identify effects resulting from question format and wording and also to help disentangle possible artefactual and substantive explanations.

References

Alwin, D.F., Braun, M. and Scott, J. (1992). The Separation of Work and the Family: Attitudes towards women's labour-force participation in Germany, Great Britain, and the United States. *European Sociological Review 8*: 13-37.

Alwin, D.F., Braun, M., Harkness, J.A. and Scott, J. (1994). Measurement in Multi-National Surveys. In: I. Borg and P.Ph. Mohler (eds.), *Trends and Perspectives in Empirical Social Rresearch* (pp. 26-39). New York: de Gruyter.

Borg, I. and Groenen, P. (1997). *Modern Multidimensional Scaling*. New York: Springer.

Borg, I. and Shye, S. (1995). *Facet Theory: Form and Content*. Newbury Park: Sage.

Braun, M. (1993). Potential Problems of Functional Equivalence in ISSP 88 (Family and Changing Gender Roles). Paper presented at the International Social Survey Programme Scientific Meeting, Chicago.

Braun, M. (1994). The International Social Survey Programme (ISSP). In: P. Flora, F. Kraus, H.-H. Noll and F. Rothenbacher (eds.), *Social Statistics and Social Reporting in and for Europe* (pp. 305-311). Bonn: Informationszentrum Sozialwissenschaften.

Braun, M. and Bandilla, W. (1992). Familie und Rolle der Frau. In: Statistisches Bundesamt (ed.), *Datenreport 1992. Zahlen und Fakten über die Bundesrepublik Deutschland.* Bonn: Bundeszentrale für politische Bildung.

Braun, M., Scott, J. and Alwin, D.F. (1993). Attitudes on Marriage and Divorce in Comparative Perspective. Paper presented at the International Social Survey Programme Scientific Meeting, Chicago.

Braun, M., Scott, J. and Alwin, D.F. (1994). Economic Necessity or Self-Actualization? Attitudes toward Women's Labour-Force Participation in East and West Germany. *European Sociological Review 10*: 29-47.

Davis, J.A. and Jowell, R. (1989). Measuring National Differences - An Introduction to the International Social Survey Programme. In: R. Jowell, S. Witherspoon and L. Brook (eds.), *British Social Attitudes - special international report* (pp. 1-13). Aldershot: Gower.

Hulin, C.L. (1987). Psychometric Theory of Item and Scale Translations: Equivalence across languages. *Journal of Cross-Cultural Psychology 18*: 115-142.

Küchler, M. (1987). The Utility of Surveys for Cross-National Research. *Social Science Research 16*: 229-244.

Scheuch, E.K. (1968). The Cross-Cultural Use of Sample Surveys: Problems of comparability. In: S. Rokkan (ed.), *Comparative Research across Cultures and Nations* (pp. 176-209). Paris: Mouton.

Smith, T.W. (1988). The Ups and Downs of Cross-National Survey Research. GSS Cross-National Report No. 8. Chicago: National Opinion Research Center.

Van de Vijver, F.J.R. and Leung, K. (1997). *Methods and Data Analysis for Cross-Cultural Research*. London: Sage.

Van de Vijver, F.J.R. and Poortinga, Y.H. (1997). Towards an Integrated Analysis of Bias in Cross-Cultural Assessment. *European Journal of Psychological Assessment 13*: 29-37.

Watkins, D. (1989). The Role of Confirmatory Factor Analysis in Cross-Cultural Research. *International Journal of Psychology 24*: 685-701.