Cross-cultural comparison with quantitative and qualitative methods

Wolfgang Wagner
Nicole Bronberger

Paradoxically, the issue of psychological comparisons across cultures becomes the more pressing, the more the modern world undergoes processes of economic, political and social unification. Globalisation, as it is called nowadays, highlights cultural differences and draws attention even to minor variations in understanding and behaviour between people from different countries. As long as the representatives of each culture acted primarily within their geographically limited areas, it was of little importance that Austrians exhibit a different mentality than Japanese and that the work values of Egyptian Arabs differ from Brazilian values. It is under the conditions of a global market and the earth-spanning Internet that Austrian, Japanese, Egyptian and Brazilian mentalities, you name them, meet on a daily and regular basis and challenge psychology's understanding of cultural differences.

In the following it shall be argued that straightforward comparison of measurements, in the quantitative domain, and of semantic interpretations, in the qualitative domain, across cultures easily leads to inadequate results. This is due to the fact that scales, questionnaire items and text produced through interview techniques or open-ended questions have culturally specific meanings, that is, they cannot be mapped onto the same semantic metric unless otherwise proven. These culture-specific structures are called, metaphorically, cultural metrics. The claim shall be illustrated by examples from quantitative and qualitative research.

Conceptual ethnocentrism in cross-cultural comparison

Researching culture has never been an easy exercise for psychology. There are now three distinct approaches that deal with this issue: Cultural psychology, indigenous psychology and cross-cultural psychology. Cultural psychology is informed by anthropology and maintains:

- that psychological pluralism exists despite many universals, where researching the latter is not its aim;
- that it is interested in studying ethnic and cultural sources of psychological and social diversity;
- that thick description is needed to understand local stimulus conditions;
- that local sense-making and action patterns of intentional actors can be seen as a consequence of socially inherited values and representations;
- and that mentalities rather than the mind is the subject matter of cultural psychology. (Shweder, 2000, p. 209f)

A mentality in the present understanding is "the actual cognitive functioning of a particular person or people. To describe a 'mentality' ... is to get specific about the particular conceptual contents (the 'ideas') that have actually been cognised and activated by that person or people. To describe a 'mentality' is also to get specific about the particular mental processes (the particular senses, feelings, memories, desires, inferences, imaginings, etc.) that have been recruited by this or that person or people to make their cognising and activation of 'ideas' (conceptual contents) visible." (id., ibid., p. 210) Cultural psychology, hence, attempts to give justice to the particularities of cultures and their capacity to shape the workings of the psychological faculties. Being a descriptive approach to local worlds, cultural psychology does neither assume that psychological theories ("classical" ethnocentrism) nor any dimension or construct ("conceptual" ethnocentrism) developed in the West apply to non-Western cultures unless otherwise proven. It emphasises local description and not comparison between cultures which presupposes at least one dimension onto which the cultures can be mapped.

The potential danger of conceptual ethnocentrism lingers with cross-cultural psychological comparison. Cross-cultural psychologists "find differences in the meaning of constructs annoying, since such differences make the equivalent measurement of constructs more difficult", as Triandis acknowledges in a recent paper (Triandis, 2000, p. 188). Their research depends on cross-culturally valid dimensions of measurement, but this validity cannot be established by local standardisation of scales, as often suggested. A scale can only be standardised by statistical means after the

scale has been conceptually defined, that is, after it has been established that the meaning of the underlying psychological construct is the same and that the same metric applies across cultures.

This problem has always preoccupied cross-cultural psychologists and there have been several statistical and other methods developed to check for bias and for correcting its effects. Building upon an extended literature, van de Vijver and Leung (1997) identify three biases involving items, methods and constructs. Item bias, in their understanding, is a measurement artefact at the level of scales and items. It involves inadequate translation or formulation of items as well as the fact that the real-world referent of an item might not exist in one or the other culture being compared. Method bias refers to differential tendencies of acquiescence and extremity in scale use and differential familiarity with a stimulus as well as differences in the situations where a test is being applied. Both, item and method bias, are technical problems and can be avoided by adequate assessment and measurement procedures.

The third, construct bias, results if the construct investigated in a study is not the same in each involved culture. This comes closest to what shall be called "conceptual ethnocentrism". Van de Vijver and Leung (ibid.) suggest several approaches of how to avoid this problem. One would be the "decentred approach" where in the theoretical development and design of a study the researchers employ culturally divergent perspectives and try to give equal weight to all cultures under study. This approach would yield a set of constructs and items which cannot automatically be assumed to be valid in all culture specific samples. The other, the "convergence approach", starts with local researchers from each culture developing an indigenous conceptual structure and instrument for tackling the problem under scrutiny in their culture. If, in the course of the investigation and comparison, it is found that the local results converge across cultures there is good reason to assume cross-cultural validity of the phenomenon. As ideal as the convergence approach might be to correct for ethnocentric bias, as difficult it is in practice.

The problem of construct bias is at the heart of *conceptual ethnocentrism*. This is understood as the assumption that a theoretical variable or parameter found to be a relevant characteristic of one culture can be used to map the variability of other cultures. Mind that this is not a measurement or

statistical problem and therefore the term "bias" with its strong methodological connotation seems not to be well chosen. An, in the meantime, classical example is the individualism–collectivism variable often used as an independent variable to distinguish cultures from each other. The idea to this variable and the associated scale resulted from the Western individualism trait as one pole of an assumed variable where all non-individualist cultural groups can be mapped onto a position between the poles of strong individualism and strong collectivism. It is, however, ethnocentric to suppose that such a variable can capture the complexities of non-Western cultures which happen not to fall squarely into the Western individualistic mould. While the individualism pole might capture well a Western trait, the collectivism pole is likely not to capture the varieties of non-individualist cultures (e.g. Minoura 1996). Indigenous psychological research enriched the conceptual inventory of cross-cultural psychology with variables such as "tightness", "complexity", "activity", "honour" and "verticality" (Triandis 1996, p. 408f). Their usefulness and translatability into different cultural understandings has still to be proven. Until then indigenous psychologies should not be seen as modifications of contemporary (Western) psychology, but as contemporary local psychologies by their own right and with their own variables (Wagner, 1997; Yang, 2000). Otherwise, while pretending to investigate culture, cross-cultural psychology might in fact only investigate nature as it was "discovered" in Western laboratories (Jahoda, 1986).

Cross-cultural comparison of mentalities such as values and beliefs are prime candidates for conceptual ethnocentrism. If, for example, "persistence" appears to be a cultural value in Hong Kong but not with Illinois undergraduates, and "to be well adjusted" is claimed to be a value in both samples, the question arises whether "persistence" and "being well adjusted" designates the same things in both cultures. This cannot be established by comparing scores on a common scale, even with proper translation. The results derive from a study by Triandis, Bontempo, Leung and Hui (1990) who show that what is thought of the expression "being well adjusted" to mean in Illinois English is widely shared among Illinois undergraduates and that what is thought of the Chinese language equivalent of "being well adjusted" to mean in Hong Kong is widely shared among Hong Kong

undergraduates. It is not shown, however, that the cultural "things" designated by the respective English and Hong-Kong Chinese words "being well adjusted" are socially and culturally equivalent.

The local values contained in the respective Illinois and Hong Kong understandings of "being well adjusted" would in fact only be equivalent if it meant in both cases

- either "to talk, think and behave like others in my group",
- or "to appear like talking, thinking and behaving like others in my group",
- or "to talk, think and behave as I wish as long as it does not annoy anybody else in my group".

The expression "being well adjusted" would not be equivalent if in one sample it meant one thing and in the other sample it meant another thing. For example, being well adjusted in the sense of (c) could very well be a US-American understanding; being well adjusted in the sense of (a) or (b) have more the flavour of an Asian culture. Even if all three have the same literal translation of "being well adjusted", the specific content is crucially different.

Conceptual ethnocentrism is not a matter of incorrect translation even if translation of such sensitive concepts is quite problematic. We can translate virtually every word existing in Chinese or English into any other language. But translations, though literally correct, rarely capture what an indigenous concept means in the local world. Understanding words and concepts means to define the very phenomena they designate by virtue of the specific local context of cultural practices and language use. Conceptual ethnocentrism assumes that the very psychological concepts like mind, perception, emotion, motivation, personality, etc. are valid concepts for constructing non-Western variants of general theories. Evidence suggests that local theories can only be built with local psychological concepts as the indigenous psychology and the cultural psychology program attempt to do (Jodelet, 1993; Kim, 2000; Yang, 2000).

As a final example for potential ethnocentrism serves a theory–guided cross-cultural study which investigated inter-personal processes (Wagner, Kirchler, Clack, Tekarslan & Verma 1990). The study compared spouses' interdependence in conflict in traditional cultures where a strong gender-role segregation and associated male dominance exists (India, Turkey) with

cultures characterised by gender-role integration and egalitarian values (Austria, USA). It was found that emotional interdependence of spouses is much less in traditional than in Western countries. The degree of emotional interdependence in conflict was operationalised as a variable computed from three scores. The scores were the subjects' ratings of well-being in a purchasing conflict where they were asked to imagine situations (a) where they buy a personally desired commodity despite their spouse's disagreement, (b) where they do not buy the commodity because of their spouse's disagreement and (c) where they buy the commodity with their spouse's agreement.

In this study purchasing was considered a sufficiently comparable activity in all four countries. People buy and sell commodities all over the world. By superficial appearance purchasing is the act to exchange a token (money) for a product with utility. What the authors did not consider is that purchasing under the auspices of the spouse agreeing or disagreeing may *mean* something completely different in cultures with a profound role-segregation than purchasing in cultures with role-integration. First, marriages in traditional cultures are often arranged and not autonomous decisions of spouses as is the case in the West; second, love in the Western sense of sexual attraction and shared personal interests does not necessarily characterise marriages in other regions of the world; third, culturally gender-segregated activities, responsibilities, competencies and spaces in the house already imply that the other spouse is not supposed to share in the same activities, responsibilities, competencies and spaces. By this very cultural implication alone it is clear that spouses must be much more independent in traditional cultures than in Western ones. What the authors conceptualised as conflict by virtue of our Western experience (also the Indian and Turkish collaborators can be said to be Westernised to a certain degree) very probably is no conflict at all in traditional cultures.

This study was not on mentalities but on interaction patterns and it obeyed the rules that the Laboratory of Comparative Cognition (1979) set for cross-cultural comparison, that is to statistically compare only within-country-interactions across countries instead of main effects. Nevertheless it was based on the unwarranted assumption that the cultural

meaning of the construct "marital conflict" is the same in all four countries. Mind, mentality and social interaction are more closely interdependent than usually acknowledged.

Cultural metrics

Culture as a functionally organised system

In most contexts culture must be seen as a semantic structure of meanings, "a pattern of shared attitudes, beliefs, categorisations, self-definitions, norms, role definitions, and values..." (Triandis, 1996, p. 408) and as

> an organized body of rules concerning the ways in which individuals in a population should communicate with one another, think about themselves and toward objects in their environments. The rules are not universally or constantly obeyed, but they are recognized by all and they ordinarily operate to limit the range of variation in patterns of communication, belief, value, and social behaviour in that population. (LeVine, 1982)

This system cannot be divided into rules, attitudes, beliefs, categorisations, self-definitions, norms, role definitions, or values without loosing the essential meaning inherent in their delicate cross-reference with other meanings, cognitions and feeling. Each of these is functionally related to many others.

Consider the following metaphorical illustration:[1] If you asked chemists what a hormone is, they will tell you its chemical composition and molecular structure as a polypeptide, a steroid or an amino-acid derivative. Note that this is a chemical characterisation of a hormone, and that the chemical character is not a sufficient definition for a hormone (e.g. there are polypeptides which are not hormones and hormones belong to different chemical classes). Instead a hormone is a substance which is released from gland cells under certain physiological conditions and acts on receptors which

1 Personal communication by Günter P. Wagner, New Haven.

trigger a reaction functionally related to the trigger of the hormone release. A hormone is thus defined by the way the cells "use" it in their physiological activities.

In the same way, the simple clause "He did x because of y" can either be an attribution statement or something else depending on the context. In any case a linguistic analysis of such a clause is insufficient. For it to be an attribution it is necessary to prove that the sentence was uttered in a context which called for an explanation and not just for a free association or a recital of a text; that "...did x" is a salient activity in the culture, otherwise it would not call for an attribution; that "...because of y" gives a sensible reason in that culture; and that the reason given can be classified as internal or external, stable or unstable. A reason like for example "...because he needed to win" can be internal if conceived as a "need" or intention, or external, if conceived as the pressures of his trainer in a sports competition. One needs to connect a complex clause such as this one to a whole range of contexts which make up the respective culture before one can call it an interpersonal attribution. In other words, just as a hormone is defined by the functional role it plays in certain biological contexts, an interpersonal attribution is defined by the functional role it plays in certain cultural contexts and both, the definitions of hormones and attributions, are thus structurally similar to the meaning of words being defined by the way of its situated use (Wittgenstein 1958).

Cultural metrics in quantitative methods

Given the importance of functional and other relationships between variables in characterising cultures, Triandis (1996, p. 407), suggests to consider "cultural syndromes". "Cultural syndromes are conceived as dimensions of cultural variation that can be used as parameters of psychological theories ... In that way, the current psychological theories will become special cases of the universal theories." Although the author's latter claim may be a overly optimistic, the concept of cultural syndromes is an important one. If culture comprises systems of beliefs and practices which belong together and which are meaningfully interrelated, dependent measures can only be sets of overt and verbal behaviours. It is the "...use of theoretically motivated, *within-group* (emphasis in the original) observation as

a means of specifying culturally patterned activities that can be used as 'measures' by procedures which maximise representativeness" (Laboratory of Comparative Human Cognition 1979, p. 168).

Using bundles of overt and/or verbal behaviour variables as dependent measure addresses two issues simultaneously: First it allows to compare statistical interactions between variables instead of the main effect of single variables across cultures. ... Using bundles of variables obliges the researcher to look at interactions within cultures and comparing these interactions between cultures. In the statistical sense only an interaction found in one culture which is replicated in another culture allows to conclude that the effect is shared by those cultures (Campbell 1961). But, as Amir and Sharon (1987) have demonstrated, interactions can rarely be replicated across cultures. Second, bundles of variables allow to pin down the local interrelated meanings within a semantic field of cognitive and/or evaluative behaviours. Bundles of dependent measures also address the issue of semantics. It is very difficult if not outright impossible to assess the local meaning of the response to an attitude, value or belief item without reference to other attitudes, values, beliefs or practices. A set of responses on continuous or categorical scales, be they answers to closed questions or word associations to stimuli, can be analysed by non-linear multivariate statistics. The resulting pattern of the responses then gives an impression of the semantic relationships – or *semantic metric* – in each culture.

A "semantic metric" shall be defined here as *the pattern of implicit meanings respondents attribute to a questionnaire item or to a word or proposition in talk and writing*. These meanings determine the relationship of one measure to another measure, of one scale difference to a difference on another scale, and of one proposition to another proposition in text. Thereby respondents pertaining to the same culture define a metric – in a loose sense of the term – within which all their measures and text are defined. That is to say that measurement patterns are semantically mediated by the culture of the respondents. As Berry (2000, p. 197) puts it: "in studying behaviour one has to be 'cultural' before being 'cross'".

What applies to quantitative research and statistical comparison applies to some sorts of qualitative comparison as well. Discourse and text are as much embedded in local context as responses to questionnaires. But while quantitative data easily evoke the illusion of being decontextualised and therefore objective and equivalent across cultural samples, qualitative researchers are acutely aware of the potential fallacies of context and interpretation (cf. Straub, 1999). This does not mean, however, that the problem of semantic metrics does not apply to the comparison of qualitative data.

Interview transcripts and other qualitative material can be compared across languages and cultures using two approaches: one is to translate the transcripts from the cultural groups, to pool them and to analyse them jointly. The second is to analyse and interpret the qualitative material locally through researchers native to the respective culture or language and to compare the results in a second step. Both are being used in the not too numerous cross-cultural qualitative research literature.

An example of the first approach, pooling and joint analysis, is Dahlin and Watkins (2000) work on the views of Chinese and German students' views on the role of repetition and memorising in understanding and learning. The interviewees from both samples were living, or, in the case of the Germans, had been living in Hong Kong for some years. The Chinese were interviewed in Cantonese and the interview recordings were translated and transcribed verbatim into English by the respective interviewer. The German students were interviewed in English. Finally, the pooled English transcripts were analysed and interpreted by two trained researchers who maintained close communication during this work.

Gibbons et al. (1993) provide a good example of the second approach, that is local interpretation and subsequent comparison. They used pictures of women doing housework and office work drawn by adolescents from three cultures, Guatemalan, Phillipino and US-American, and had other adolescents from the three cultures content analyse the drawings. Naturally, using pictorial material, complicates the comparative problem further, but by

having indigenous coders doing the content analysis of the drawings they first established a local frame of category interpretation that was later used for comparison.

The two approaches are quite different in their implications because straightforward translation of the original text is a completely different business than translating interpretations. Translating a natural text from one language to another, even if done with all precautions, cannot guarantee that the translation leaves the implicit and contextual meanings of propositions, phrases and paragraphs unaltered, besides being uneconomical and laborious. This is the problem professional translators of novels and poems face in their daily work. Local qualitative analysis of text is less affected by this problem. Linguistically and culturally competent researchers doing text analysis on the spot are more likely to make implicit and contextual meanings explicit in their interpretations. These interpretations are supposed to capture structural semantic features of text as well as its content on a more general level. That is, in a metaphorical sense, the interpretation is done within the local semantic metric.

Translating interpretations, particularly if the researchers cooperate closely in a face to face situation, is a much better warrant of comparability or, in the alternative case, of incomparability, than translation of the original material. Even the failure to prove equivalence of semantic metrics and spaces is an interesting result in its own right and can enrich psychology's insight into the variability of mentalities and the functioning of culturally diverse minds. These issues are being illustrated in the next sections.

Examples of research

In this section I will present four examples of research that span from using quantitative to qualitative methods in cross-cultural comparison and which first investigated the local metrics before proceeding to comparison. The first and straightforward quantitative study was published by an international research team on the cross-cultural understanding of work values (Meaning of Work Research Team, 1987), in the others the present author has been involved. These are a study on the understanding of war and peace in Spain and Nicaragua (Wagner, Valencia & Elejabarrieta, 1996) and a study on the meaning of biotechnology in six European countries (Wagner

et al., 2002). These two can be labelled "semi-quantitative" since they are using qualitative data, i.e. word-associations and free responses, that are statistically analysed. The fourth study is about people's concerns about biotechnology and is a qualitative analysis of focus-group discussions in ten European countries (Wagner et al., 2001).

Quantitative comparison of questionnaire data

The "Meaning of Work" research team conducted an international comparative study on the meaning of work. The key variable in the research was the "work centrality score" which was obtained by having respondents indicate the following: "Assign a total of 100 points to indicate how important areas are in your life at the present time – Leisure, Community, Work, Religion, and Family." The points assigned to Work constituted the work centrality score. It was thought to reflect the relative significance of work in a respondent's life space.

The authors state that "when making international comparison, it is important to consider response frequency distributions and averages. This, however, is not sufficient, and in some instances, may lead to erroneous or incomplete interpretations" (Meaning of Work Research Team, 1987, p. 221). Consequently they not only compared the scores of work centrality across several countries but additionally conducted a non-linear multivariate analysis called "quantification on response pattern" within each sample (Hayashi, 1950). This technique is the same as correspondence analysis and allows to analyse the interrelationship among multivariate categorical data. In the present research it was used to elucidate interrelationships between work centrality and more than 30 items about the respondents' definition of work, their reasons to work, how they feel at work, etc. The result of this procedure is the position of each variable category in a multidimensional space. The closer two categories are mapped in the space, the more highly the two categories are related. By looking at the trace of the respondents' work centrality scores within each country's space the meaning of a high or low score can be determined for each culture (Figure 1).

Note: Graphs adapted from Meaning of Work Research Team (1987), pp 230, 232, 234. Note that the axes were rotated to match the work centrality scale as closely as possible among the three countries.
Points with little arrows at the frame border indicate the point being positioned outside of the drawing.

Figure 1 – Trace of 4-point work centrality scale (bold line, W1-W4) among trace of 7-point work centrality (trichomised, X1-X3) and lottery item (L1-L3) for Germany, Japan and USA.

Note that Figure 1 depicts only three variables of the study to allow easy visual inspection. Further, the graphs were rotated to match the general orientation of the work centrality scores in all three countries. It contains an alternative scale of work centrality (L1 to L3) and a "lottery item" (X1 to X3) asking subjects what they would do if they won a lot of money: to stop working (L1), to continue working in the same job (L2) or to continue work but with changed conditions (L3). Visual inspection of the simplified graphs makes it immediately clear that the meaning of work as expressed by the trace of the centrality index is not the same in Japan as in Germany or in the USA.

Using more variables from the original data, the authors further analyse that the Japanese scored highly and the US lowest on the work centrality scale. This might lead to the mislead conclusion that US workers are less motivated to work than the Japanese. In reality, the response pattern analysis reveals that the meaning of a US-score between 20 and 39 points (i.e.

moderately high; people who define work positively as something that produces social value) is to place importance on expressive aspects of work. Japanese workers scoring between 20 and 39 are characterised by trust in others and seeking good interpersonal relations; they put interpersonal matters in their workplace first and work itself second. Further on, the meaning of a 20 to 39 points score in the US is equivalent to the meaning of a score between 40 and 59 in Japan. Only the meaning of scores less than 20 (i.e. a negative view of work which one is forced to do, and an emphasis on economic and material conditions of work) is the same in the two countries.

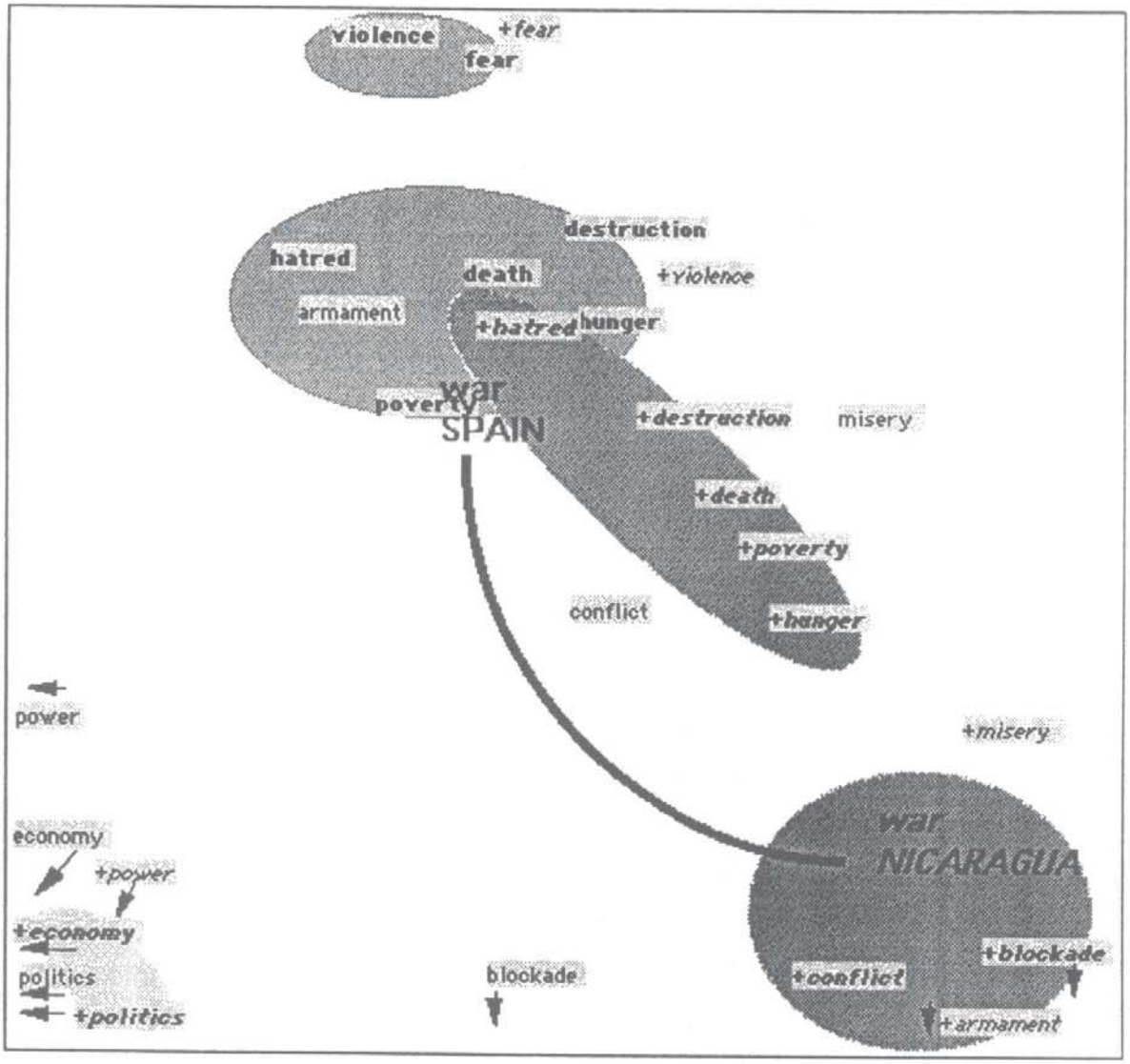
In sum, the analyses reveal that the work centrality scale spans the same semantic metric in the German and US-American sample, but a different one in the Japanese sample. One and the same work centrality score may indicate a completely different attitude in different cultures while two different scores may be an expression of the same attitude. Purely quantitative scores are hard to compare and interpret if the researcher cannot ascertain that they have the same meaning in the compared cultures.

Comparing the structure of word associations

A study by Wagner, Valencia and Elejabarrieta (1996) illustrates a similar problem. The authors investigated the structure of word associations dependent on the context in which they are assessed. Respondents from Spain and Nicaragua produced free associations about international conflict and peace.

While the goal of the original research does not concern us here, the data shall serve to illustrate a method which allows to deal with bundles of variables, even if they are free associations and therefore different in the two cultures. The variables were the words which the subjects associated with the stimulus words "international conflict". Figure 2, which is not included in the original article, shows how the associations about international conflict in Spain and Nicaragua are related to each other (a) within and (b) between the two countries.

A correspondence analysis of the stacked co-occurrence matrices of the 15 most frequent words in each country yields a multidimensional space of which the first two dimensions are depicted. They explain about 50% of



Note: Data from Wagner, Valencia and Elejabarrieta (1996). Words from Nicaraguan subjects have a "+" in front and are in italics. Light grey clusters: Nicaraguan. Dark grey clusters: Spanish. Bold type: Words pertaining to the enveloping cluster. The arc points to the relevant word "war".

Points with little arrows at the frame border indicate the point being positioned outside of the drawing.

Figure 2 – Semantic space of correspondences between associated words about war and peace for Nicaragua and Spain.

variance. This space can be interpreted as the semantic space of the word associations. A cluster-analysis yields two well connected clusters for each country (in the centre of the figure). Nicaragua contains 5 words

(destruction, death, hatred, poverty, hunger), Spain 6 (the same plus "war"). They indicate that the majority of subjects from both countries have a similar lexicon of proximal and affectively laden word associations about international conflict. Spanish subjects also exhibit a loose cluster connecting "violence" and "fear" (upper centre of Fig. 3). Other Nicaraguan subjects produce a loosely connected cluster connecting "economy" and "politics" (lower left corner of Fig. 3) and still others a well connected cluster encompassing "conflict", "blockade" and "war" (lower right corner of Fig. 3).

The point in case is the position of the word "war". While there is no doubt that most Nicaraguans and Spanish share some basic understanding of international conflict as indicated by the two central clusters, Nicaraguan subjects do not place the word "war" in this central cluster. It is a sub-sample of Nicaraguans who associate it together with "blockade" and "conflict" (see the two-pointed arrow in figure 3). This is a semantic complex of more "intellectual" words produced by a sub-sample which can easily be interpreted as resulting from their – then recent – experience of unrest, civil war and US-intervention. The example shows that Nicaraguans have a differently patterned perception of "international conflict" than Spaniards. Hence, their scores on a "conflict scale" and the resulting "conflict score" would be situated within a different semantic metric than the score of Spanish respondents.

Comparing text through automatic analysis

The comparison of text across cultures or groups speaking different languages constitutes a particularly difficult task and is relatively rarely done. The present example is taken from an international research group investigating the perception of biotechnology in various European countries (Wagner et al., 2002). An open-ended question in a Eurobarometer survey covering all member countries of the European Union, asked respondents to write down what comes to their mind when thinking of modern biotechnology. The respondents produced anything between no response and several fully formulated sentences.

This kind of data is influenced by several conditions: First, data collection was run in each country by different sub-contractors and within each country many different interviewers conducted the interviews. Some

may have let the respondents write their comment themselves, others may have summarised only the gist of the response themselves. Second, the complete sample comprises responses in 13 separate languages and, if language and nationhood has anything to do with culture, the national samples constituted 15 different cultural sub-groups of what might be called the common European cultural heritage. The shortcomings in data quality that one might expect in the present data were at least in part alleviated by the sample size which was statistically representative within each country.

According to the available groups of collaborators in various countries, the open-ended responses from six countries, Austria, France, Germany, Norway, Sweden und the United Kingdom, were analysed using ALCESTE (Reinert, 1983; 1990). This program allows to analyse text data automatically and uses descendent hierarchical classification, segmentation, correspondence analysis and the theory of dynamic clouds in its procedure. The algorithm produces matrices of co-occurrences of all words which are then decomposed and the words descendingly clustered according to their occurring in proximity or not in the text. If the size of the text-corpus is sufficiently large, the program allows to identify "discursive spaces" that describe the principal topics being mentioned in the text corpus (for an overview on the procedure see Kronberger & Wagner, 2000).

To preserve local meanings and linguistic idiosyncrasies in the data, the texts entering the analysis could, of course, not be translated and merged to a single data file. The principle of maintaining the local semantic metrics demands that each national sample be analysed independently. Hence, in the first step, each of the six national corpuses of text data were analysed separately and, in the second step, the resulting cluster solutions were collected and interpreted in comparison. All this was done in close collaboration with the national and language-native researchers. The over-all result is presented in Table 1.

Table 1 cross-tabulates the countries in the rows and the discursive clusters in the columns. The cells contain a short description of each cluster in each country and is left blank, if the cluster was not found.

It can be seen that the discursive clusters found independently in each country's text data match quite consistently across countries. Besides this surprising match of the cluster solutions, particularly the cluster comprising ideas of meddling and interfering with nature re-appears in each country and

Table 1 – Lexical classes of ALCESTE crosstabulated by countries.

	WHAT IS BIOTECHNOLOGY? FOCUS ON CONTENT					IS BIOTECHNOLOGY GOOD OR BAD? FOCUS ON EVALUATION					COUNTRY SPECIFIC			
	General (rather neutral)		Specific: Domains of Application (evaluation involved)			Positive	Ambivalent	Negative Evaluation				Lacking knowledge		
	Research (progress)	*Manipulation/ Alteration*	*Food*	*Reproduction*	*Medicine*	*Good*	*Good but risky*[1]	*Risky/ Dangerous*[2]	*expression of fear*	*Interfering with nature*		*Echo*[3]	*Guessing*[4]	*Don't know*
AUSTRIA	Biotechnology is a scientific activity applicated to plants, animals and humans (food, reproduction, medicine) (27%)						Good but risky/ dangerous (fear) (22%)	Unknown effects/ dangerous (16%)		Interfering with nature STOP! (36%)		see Interfering with nature		see unknown effects
FRANCE	Research (11%)		Food/ Agricul ture (15%)	Reproduction (2%)	Medicine (14%)	Improve- ment (10%)	Dangerous/risky although there can be good effects (also morally dangerous) (8%)		Fear Against nature (18%)			Echo (3%)	Guessed (16%)	Don't know (3%)
GERMANY		Manipulation of plants, animals, humans/ Agriculture (16%)	Food (also medicine and reproduction) (15%)		Medicine (12%)		Good but risky risky/dangerous (fear) (37%)			Interfering with nature STOP! (11%)		see Medicine Good but risky		Don't know (10%)
NORWAY	Research (8%)	Alteration of plants, animals, humans(21%)	Food (8%)		Medicine (14%)	Good but frightening Unspecific worry (22%)				Interfering with nature (10%)		see Medicine Good but frightening		
			Food and Reproduction (16%)											
SWEDEN	Research (19%)	Manipulation of plants and animals (11%)	Food and Reproduction (7%)				Good if used the right way/ dangerous (15%)		Fear too fast (19%)	Interfering with nature (21%)	Belgian Blue (9%)	see Research		
UK			Food (21%)	Reproduction (7%)	Medicine (21%)			Unspecific worry/dangerous (fear) (16%)		Interfering with nature (18%)		see Medicine		Don't know (17%)

[1] Good but risky: may have good effects but is risky and dangerous, therefore must be applied properly, demand for control

[2] Risky and Dangerous: biotechnology is unpredictable and therefore dangerous, fear of loss of control

[3] Respondents repeat technologies mentioned in the preceeding question ("telecommunication", "solar energy", etc.)

[4] Associations evoked by the terms "bio", "gene" and "technology" (mostly positive: e.g. ecologically beneficial or optimistic view of science)

Numbers in parenthesis indicate for each country the percentage of responses being classified to a specific discourse

can be said to be a shared concern in all six countries. This finding is independent of any bias that might have been introduced by translating the original responses in a single language and also independent of any biased interpretations that might ensue when researchers interpret responses from cultures where they are not native. Both of these problems are frequently introduced in cross-cultural research.

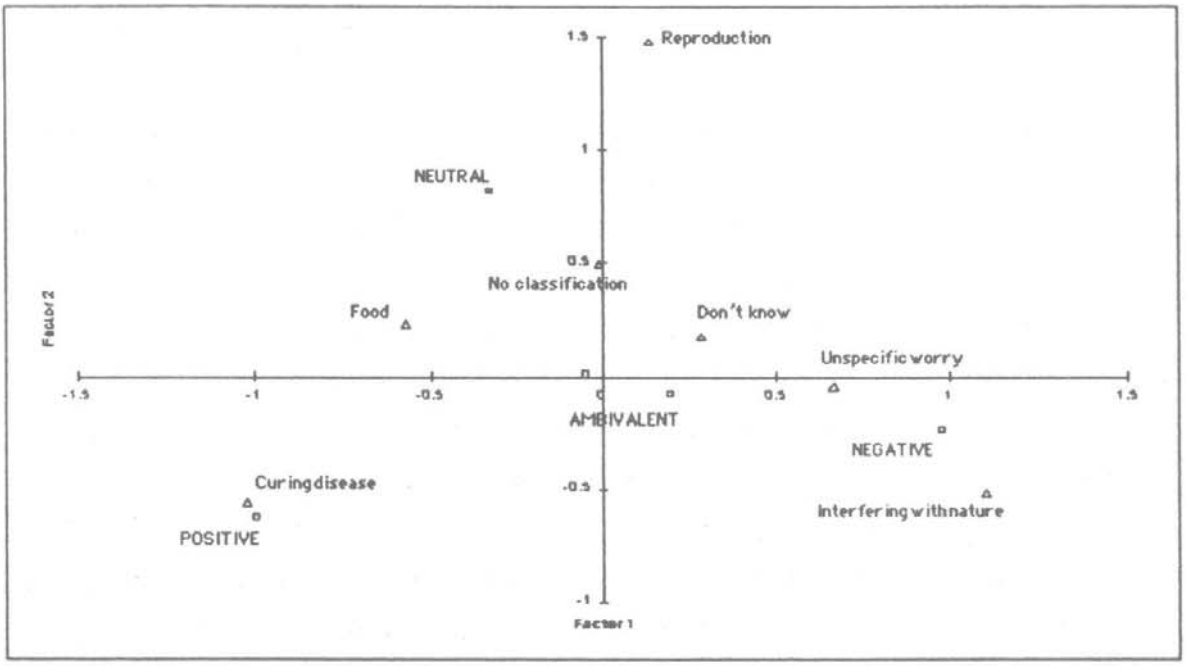
"Qualitative purists", it must be admitted, must view automatic analysis of text with suspicion since the method uses structural features of text, that is proximity of words, as a means to reconstruct classes of meaning. To check whether the assumption that structure allows to extract meaning is valid, Allum (1998) compared two independent analyses of the UK-data set: one was the ALCESTE solution used in the present example and the other was a manual content analysis and categorisation of the same response set. Figure 3 depicts the automatically derived classes and the manually derived categories projected into the same correspondence analytic space.

Figure 3 depicts the categories found in a classical content analysis of the evaluative tone of the open ended responses (little squares, capital letters) and some of the discursive classes found in the automatic analysis (little triangles). The surprisingly good coincidence of the two independently derived results corroborates the validity of ALCESTE's automatic analysis.

Comparison of focus-group discourse

This last example is from a research that derives from the aforementioned investigation. The European research group on public concerns about biotechnology attempted to corroborate and to understand more profoundly the aforementioned results from an automatic text analysis by conducting focus-group discussions on people's concerns about biotechnology in various European countries (Kronberger et al., 2001; Wagner et al., 2001). Here, as with other methods, it was necessary to observe local meaning systems in the analysis. Therefore translation of the focus-group transcripts and simultaneous qualitative analysis was impossible.

The study involved ten countries: Austria, Denmark, Finland, France, Germany, Great Britain, Italy, Portugal, Sweden and Switzerland. In each country there existed a local team of researchers responsible for this task. The research was conducted in seven steps:



Note: Reproduced with friendly permission by Nick Allum (Allum, 1998).

Figure 3 – ALCESTE classes (triangles) and evaluative categories from manual analysis (squares, capital letters) of open responses in a questionnaire about biotechnology in the UK projected into same correspondence space.

(a)The research teams of all ten countries convened and agreed on shared interview guidelines, the sampling rules of the focus-group participants and the general procedure. It was agreed, for example, that the groups should be homogeneous with regard to education levels, because more ignorant subjects are likely to mute in the presence of better educated participants.

(b)The researchers ran the focus-groups in their own countries according to the agreed guidelines approximately in the same period of the year. In most countries the focus-groups were recorded using both, audio and video-recording and their talk and discussion subsequently transcribed. The video-recording was useful in determining who said what in the focus-groups, particularly in the case of simultaneous speech.

(c)The transcript was analysed locally, that is in each original language and culture by experienced qualitative researchers using ATLAS/ti or

NUDIST. This analysis intended to reveal the most prominent features appearing in the material. Besides topical content, particular attention was paid to metaphors used and the way focus-group participants referred to different applications of biotechnology.

(d)The results of the first analysis were brought to a joint meeting of all research groups and they were presented and discussed. This workshop allowed to identify universal topics, images and metaphors as well as features of the transcripts that had only local significance. Subsequently the workshop participants developed a grid of those categories and features that were deemed relevant to the research, be they universal or local. The grid allowed to cross-tabulate content categories and discursive features. It was to be used in a second local analysis of the texts and allowed to enter any local content that was considered relevant by an analyst.

(e)In a second analysis of the texts the researchers were supposed to search for the categories and features constituting the grid in their own material. For the task of comparison, the principle content of the grid were two or three examples for each feature and content category. Table 2 presents the general format of this grid.

(f)Once each research group had completed the grids, a final joint meeting of the researchers established the bases for comparison, such that researchers checked their own grid against the background of other grids. This procedure helped to correct interpretational biases and resulted in minor corrections.

Table 2 crudely shows the scheme of analysis and comparison. For each country a separate crosstabulation of interpretive category and associated discursive features was constructed. The cells contained examples of focus-group text that local researchers considered a typical illustration of the way a certain topic was talked about in the focus-group sample. No need to say that this procedure resulted in rather extensive cross-tabulations which were subsequently used for comparison.

A qualitative research as the present one is, of course, prone to attract many problems. First, the selection of the samples is hard if not impossible to parallelise in different countries. Second, qualitative analysts are likely to have their own styles of analysing text even if a shared method is agreed on as the grounded theory approach (Strauss, 1987) in the present case. Since qualitative researchers can hardly be "parallelised", this approach appeared to

Table 2 – Scheme for comparing focus-group results across countries

	Situational or discursive feature A	Situational or discursive feature B	Situational or discursive feature C	etc.
		COUNTRY 1		
Interpretative Category 1		Specific examples of text for country 1, category 1 and feature B		
Interpretative Category 2				
Etc.				
		COUNTRY 2		
Interpretative Category 1				
Interpretative Category 2			Specific examples of text for country 2, category 2 and feature C	
etc.				
		COUNTRY X		
Interpretative Category 1				
Interpretative Category 2				
Etc.				Specific examples of text for country X, category Y and feature Z

be the best common denominator, although it allows a very high degree of freedom. This freedom, however, was kept under control through the regular meetings. Finally, the biggest problem for professional qualitative researchers is perhaps the fact that a comparative analysis such as this one prohibits analysing the "deep structure" of the texts. Because of the comparative goal, the grids used to present the results were a methodological compromise and

automatically lead to simplifications and polishing idiosyncratic edges of focus-group discourse. This contrasts with much of qualitative analysis which usually strives for an in-depth understanding of texts beyond mere content analysis. Such an in-depth analysis is favoured by cultural psychologists but it would probably not allow cross-country comparison.

Conclusions

The examples presented in the foregoing sections provide an illustration of how the local semantic metric of cultures and language groups can be respected in comparative research. In the quantitative domain a social psychological scale needs to be based on an equivalent semantic metric if it is used across cultures. Without this warrant scale scores and what they mean for the respondents cannot be compared. There exist statistical methods to check for comparable metrics in such data. Van de Vijver and Leung (1997), for example, suggest parametric methods, such as exploratory factor analysis and subsequent target rotation among others. While such methods may yield reliable results if the data can legitimately be considered to be of a parametric quality; But this is a big "if", given the potential biases such as item and method bias introduced by different research teams collecting the data under varying circumstances. An alternative approach are non-parametric methods such as the one illustrated by the "Meaning of Work" research team (Meaning of Work Research Team, 1987) as well as in the example of word associations using data from the Wagner, Valencia and Elejabarrieta (1996) "war-peace" study. The validity of these methods does not depend upon precarious parametric assumptions and they allow to visualise the internal relationships among a set of variables. Therefore they are more intuitive to the researcher (cf. van de Geer, 1993).

Automatic classification of text through ALCESTE (Reinert, 1990) is also a non-parametric structural method. The resulting clusters describe discursive classes, that is, words and phrases that occur in context in a large corpus of text. Using this method on culturally homogeneous text and comparing the obtained cluster solutions in a second step allows to establish similarities and discrepancies of discourse between culture and language groups.

It might appear as inappropriate to talk of semantic metrics in the example of cross-country qualitative analysis of focus-groups but there is some justification for it. In principle, the basic problem in qualitative comparison is the same as in the quantitative domain. Comparing two data sets presupposes identical meaning of items and scores, that is, a comparable semantic metric. With qualitative material this can only be established by doing content-analyses and interpretations locally without prior translation. Only in a second step interpretations and category systems can be translated and brought to bear on a comparative perspective.

Pike's (1967) idea of distinguishing the emic from the etic approach in cross-cultural psychology was a fruitful one and instigated decades of methodological discussions. Nowadays none of the two is exclusively favoured and the most promising methodological developments have been combinations or integrations of both (e.g. Berry, 1989; Helfrich, 1999; Van de Vijver & Leung, 1997, to name but a few) as well as Valsiner's attempt to keep the journal "Culture and Psychology" free of simple numerical cross-cultural comparison research (Valsiner, 2001).

By the same token, the concept of a cultural metric allows to view quantitative and qualitative comparative methodology within a framework integrating emics and etics. This perspective needs further analysis as to what degree it may allow to simultaneously consider quantitative and qualitative material bearing on the same phenomenon. Attempts to use both data sources have been made but found to be notoriously difficult. This issue is being discussed under the heading of validity and "triangulation" of qualitative methods (c.f. Fielding & Fielding, 1986; Flick, 1992). In any case, the debate is far from a close and promises more exciting "culture's adventures in psychology" (Valsiner, 2001, p. 5).

Resumo

Este artigo propõe-se a discutir metodologias qualitativas e quantitativas utilizadas para apreender diferenças culturais e a decorrente comparação entre mentalidades e valores nas diversas culturas.

Palavras-chave: representações sociais; métodos qualitativos e quantitativos; *cross-cultural.*

Abstract

This article discusses qualitative and quantitative methodologies used to study cultural differences and the consequent comparison between mentalities and moral values in several cultures.

Key-words: *social representations; qualitative and quantitative methodologies; cross-cultural.*

Resumen

Este artículo discute las metodologías cualitativas y cuantitativas utilizadas para aprender las diferencias culturales, y la consecuente comparación entre mentalidades y valores en las distintos culturas.

Palabras claves: *representaciones sociales; metodologías cualitativas y cuantitativas; cross-cultural.*

References

Allum, N. C. (1998). *A social representations approach to the comparison of three textual corpora using ALCESTE*. Unpublished Thesis. London School of Economics and Political Science, UK.

Amir, Y. & Sharon, I. (1987). Are social psychological laws cross-culturally valid? *Journal of Cross-Cultural Psychology,* 18, pp. 383-470.

Berry, J. W. (1989). Imposed etics, emics, derived etics. *International Journal of Psychology*, 24, 721-735.

____ (2000). Cross-cultural psychology: a symbiosis of cultural and comparative approaches. *Asian Journal of Social Psychology*, 3, pp. 197-205.

Campbell, D. R. (1961). "The mutual methodological relevance of anthropology and psychology". In: Hsu, F. L. K. (ed.). *Psychological Anthropology*. Homewood, Ill, Dorsey.

Dahlin, B. & Watkins, D. (2000). The role of repetition in the processes of memorising and understanding: A comparison of the views of german and chinese secondary school students in Hong Kong. *British Journal of Educational Psychology*, 70, pp. 65-84.

Fielding, N. G. & Fielding, J. L. (1986). *Linking data*. London, Sage.

Flick, U. (1992). Triangulation revisited: strategy of validation or alternative? *Journal for the Theory of Social Behaviour*, 22, pp. 175-198.

Gibbons, J. L.; Lynn, M.; Stiles, D. A.; Jerez de Berducido, E.; Richter, R., Walker, K. & Wiley, D. (1993). Adolescents' images of women as office workers and homemakers. *Psychology of Women Quarterly*, 17, Guatemalan, Filipino, and U.S.A., pp. 373-388.

Hayashi, C. (1950). On the quantification of qualitatice data from the mathematics-statistical point of view. *Annals of Statistical Mathematics*, 2.

Helfrich, H. (1999). Beyond the dilemma of cross-cultural psychology: resolving the tension between etic and emic approaches. *Culture and Psychology*, 5, pp. 131-153.

Jahoda, G. (1986). Nature, culture and social psychology. *European Journal of Social Psychology,* 16, pp. 17-30.

Jodelet, D. (1993). "Indigenous psychologies and social representations of the body and self". In: Kim, U. & Berry, J. W. (eds.). *Indigenous Psychologies.* London, Sage

Kim, U. (2000). Indigenous, cultural, and cross-cultural psychology: a theoretical, conceptual, and epistemological analysis. *Asian Journal of Social Psychology*, 3, pp. 265-288.

Kronberger, N. & Wagner, W. (2000). "Keywords in context: Statistical analysis of text features". In: Bauer, M. & Gaskell, G. (eds.). *Qualitative Researching with Text, Image and Sound. A Practical Handbook*. London, Sage.

Kronberger, N.; Dahinden, U.; Allansdottir, A.; Seger, N.; Pfenning, U.; Gaskell, G.; Allum, N.; Rusanen, T.; Montali, L.; Wagner, W.; Cheveigné, S.; Diego, C. & Mortensen, A. (2001). "The train departed without us" – Public perceptions of biotechnology in ten European countries. *Politeia*

Laboratory of Comparative Human Cognition (1979). What's cultural about cross-cultural cognitive psychology? *Annual Review of Psychology*, 30, pp. 145-172.

LeVine, R. A. (1982). *Culture, behaviour and personality*. New York, Aldine.

Meaning of Work Research Team (1987). "International comparison of the relationships between MOW variables". In: Meaning of Work Research Team (eds.). *The meaning of working*. New York, Academic Press.

Minoura, Y. (1996). A plea for a hypothesis-generating approach to link the individual's world of meaning and society's cultural orientation. *Culture and Psychology*, 2, pp. 53-62.

Pike, K. L. (1967). *Language in relation to a unified theory of the structure of humna behavior*. The Hague, Mouton.

Reinert, M. (1983). Une méthode de classification descendante hiérarchique: application a l'analyse lexicale par contexte. *Les Cahiers de l'Analyse des Données*, v. 8, n. 2, pp. 187-198.

Reinert, M. (1990). ALCESTE. Une méthodologie d'analyse des données textuelles et une application: Aurélia de Gérard de Nerval. *Bulletin de méthodologie sociologique*, 26, pp. 24-54.

Shweder, R. A. (2000). The psychology of practice and the practice of the three psychologies. *Asian Journal of Social Psychology*, 3, pp. 207-222.

Straub, J. (1999). *Handlung, Interpretation, Kritik. Grundzüge einer textwissenschaftlichen Handlungs- und Kulturpsychologie (Action, interpretation, critique. An outline of a text-scientific psychology of action and culture.* German, Berlin, de Gruyter.

Strauss, A. L. (1987). *Qualitative analysis for social scientists.* Cambridge, Cambridge University Press.

Triandis, H. C. (1996). The psychological measurement of cultural syndromes. *American Psychologist*, 51, pp. 407-415.

_____ (2000). Dialectics between cultural and cross-cultural psychology. *Asian Journal of Social Psychology*, 3, pp. 185-196.

Triandis, H. C.; Bontempo, R.; Leung, K. & Hui, C. K. (1990). A method for determining cultural, demographic, and personal constructs. *Journal of Cross-Cultural Psychology*, 21, pp. 302-318.

Valsiner, J. (2001). Editorial: the first six years: culture's adventures in psychology. *Culture and Psychology*, 7, pp. 5-48.

Van de Geer, J. P. (1993). *Multivariate analysis of categorical data: applications*. London, Sage.

Van de Vijver, F. & Leung, K. (1997). *Methods and data analysis for cross-cultural research*. London, Sage.

Wagner, W. (1997). "Local knowledge, social representations and psychological theory". In: Leung, K.; Kim, U.; Yamaguchi, S. & Kashima, Y. (eds.). *Progress in asian social psychology*, Vol. 1. Singapore, Wiley.

Wagner, W.; Valencia, J. & Elejabarrieta, F. (1996). Relevance, discourse and the 'hot' stable core of social representations – A structural analysis of word associations. *British Journal of Social Psychology*, 35, pp. 331-351.

Wagner, W.; Kirchler, E.; Clack, F.; Tekarslan, E. & Verma, J. (1990). Male dominance, role segregation and spouses' interdependence in conflict. *Journal of Cross-Cultural Psychology*, 21, pp. 48-70.

Wagner, W.; Kronberger, N.; Gaskell, G.; Allum, N.; Allansdottir, A.; Cheveigné, S.; Dahinden, U.; Diego, C.; Montali, L.; Mortensen, A.; Pfenning, U.; Rusanen, T. & Seger, N. (2001). Nature in disorder: the troubled public of biotechnology. In: Gaskell, G. & Bauer, M. (eds.). *Biotechnology 1996-2000: the years of controversy*. London, Museum of Science and Industry.

Wagner, W.; Kronberger, N.; Allum, N.; Correia Jesuino, J.; De Cheveigné, S.; Diego, C.; Gaskell, G.; Heinßen, M.; Midden, C.; Odegaard, M.; Olsson, S.; Rizzo, B.; Rusanen, T. & Stathopoulou, A. (2002). Pandora's genes – Images of biotechnology and nature in Europe. In: Gaskell, G. et al. (eds.). *Biotechnology – the making of a global controversy*. Cambridge, Cambridge University Press.

Wittgenstein, L. (1958). *Philosophical investigations*. Oxford, Blackwell.

Yang, K.-S. (2000). Monocultural and cross-cultural indigenous approaches: The royal road to the development of a balanced global psychology. *Asian Journal of Social Psychology*, 3, pp. 241-264.

Wolfgang Wagner
Nicole Bronberger
Dept of Education and Psychology
University of Linz, A–4040 Linz, EU-Austria.
E-mail: w.wagner@jku.at