Developments in Survey Methodology
Anuška Ferligoj and Andrej Mrvar (Editors)
Metodološki zvezki, 15, Ljubljana: FDV, 2000

How to Develop an Algorithm in Order to Choose the Appropriate Measurement Instrument

Jasna Horvat and Nataša Šarlija[1]

Abstract

The purpose of this paper is to create an algorithm for the construction of quality measurement instruments. The algorithm contains: the measurement instruments base, scale base, and the base of quality criteria for the survey.

In order to use the suggested algorithm, the researcher should first establish the general field of his/her research, then choose the specific branch in the field, then sub-branch and finally the phenomenon in the chosen sub-branch. Then, (s)he should select the appropriate measurement instrument, which (s)he would be free to modify. The researcher could browse through scientific fields, branches and sub-branches, and select from different measurement instruments suitable for the phenomenon, in such a way that (s)he could not only select from predefined sets of measurement instruments, but also create new ones. Every intervention would lead to new testing of measurement instruments in order to ensure stability of the given instrument characteristics.

The ability to rebuild the existing databases, their selectivity, their ability to interchange among themselves and to be tested, are the foundation for the process of building the measurement instruments and scale bases. In order to rebuild the quality criteria base, the process of permanent monitoring and implementation of new theoretical achievements is essential, as is the ability to define new criteria and redefine the existing ones.

The suggested algorithm gives the researcher the ability to select the measurement instrument from the existing bases, or to create a new one. The precondition for using the measurement instrument is running a pilot survey and correcting it with the help of the algorithm, according to the information contained in the knowledge base of quality criteria for the measurement instrument. To bring the suggested algorithm to life, and to get it to regulate itself, the system should be placed on the Internet, and the researchers should be encouraged to start using it.

[1] Faculty of Economics in Osijek, Gajev trg 7, Osijek, Croatia.

1 Basic structure of the survey construction algorithm

Every quality measurement instrument has certain common characteristics. They are developed in such a way that they can be applied as criteria of efficiency for any existing or new measurement instrument.

Numerous authors (e.g., Alwin, 1991; Bailey, 1982; Bearden et al., 1993; Biemer et al., 1991; Cronbach, 1949; Lansing et al., 1971) pledge for evaluation of the quality of a measurement instrument through its ability to fulfil the following requirements:

1. validity
2. reliability
3. ease of testing
4. ability to interpret and compare
5. economy

If the problems concerning the selection of the basic set units into the sample, determining the scope of the sample and the basic set units co-operation level are disregarded, it becomes possible to create an algorithm which would ensure that the quality measurement instrument (one or several questions within a survey) or instruments (various number of questions) are developed.

The fundamental idea to create an algorithm for a survey arises from the following:

1. measures and criteria used to evaluate the validity of a particular measurement instrument,
2. tested and classified measurement scales and measurement instruments for the specific phenomena in social sciences,
3. highly developed statistical applications which ensure quality data processing, but also the analysis of the measurement instruments,
4. applications with the ability to easily transform data (from one known database into the other) and data files (from one statistical computer application into the other),
5. computer assisted personal interviewing applications, which prevent the interviewer reverting to any prior question and give the interviewer the ability to measure the relevance of certain items by means of control variables,
6. highly developed computer decision support system tools.

Obviously, the availability of these items enable a certain level of generalisation of the rules and procedures which must be fulfilled for the particular measurement instrument to be sufficient for measuring the particular social phenomenon. In such generalisations one should always keep in mind the need to foresee as many limitations as possible and the need to ensure the ability to upgrade as well as the flexibility of the set of instruments itself.

2 Stages in the creation of the quality measurement instrument construction algorithm

The main idea for the creation of the survey construction algorithm is to include related fields of all social sciences. As a method, the survey is a common feature of all social sciences of our time. There is almost no area of social life, either of an individual or a whole social group (economical, social, educational, etc.) that doesn’t influence some other area.

Since a number of tested measurement instruments already exists for the majority of the social sciences’ fields, it is possible to create such an algorithm for constructing a survey, which would include tested measurement instruments applicable to one phenomenon for various fields of social sciences exploring that same phenomenon. Search levels of the measurement instrument base during the algorithm construction is drawn out in Figure 1.

Figure 1 shows one of the possible ways to create the measurement instruments base. The field classification of some scientific area is in its reduced form shown on the example of economy. Since the classification of scientific fields is the problem itself, Figure 1 tried to show the idea for creation of the measurement instruments base that contains all the tested instruments of the particular phenomenon regardless the scientific discipline of the earlier research on whose results the instrument was created.

In other words, by retrieving fields, branches and sub-branches within the measurement instruments base the researcher comes closer to the list of the tested measurement instruments for the phenomenon in question.

As Figure 1 shows, the user of a quality measurement instrument construction application should in the first phase of his/her work search through all the fields of social sciences concerning the phenomena of interest. This means that if we observe the marketing, the measurement instruments base will include all the measurement instruments from that area, regardless whether they come from the area of microeconomy, management, consumer psychology or some other fields concerned with market research.

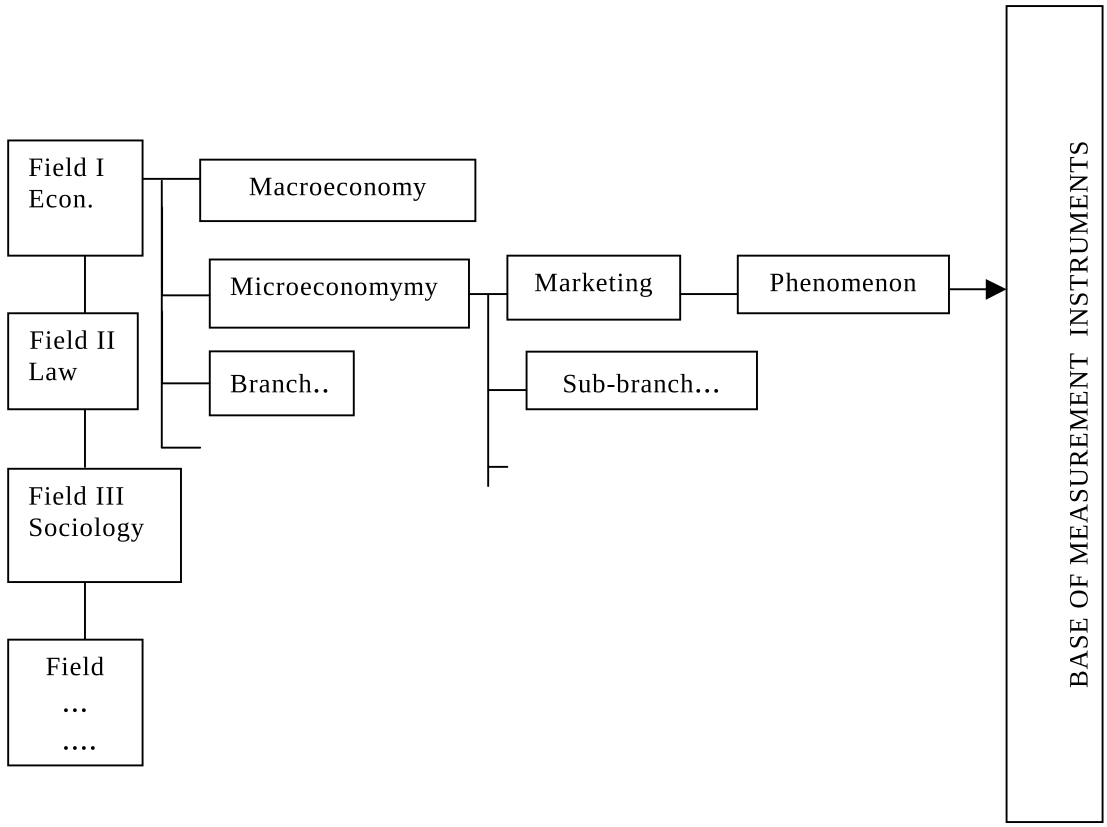

Figure 1: Search levels of measurement instrument base at constructing an algorithm for creating a quality measurement instrument.

3 Selection of measurement instruments

“Measurement means assigning numbers to objects or events, based on some chosen rules. The basis for every measurement is a theory. There has to be a rule of correspondence between the theoretical concept which is to be measured and the measurement instrument which provides the measures.” (Ferligoj et al., 1995)

Measurement could be seen as a three-stage-process, where on the first stage the concept is theoretically defined (conceptualisation), on the second stage the operational variable is defined for each theoretical one (operationalisation), and on the third stage each variable is measured, which results in a set of measured variables.

Recognition and evaluation of errors in the measurement process and the measurement results are the starting point for evaluating measurement quality as regards the measurement procedure. The measurement quality conveys the objectivity of the research through the measures of reliability and validity.

Each field in social sciences is divided into branches, and further into sub-branches. Every sub-branch should contain two types of information in its base (measurement instruments base, measurement scales base):

1. published tested measurement instruments,
2. measurement instruments already used in research (therefore tested), but not published.

This means that, in order to develop a survey construction algorithm, there emerges a need for a large database developed by the means of:

a) permanently monitoring for publicly available tested measurement instruments for particular fields of social sciences, and their inclusion into the database,
b) gathering measurement instruments (along with the databases, i.e. research results) from:
    - scientific institutions,
    - government institutions in charge of conducting statistical research,
    - research associations (e. g., Association for Consumer Research, Graduate School of Management, 632 TNRB, Brigham Young University, Provo, UT (84602); American Marketing Association, AMA Proceedings, 250 South Wacker Drive, Suite 200, Chicago, IL 60605-5819),
    - companies conducting research on a commercial basis,
    - companies conducting research for their own needs,
    - other sources.

Such large, permanently revised, databases are therefore the basic precondition of the survey construction algorithm. If a user wishes to research a specific phenomenon (s)he should first select the field the research belongs to, then the branch in the field of the social sciences and then the problem, i.e. phenomenon being researched within the primary field.

The second step is to select one from the offered phenomenon bases containing the tested measurement instruments. In the process of selecting a measurement instrument the user must be able to apply changes to it as well as to the previously used measurement scale. After the questions from the analysed base have been exhausted, the user should be able to refer to some other base in the same field of social sciences.

After the particular – primary – field of research has been exhausted, the researcher should be given the possibility to search through any other field of social sciences, including branches, sub-branches, and the phenomena for which the tested measurement instruments had been gathered.

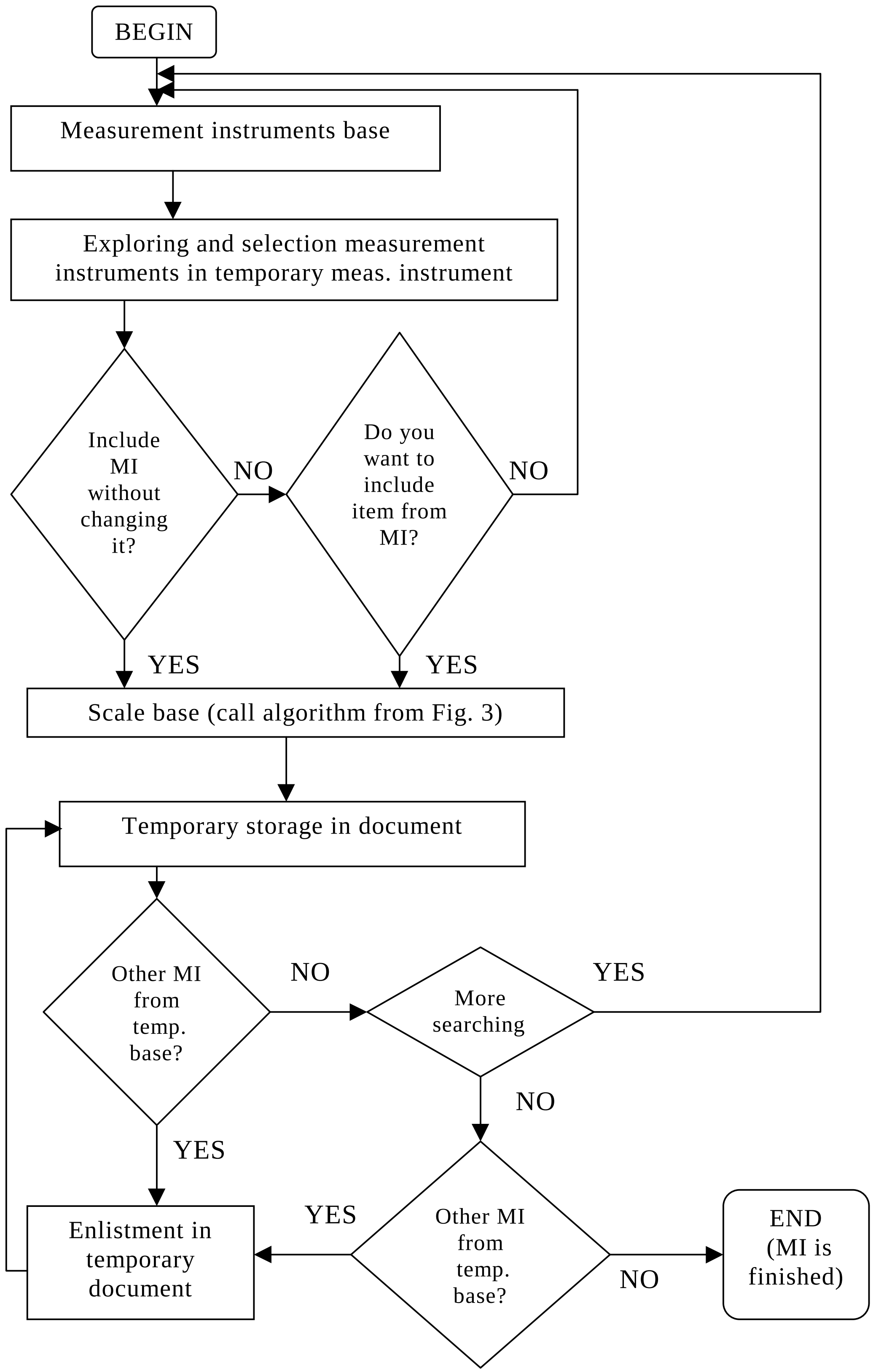

Figure 2: Algorithm for selecting measurement instruments and scales.

In this way the researcher creates the measurement instruments set, not necessarily a new one, which (s)he may further select or change in some other way. All the selected questions, along with the scales, are saved into a text file, which could be easily manipulated with any word processor program.

The process of selecting questions into the new survey is shown in Figure 2.

Users should be able to search through bases, both on measurement instruments and scales, using AND, NOT, and OR logical operators and the combinations of those. It would give them the opportunity to search not only through the hierarchy of fields, branches, sub-branches and phenomena, but also by keywords used in measurement instruments and scales (especially if those scales are nominal).

According to Figure 2, the first thing a researcher should do is to search through the measurement instruments base looking for a specific phenomenon, the result of which would be the instrument list. After close investigation of the list some measurement instruments could be taken over (partly or completely). The next step of the algorithm is to choose a scale (the procedure is explained in Figure 3) from the offered scales base. This completes the first phase which could be called ‘measurement instruments (or their items) and scales selection’.

The algorithm allows the addition of new items that are not among the stored measurement instruments as well as the combination of various measurement instruments from the whole base.

4 Measurement instruments base

One of the basic preconditions for developing the survey construction algorithm is to create the measurement instruments bases. Those bases should contain tested surveys for the fields, branches, sub-branches and particular phenomena in the social sciences.

For instance, the field - economics -, the branch - microeconomics -, the sub-branch - marketing -, the base could contain the following six groups of phenomena (Bearden et al., 1993):

1. consumers individual characteristics,
2. consumers system of values,
3. consumers participation in the informational process,
4. consumers reactions to commercial propaganda,
5. consumers opinions on companies, agencies and the market
6. sales, sales management, external and internal influences on sales department.

5 Scale base

Beside the measurement instruments base, the survey construction algorithm contains the base of the applicable scales. The scale base contains well-known scales classified according to the levels of measurement. One of the possible scale classifications (according to van Logchem et al., 1996) can be seen in Table 1.

Scale is an item or set of items (generally more than one) for measuring some property features such as attitude, for example. The property is generally considered to be unidimensional and a quantitative score is usually derived (for example, an IQ score or an alienation score).

Figure 3 presents an algorithm for the selection of scales from the scale base. Within that base, beside the types of scales, the ability for the particular changes to be made while redefining the format modality of particular types of scales must be ensured. Particular changes that redefine the format modality of certain types of scales must be allowed to happen in the base. It is important to point out that it is possible to extend the used classification of measurement and scale levels according to the opinion of other authors.

Table 1: Levels of measurement and scales (van Logchem et al., 1996).

Levels of measurement	Scales
Non-metric measurement	non-metric scales categorical scales measurement of discrete variables
Nominal measurement	nominal scales measurement of attributes
Partially ordered measurement	partially ordered scales
Ordinal measurement *ordered metric scales	ordinal scales measurement of rank-order data metric analysis of ordinal data
Metric measurement	metric scales measurement of continuous variables
Interval measurement	interval scales log-interval scales difference scales
Ratio measurement	ratio scales cardinal measurement absolute scales similarity scales

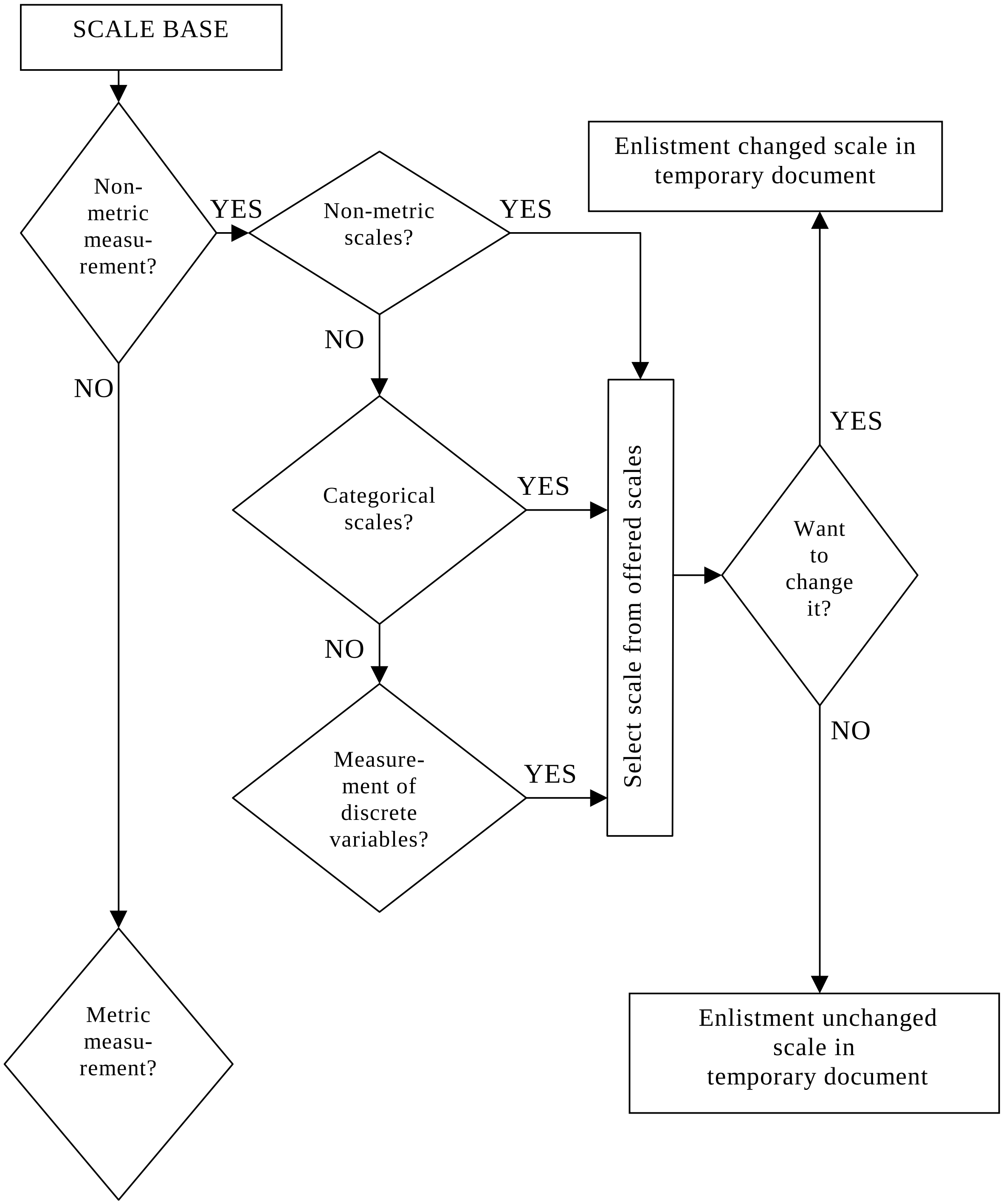

Figure 3: Algorithm for selecting a scale from the scale base.

As seen in Figure 3, the researcher can choose the level of measurement. The next step is to choose the scale. In other words, the database is assisting the edition of the questionnaire according to the measurement level selected. The precondition for using the complete algorithm is the vast knowledge and rich experience of the researcher, since it can be assumed that the measurement quality will be different from that reported in the database if the researcher manipulates the type of measurement level. Therefore it is decided that the algorithm for the selection of a scale from the scale base should contain a very detailed help function.

6 Defining criteria for the quality of the research

As the individual surveys differ by their purpose, the assumption must be made that the quality criteria for the particular survey differ accordingly. Therefore, the algorithm for defining the criteria for the quality of the research should contain:

- base of the criteria for the quality of the research, with different levels of criteria for evaluation,
- the ability to redefine and modify that base.

The researcher should first select the criteria for the evaluation of the survey quality, paying attention to characteristics of the survey itself, and also to the level of the criteria for evaluation. Figure 4 shows the algorithm for assessing the criteria for the quality of the research.

7 Base of criteria for evaluating the quality of research

Base of criteria for evaluating the quality of research could contain the following constant elements (Robinskon et al., 1991), i.e. necessary for its existence:

- respect for theoretical development, i.e. theoretical structure of the research;
- number of pilot-tests, which means revising, i.e. developing items (following the need for exclusion of unreliable items);
- presentation of measures of descriptive statistics (i.e. arithmetical mean and standard deviation) for the whole sample and sub-samples (i.e. homogenous groups of interviewees);
- analysis of the sample (type of sample, rate of participating interviewers in the survey);
- evaluating inter-item correlation;
- values of the α coefficient;
- results of the factor analysis;

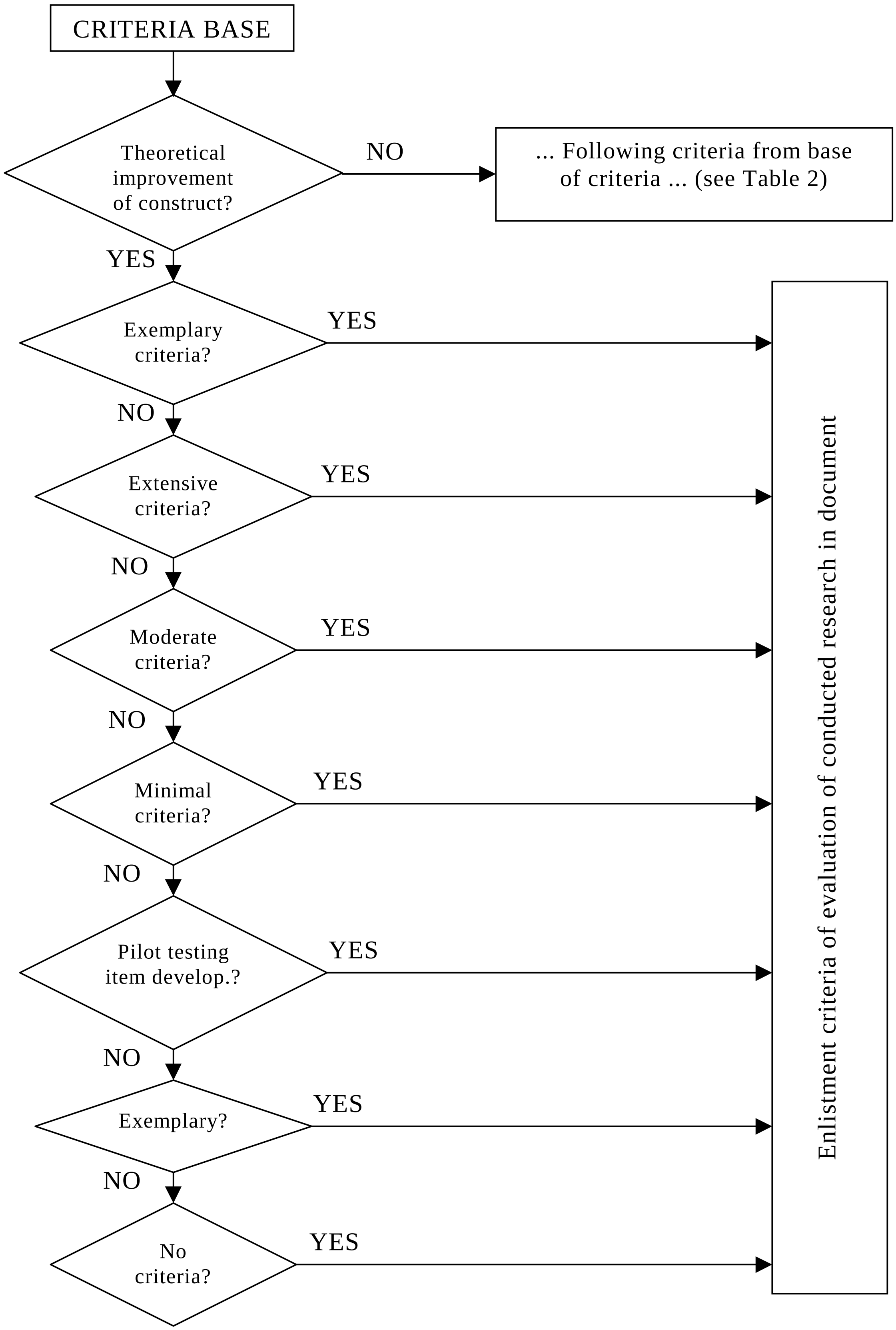

Figure 4: Algorithm for assessing quality of research.

- characteristics of the results of retests and their comparability (test-retest);
- evaluating groups validity;
- convergent validity;
- discriminant validity;
- analysis of independent studies.

As already said, the researcher should be allowed to select the level of criteria for evaluating the quality of the research:

- exemplary level of criteria,
- extensive level of criteria,
- moderate level of criteria,
- minimal level of criteria,
- no criteria.

To achieve greater flexibility of the system, the base should have the ability to expand and to be revised. Furthermore, it would be useful to make some recommendations for further analysis in order to improve the quality of the research. In such a way, the error of excluding or neglecting certain fields in statistical analysis of functional paradigms of the models, as the dominating methodology for evaluating the hypotheses by analysing the variables can be avoided. The emphasis should be put on warning the researcher from excluding unnoticed, latent correlation between variables and their values.

Table 2 contains some general criteria and also levels for evaluating attitude measures.

8 Conclusion

The purpose of the paper is to explain the algorithm for choosing an appropriate measurement instrument. The paper suggests the creation of three bases: the measurement instruments base, the scale base, and the base of the research quality criteria. Measurement instruments and scale bases are included into the algorithm for the selection of measurement instruments and scales, in order to suggest the researcher the way to create his own measurement instrument or to select one from the existing, already tested instruments. All the used base-founded classifications (measurement instruments base, scale base, criteria base) can be treated as temporary. The ability to be redefined and modified is a necessity for all three bases. However, only qualified scientists can be allowed to make such modifications.

Table 2: Some general rating criteria for evaluating attitude measures (Robinskon et al., 1991).

Criterion rating	Exemplary	Extensive	Moderate	Minimal	None
Theoretical develop-ment structure	Reflect several important works in the field plus extensive face validity checks	Either reviews several works or has extensive face validity	Reviews more than one source	Reviews one (no sources)	Ad hoc
Pilot testing/item develo-pment	More then 250 items in the initial pool; several pilot studies	100-250 items in initial pool; more then two pilot studies	50-100 items in initial pool; two pilot studies	Some items eliminated; one small pilot study	All initial items included; no pilot study
Means for some available norms	Means and SDs for several sub-samples and total sample; extensive information for each item	Means and SDs for total and some groups; some item information	Means for some sub-groups; information for some items	Means for total group only; information for 1-2 items	None; no item information
Samples of respon-dents	Random sample of nation/ community with response rate over 60%	Cross-sectional sample of nation/ community; random national sample of college students	Some representa-tion of non-college groups; random sample of college students in same departments or colleges	Two or more college classes (some hetero-geneity)	One classroom group only (no hetero-geneity)
Inter-item corre-lations	Inter-item correlation average of .30 or better	Inter-item correlation average of .20-.29	Inter-item correlation average of .10-.19	Inter-item correlation below .10	No inter-item analysis reported
Coefficient alpha	.80 or better	.70-.79	.60-.69	< .60	Not reported
Factor analysis	single factor from factor	single factor from factor	single factor from factor	Some items on some factors	No factor structure

	analysis	analysis	analysis		
Test-retest	Scale scores correlate more than .50 over at least a 1-year period	Scale scores correlate more than .40 over a 3-12 month period	Scale scores correlate more than .30 over a 1-3 month period	Scale scores correlate more than .20 over less than a 1-month period	No data reported
Known groups validity	Discrimina-te between known groups highly significant; groups also diverse	Discrimina-te between known groups highly significant	Discrimina-te between known groups significant	Discrimina-te between known groups	No known groups data
Convergent validity	Highly significant correlations with more than two related measures	Significant correlations with more than two related measures	Significant correlations with two related measures	Significant correlations with one related measure	No significant correlation reported
Discriminant validity	Significan-tly different from four or more unrelated measures	Significan-tly different from two or three unrelated measures	Significan-tly different from one unrelated measures	Different from one correlated measure	No difference or no data
Freedom from response	Three or more studies show indepen-dence	Two studies show indepen-dence	One study shows independ-ence	Some show independence, others do not	No test of indepen-dence

On the other hand, all researchers should be in the position to use the suggested algorithm. In order to achieve the widespread acceptance, the algorithm should be made accessible through the Internet. Since the tested measurement instruments, scales and the quality criteria already exist and are published, there is a possibility of unifying that. The algorithm allows the researcher to use the unity with the aid of defined cognitive mechanisms. In other words, it contains the cognitive strategy based on the expert (methodologist, statistician) usage when evaluating the quality of measurement instruments. In that way, the return of information to the level of quality of the research would also be provided to users who may not necessarily be methodologists.

The use of the algorithm for evaluation of the measurement instrument quality should allow more quality interpretation of the research results, as well as the conceptual changes if the research is to be repeated.

References

[1] Alwin, D.F. (1991): Research On Survey Quality. *Sociological Methods and Research*, **20**, 3-29.

[2] Bailey, Kenneth D. (1982): *Methods of Social Research*. The Free Press, A Division of Macmillan, Inc.

[3] Bearden, W.O., Netemeyer, R.G., and Mobley, M.F. (1993): *Handbook of Marketing Scales - Multi Item Measures for Marketing and Consumer Behavior Research*. Sage Publications, Newbury Park, London, New Delhi.

[4] Biemer, P.P., Groves, R.M., Lyberg, L.E., Mathiowetz, N.A., and Sudman, S. (1991): *Measurement Errors in Surveys*. New York: Wiley.

[5] Cronbach, L.J. (1949): *Essentials of Physiological Testing*. New York: Harper & Brothers.

[6] Ferligoj, A, Leskošek, K., and Kogovšek, T. (1995): Zanesljivost in veljavnost merjenja. *Metodološki zvezki*, **11**, Ljubljana: FDV.

[7] Lansing, J.B. and Morgan, J.N. (1971): *Economic Survey Methods*. Ann Arbor: The University of Michigan.

[8] Lessler, J.T. and Kalsbeek, W.D. (1992): *Nonsampling Error in Surveys*. New York: Wiley.

[9] Logchem, M., Kantebeen, G.W., Reenders, A.H., and Hoogenboom, J.M.A. (1996): *Thesaurus of Social Research Methodology*. SRM-Documentation Centre, Rotterdam.

[10] Robinskon, J.P., Shaver, P.R., and Wrightsman, L.S. (1991): *Measures of Personality and Social Psychological Attitudes*. San Diego: Academic Press.