Bulletin of the *Transilvania* University of Braşov • Vol. 5 (54) No. 1 - 2012
Series VII: Social Sciences • Law

DATA MANAGEMENT IN META-ANALYTICAL PROCEDURES: MYTHS, REALITIES AND BIASES

Radu I. POPA[1]

Abstract: *The aim of the present paper is to highlight the importance of a correct data management and the preparation of a proper statistical value base in developing a meta-analysis. When dealing with meta-analytical procedures, research studies encounter a wide range of challenges, ranging from the classical value losses towards more modern coding or sampling biases. The theoretical and practical examples of this paper indicate that a meta-analytical approach is most useful and applicable when related to a clear problem specification, a correct coding scheme and the use of eligible study reports.*

Key words: *meta-analysis, coding protocol, study descriptors, bias.*

1. Introduction

Across the long history of methods inside the psychological field of research, many a time we have witnessed the lack of a quantitative procedure which could revise a research theme or phenomenon, in order to determine in a certain degree the result of a study, which can be replicated by other studies with success.

In 1952, Hans Eysenck [6] conducted a challenging debate upon the beneficial effects of psychotherapy on patients, while Gene V. Glass offered a statistically standardized response called 'meta-analysis' in order to present the effectiveness of psychotherapy across a wide range of studies [6]. Since 1970 meta-analysis has been developing as a collection of methods and techniques for quantitative research synthesis.

In time, the meta-analysis approach has been refined, its methodology being restructured and applied to various subjects. This form of survey or integrated analysis still offers answers to research reports on psychological processes at a complex level, with a strong possibility to generalize results, theories and paradigms.

2. A theoretical view on meta-analysis

Meta-analysis procedures impose a structured approach on the evaluation of research findings. It brings out a much more complex and strong apparatus, resolving a wide range of issues, which are dealt in the conventional research review techniques [5]. Specification of criteria, coding strategies for eligible studies, data analysis and protocols, sampling information and possible biases are a few elements which recommend the meta-analysis as a proper method when

[1] Faculty of Psychology and Education Sciences, University of Bucharest.

researchers' aim is to offer the 'meta-image' upon a well-debated subject inside the research field, across studies and time [2], [6], [9].

The database of a meta-analysis can provide multiple ways to work with the collection of studies, results and statistical frameworks [6].

2.1. Myths in meta-analytical procedures

At present, this type of analysis still tends to challenge its own users. By being a complex method, which many times imposes more questions than answers, the meta-analysis restricts the wide range of usage among researchers, by having a sophisticated technical and statistical apparatus [4]. Other studies show that there is an increase in the interest of meta-analysis development, eradicating the myth of the 'not so user friendly' method which consumes a lot of time and effort [6]. From another point of view, by selecting the proper statistical tools, when conducting such types of analysis may be the most proper and correct way to surpass this myth and benefit on the wide range of advantages meta-analysis has to provide.

Through meta-analysis, research reports can be summarized, coded, integrated and interpreted at a complex level, confirming the psychological topic in its domain of applicability [6], [7].

In this manner, another myth, wrongfully attributed, is to associate the meta-analysis with the evaluation of theoretical papers, qualitative research reports with no quantitative data, policy reviews [5], [6] etc. The meta-analysis can provide a statistical correct and strong support only for quantitative research studies which provide complete data and results [6], [7], [9]. At this point, studies which do not report their descriptive or inferential statistics upon the resulting data or present in a synthetic way their reports will not be able to be comprised in a meta-analytic evaluation, due to the lack of statistical information, which can bias the meta-analysis [2], [4], [5].

The third myth consists on the output of meta-analysis, seen sometimes as the general perspective, which unifies the results of the research report upon a certain psychological process or phenomenon [2], [5], [6].

Instead, practice shows that many studies report heterogeneous results and conclusions, affecting the meta-analysis procedure and evaluation [6], [8].

Even so, a correct meta-analysis will point out the general result, the main trend and conclusion upon the collection of research studies on a certain topic, showing also, where needed, the differences between the various results of studies and explaining the intensity of the phenomenon at general level of population and showing even obscure or hidden relations between variables.

Other authors suggest that even at this level, individual research studies may sometimes provide a proper view on specific topics when evaluating them [1], [3], [10].

The fourth myth refers to data management inside meta-analysis, where many a time the researcher will deal with a mix of studies, when comparing results which differ a lot from each other.

By carefully choosing the eligible studies, variables and data that provide the correct information upon the investigated process or phenomenon, a meta-analysis will evaluate studies that are comparable and with no extreme or disparate results [5], [6], [7].

In this manner, testing the variance of effect size distributions and analysing correctly the study result subgroups will show the sources of differences.

3. Data management inside meta-analysis

When dealing with a wide range of data collections in meta-analysis, different software programs have been developed over time, from the first MetaDOS by Stauffer in 1996 or Metawin by Rosenberg in 1997, RevMan by Cochrane in 2000, Comprehensive Meta-analysis (CMA), MetAnalysis, MIX or WEasyMA, to the well known applications for general statistics calculations in social sciences such as SPSS, SYSTAT, SAS, PASW Statistics or IBM SPSS Statistics which can be appropriate platforms for meta-analysis evaluations [6], [7], [9].

After gathering the eligible studies, the coding process of research papers is the new vital point. In this manner all the coding procedures are related to a coding protocol that indicates what type of data is about to be evaluated and processed inside the meta-analysis [8].

There are two distinct parts when considering the coding process [6]. The first one refers to the selection of study descriptors and encoding the selected features. The second part imposes the encoding of research results or findings under the form of effect sizes. In this view, data from effect sizes values becomes the dependent variable while the features of the research studies form the independent variable (e.g. measures, samples, constructs, treatment, context, designs etc.).

Still, data management inside the coding process may pose a lot of challenges at practical level when values are missing, the items to be evaluated are not defined properly or other values are unusual and ambiguous [9], [10]. Moreover, a lot of studies do not present the complete results or data in order to be integrated inside the meta-analysis. From this perspective, a correct and strong coding form and guidelines may spare the researcher of the troubles in data management, loss of data, difficult coding or biases.

Many authors recommend the usage of coding units and modules for the coding to be complete [2], [4-6].

Other authors present several items that can be used in order to facilitate the coding process and support data for effect sizes values such as: sample size, means and standard deviations, proportions, estimation methods, confidence rating, amount of attrition, reliability of variables, range restriction [6], [7], [9], [11].

Another dimension of data management in meta-analysis consists of the treatment in study descriptors.

Some authors categorize those values in substantive elements (e.g. sample source, sample descriptors, independent variables etc.), study methods (e.g. sampling methods, research design, statistical power, attrition, data analysis, external validity etc.) and source descriptors (e.g. publication form, year of publication, country of publication etc.) [6], [7], [9].

By sorting the valuable data, after strong criteria, may be one of the successful tools in meta-analysis.

Studies which provide clear distinguished features, key variables or research subjects and methods, coded under quantitative data, offer the means to build up a relevant synthetic analysis [2], [11].

Sorting the eligible studies can be done while using an identification value, in order to specify the date of entry in the database [8].

This function allows a complex search when dealing with a huge collection of studies inside a meta-analysis, being useful also in the establishment of a certain hierarchy of studies, depending on the features the researcher wants to outline across the evaluation.

3.1. Bias corrections and adjustments

Hunter and Schmidt proposed in 1994 several methods for adjusting the effect sizes of meta-analysis [5], [6]. Among these, the adjustments of unreliability of variables, range restrictions or dichotomization are the most frequently used.

In this manner, the researcher has to decide which method to apply, what variables are about to be adjusted in order to estimate as correct as possible the magnitude, contained by the effect size.

For example, when adjusting the unreliability of a variable, the un-attenuated effect size can be done by using the formula [6], [7-9]:

$$ES' = \frac{ES}{\sqrt{r_{yy}}}. \tag{1}$$

In the above case, the *ES* is the observed effect size, while the r_{yy} is the reliability coefficient.

In the case of two variables, the formula becomes as follows, r_{xx} being the reliability coefficient for the first variable, while r_{yy} is the reliability coefficient for the second variable:

$$ES'_r = \frac{ES_r}{\sqrt{r_{xx}}\sqrt{r_{yy}}}. \tag{2}$$

When dealing with a case of range restriction the formula is:

$$ES'_r = \frac{(U)(ES)}{\sqrt{(U^2-1)ES^2+1}}. \tag{3}$$

In this case, the U value is the ratio of the study standard deviation to the unrestricted target standard deviation [5].

A good bias prevention starts with the correct selection of studies for the meta-analysis.

Many reports do not state the values needed for meta-analysis calculations and procedures, others present them in a synthetic way and others report only some data about their research sample, design, instruments, measures or full results [7], [9].

The lack of data may lead to serious biases inside the meta-analysis. This is why the researcher must decide and select the eligible study reports for analysis.

In some cases, depending on the study aim and objectives, the researcher may decide to take into evaluation studies which present partial data on variables (e.g. correlation coefficients, means, ranges etc.) in their report, with the condition that this does not affect the interactions and data management in the meta-analysis [8], [10], [11].

Meta-analysis, by being a complex method, which deals with major collections of data, imposes frequently a greater attention for value loss, when building a database.

From this view, by using a coding guide and protocol may prevent the absence of some results, values or even the incorrect completion of the database which result in a bias (see Figure 1).

Another bias of results in meta-analysis may be caused by extreme data, contained in various studies [5], [6]. The decision to keep or eliminate these values is highly dependent on the study aim and the researcher's objectives. In some cases, if studies have the tendency to present opposite results, even if extreme, may be beneficial for the meta-analysis to evaluate the reports [1], [3], [7]. The result can indicate and clarify different tendencies, confirm a new variable, hidden until then to individual research studies or even 'rewrite' the paradigm, defined in a new way, with new interactions and values.

Study ID	Publication reference		Statistical data							
	Publication data	*Publication year*	*Sample size*	*Gender*	*Mean age*	*Standard deviation*	*Type of organizational culture*	*Variables evaluation and instruments*	*Significance tests values*	*Correlations values*
001										
002										
003										

Source: A model for coding data inside a meta-analysis study

Fig. 1. *Example of a coding form for data management inside meta-analysis*

3.2. Coding protocol

The coding form, as discussed before in the paper, organizes the database of the meta-analysis. Clear and efficient problem specifications with a correct coding scheme improve the data management process and prevent biases at a starting point in meta-analytical procedures [7], [8]. In other words, the researcher must choose the topic to be investigated, while describing the main associations, relations or links between variables, which are to be evaluated inside the meta-analysis. For example, a meta-analytic study might decide to focus on the associations between personality dimensions and stress at work, while evaluating the coping strategies. At the same time, another approach may focus on personality dimensions and stress at work, while evaluating the burn-out process. In this manner, the study must specify and describe in a clear way, from the start the problem to be evaluated, in order to select the proper reports. As for coding the articles, one criterion can consist at first on their publication information, which may lead to forming a collection, specified between a certain period of time, theme and publication type (see Figure 1). Secondly, cells containing data about the sample size, gender proportion and specifications, mean age and its standard deviation, support organizing a primary database needed in future calculations [4-6]. Third point brings out the need for effect size data and measure descriptors in order to evaluate and calculate further on the interactions inside the meta-analysis [2], [6]. The researcher may adapt in any way possible, according to the aim and variables of the meta-analytic approach, the coding form, categories and data to be managed in his analysis [5], [6]. For example, in a meta-analysis that focuses on personality dimensions, according to a theoretical model, coders can be given brief definitions of each personality dimension, along with a set of descriptive adjectives reflecting high and low scores and a list of subscales, if necessary, depending on measures that were used [1], [3], [6]. In addition to this example of coding, the researcher may include coders for different instruments for evaluating the personality dimensions, which focus on positive affectivity, behavioural activation, sociability, and high activity level, emotional stability, sensitivity to threat, and trait anxiety, task orientation, persistence, impulsivity or cooperativeness [1], [3], [7] etc. Moreover, coding will help organize the studies which are comparable, using if possible similar designs and forms, samples and procedures.

4. Conclusion

Meta-analysis still improves and develops in a challenging framework, being many a time underused, researchers fearing its complex apparatus and methods.

On the opposite, this procedure, if properly applied, can lead to a better understanding and a specific perspective on the situational field of investigation. A meta-analysis will provide every user with a profound analysis of the results, underlining the general pattern of research findings, many times hidden to the evaluation and promising a strong degree of applicability.

Acknowledgements

This article would not be possible without the help, support, teachings and good will of Ph.D. reader Florin Alin Sava and Ph.D. senior lecturer Laurenţiu Maricuţoiu at the West University of Timişoara, their expertise and studies in understanding meta-analysis being of a vital value. My appreciation goes also to Ph.D. professor Marcela Luca at the Transilvania University of Braşov, and Ph.D. professor Mihai Aniţei at the Bucharest University for being a vivid inspiration and example in the study of organizational research field.

Other information may be obtained from the address**:** radupopaioan@yahoo.com.

References

1. Bono, J.E., Judge, T.A.: *Personality and transformational and transactional leadership: A meta-analysis.* In: Journal of Applied Psychology (2004) Vol. 89 (5), p. 901-910.
2. Borenstein, M., Hedges, L.V., Higgins, J.P.T., Rothstein, H.R.: *Introduction to meta-analysis*. West Sussex. John Wiley & Sons, 2009.
3. DeGroot, T., Kiker, S., Cross, T.: *A meta-analysis to review organizational outcomes related to charismatic leadership*. In: The Canadian Journal of Administrative Sciences (2000), Vol. 17 (4), p. 356-371.
4. Eden, D.: *Replication, Meta-Analysis, Scientific Progress and AMJ's Publication Policy.* In: Academy of Management Journal (2002), Vol. 45 (5), p. 841-846.
5. Hunter, J.E., Schmidt, F.L.: *Methods of Meta-Analysis. Correcting Error and Bias in Research Findings*. Thousand Oaks. Sage Publications, 2004.
6. Lipsey, M.W., Wilson, D.B.: *Practical meta-Analysis*. Thousand Oaks. Sage Publications, 2001.
7. Maricuţoiu, L.P.: *Introducere în metaanaliză. Metaanaliza corelaţiilor (Introduction in meta-analysis. The meta-analysis of correlations)*. In: Human Resources Psychology (2008), Vol. 6 (2), p. 112-127.
8. Popa, R.I., Negoescu, L., Pleşea, S. Maricuţoiu, L.P.: *Personalitate şi leadership transformaţional: o metaanaliză (Personality and the transformational leadership: a meta-analysis).* In: *Psychology in modern organizations*, Avram E. (ed.). University Press, Bucureşti, 2008, p. 138-152.
9. Rosenthal, R., DiMatteo, M.R.: *Meta-analysis: Recent Developments in Quantitative Methods for Literature Reviews*. In: Annual Review of Psychology **52** (2001), p. 59-82.
10. Shao, L., Webber, S.: *A cross-cultural test of the "five-factor model of personality and transformational leadership"*. In: Journal of Business Research **59** (2006), p. 936-944.
11. Sulea, C., Maricuţoiu, L., Zaborilă, D.C., Pitariu, H.D.: *Predicting counterproductive work behaviours: A meta-analysis of their relationship with individual and situational factors*. In: Human Resources Psychology (2010), Vol. 8 (1), p. 66-81.