# Doing and reporting a meta-analysis

Juan Botella[1] and Hilda Gambara (*Universidad Autónoma de Madrid, España*)

**ABSTRACT.** Meta-analysis, a methodology for reviewing, integrating and analyzing research results has spread by all the fields of Psychology. But because of its relative youth and the relative heterogeneity of practices appeared under this label, many researchers have not get a good background on it. In the present theoretical study it is described the whole process, beginning with the preliminary questions and the motivations justifying it. Second, the details of how to work out each of the tasks, ordered as stages, in a meta-analysis. Finally, the way to report the results of a meta-analysis in a scientific journal as International Journal of Clinical and Health Psychology. Besides highlighting some frequent problems that arise during the process, a number of practical advises for the novices are also given. Finally, it is proposed a guide of relevant issues to be taken in account for writing a meta-analytic report. The guide is addressed both to the authors and the reviewers of this type of papers.

**KEY WORDS.** Meta-analysis. Research synthesis. Research reports. Research methods. Theoretical study.

**RESUMEN.** El meta-análisis, una metodología para revisar, integrar y analizar resultados de investigación, se ha extendido por todos los campos de la Psicología. Sin embargo, debido a su relativa juventud y a la relativa heterogeneidad de las prácticas aparecidas bajo este epígrafe, muchos investigadores no han adquirido una buena formación básica sobre ella. En el presente estudio teórico se describe el proceso completo, comenzando por las preguntas preliminares y las motivaciones que lo justifican; en segundo lugar, se detalla cómo afrontar cada una de las tareas, ordenadas como fases,

---

[1] Correspondence: Facultad de Psicología. Universidad Autónoma de Madrid. C/ Ivan Pavlov, 6. Campus de Cantoblanco. 28049 Madrid (España). E-mail: juan.botella@uam.es

que integran un meta-análisis; por último, se expone la forma como se debe informar de los resultados de un meta-análisis en una revista científica como *International Journal of Clinical and Health Psychology*. Además de destacar algunos problemas que surgen con frecuencia durante el proceso, se presentan también algunos consejos dirigidos a los novatos. Para terminar, se propone una guía de aspectos relevantes que se deben tener en cuenta para la elaboración de un informe meta-analítico. Esta guía está dirigida tanto a los autores como a los revisores de un meta-análisis.

**PALABRAS CLAVE**. Meta-análisis. Integración de investigaciones. Informes de investigación. Metodología de investigación. Estudio teórico.

**RESUMO.** A meta-análise, uma metodologia para rever, integrar e analisar resultados de investigação, estendeu-se por todos os campos da Psicologia. No entanto, devido à sua relativa juventude e à relativa heterogeneidade das práticas que surgiram sob esta epígrafe, muitos investigadores não adquiriram uma boa formação de base sobre ela. No presente estudo teórico descreve-se o processo completo, começando pelas perguntas preliminares e as motivações que o justificam; em segundo lugar, detalha-se a forma de enfrentar cada uma das tarefas, ordenadas como fases, que integram uma meta-análise; por último, expõe-se a forma como se deve informar acerca dos resultados de uma meta-análise numa revista científica como o *International Journal of Clinical and Health Psychology*. Para além de destacar alguns problemas que surgem com frequência durante o processo, apresentam-se também alguns conselhos dirigidos aos iniciados. Para terminar, propõe-se um guia de aspectos relevantes que se devem ter em conta para a elaboração de um relatório meta-analítico. Este guia está dirigido tanto aos autores como aos revisores de uma meta-análise.

**PALAVRAS CHAVE.** Meta-análise. Integração de investigações. Relatórios de investigação. Metodologia de investigação. Estudo teórico.

## Introduction

In the 80s and 90s growth and spread "meta-analysis", a methodology for reviewing and synthesizing research results (Chalmers, Hedges, and Cooper, 2002). Its present importance is reflected, for example, in the recommendations for the primary research reports of including the effect size along with the statistical tests, or at least enough information to calculate it, as a way to allow its integration in future meta-analysis (Wilkinson and Task Force on Statistical Inference, 1999). But as a consequence of this rapid rise many researchers have learned about it without any assumed standard both for doing and reporting them. The heterogeneity of practices appeared under this label has yielded several attempts to make them more homogeneous in Psychology (Halvorsen, 1994; Rosenthal, 1995), Medicine (Quorum Group, 1999), and other fields.

In the present theoretical study (Montero and León, 2005) it is explained briefly how to do a meta-analysis and how to report its results for publication by a scientific journal of psychology as *International Journal of Clinical and Health Psychology* (see a more extended treatment in Botella and Gambara, 2002). It is addressed to several

groups of readers: a) those that already know the theoretical advantages of this methodology and want to learn practical details about it; b) those that are going to do their first meta-analysis; c) those that once done a meta-analysis want to know how to write the report in a format more acceptable by a scientific journal; d) those that want to get criteria of quality to read and assess meta-analytic reports. The paper is organized in several sections. The first one is devoted to the questions that from our point of view must be answered before making the decision of doing a meta-analysis. In the second one, we will give the details of how to work out each of the stages in a meta-analysis, indicating also how they must be reflected in the report. In the third one, several general recommendations about the report are provided. In the following, we will give some practical advises that will help to the novices to give their first steps with confidence. The last one is devoted to the conclusions; includes an ordered list of relevant issues to be taken in account, directed both to the authors and the potential reviewers of a meta-analysis.

## Preliminary questions

Before making the decision of doing a meta-analysis it is advised to stop to thinking about what is going to be done. Specifically, it is advised to gain clear answers for two preliminary questions. The first one has to do with the motivation for doing that specific meta-analysis and the goals pursuit; the second one is whether the meta-analytic methodology is the most appropriate and why. Related with the first one, it is important to have an answer and communicate it in the report. Complicity with the reader is an excellent starting point to transmit the message. The following list of motivations is not exhaustive, but probably in most cases the starting point is one or several of them. Any of these motivations justify carrying out a meta-analysis:

– Collecting and integrating the evidence related with the degree of association between variables. For example, the American Psychological Association Division 12 Task Force on Psychological Interventions (Chambless *et al.,* 1998) cites Dobson's meta-analysis (1989) about the efficacy of Beck's Cognitive Therapy for depression. They consider this source as enough for including this technique in the list of Empirically Validated Therapies.

– Looking for some order in a contradictory or confusing empirical basis. For example, in the meta-analysis done by Vargas, Gambara, and Botella (in press) 38 studies relating self-esteem with the age for initiating sexual activity in adolescence were integrated. Until then the researches carried out around this question had not reached a consistent conclusion; whereas some studies concluded that low self-esteem was related with earlier sexual activity, others concluded just the opposite. This meta-analysis indicates that, although self-esteem apparently has a small but significant association with the age beginning sexual activity in adolescence, the direction of that association is related with the culture of reference of the participants in the study (Anglo-Saxon vs. Latin).

– Estimating the effect size or degree of association between variables, for example for studying its evolution or for setting quantitative goals of applied interventions. Thus, Dowden and Brown (2002) did a meta-analysis for estimating the predictive capacity of different categories of substance abuse for recidivism in delinquency. They concluded that the highest capacity was for a combination of drugs abuse and/or alcohol, then only drugs abuse, parental drugs abuse, and alcohol abuse, in that order.

– Giving arguments or evidences to shed light on debates between alternative positions. In those cases there are usually results that support each position, but there are not studies that directly address differential predictions from both positions. For example, there are inconsistent results for the issue of whether there is impairment in the ability for doing two tasks at the same time in the elderly. Riby, Perfect, and Stollery (2004) have proposed to classify the tasks according to the fact that they convey controlled processing or have a motor component. According to those authors the impairment appears only in those tasks, but not in more simple tasks or in those relying in automatic processing. Nevertheless, it should not be forgotten that even when the goals are clearly specified, the meta-analysis is always exploratory in some degree; the analysis of the relationship between variables for which the authors do not have any hypothesis is almost routine.

– Updating a meta-analysis already published, given that since then a significant number of new studies have been published; that is the case, for example, of the meta-analysis of Gil-Gómez de Liaño and Botella (2003) about the changes in the "negative priming" effect associated with age.

– Reformulating a meta-analysis. Sometimes an author considers that the way a given meta-analysis has been carried out is not correct or let out important aspects of the problem. It is proposed an alternative way and is applied to the same empirical base to show how the estimates and the conclusions change. Thus, Del Vecchio and O'Leary (2003) maintain that previous meta-analysis about the effects of anger treatments did not exert good controls, particularly in the inclusion and exclusion criteria, besides the fact that no distinction was made between several features considered important for this problem; they justify their meta-analysis in the necessity of improving the previous.

– Developing a statement, separating it in more specific statements related with more detailed aspects. For example, although the relationship between self-focused attention and negative affect is well known, Mor and Winquist (2002) show in their meta-analysis that this relationship must be re-stated as a function of gender, whether it refers to clinical populations, whether self-attention is focused on private or social aspects, and several diagnostic issues.

For the question about the choice of meta-analysis as the methodology to be applied, we find few circumstances in which other methodologies for review and synthesis should be preferred. In fact, Cooper and Rosenthal (1980) found that the degree of agreement between reviewers was larger when using meta-analysis than when doing a
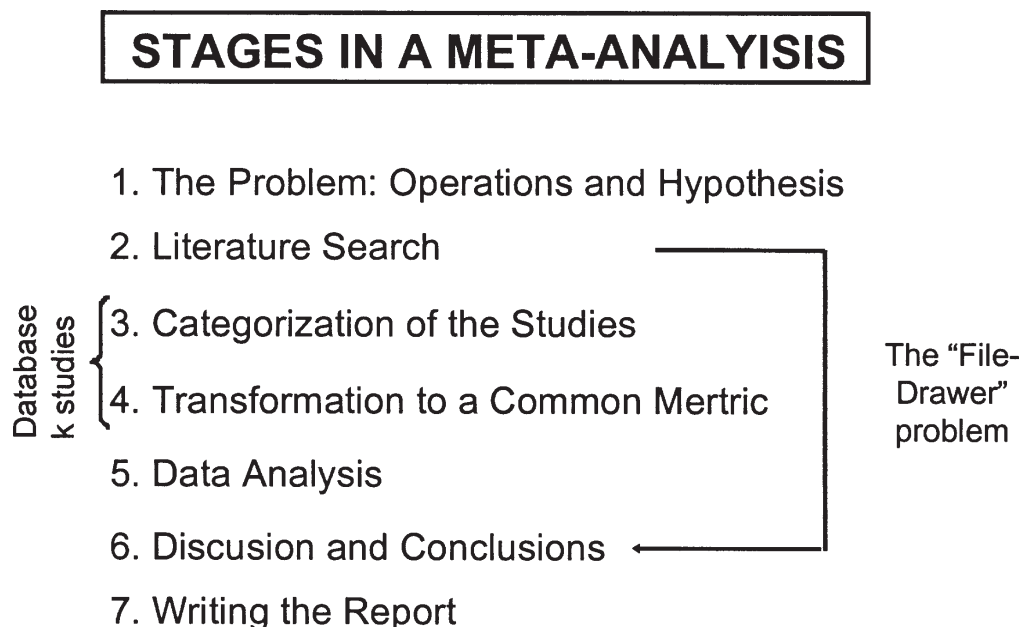
narrative review. But even when the own nature of the problem faced advise to keep some margin for narrative integration, this do not implies to replace meta-analysis, but to complement it. But then, when the meta-analysis is not the appropriate methodology? First, it is obvious that if the studies to be integrated have been done from a qualitative perspective it is impossible to use it. In other cases what is intended is to shed light on a debate related with the causal direction of the relationship between two variables. Then maybe the best option is not to make a review, but to design a primary research to reach an answer for the question. Specifically, when asking questions about the causal relationships we should not forget that meta-analysis is a correlational, not manipulative, methodology. Even although the primary studies are manipulative in nature, their secondary integration can not be treated as if it is (Matt and Cook, 1994). Anyway, although a new primary research is done to face the question about the causal direction, a contradictory empirical base is always a challenge. Meta-analysis is the ideal methodology to discover the keys that explain that variability. Another situation in which perhaps meta-analysis is not the more appropriate methodology is when primary researches are not enough homogeneous for an integrative review. Of course, there is not an objective criterion to decide that the studies reach the proper level of homogeneity (Cooper, 2003); however, unquestionably this is an argument that deserves to be considered. That situation shows up above all when a research problem is studied from different fields or perspectives, each with their own methodological and conceptual frameworks. It will show up also when the way to understand the problem has changed very much along the years and the empirical base contains results from studies done in those different epochs. Finally, a question frequently asked has to do with the minimum number of studies for doing a meta-analysis. When the goal is only a quantitative integration there is no minimum; two studies is enough, although with a small number of studies the statistical analyses are rather unstable. However, when the goal is to summarize and communicate the "state of the art" and only 3 or 4 studies are available, then it is probably better to make a description and assessment of the studies one by one, indicating the differences and similarities.

## Stages in a meta-analysis

What follows is not a list of stages, as they do not necessarily follow a linear sequence, each beginning after the other finishes. The practice of meta-analysis teaches that often the author realizes that one of the previous decisions should be reviewed and goes back to a previous task. The prescription of a strict sequence would not be realistic. Let us see an example. To integrate the studies involving negative affect we can assume a coding system with 3 categories, according to the type studied: a) mood, b) depression, and c) anxiety. Suppose that after the process of coding is finished it is found that category Anxiety contains 80% of the studies and category Mood is empty. The meta-analyst can then modify this criterion of classification, eliminating category Mood and split category Anxiety in other categories that are also significant for what is being studied (e.g., generalized anxiety, social anxiety, specific phobia, panic, and assessment anxiety).

Having in mind all this, we are going to expose the stages understanding them more as the tasks that have to be done, presented in a logical order, although the real practice will give place to overlappings, coming backs and other alterations of that order (Table 1).

**TABLE 1.** Scheme of the stages (or tasks; see text) in carrying out a meta-analysis.

## STAGES IN A META-ANALYISIS

1. The Problem: Operations and Hypothesis

2. Literature Search

**Database k studies** {

3. Categorization of the Studies

4. Transformation to a Common Mertric

The "File-Drawer" problem

5. Data Analysis

6. Discusion and Conclusions

7. Writing the Report

*Specifying the problem: Operationalism*

As with any other research methodology, the general goal assumed in a meta-analysis is only useful when transformed to one or several specific goals, operationally defined. For example, Mintz, Dobson, and Romney (2003) have published a meta-analysis about the role of insight in schizophrenia. This complex construct is operationally defined according to five dimensions: patient's awareness of mental disorder, awareness of the social consequences of disorder, awareness of the need for treatment, awareness of symptoms, and attribution of symptoms to disorder. Given the variety of ways this construct can be defined along the studies, they decided to accept for the meta-analysis those studies in which it is operationalized according to any of these dimensions. Another positive item for this section is an explicit anticipation of the results that would be expected from the several hypotheses managed. Related with the report, the Introduction section must clarify both the motivation for that meta-analysis and its goals. Then the operationalisms must be specified, so that the reader can assess them and decide if is or not in agreement with them. After reading this part must know what is going to find in the meta-analysis and why is important to read it. The first paragraphs must become an invitation, generating the complicity of the reader with the goals of the author.

It is frequent to present the "state of the art" in a way that there is no doubt about the motivation for doing it and also that the meta-analysis is pertinent. Reviewing the list of the previous section, the motivation could be argued, for example, in any of the

following ways: a) summarize the studies with results favoring different (or even in-compatible) points of views, making obvious the necessity for clarification; b) highlight the lack of combined estimations (global or partial) of the effect size that can be used for the design and assessment of intervention programs; c) indicate the amount of studies published since the last meta-analysis and the need to update it; d) argue why the operations, the selection of studies or, in general, the decisions made in other meta-analysis about the same question are not the more appropriate or useful and an alternative formulation is needed; e) collect and integrate the evidence supporting the efficacy of an intervention.

In short, in this section is where the conceptual and argumentative nucleus must be explained, using above all enumeration, comparison and exposition paragraphs (León, 2005; Ramos-Álvarez and Catena, 2004).

*Procedures for searching and collecting information*

What is intended is to locate all the studies done that contain data relevant for the goal of our meta-analysis, whenever they are enough homogeneous with the other as to be integrated. There are many sources where the several procedures for the search are described (Botella and Gambara, 2002; Cooper and Hedges, 1994). However, the tendency is to rely more and more on the electronic bases, as they are increasingly wide in the amount and the variety of their input sources. If we add the increasing ease of access and the more and more sophistication of the criterion for searching, the result is that today the tendency is to put the main accent on that source. Even so, we should not forget the unpublished studies (e.g., doctoral dissertations, congress abstracts, etc.) in order to minimize the publication bias (Cooper, 2003; Lipsey and Wilson, 2001).

Once the base of studies is located, we will find that not all them can be integrated in the meta-analysis. In some cases because they lack some crucial information, either of the statistical results or of important characteristics of the study. In other cases the problem will be a specific characteristic that is not shared with the others and makes it non-homogeneous for a synthesis. The authors must devote the time necessary to establish well defined inclusion and exclusion criteria.

The report must include the search procedures, specifying the bases employed, the date limits of the searches, the key words, the date of the search and, in general, all that would allow a replication of the search. It is also essential to inform of the inclusion and exclusion criteria of the studies for the meta-analysis. The most usual criteria only need to be named, but the most new ones should be explained and justified. In any case, any not obvious operation must be made explicit in the report.

*Adopting a common metric, specifying an effect size index*

One of the elements that allowed the beginning and development of meta-analysis was the adoption of a strategy that allow transforming to a common metric the results of studies done with different methodologies and/or distinct measures. Although during several years the *p*-value from the statistical test was an alternative, finally the effect size has become the generalized way to compare, analyze and integrate studies. There are many indexes of effect size; each one can be more appropriate for a different

measurement scenario or for some specific goals (Lipsey and Wilson, 2001). For measuring the relationship between one dichotomic or dichotomized variable and one quantitative, the most appropriate index is the standardized mean difference (*d*). For two quantitative variables is more appropriate the Pearson's correlation coefficient (*r*). Between two dichotomic variables it is usual to use the *odds ratio*.

The choice of the index can have important consequences, as they differ in characteristics as bias, consistency and efficiency. We recommend to consult specific reviews in which the advantages and difficulties of each are discussed (e.g., Lipsey and Wilson, 2001; see the special section about this issue in *Psychological Methods*, 8, 4, 2003).

This aspect of meta-analyis is increasingly normalized to the point that is not necessary to justify in the report the effect size index chosen; justification is only necessary for the index when the chosen is not one of the most frequently used and for the way the groups are defined when this is relevant.

*Preparing the database*

Once collected the studies that are going to be part of the meta-analyis and once adopted an effect size index, a detailed exploration of the studies is in order. The objectives are to obtain from each study the information needed to calculate the effect size and code it according to a series of categories previously defined. The main question is, of course, how many and which must be those characteristics. This question is no way innocent. It is only possible to arrive to a given answer if a question that can lead to it is asked. That is, it is only possible to find any relationship between the effect size and a specific characteristic if it is measured and coded. In the set of characteristics we are in fact setting the limits of the explanations that we are willing to manage. Of course, it must include all the characteristics suspicious of having something to do with what we are studying. But if we only study those characteristics it is possible to let out the variable with the largest covariation with the effect size.

The characteristics quantified and coded are usually classified in three groups: substantive, methodological and extrinsic (Lipsey, 1994). Once set the characteristics, the reports must be studied and coded one by one. The best way to understand this process is that proposed by Cordray (1990). This author suggests behaving as whether we are carrying out a survey to which the answers are given by research reports, instead of persons. For each question the researcher looks in the report for what is needed to give an unequivocal answer. Ideally, the process of coding and classification must be done independently by two or more coders, quantifying the interrater reliability.

With the results of this process the database will be ready for the analyses. In general, this database will have one study or estimate of the effect size on each raw and three groups of columns. The first group is a single column identifying the study. The second group is formed by the variables with which the characteristics of the study are coded, including the three types above mentioned. In the third one, they are included the estimates of the effect size and the sample sizes. The heading of the database will look as that of Table 2.

**TABLE 2.** Example for the heading of the database for a meta-analysis.

| STUDY | CHARACTERISTICS | | | | | | | RESULTS | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | SUBSTANTIVE | | | METHODOLOGICAL | | | EXTRINSIC | $N_1$ | $N_2$ | ES | p |
| | Type of treatment | Diagnostic | … | Design type | Control group | … | Year of public. | ... | | | |
| 1 | | | | | | | | | | | |
| 2 | | | | | | | | | | | |
| 3 | | | | | | | | | | | |
| 4 | | | | | | | | | | | |
| 5 | | | | | | | | | | | |
| 6 | | | | | | | | | | | |
| … | | | | | | | | | | | |
| k | | | | | | | | | | | |

Let us see what should appear in the report related with this stage. If the editor does not impede it arguing a problem of space, it is good to include the database in the report. If it is too long is better shortening that taking it off, including at least the first and third part and excluding from the second one some of the characteristics coded. Furthermore, the author must put at the disposal of the scientific community the so-called "code book"; this shows how the characteristics of the studies have been classified in the database. A footnote must make available the code-book for anyone. Transparency is essential in this stage of a meta-analysis. Do not forget that from here the meta-analysis will be done on this database. What is not included on it will keep definitively out, and what is included will be managed in the way it has been coded and quantified. The degree of agreement reached between the coders must also appear in the report.

*Data analysis*

This stage includes two different parts. The first one has to do with the information related with the characteristics describing the primary studies, the second one with the analysis of the results of the synthesis properly. The information of the characteristics must include the frequencies with which the categories of each main characteristic has been observed (e.g., frequencies of studies in which the participants were internal, external or a mixed sample) or its average (e.g., the average age of the participants or the average interval of time for the follow-up). This information allows making an idea of the "sociology" of the base of studies, its size and composition, with all the biases conveyed on it. The main part of this section will refer to the effect size estimates. Remember the words from Rosenthal (1995) when says that "The heart of a meta-analytic review is a description of the obtained effect sizes" (p. 185). The analyses of the effect size data to be integrated combined and compared are usually organized in three sets. First, the single joint estimate of the effect size, is usually done by weighting the studies by the inverse of the variance of the estimate; as this variance is an inverse function of the effect size, the studies with larger samples have a heavier weight on the global estimate. Second, it is almost routine to study and test the homogeneity of the estimates, for which we have available several statistic and tests (Huedo-Medina, Sánchez-

Meca, Marín-Martínez, and Botella, in press), although the most employed is Hedges and Olkin's $Q$ (1985). This test allows making a decision as to whether the observed results are better explained by a fixed effects or a random effects model (Cooper and Hedges, 1994; Hedges and Olkin, 1985; Lipsey and Wilson, 2001). Third, the analyses guided to explain the heterogeneity observed in the estimates. Again, the framework more frequently used is, undoubtless, that proposed by Hedges and Olkin (1985). However, the field of psychometrics has their own traditions. In integrating validity coefficients the disatenuation techniques proposed by Hunter and Schmidt (2004) are the most popular. For integrating reliability coefficients there are yet few examples in the primary literature (Thompson, 2003).

Related with the report, we are again in a stage with enough normalization as not to need to give too many details about how things have been done. It will be only necessary an explanation of the decisions or procedures that are less usual. Related with the characteristics of the studies, these can be summarized in a single table showing the frequencies of the categories coded, the averages of quantitative characteristics (as the average size of the samples), etc. Sometimes this information is not included in the Results section, but in the Method section.

What are essential for this section are the results of the combination and comparaison processes. First, the global estimate of the effect size including the standard error. The most usual and informative format is that of a confidence interval. Second, the results of the test for homogeneity. From this result several things will derive. On one side, the basis to assume a fixed effects or a random effects model. On the other side, the need to make partial estimations of the effect size. Third, most part of this section is usually devoted to fit models designed to explain the variability observed in the effect size estimates. Here, the effect size becomes the dependent variable and the studies characteristics become the explanatory factors. For each one, it must be conveyed information related with the statistical decision, together with the new tests for homogeneity within the category.

Given that the meta-analysis has always something of exploratory, it is good to display the observed effect size in a descriptive format that allows to the reader to make their own conjectures. Steam-and-leaf diagrams, Box-and-Whiskers graphics, Forrest Plot figures and, in general, the tools developed from the exploratory framework (Tukey, 1977) are ideal for this function.

*Discussing and reaching conclusions*

As with any other evidence-based methodology, the Discussion stage is that devoted to assess the hypotheses taking as a basis the results. It is also the right place to assess the degree in which the goals have been overtaken. They are also interpreted other unexpected results, whenever they are relevant for a better understanding of what is under study. This is also the moment to assess the degree in which the conclusions are threatened by the "file-drawer problem" (Rosenthal, 1979). As is well known, in accepting for publication research reports there is a bias against the statistically non-significant results. This bias has as a consequence an overestimation of the joint estimate of the effect size. Many sources recommend procedures to assess the degree in which a meta-

analysis is threatened by this problem (see in Botella and Gambara, 2002). In the report of a meta-analysis this section is crucial, as many readers will make a partial or quick reading, skipping directly from the Introduction, or even from the Abstract, through the Discussion.

Those ideas related with the characteristics of the studies and those specifically related with the integration of the results must be presented separately. The inclusion of the first in this section will allow making a more balanced assessment of the conclusions, as the reader knows the potential biases and pitfalls. Related with the second, we echo León (2005) in that in scientific texts a certain degree of redundancy both in the ideas and in the terms used, instead of synonymous, not only is not a fault but a virtue that improves communication. That is why we follow his recommendation of employing in this section the same terms used in the Introduction to state the problem and formulate the hypotheses. This helps to make obvious the logic relationship between the points of start and arrival. It is also the appropriate place to point future developments, applications or, even, shortcomings detected in the field of study. At the end of the Discussion section or in a specific one for Conclusions, the bottom line of the study should be summarized again.

## Writing the report

When describing each stage we have already pointed out how should be reflected on the report. However, it is worth to comment some more general aspects, and a few principles that must guide the decisions of what to include on it. The Manual of the American Psychological Association (2001) for publishing scientific papers in psychology assumes a structure based on several sections with homogeneous headings and contents (see a very recommendable summary in León, 2005 or the articles published in *International Journal of Clinical and Health Psychology* for by Ramos-Álvarez and Catena, 2004 or Carretero-Dios and Pérez, 2005). In writing a report for a journal that assumes those criteria what is most suitable is an adaptation of that structure to the contents of a meta-analytic report. In the Method section will not appear headings as "Participants", "Apparatus and Stimuli", or "Procedure", but others as "Literature Search", "Inclusion and Exclusion Criteria", "Coding of Characteristics" or "Data Analysis". On the contrary, the own section "Introduction" (if it appears with that label), besides those of "Methods", "Results" and "Discussion", must be kept with the same contents as in the reports of other types of researches. Even realizing we are redundant, we reiterate the importance of trying to include in the report the database and the sources where the primary studies that form part of the meta-analysis can be located. The editors try to save space eliminating just this. The authors must try to negotiate a format for preserving it, even dropping some characteristics to make it thinner. The sources of the studies have the specific solution of marking them with asterisks in the references (American Psychological Association, 2001).

The more important aspects for assessing the quality of a meta-analytic report are the principles of transparency in making decisions and procedures, the degree of replicability and the clarity of the message. When doubting about any aspect of the

report we recommend ask the question of where the application of these principles will take us. The report must include enough information about how the critical choices have been made when carrying out the meta-analysis. When there are problems of space at least the reader must be informed of how can reach that information (e.g., the code book). But besides being transparent, the reader must gain an opportunity to replicate the meta-analysis. If the report does not have enough information for replication, then it is not good enough.

Another aspect that generates confidence in the validity of the conclusions of a meta-analysis, first in the reviewer and then in the reader, is to confirm in the report that the authors have given the importance due to the control of the sources that threat validity. It will be a positive aspect to inform of the caution measures taken for the threats that are better known. Between them we highlight the problem of the independence between the estimates contained in the database, the assessment of the interrater agreement or the specific threat of the "file-drawer problem" (Rosenthal, 1979).

Finally, a few lines about the Title and the Abstract. It is a good practice that the title includes, besides a few words indicating the issue of the basic question, the own term "Meta-Analysis", as this makes much easier its location and screening in electronic searches (without forgetting that sometimes other terms are used to refer to the same methodology, as for example "Research Synthesis"). The abstract must include at least four items: a sentence stating the problem, the number of studies located and integrated, the joint estimate of the effect size and the main conclusion.

## Some practical advises

In any meta-analysis must participate an expert on the substantive field of inquire under study. When carrying out a meta-analysis it is necessary to make many critical choices for which the better basis is a deep knowledge of the field from inside. However, it is also advisable to work together with someone with experience in research methodology and statistical analysis. The procedures are not sophisticated, but it is frequent to find in the primary studies a large variety of situations, that require different solutions. If the expert in the field has a good background in applied statistics, this can be enough. However, many authors without experience in meta-analysis will feel more comfortable working with a methodologist or at least someone that supervise the most sensitive decisions.

In many occasions there are doubts as to how to do something. In those cases the better advice is "do it in both ways" (Greenhouse and Iyengar, 1994). In many cases will not make any difference, but in others will do. It is good to make these sensitivity analyses for knowing the consequences that doing it in each way have on the conclusions, giving the readers the opportunity of achieving their own criterion. Obviously, these alternatives and their consequences must appear on the report.

It is not necessary to persuade the reader of the hardness of the meta-analytic methodology, neither to teach to use it. It is assumed that is an appropriate methodology and that is not necessary a justification each time it is applied. The sources to learn how to do a meta-analysis are not the own meta-analysis. The same happens with data

analysis. There is no need to explain it again, neither to justify the procedures for analysis. It is enough to make reference to a textbook.

As in the reports of researches done with any other methodology, speculation is allowed, but must be pointed out as such. The central body of a meta-analysis must be the analysis of the results, the assessment of the hypotheses and the discussion of the problems that have arisen. But it is legitimate to give a further step and make some speculation, whenever it is recognized that is what is being done, and it is clearly separated from direct interpretations of the data. A meta-analysis is a place especially prone for a specific type of implicit speculation: the inferences about causal relationships. We should not forget that meta-analysis is not a manipulative methodology. Even when they are synthesized studies that are manipulative in nature, the validity of the inferences related with causal direction is doubtful (see Cooper, 1998 and Matt and Cook, 1994, respectively, as examples of different points of view).

In Table 3 we propose a list of the more relevant issues aforementioned in this paper. Both the author and the reviewer of a meta-analysis can employ it as a checklist; the first to write the report and the second to review it.

**TABLE 3.** Relevant issues for writing and assessing a meta-analytic report.

TITLE AND ABSTRACT

| | Yes | Doubtful | Not |
|---|---|---|---|
| The label "meta-analysis" or another identifying the methodology applied appears on it. | | | |
| The problem under study is precisely summarized. | | | |
| The abstract includes the number of studies integrated, the joint estimate of the ES and the main conclusions. | | | |
| The key words allow recovering the report both by way of the methodology and of the substantive problem studied. | | | |

INTRODUCTION

| | Yes | Doubtful | Not |
|---|---|---|---|
| The motivation for doing this MA is specified and properly argued, with spetial attention to the empirical basis available in the literature. | | | |
| The theoretical and empirical domains of reference are clearly established. | | | |
| The "state of the art" is well sumarized. | | | |
| It is put forward how the questions will be answered. | | | |
| They remain explicit how the constructs and variables are operationalized. | | | |
| The expected results are anticipated. | | | |
| The references are relevant and appropriate. | | | |

METHOD

| | Yes | Doubtful | Not |
|---|---|---|---|
| The section is separated in headings that make easier the reading, such as "Search of Studies", "Inclusion and Exclusion Criteria", "Coding of Characteristics" or "Statistical Analysis". | | | |
| The bibliographical sources of the search are detailed, along as the key words employed and the dates. | | | |
| The procedures used to locate the unpublished studies are reported. | | | |
| The criteria for inclusion and exclusion of the studies are adequately highlighted; those less known are explained and the number of studies excluded and the reasons for exclusion are reported. | | | |
| Any difficulty with the independence of the studies is correctly solved. | | | |
| A list with the main characteristics coded and measured is showed. | | | |
| The code book is available for the reader. | | | |
| The database appears in an acceptable format. | | | |
| The procedures for quantifying the effect size are made explicit. | | | |
| There is information about the procedures employed to improve reliability in the process of coding, including the value of the index of interrater agreement obtained. | | | |
| The techniques used for Data Analysis are correct. | | | |

RESULTS

| | Yes | Doubtful | Not |
|---|---|---|---|
| Tables and/or graphics with appropriate numeric summaries of the characteristics of the studies are offered. | | | |
| The obtained effect size's are summarized in descriptive/exploratory formats. | | | |
| The problem of the outliers is adequately faced and sensitivity analyses are made. | | | |
| It is distinguished clearly between the joint estimate of the effect size and other analysis. Confidence Intervals for estimates are offered. When groups of studies are classified acording to a given criterion, the criterion is relevant and partial estimates of ES for the groups are also offered. | | | |
| They are indicated the weighting procedure, the number of point estimates of the effect size integrated and the total number of participants in the global estimation. | | | |
| There is enough information of the homogeneity tests done. | | | |
| It is studied how the several models designed from the characteristics that are relevant for the theoretical analysis managed fit the data; there is information of the values of the statistical tests and the degrees of freedom. | | | |

DISCUSSION

| | Yes | Doubtful | Not |
|---|---|---|---|
| They are offered appropriate interpretations and assessments of the initial positions, closely linked to the results. | | | |
| The authors do not go further from where the results point. | | | |
| The "file-drawer" problem is worked in some way, offering a quantification of the degree in which threats the conclusions. | | | |
| The implications of the results are discussed, suggesting future developments both for research and applications. | | | |

GENERAL CRITERIA OF QUALITY OF THE REPORT

| | Yes | Doubtful | Not |
|---|---|---|---|
| In general, the procedures and criteria to make decisions are clear. | | | |
| The report allows an easy replication. | | | |
| The studies accepted are individually identified. | | | |
| The report fits to the structure assumed by the journal (e.g., APA norms). | | | |

## Conclusions

Meta-analysis is a complex methodology, composed by multiple tasks in which many small choices can determine the results and the conclusions. We do not recommend make the decisions merely by inertia, because "that is the way they are always done". Do everything by a good reason, take appropriate measures to give your decisions the maximum transparency and be prepared to defend them; perhaps will have to do with the reviewers. The same must be said about the report, for which replicability is it's maximum value. The final message of a meta-analysis must be clear. Even if the empirical base available is confusing, the message must be clear and balanced for that lack of consistency. Transparency and replicability will favor the process of identification between the author and the reader. They facilitate the complicity in the process and, thus, increase the receptivity for the messages derived from the study.

## References

American Psychological Association (2001). *Publication manual of the American Psychological Association* (5th ed.). Washington, DC. Author.

Botella, J. and Gambara, H. (2002). *¿Qué es el meta-análisis?* Madrid: Biblioteca Nueva.

Carretero-Dios, H. and Pérez, C. (2005). Normas para el desarrollo y revisión de estudios instrumentales. *International Journal of Clinical and Health Psychology, 5,* 521-551.

Chalmers, I., Hedges, L. V., and Cooper, H. (2002). A brief history of research synthesis. *Evaluations and the Health Professions, 25,* 12-37.

Chambless, D. L., Baker, M., Baucom, D. H., Beutler, L. E., Calhoun, K. S., *et al.* (1998). Update on empirically validated therapies, II. *The Clinical Psychologist, 51,* 3-16.

Cooper, H. (1998). *Synthesyzing research: A guide for literature reviews* (3ª ed.). Beverly Hills, CA: Sage.

Cooper, H. (2003). Editorial. *Psychological Bulletin, 129,* 3-9.

Cooper, H. and Hedges, L. V. (1994). *The handbook of research synthesis*. New York: Russell Sage Foundation.

Cooper, H. M. and Rosenthal, R. (1980). Statistical versus traditional procedures for summarizing research findings. *Psychological Bulletin, 87,* 442-449.

Cordray, D. S. (1990). Strengthening causal interpretations of non-experimental data: The role of meta-analysis. In L. Sechrest, E. Perrin, and J. Bunker (Eds.), *Research methodology: Strengthening causal interpretations of non-experimental data* (pp. 151-172). Washington, DC: U. S. Department of Health and Human Services, Agency for Health Care Policy and Research.

Del Vecchio, T. and O'Leary, K. D. (2003). Effectiveness of anger treatments for specific anger problems: A meta-analytic review. *Clinical Psychology Review, 24,* 15-34.

Dobson, K. S. (1989). A meta-analysis of the efficacy of cognitive therapy for depression. *Journal of Consulting and Clinical Psychology*, *57,* 414-419.

Dowden, C. and Brown, S. L. (2002). The role of substance abuse factors in predicting recidivism: A meta-analysis. *Psychology, Crime & Law, 8,* 243-264.

Gil-Gomez de Liaño, B. and Botella, J. (2003). Diferencias en el efecto de *priming* negativo en la edad: un meta-análisis. *Revista Española de Geriatría y Gerontología*, *38,* 161-169.

Greenhouse, J. B. and Iyenger, S. (1994). Sensitivity analysis and diagnostics. In H. Cooper and L. V. Hedges. *The handbook of research synthesis*. New York: Russell Sage Foundation.

Halvorsen, K. T. (1994). The reporting format. In H. Cooper and L. V. Hedges (Eds.), *The handbook of research synthesis*. New York: Russell Sage Foundation.

Hedges, L. V. and Olkin, I. (1985). *Statistical methods of meta-analysis*. Orlando: Academic Press.

Huedo-Medina, T.N., Sánchez-Meca, J., Marín-Martínez, F., and Botella, J. (in press). Assessing heterogeneity in meta-analysis: $Q$ statistic or $I^2$ index? *Psychological Methods*.

Hunter, J. E. and Schmidt, F. L. (2004*). Methods of meta-analysis: Correcting error and bias in research findings* (2nd ed.). Newbury Park, CA: Sage.

León, O. G. (2005). *Cómo redactar textos científicos en psicología y educación*. Coruña: Netbiblo.

Lipsey, M. W. (1994). Identifying potentially interesting variables and analysis opportunities. In H. Cooper and L. V. Hedges (Eds.), *The handbook of research synthesis*. New York: Russell Sage Foundation.

Lipsey, M. W. and Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.

Matt, G. E. and Cook, T. D. (1994). Threats to the validity of research syntheses. In H. Cooper and L. V. Hedges (Eds.), *The handbook of research synthesis*. New York: Russell Sage Foundation.

Mintz, A. R., Dobson, K. S., and Romney, D. M. (2003). Insight in schizophrenia: A meta-analysis. *Schizophrenia Research, 61*, 75-88.

Montero, I. and León, O.G. (2005). Sistema de clasificación del método en los informes de investigación en Psicología. *International Journal of Clinical and Health Psychology, 5*, 115-127.

Mor, N. and Winquist, J. (2002). Self-Focused attention and negative affect: A meta-analysis. *Psychological Bulletin, 128*, 638-662.

Quorum Group (1999). Improving the quality of reports of meta-analyses of randomised controlled trials: The Quorum statement. *The Lancet, 354,* 1896-1900.

Ramos-Álvarez, M.M. and Catena, A. (2004). Normas para la elaboración y revisión de artículos originales experimentales en Ciencias del Comportamiento. *International Journal of Clinical and Health Psychology, 4*, 173-189.

Riby, L. M., Perfect, T. J., and Stollery, B. T. (2004). The effects of age and task domain on dual task performance: A meta-analysis. *European Journal of Cognitive Psychology, 16*, 863-891.

Rosenthal, R. (1979). The "file-drawer problem" and tolerance for null results. *Psychological Bulletin, 86*, 638-641.

Rosenthal, R. (1995). Writing meta-analytic reviews. *Psychological Bulletin, 18,* 183-192.

Thompson, B. (2003). *Score reliability. Contemporary thinking on reliability issues*. Thousand Oaks, CA: Sage.

Tukey, J. W. (1977). *Exploring data analysis*. Reading, MA: Addison-Wesley.

Vargas, E., Gambara, H., and Botella, J. (in press). Autoestima e inicio de actividad sexual en la adolescencia: un estudio meta-analítico. *International Journal of Clinical and Health Psychology*.

Wilkinson, L. and Task Force on Statistical Inference (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist*, *54*, 594-604.