

SYSTEMATIC REVIEWS AND META-ANALYSES: TOOLS FOR PROFESSIONAL PRACTICE*

Julio Sánchez-Meca* and Juan Botella**

*University of Murcia and **Autónoma University of Madrid

Traditionally there has been a degree of separation between professional practice and psychological research. However, over the last two decades, Evidence-Based Psychology (EBP) has emerged as a very useful methodological tool for linking the two, the aim of EBP being to base professional psychological practice on the best evidence obtained from research. Systematic reviews (SRs) and meta-analyses (MAs) are considered the best tools for synthesizing the scientific evidence as to which treatments, interventions, or prevention programmes should be applied for a given psychological problem. Thus, psychologists need to know what SRs and MAs are, how they are carried out and, most importantly, how SRs and MAs can be critically appraised. The purpose of this article is to present SRs and MAs, together with some guidelines on how to make a critical reading of them.

Key words: Evidence-based psychology, Systematic reviews, Meta-analysis, Intervention evaluation.

La práctica profesional y la investigación psicológica han estado demasiado separadas. Sin embargo, con la llegada de la Psicología Basada en la Evidencia (PBE) esta herramienta metodológica se ha convertido en el mejor modo de unir las mejores pruebas con la práctica psicológica. La PBE preconiza que la práctica profesional esté basada en las mejores pruebas obtenidas desde la investigación psicológica. Las revisiones sistemáticas (RSs) y los meta-análisis (MAs) se consideran actualmente como las mejores herramientas para sintetizar las pruebas científicas respecto a qué tratamientos, intervenciones o programas de prevención deberían aplicarse para un determinado problema psicológico. Así pues, los psicólogos tienen que saber qué son las RSs y los MAs, cómo se hacen y, lo que es más importante, cómo podemos hacer valoraciones críticas de ellos. El propósito de este artículo es presentar las RSs y los MAs, así como alguna guía orientativa sobre cómo hacer una lectura crítica de ellos.

Palabras clave: Psicología basada en la evidencia, Revisiones sistemáticas, Meta-análisis, Evaluación de intervenciones.

There is no doubt that the professional exercise of psychology, in whatever field, must be based on the best scientific evidence and proof. However, the techniques applied in professional practice at a given moment are out of step with the achievements of research up to that same point in time. This discrepancy has two main causes. On the one hand, there is no strong connection between the world of professional practice and that of research, which is carried out largely in universities. On the other hand, until around twenty years ago, the behavioural sciences were characterized by poor compilation of research data, which meant that scientific advances were slow to be implemented in the

routine practical context. These factors have led to psychology professionals seeing the world of research as something divorced from their usual practice, and without utility in the form of results that can be applied rapidly and directly to their everyday activity.

If this is what happens in the case of psychologists themselves, the picture is worse still in that of professionals from other disciplines or politicians and other decision-makers who request the expert opinion of psychologists. As we have already pointed out elsewhere (Botella & Gambara, 2006a), the problem is that the data generated by psychology are often confused and contradictory. Recently, we have attended debates on highly ideologically charged issues, such as adoption by homosexual couples, in which the politicians defending each position were accompanied by psychologists, supposedly experts, each saying the opposite to the other, but presenting it as though it were the position that emerged from the evidence gathered by psychology.

This situation, disheartening as it may be, is changing thanks to two methodological advances that address this

Correspondence: Julio Sánchez-Meca. Dpto Psicología Básica y Metodología. Facultad de Psicología. Campus de Espinardo. 30100-Murcia. España. E-mail: jsmecca@um.es
www.um.es/metaanalysis

.....
(*) Este artículo ha sido financiado al primer firmante por el Fondo de Investigación Sanitaria, convocatoria de Evaluación de Tecnologías Sanitarias (Proyecto No: PI07/90384) y al segundo firmante por el Ministerio de Educación y Ciencia (Proyecto No: SEJ2006-12546/PSIC).

lack of fit between professional practice and research: the Evidence-Based Psychology (EBP) approach, on the one hand, and systematic reviews and meta-analyses, on the other.

The EBP approach constitutes a methodological tool intended to modify the way psychology professionals work so that they take into consideration in their everyday decisions the best scientific evidence or proof in relation to a given problem. The problem could involve deciding which treatment technique to use with a patient suffering from a particular psychological disorder; which intervention programme is the most appropriate for preventing certain maladaptive behaviours; which is the best diagnostic method for a psychological disorder; and so on. Once the problem has been adequately formulated, the EBP approach consists in carrying out a search for the evidence or proof supporting the best course of action. This type of information search requires the use of new information and communication technologies, and especially Internet resources. Once the scientific proof has been located, in the specialist psychology publications, the next step in the EBP approach is to make a critical analysis of that proof,

which requires psychology professionals to put into practice their knowledge and expertise on research methods, designs, data analysis and measurement instruments. And in a final phase, EBP involves applying its findings to professional practice.

And what are the best scientific findings or evidence that can endorse the routine application of a treatment, a prevention programme or an assessment or diagnostic technique? It is widely accepted within the social and health sciences community that the most reliable scientific evidence is that provided by primary studies based on randomized clinical trials (RCTs), which involve the random assignment of participants to the experimental conditions (Nezu & Nezu, 2008).

Nevertheless, it is common, on trying to select the evidence about a particular problem, to find numerous empirical studies, all of which have dealt with this same issue. This accumulation of information may impede the putting into practice of the EBP approach, on making it unviable for professionals with heavy workloads to select the relevant studies and to make a critical appraisal of each one within a sufficiently short time period. And this is where systematic reviews (SRs) and meta-analyses (MAs) come into play. As a way of overcoming the problems deriving from the poor compilation of data in the social sciences, SRs and MAs constitute a research methodology whose objective is to collate in a systematic and objective fashion the evidence obtained in empirical studies on a single problem. Thus, reading an SR or an MA about the problem in question permits professionals to save time and offers them an overview of the scientific evidence in relation to that problem. Moreover, and as summarized in Table 1, the scientific community has accepted MAs as the methodology that can provide the best proof or evidence about a problem, when the accumulated empirical research consists of experimental studies (or RCTs).

Papeles del Psicólogo has covered the EBP approach previously in the form of the excellent articles by Frías Navarro and Pascual Llobell (2003) and by Pascual Llobell, Frías Navarro and Monterde (2004). There are also other fine presentations of the Evidence-Based Practice approach in the Spanish language (Grupo de Atención Sanitaria Basada en la Evidencia, 2007; Navarro, Giribet & Aguinaga, 1999; Vázquez & Nieto,

TABLE 1
QUALITY LEVELS OF THE SCIENTIFIC EVIDENCE OR PROOF

Level	Type of proof
1	(a) Meta-analysis (homogeneous) ¹ of RCTs (b) An RCT (with narrow CI) (c) All or none ²
2	(a) Meta-analysis (homogeneous) ¹ of cohort studies (b) A cohort study (including a low-quality RCT) (c) Study "of results" ³ ; ecological studies
3	(a) Meta-analysis (homogeneous) ¹ of case-control studies (b) A case-control study
4	(a) Series of cases (and cohort studies and case-control studies, poor quality)
5	(a) Experts' opinions without explicit critical assessment, or based on physiology or "first principles" ⁴ , or research based on the "authorities" criterion

¹The results of the RCTs are homogeneous.
²This is fulfilled when all patients used to die before the treatment was available, but now some survive on being treated; or when some patients used to die before the treatment was available, but now none die on being treated.
³Studies a cohort of patients with the same diagnosis and relates its clinical results to the care they have received.
⁴The pathophysiological principles used for determining the clinical practice.

¹ The reader can consult the respective websites of the Cochrane Collaboration (www.cochrane.org) and the Campbell Collaboration (www.campbellcollaboration.org), or indeed the *Centro Cochrane Iberoamericano* site based in Barcelona (www.cochrane.es).



2003). Therefore, this article focuses on the other methodological advance which, in our view, has become a key element in the quest to connect research with professional practice: meta-analysis. It is our understanding that psychology professionals should be familiar with this methodology, since the critical reading of meta-analytic studies can be of great use to them in their everyday decision-making about which treatments or diagnostic techniques to apply. At the same time, the reading of meta-analytic studies facilitates implementation of the EBP approach on offering in an integrated way the best scientific evidence or proof about a given problem, with the consequent time savings for the professional. In any case, the proliferation of meta-analytic studies in the field of psychology guarantees that, sooner or later, all psychology professionals will have to face up to the task of critically reading studies of this type, and consequently familiarizing themselves with this methodology which is now emerging as essential.

In the following pages we shall look at what MAs and SRs are, we shall consider the stages involved in carrying out an MA, we shall illustrate this methodology with a real example and we shall present a guide to aid the critical reading of meta-analytic studies. We shall conclude with some final reflections and some suggested reading. For a fuller treatment of SRs and MAs we refer the reader to Borenstein, Hedges, Higgins and Rothstein, 2009; Botella and Gambará, 2002; Cooper, 2010; Cooper, Hedges and Valentine, 2009; Littell, Corcoran and Pillai, 2008; Marín Martínez, Sánchez Meca, Huedo and Fernández, 2007; Marín Martínez, Sánchez Meca and López López, 2009; Petticrew and Roberts, 2006; Sánchez Meca, 1999, 2003, 2008; Sánchez Meca and Ato, 1989; and Sánchez Meca and Marín Martínez (in press).

SYSTEMATIC REVIEWS AND META-ANALYSES

An SR is a review of a question that is clearly formulated, and which uses systematic and explicit methods for the identification, selection and critical appraisal of studies relevant to that question, as well as the collation and analysis of data from the studies included in the review (Martín, Tobías & Seoane, 2006). SRs emerge as an attempt to remedy the limitations of traditional reviews, characterized by their qualitative nature and lack of adequate systematization. Those of us who stress the advantages of SRs versus traditional reviews base our arguments on the premise that the process of the review of scientific literature on any matter should be subject to

the same criteria of scientific rigour as empirical research: objectivity, systematization and replicability of results. That is, the review of empirical studies on a given question is just as much a scientific task as the carrying out of an empirical study itself.

If in an SR we are capable of quantifying, by means of some statistical index of effect size, the results of each empirical study included and of applying statistical analysis techniques to extract the essence of those studies, then the SR becomes a meta-analysis (MA). An MA is, then, an SR in which statistical methods are used for analyzing the results of the studies included in it (Littell et al., 2008). From this it follows that all MAs are SRs, but not all SRs are necessarily MAs. There are qualitative SRs in which statistical methods are not applied to the study results, which are instead subjected to qualitative assessment.

Given their greater level of quantification and rigour, within SRs, MAs are the types of review that provide the most valid evidence about a problem (Cooper, 2010). It is for this reason that the present article focuses specifically on how an MA is carried out and interpreted. Moreover, while a great diversity of problems can be studied by means of MA, the most useful type for psychology professionals is probably that which sets out to examine the efficacy of different treatments or intervention/prevention programmes for psychological, psychosocial or behaviour disorders. We shall therefore focus on this type of MA.

What can an MA offer us? On applying statistical techniques for integrating the results of a set of empirical studies about the efficacy of treatments or intervention programmes, an MA permits us to answer questions such as: (a) what is the magnitude of the global effect of the different treatments?; (b) are the efficacy results of the different treatments homogeneous?; (c) in the case that they are not homogeneous, what factors can explain this heterogeneity of results?; and (d) is it possible to formulate an explanatory model capable of accounting for such heterogeneity in the results? To respond to these questions, an MA involves following the typical steps of an SR and applying statistical techniques of integration.

Where can we find MAs? In practically any psychology journal it is possible to find some meta-analytic study on an issue of potential interest to professional practitioners. Specifically, journals that regularly publish these types of study include *Psychological Bulletin*, *Clinical Psychology Review*, *Journal of Applied Psychology* or *Journal of*



Consulting and Clinical Psychology. As far as Spanish-language publications are concerned, they can be found in journals such as *Psicothema* or the *International Journal of Clinical and Health Psychology*, both of which are bilingual. Moreover, meta-analytic studies can be easily located via the Google Scholar search engine. Finally, special mention should be made of the Cochrane Collaboration and the Campbell Collaboration, two international organizations whose aim is to promote high-quality meta-analyses on the efficacy of interventions in different fields related to psychologists' professional activity. Thus, the Cochrane Collaboration's website contains numerous SRs and MAs in the field of clinical psychology, whilst the Campbell Collaboration focuses on these types of analysis in the areas of Education, Social Services and Criminology (Sánchez Meca, Boruch, Petrosino & Rosa Alcázar, 2002)¹.

PHASES OF A META-ANALYSIS

Carrying out an MA involves following the same stages as in any empirical study, even if some of them have certain peculiarities that should be clarified. Basically, an MA can be considered to comprise five phases or stages:

- (1) Formulation of the problem
- (2) Selection of the studies
- (3) Coding of the studies
- (4) Statistical analysis and interpretation
- (5) Publication

(1) *Formulation of the problem.* The first stage consists in formulating clearly and objectively the question to which we want to respond. This implies defining theoretically and operationally the psychological constructs under study. For example, in a meta-analysis on the efficacy of psychological treatments for panic disorder with or without agoraphobia (Sánchez Meca, Rosa Alcázar, Marín Martínez & Gómez Conesa, 2010), key concepts were defined, such as the nature of the psychological treatments under study and of panic disorder with or without agoraphobia, and which measurements of efficacy results would be admitted in the MA.

(2) *Selection of the studies.* The next step consists in defining the selection criteria for the studies. It should be borne in mind that carrying out an MA involves selecting

empirical studies with certain similar characteristics as regards the research design (e.g., all the studies should include at least a treatment group and a control group, both with pre-test and post-test measures), so that it is possible to apply to all of them a single effect size index that permits their metric comparison. Thus, although the selection criteria will depend on the MA in question, there must necessarily be some specifications in relation to the type of study design admissible, the way the outcome variables were measured, the participant characteristics and the characteristics of the treatments. For example, for inclusion in the MA cited above, the empirical studies had to include at least a treatment group and a control group made up of adults diagnosed with panic disorder with or without agoraphobia (both with pre-test and post-test measures), and the treatment applied had to be purely psychological, with no involvement of psychoactive drugs. Moreover, the studies must have been carried out between 1980 and 2006.

Once the selection criteria for the studies have been fixed, the search for them can be carried out, and for this it is necessary to use electronic databases (e.g., PsycInfo, MedLine, ERIC), consult specialist journals and contact recognized authors on the issue to request from them studies that are difficult to locate. The combination of formal and informal sources in the search process should guarantee maximum comprehensiveness in this process, as well as the location of published and unpublished studies, so that publication bias can be examined. In the MA on panic disorder the electronic databases PsycInfo, Medline, Psycodoc and the Cochrane Library were consulted, in addition to previous MAs, articles, books, review chapters in books and specialist clinical psychology journals; furthermore, recognized authors in the field were contacted, all with the object of locating the highest possible number of empirical studies that met the selection criteria.

(3) *Coding of the studies.* Once all the selected empirical studies have been located and accessed, the next stage consists in recording the characteristics of those studies. To this end, a Coding Manual is drawn up of all the characteristics of the studies that could act as moderators of the efficacy results of the treatments analyzed. On the basis of the Coding Manual, a Coding Form for the

² The results of the MA presented in this article represent only a small sample of all the evidence it provided. Thus, it should be made clear that for each study an effect size was calculated for each different result measure (panic attacks, agoraphobic behaviours, general level of anxiety, of depression, of global adjustment, etc.). In the present article we refer solely to the results obtained with the measures of panic, which are the most relevant for this disorder.

moderator variables is drawn up. Although the study characteristics to be coded will depend on the purpose of each MA, they can be classified in several clusters or categories. Thus, we speak of treatment variables, participant variables, context variables, methodological variables and extrinsic variables.

Treatment variables are those related to the treatment applied in the research. They therefore include the type of treatment itself (e.g., cognitive therapy, in vivo exposure, deep relaxation), the duration of the treatment, its intensity, its form of application (individual versus group-based), and so on.

Participant variables have to do with participants' characteristics. Thus, participant variables would include mean age of the sample, its make-up by gender, social background or seriousness of the disorder, etc.

Context variables refer to the location in which the intervention took place. This could be a hospital, a private clinic, a psychology consulting room, a school, the home, and so on. Also considered as a context variable would be the fact of whether those receiving the treatment are inpatients or outpatients.

Methodological variables are those related to the design and instrumentation of the empirical study. Thus, highly relevant methodological variables in an MA would be type of design (experimental versus quasi-experimental), sample size, experimental mortality, the inclusion of pre-test and post-test measures (or only of post-test), the carrying out of statistical analyses by intention-to-treat or only with those completing the treatment, the use of "blind" evaluators – that is, without knowledge of the treatment patients are receiving –, or the diagnostic criterion used in the study to assess participants. All such characteristics allow us to rate the methodological quality of the studies, and consequently the likelihood of bias in the results that emerge.

Finally, *extrinsic variables* are also usually coded. These variables are so called because they represent characteristics of the studies which, in principle, should not have anything to do with the scientific process of the study, but which nevertheless can affect its results. This category would include variables such as source of the report (published versus unpublished), authors' background (psychologist, psychiatrist, etc.) or year in which the study was carried out.

The purpose of this phase involving the coding of study characteristics is none other than to identify a set of variables that can explain the variability of the efficacy

results of the different studies, a question that is statistically analyzed in the following phase. In the MA on panic disorder, the heterogeneity in the efficacy results displayed by the different empirical studies could be due to the fact that those studies had applied different psychological treatment techniques, of different durations, on patient samples with variable age, gender distribution and seriousness of the disorder, and using a variety of designs and methodological characteristics. The coding of all these variables in the studies is precisely intended to reveal which of them may be related to the results on efficacy.

In this phase it is highly important to check the reliability of the characteristics coding process. Therefore, it is customary for two or more researchers to code independently all or some of the empirical studies and for their degree of agreement to be checked. Only in this way can we know whether the MA has applied objective and systematic norms in the coding process.

(4) *Statistical analysis and interpretation.* In addition to considering the moderator variables of the studies, carrying out an MA requires the calculation of a quantitative index that makes it possible to measure all the study results using the same metric. This is due to the fact that the studies will measure the effects of the treatments with different psychological tests and scales, so that the results are not directly comparable, on being expressed in different units of measurement. This homogenization of the results is achieved through the application of some effect size index. The effect size, then, reflects the degree to which the results of the treatment group participants differ, on average, from those of the control group participants. Although there is great variation in the effect size indices we can find in MAs, the most widely used is the standardized mean difference, defined as the difference between the means of the two groups divided by their pooled standard deviation. On dividing by the standard deviation we obtain a homogeneous and comparable quantitative index regardless of the tests or scales used in the different studies, since they can be interpreted as standard units of separation between the means of the two groups. Apart from the standardized mean difference, it is also common to find in MAs effect size indices for dichotomized outcome variables, such as difference of proportions, relative risk or odds ratio, with one index being able to be transformed into another (Sánchez Meca, Marín Martínez & Chacón Moscoso, 2003).

Having recorded for each study its characteristics (moderator variables) and its effect size, the resulting database can be subjected to statistical analyses that allow us to respond to the key questions faced by an MA: (a) what is the magnitude of the mean effect of all the studies?; (b) are the effect sizes of the studies homogeneous?; (c) in the case that they are not homogeneous, what study characteristics might account for this heterogeneity?; and (d) is it possible to formulate an explanatory model of this heterogeneity of effect sizes based on a sub-set of the moderator variables coded?

In order to respond to these questions, statistical analysis techniques are applied in which the weight of each study in the meta-analytic computations depends on the precision displayed by its effect size, and the precision depends on the sample size: the larger the sample size, the greater the precision, and hence, the greater the weight in the analyses. In this way a weighted mean of the effect sizes is calculated, together with the confidence interval; the degree of heterogeneity of the effect sizes is assessed, and if the effect sizes are not homogeneous, the influence of the moderator variables on the effect sizes is analyzed. This last phase of the analysis is carried out through the application of weighted procedures based on the analysis of variance (or analysis by subgroups) and on the regression models (meta-regression), so that the dependent variable is constituted by the effect sizes obtained in the studies, whilst the independent or predictor variables are the characteristics of the studies.

(5) *Publication*. The final phase of an MA, as with any other type of study, consists in disseminating its results. The publication of an MA is governed by the same norms as any other empirical study (Botella & Gambara, 2006b). Thus, the sections of a meta-analytic study are customarily: Introduction, Method, Results, and Discussion and Conclusions.

In the introduction the issue under study is reviewed, the relevant psychological constructs are defined and the objectives of the MA are formulated. The method section includes different subsections. First of all, the subsection 'search for studies' has the aim of specifying the selection criteria for the studies and the literature search procedures used. Secondly, there is a subsection outlining the coding process, in which the characteristics of the studies coded are also described. The third and final subsection, usually entitled 'statistical analysis', serves to define the effect size index used in the MA, as well as the statistical integration techniques applied. The object of the

method section is to permit the MA to be replicated by other researchers, so that it should make as explicit as possible all the decisions taken during the MA process.

The results section presents the results of the statistical analyses applied in the MA, attempting to respond to the four basic questions posed by an MA and listed above; in the discussion section the results of the MA are considered within the context of the previous literature on the issue, their practical relevance and implications for professional practice are discussed, and future research lines are suggested.

AN EXAMPLE

To continue with the example referred to above on the efficacy of psychological treatments for panic disorder with or without agoraphobia (Sánchez Meca et al., 2010), in that MA the authors succeeded in selecting 65 studies which met the selection criteria, and in each one of them a standardized mean difference (d) was obtained that compared the mean results achieved by the treatment and control groups in the post-test². The results of the MA are based on a total sample of over 2300 patients with this psychological disorder, which gives an idea of the degree of generalization the MA results can offer us.

To the question "What is the overall degree of efficacy obtained with the whole set of studies?", this MA offered a mean effect size $d_+ = 1.015$, with a statistically significant confidence interval, which took values of between 0.855 and 1.175. The value 1.015 can be interpreted in standard units and, following Cohen's criterion (1988), values of around 0.2, 0.5 and 0.8 can be considered as reflecting a practical significance of low, medium and high magnitude, respectively. Therefore, the value 1.015 implies high efficacy of psychological treatments, in general, for panic disorder. Another practical interpretation of the mean effect can be made by assuming that the reduction in panic attacks in the treatment and control groups follows a normal distribution, and that the value $d = 1.015$ represents in standard units the separation between mean levels of the two groups. Thus, taking as the population of reference the control group, the effect $d = 1.015$ would indicate that, on average, patients who had received psychological treatment are situated in percentile 84.4% of the distribution of the controls – in other words, that the psychological treatments had succeeded in reducing panic attacks by 34.4% with respect to the controls. Figure 1 shows in graphic form this practical interpretation of mean effect size.

The second question – closely related to the previous one – requiring a response from an MA is whether the effect sizes are homogeneous around their mean, or whether, on the other hand, they show such heterogeneity that the mean fails to accurately represent the set of studies. Using the appropriate statistical tests, such as the Q statistic and I^2 index (Borenstein et al., 2009), MA permits a response to this question which, in the case of the MA in our example, led to the conclusion that the studies reflected efficacy results (quantified in terms of the effect sizes) that were strongly heterogeneous.

As a consequence of the heterogeneity shown by the effect sizes, it becomes necessary to answer a third question: “What characteristics of the studies might be affecting the heterogeneity? It is in this phase in which techniques of variance analysis and regression analysis are applied to identify the moderator variables of efficacy. In an MA on psychological treatments, the most important moderator variable is the type of treatment applied in the studies. On classifying the studies according to treatment modality and calculating the mean effect size obtained in each of them, it is possible to compare their efficacy results. A very useful way of presenting the comparison between treatments is through the construction of a graph called a ‘forest plot’, in which the mean effect and confidence interval for each treatment are presented in graphic form. Figure 2 presents the ‘forest plot’ for this MA, which shows the mean effect sizes obtained by the different treatment techniques and the combinations between them. Thus the graph reveals how some treatment techniques have obtained a mean effect that does not differ significantly from the null effect, since their confidence interval includes the 0 value (for example, cognitive therapy alone), and how other techniques have indeed obtained a statistically significant mean effect (for example, exposure with relapse prevention, or the latter combined with relaxation/breathing).

GUIDE TO THE CRITICAL READING OF META-ANALYSES

Examination of the results provided by an MA on the efficacy of treatments, interventions or prevention programmes permits readers to rate the differential efficacy of different treatments, and thus helps them to decide which treatment to apply in a particular case. Nevertheless, the critical reading of MAs necessarily requires that the professional has adequate knowledge of what an MA is, how it is carried out and to what biases its results are exposed. Conscious of this problem, experts

in MA have made considerable efforts to draw up orientative guides to the critical reading of meta-analytic studies, which have resulted in a range of such publications. Rather than reproducing any of these guides, we propose one based fundamentally on the two most recently proposed in the literature: PRISMA (‘Preferred Reporting Items for Systematic Reviews and Meta-Analyses’; Moher et al., 2009), which is itself an improvement of QUOROM (‘Quality Of Reporting Of Meta-analyses’; Moher et al., 1994), and AMSTAR (Shea, Grimshaw, Wells et al., 2007; Shea, Hamel, Wells et al., 2009), which includes 11 questions on the implementation and publication of an SR or an MA.

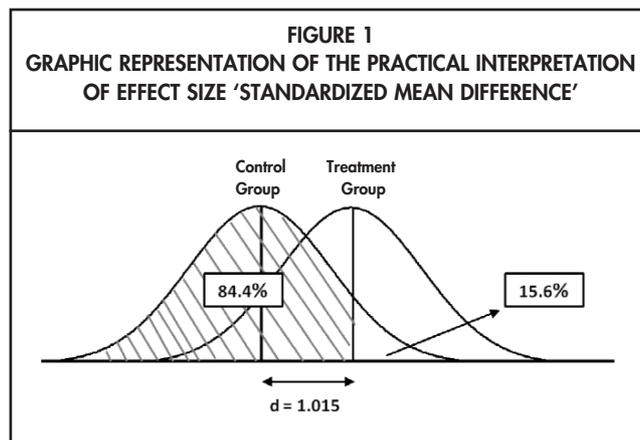


FIGURE 2
 “Forest plot” graphic representation of the mean efficacy results obtained by different psychological treatments for panic disorder with or without agoraphobia. RB: Relaxation/Breathing techniques. ERP: Exposure with relapse prevention (basically, in vivo exposure). CT: cognitive therapy (basically, cognitive restructuring techniques). EMDR: Eye Movement Desensitization Reprocessing. d_+ is the mean effect size obtained by each treatment. CI: Confidence interval

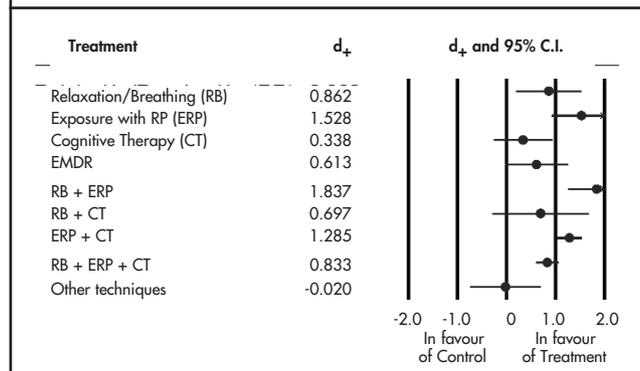


TABLE 2
LIST OF QUESTIONS ORIENTED TO THE CRITICAL READING OF SRs AND MAs

1. Is the study identified as an MA?	yes
	No
	Not available
2. Does it include an Abstract that presents the objectives, method, results and principal conclusions? A structured abstract should be provided, covering: justification; objectives; data sources; study selection criteria, participants and interventions; quality rating of the studies and synthetic methods; results; study limitations; conclusions and implications of the principal results.	yes
	No
	Not available
3. Does the Introduction section describe in an explicit way the questions and the objectives of the MA? There should be an explicit declaration of the questions intended to be answered, with reference to the participants, the interventions, the comparisons, the outcome variables and the design of the studies (PICOS: Participants, Interventions, Comparisons, Outcomes, and Study design).	yes
	No
	Not available
4. Does the Method section specify the inclusion criteria for the studies? There should be specification of the characteristics of the studies (e.g., PICOS, duration of follow-up period) and the characteristics of the studies used as eligibility criteria, stating the reasons for their consideration (e.g., years considered, languages, publication status).	yes
	No
	Not available
5. Does the Method section indicate the procedures used for the study search? All the information sources used in the search should be described (e.g., databases with the dates they cover, contacts with authors of the studies to identify additional studies), as well as the last date of search. The complete electronic search strategy of at least one database should be presented, including possible limitations imposed, so that any researcher can repeat it.	yes
	No
	Not available
6. Does the Method section specify the study variables coded? It should include a description of the method for extracting data from the primary studies (e.g., coding forms applied independently by two or more coders), and of any processes for obtaining and confirming data employed by the reviewers. It should also include a list of all the variables recorded in the studies, as well as their definition (e.g., PICOS, sources of funding), together with information on any assumptions and simplifications made.	yes
	No
	Not available
7. Does the Method section refer to the reliability of the coding? A good MA should have analyzed the reliability of the coding of the moderator variables of the studies, and should present the results of that analysis in terms of kappa indices and intra-class correlations.	yes
	No
	Not available
8. Does the Method section specify the effect size index/indices? The effect size index or indices used in the MA should be specified (e.g., standardized difference of means, odds ratio).	yes
	No
	Not available
9. Does the Method section describe the statistical methods used in the MA? There should be a description of the data treatment methods, and of how the results of the studies were combined (e.g., fixed effects model, random effects model or mixed effects model). Reference should also be made to the measures of consistency employed for analyzing the heterogeneity of the effects (e.g., Q and I^2). There should be some assessment of the risk of bias that might affect the accumulated evidence (e.g., publication bias, selective reporting within the studies). Additional analysis methods should be described (e.g., sensitivity analyses, analyses by subgroups, meta-regression).	yes
	No
	Not available
10. Does the Results section present the characteristics of the studies? There should be a description of the characteristics of the studies included; a table should be provided showing these characteristics on an individual basis, or the reader should at least be given the possibility of access to such a table.	yes
	No
	Not available
11. Does the Results section include analysis of the studies according to their quality? The methodological quality of the studies should have been coded and its relation to the effect sizes considered, with the aim of identifying possible biases due to poor quality. If both randomized and non-randomized studies have been included, their results should be compared.	yes
	No
	Not available

The aim of the protocol presented in Table 2 is to set out the key aspects on which we should focus when reading an SR or an MA, with a view to being able to critically appraise the results it offers and its relevance for clinical practice. The questions are oriented to the different sections into which a published MA is divided: title, abstract, introduction, method, results and discussion. Basically, the questions aim to check whether meta-analyses have made explicit all the decisions that had to be taken during the performance of the MA, and this is a fundamental issue with regard to the possibility of rating its critical quality and of

guaranteeing that other researchers can replicate the analysis.

FINAL REFLECTIONS

Despite the separation of professional practice and research, the EBP approach is generating meeting points, closing the gap between these two worlds which should always go hand in hand. Moreover, SRs and MAs constitute a quick and sure way of discovering the latest scientific evidence and proof on any topic related to professional practice. It is for this reason that psychologists should familiarize themselves with this

TABLE 2
LIST OF QUESTIONS ORIENTED TO THE CRITICAL READING OF SRs AND MAs (Continuation)

12. Does the Results section present the mean effects and the consistency measures? The results of each MA carried out should be presented, including the mean effect sizes with their confidence intervals and the measures of consistency or heterogeneity (e.g., Q , I^2). Optionally, the results of the individual studies and of each MA can be presented by means of a 'forest plot'.	yes
	No
	Not available
13. If there has been heterogeneity, does the Results section present the analysis of moderators? In the case that there is heterogeneity between the effect sizes, mixed effects models should be applied, such as analyses by subgroups (ANOVAs) and meta-regression (regression analysis) to identify characteristics that moderate the results.	yes
	No
	Not available
14. Does the Results section include any sensitivity analysis? If the design included sensitivity analyses for assessing the consistency and robustness of the MA results, they should be described in this section.	yes
	No
	Not available
15. Does the Results section include a publication bias analysis? The MA should have carried out some publication bias analysis to check whether it could represent a threat to the validity of the results.	yes
	No
	Not available
16. Does the Discussion section summarize the evidence? It should include a summary of the principal results, including a reference to the evidence obtained for each principal outcome variable; there should also be some consideration of the relevance for different groups (e.g., healthcare professionals, users and politicians).	yes
	No
	Not available
17. Does the Discussion section consider the limitations of the MA? Limitations should be discussed at the level of the studies, at that of outcome variables (e.g., risks of bias) and at that of the review (e.g., incomplete recovery of studies, reporting bias).	yes
	No
	Not available
18. Does the Discussion section consider the implications for professional practice? There should be some discussion of the implications of the main results of the MA for professional clinical practice, managers, and political decision-makers.	yes
	No
	Not available
19. Does the Discussion section consider the implications for future research? It should provide a general interpretation of the results in the context of other proof and evidence, as well as discussing the implications for future research.	yes
	No
	Not available
20. Are the funding sources specified? There should be a description of the sources of funding of the SR or MA, as well as of other assistance received (e.g., provision of data) and of the role played by the funders in the systematic review, with a view to assessing possible conflicts of interest.	yes
	No
	Not available

methodology and know how to undertake a critical reading of SRs and MAs, as well as other types of studies that contribute evidence.

In a similar line, it would be advantageous for graduate and post-graduate psychology study programmes to include courses providing explanations of the EBP approach and the critical reading of not only SRs and MAs but also other types of research, such as randomized clinical trials, cohort studies or observational and correlational studies. Only in this way shall we arrive at a situation in which all psychologists realize the relevance of this methodology and its practical utility for their professional activity.

The critical reading of research should guide psychologists' practice not only in their activities directly involving people, but also in their decision-making when they occupy positions of authority in companies and institutions responsible for the management of social, health and educational services. And indeed, there should also be a tendency for Evidence-Based Policy, in the case of psychologists in public positions of responsibility or when decision-makers request our opinion based on the evidence accumulated by psychology.

Finally, to explore the methodology of MA and SRs in more depth, we suggest some reading. In Spanish, the reader might consult the work by Botella and Gambara (2002) or the chapter by Sánchez Meca (2008). In English, the texts by Cooper (2010) and Borenstein et al. (2009) are highly relevant. Moreover, there is a range of software programs designed for carrying out the statistical analyses typically involved in an MA. David B. Wilson has developed some MA macros for use in the statistical packages SPSS, SAS and STATA, and which can be obtained free of charge at: <http://mason.gmu.edu/~dwilsonb/ma.html>. Furthermore, the Cochrane Collaboration has developed the RevMan 5.0 program for carrying out MAs, which can also be obtained free of charge at that organization's website (www.cochrane.org). Nor should we omit to mention the commercial program *Comprehensive Meta-analysis 2.0*, developed by Borenstein, Hedges, Higgins and Rothstein (2005; www.meta-analysis.com).

ACKNOWLEDGEMENTS

This article was financed through grants to the first author from the Healthcare Research Fund (*Fondo de Investigación Sanitaria*), within the section "Assessment of Healthcare Technologies" (*Evaluación de Tecnologías Sanitarias*)

(Project: PI07/90384), and to the second author from the Ministry of Education and Science (*Ministerio de Educación y Ciencia*) (Project: SEJ2006-12546/PSIC).

REFERENCES

- Borenstein, M.J., Hedges, L.V., Higgins, J.P.T. & Rothstein, H.R. (2005). *Comprehensive Meta-analysis* (Vers. 2). Englewood Cliffs, NJ: Biostat, Inc.
- Borenstein, M.J., Hedges, L.V., Higgins, J.P.T. & Rothstein, H.R. (2009). *Introduction to meta-analysis*. Chichester, UK: Wiley.
- Botella, J. & Gambara, H. (2002). *Qué es el meta-análisis [What is meta-analysis?]*. Madrid: Biblioteca Nueva.
- Botella, J. & Gambara, H. (2006a). El meta-análisis: una metodología de nuestro tiempo [Meta-analysis: a methodology for our times]. *Infocop*, 29 mayo.
- Botella, J. & Gambara, H. (2006b). Doing and reporting a meta-analysis. *International Journal of Clinical and Health Psychology*, 6, 425-440.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd. ed.). Hillsdale, NJ: Erlbaum.
- Cooper, H. (2010). *Research synthesis and meta-analysis: A step-by-step approach* (3rd. ed.). Thousand Oaks, CA: Sage.
- Cooper, H., Hedges, L.V. & Valentine, J.C. (Eds.) (2009). *The handbook of research synthesis and meta-analysis* (2nd. ed.). New York: Russell Sage Foundation.
- Frías Navarro, M.D. & Pascual Llobell, J. (2003). Psicología clínica basada en pruebas: Efecto del tratamiento [Evidence-based clinical psychology: effect of treatment]. *Papeles del Psicólogo*, 85.
- Grupo de Atención Sanitaria Basada en la Evidencia (2007). *Atención sanitaria basada en la evidencia: Su aplicación a la práctica clínica [Evidence-based healthcare: its application to clinical practice]*. Murcia: Consejería de Sanidad de la Región de Murcia.
- Littell, J.H., Corcoran, J. & Pillai, V. (2008). *Systematic reviews and meta-analysis*. Oxford, UK: Oxford University Press.
- Marín Martínez, F., Sánchez Meca, J., Huedo, T. & Fernández, I. (2007). Meta-análisis: Dónde estamos y hacia dónde vamos [Meta-analysis: where are we and where are we going?]. In A. Borges & P. Prieto (Eds.), *Psicología y ciencias afines en los albores del siglo XXI [Psychology and related sciences at the dawn of the 21st century]* (pp. 87-102). La Laguna, Tenerife: Grupo Editorial Universitario.



- Marín Martínez, F., Sánchez Meca, J. & López López, J.A. (2009). El meta-análisis en el ámbito de las Ciencias de la Salud: Una metodología imprescindible para la eficiente acumulación del conocimiento [Meta-analysis in the Health Sciences: an essential methodology for efficient knowledge gathering]. *Fisioterapia*, 31, 107-114.
- Martín, J.L.R., Tobías, A. & Seoane, T. (Coords.) (2006). *Revisiones sistemáticas en ciencias de la vida* [Systematic reviews in the life sciences]. Toledo: FISCAM.
- Moher, D., Cook, D.J., Eastwood, S., Olkin, I., Rennie, D., et al. (1994). Improving the quality of reporting of meta-analysis of randomized controlled trials: The QUOROM statement. *Lancet*, 354, 1896-1900.
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Medicine*, 6(7): e1000097. doi:10.1371/journal.pmed.1000097.
- Navarro, F., Giribet, C. & Aguinaga, E. (1999). Psiquiatría basada en la evidencia: Ventajas y limitaciones [Evidence-based psychiatry: advantages and limitations]. *Psiquiatría Biológica*, 6, 77-85.
- Nezu, A. M. & Nezu, C. M. (2008). *Evidence-Based Outcome Research: A Practical Guide to Conducting Randomized Controlled Trials*. Oxford University Press.
- Pascual Llobell, J., Frías Navarro, M.D. & Monterde, H. (2004). Tratamientos psicológicos con apoyo empírico y práctica clínica basada en la evidencia [Psychological treatments with empirical support and evidence-based clinical practice]. *Papeles del Psicólogo*, 87.
- Petticrew, M. & Roberts, H. (2006). *Systematic reviews in the social sciences: A practical guide*. Malden, MA: Blackwell.
- Sánchez Meca, J. (1999). Meta-análisis para la investigación científica [Meta-analysis for scientific research]. In F.J. Sarabia-Sánchez (Coord.), *Metodología para la investigación en marketing y dirección de empresas* [Methodology for research in marketing and business management] (pp. 173-201). Madrid: Pirámide.
- Sánchez Meca (2003). La revisión del estado de la cuestión: el meta-análisis [Review of the state of the issue: meta-analysis]. In C. Camisón, M.J. Oltra & M.L. Flor (Eds.), *Enfoques, problemas y métodos de investigación en Economía y Dirección de Empresas. Tomo I* [Approaches, problems and research methods in Economics and Business Management] (pp. 101-110). Castellón: Universitat Jaume I.
- Sánchez Meca, J. (2008). Meta-análisis de la investigación [Meta-analysis of research]. In M.A. Verdugo, M. Crespo, M. Badía & B. Arias (Coords.), *Metodología en la investigación sobre discapacidad* [Methodology in research on disability]. Salamanca: Publicaciones del INICO (Colección ACTAS, 5/2008).
- Sánchez Meca, J. & Ato, M. (1989). Meta-análisis: una alternativa metodológica a las revisiones tradicionales de la investigación [Meta-analysis: a methodological alternative to traditional research reviews]. In J. Arnau & H. Carpintero (Eds.), *Tratado de psicología general. I: Historia, teoría y método* [Treatise on general psychology. I: History, theory and method] (pp. 617-669). Madrid: Alambra.
- Sánchez Meca, J., Boruch, R.F., Petrosino, A. & Rosa Alcázar, A.I. (2002). La Colaboración Campbell y la práctica basada en la evidencia [The Campbell Collaboration and evidence-based practice]. *Papeles del Psicólogo*, 83, 44-48.
- Sánchez Meca, J. & Marín Martínez, F. (in press). Meta-analysis. In B. McGaw, E. Baker & P.P. Peterson (Eds.), *International encyclopedia of education* (3rd. ed.). Oxford: Elsevier.
- Sánchez Meca, J., Marín Martínez, F. & Chacón Moscoso, S. (2003). Effect-size indices for dichotomized outcomes in meta-analysis. *Psychological Methods*, 8, 448-467.
- Sánchez Meca, J., Rosa Alcázar, A.I., Marín Martínez, F. & Gómez Conesa, A. (2010). The psychological treatment of panic disorder with and without agoraphobia: A meta-analysis. *Clinical Psychology Review*, 30, 37-50.
- Shea, B.J., Grimshaw, J.M., Wells, G.A., et al. (2007). Development of AMSTAR: A measurement tool to assess the methodological quality of systematic reviews. *Bio-Med Central*, 7(10). doi: 10.1186/1471-2288-7-10.
- Shea, B.J., Hamel, C., Wells, G.A., et al. (2009). AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. *Journal of Clinical Epidemiology*, 62, 1013-1020.
- Vázquez, C. & Nieto, M. (2003). Psicología (clínica) basada en la evidencia (PBE): Una revisión conceptual y metodológica [Evidence-based clinical psychology: a conceptual and methodological review]. In J.L. Romero (Ed.), *Psicópolis: Paradigmas actuales y alternativos en la psicología contemporánea* [Psychopolis: Current and alternative paradigms in contemporary psychology]. Barcelona: Kairós.

