Original Article

Research Synthesis Methods

Received 25 July 2012, Revised 14 January 2013, Accepted 19 January 2013 Published online 22 February 2013 in Wiley Online Library

(wileyonlinelibrary.com) DOI: 10.1002/jrsm.1076

Less is less: a systematic review of graph use in meta-analyses

Anne H. E. Schild* and Martin Voracek

Graphs are an essential part of scientific communication. Complex datasets, of which meta-analyses are textbook examples, benefit the most from visualization. Although a number of graph options for meta-analyses exist, the extent to which these are used was hitherto unclear. A systematic review on graph use in meta-analyses in three disciplines (medicine, psychology, and business) and nine journals was conducted. Interdisciplinary differences, which are mirrored in the respective journals, were revealed, that is, graph use correlates with external factors rather than methodological considerations. There was only limited variation in graph types (with forest plots as the most important representatives), and diagnostic plots were very rare. Although an increase in graph use over time could be observed, it is unlikely that this phenomenon is specific to meta-analyses. There is a gaping discrepancy between available graphic methods and their application in meta-analyses. This may be rooted in a number of factors, namely, (i) insufficient dissemination of new developments, (ii) unsatisfactory implementation in software packages, and (iii) minor attention on graphics in meta-analysis reporting guidelines. Using visualization methods to their full capacity is a further step in using meta-analysis to its full potential.

Supporting information may be found in the online version of this article.

Keywords: meta-analysis; systematic review; research synthesis; graph; forest plot

1. Introduction

Graphs are essential for effective communication in science (Few, 2004; Tukey, 1977; Tufte, 1983, 1990, 1997, 2006). The combination of the multitude of data and their growing complexity can make it difficult to keep track of all the information available, draw conclusions, and detect patterns. Those three points are key strengths of statistical graphs and figures (Egger and Carpi, 2008).

Meta-analysis, ‘the analysis of analyses’, is a statistical method used for integrating and combining results from different studies on the same topic (Glass, 1976, p. 3). Therefore, meta-analysts often deal with large and complex datasets. Typical outcomes from meta-analyses are summary effects, subgroup effects, and subgroup comparisons. These results may be complicated by large amounts of heterogeneity between studies and different types of underlying biases. It follows that meta-analyses are textbook examples of complex data structures that may support or reveal new conclusions and underlying patterns. This makes them key candidates for the application of graphs and visualizations.

Meta-analyses are applied in various disciplines (e. g., criminology, ecology, and education) and have become increasingly popular (Borenstein *et al*., 2009; Lun-dahl and Yaffe, 2007). However, medicine and psychology are still the predominant fields where this method is used, as meta-analyses are of particular importance in the application of evidence-based practice (Khan *et al*., 2011). Furthermore, meta-analytic findings are of great importance in policy and political decision making (Lavis, 2005, 2009; Lomas, 2005; Stoto, 2000). Therefore, typical recipients of meta-analytic results include not only academics but also non-researchers, such as politicians and practitioners (Lomas, 2005; Lavis, 2005, 2009). Research suggests that this particular audience shows a preference for formats that allow for easy scanning and rapid assessment (Lavis, 2005), both of which may be facilitated through graphical representations. What is more, graphics can be tailored specifically to different audiences and their respective needs (Spiegelhalter *et al*., 2011).

A number of graphs that are specific to the field of meta-analysis (see Anzures-Cabrera and Higgins, 2010 for an overview) exist. Potential areas of application of graphics in meta-analyses include visualization of processes (e. g.,

Department of Basic Psychological Research and Research Methods, School of Psychology, University of Vienna, Liebiggasse 5, A-1010 Vienna, Vienna, Austria

**Correspondence to: Anne H. E. Schild, Department of Basic Psychological Research and Research Methods, School of Psychology, University of Vienna, Liebiggasse 5, A-1010 Vienna, Vienna, Austria.*

E-mail: anne.schild@univie.ac.at

flow charts, Moher *et al*., 2009), representation of effects and uncertainty (e. g., forest plots, Lewis and Clarke, 2001), and exploration of sources of heterogeneity and diagnostics (e. g., funnel plots, Light and Pillemer, 1984).

Results of meta-analyses are most commonly displayed in forest plots (Figure 1(A)). Effect sizes from individual studies paired with their respective confidence intervals are plotted together with a pooled summary estimate for all included studies. In these graphs, the size of the plotting symbol usually corresponds to the study's weight in the analysis, and the confidence interval is depicted with a line. These two conventions have repeatedly been criticized as follows: (i) smaller studies may attract more visual attention because of the long confidence intervals; (ii) it is difficult to discern the exact point estimate of studies carrying much weight; and (iii) all points situated within the boundaries of the confidence interval appear to be equally likely. To overcome this third shortcoming, two variations of forest plots have been proposed. Density strips (Jackson, 2008), in which uncertainty is represented with shading (Figure 1(B)), may replace the traditional design of effects and their confidence intervals in forest plots. Another option are raindrop plots (Barrowman and Myers, 2003), which were originally intended to display collections of distributions and likelihoods.

Funnel plots (Figure 2(A)) are the only diagnostic plots specific to meta-analysis, and they are a common and useful tool in the exploration of sources of heterogeneity and bias. Graphical augmentations (Figure 2(B)) proposed recently (Langan *et al*., 2012) may be a vital tool in facilitating their objective interpretation (Terrin *et al*., 2005) and broaden their fields of application. Furthermore, they may also help overcome issues with unsatisfactory inter-rater

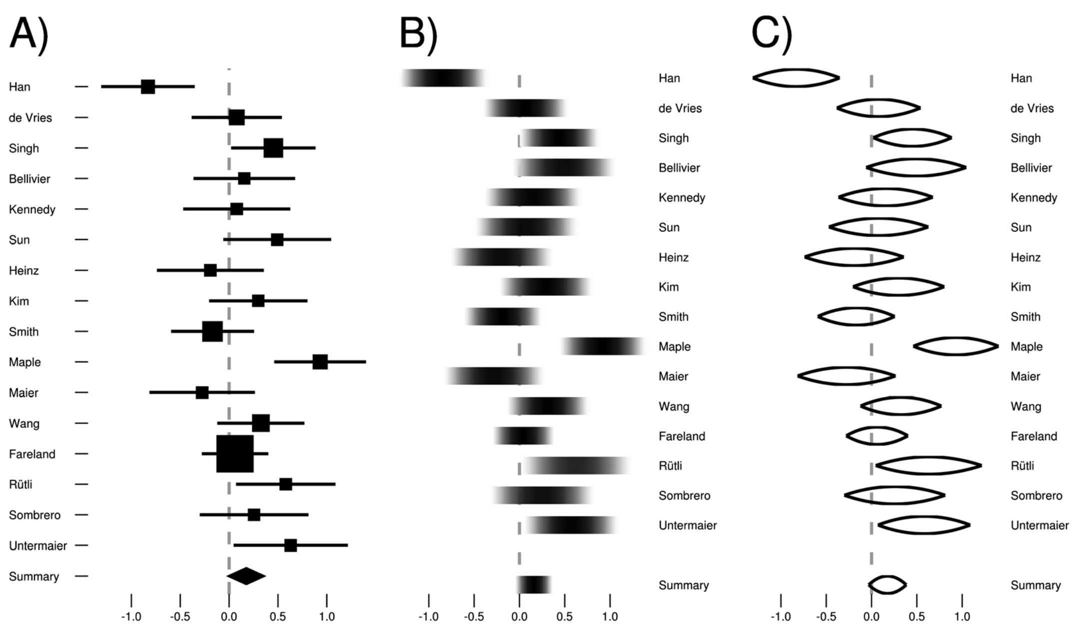

Figure 1. Variations of forest plots. (A) Conventional forest plot, (B) forest plot with density strips, and (C) forest plot with raindrop plots.

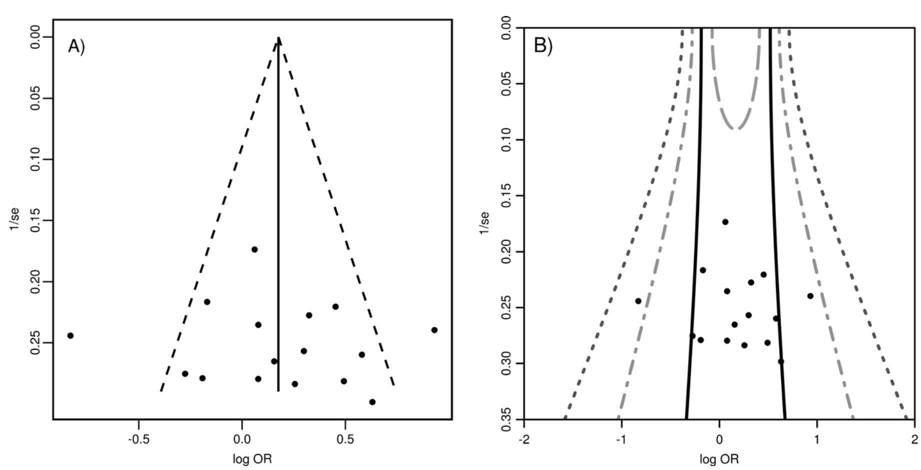

Figure 2. Variations of funnel plots. (A) Conventional funnel plot and (B) contour-enhanced funnel plot. Lines indicate different levels of between-study heterogeneity.

agreement for the assessment of funnel plot asymmetry (intraclass correlation coefficient for inverse standard error funnel plot = 0.53, 95% confidence interval, 0.412–0.632; Bax *et al*., 2009).

In addition to these more or less established plots, six new graphical methods have been introduced in the last several years (see Figure 3 for examples). This development mirrors both the increasing usage of meta-analytical methods and the corresponding need for visualizations.

The Baujat plot (Figure 3(A)) is a scatterplot in which the contribution of each study to the overall heterogeneity is plotted against the influence of this data point (Baujat *et al*., 2002). The plot allows to identify sources of heterogeneity, which may be individual studies or subgroups of studies (Baujat *et al*., 2002).

An adaptation of receiver-operator curves for assessing diagnostic test accuracy in diagnostic meta-analyses has recently been proposed (Phillips *et al*., 2010). The so-called crosshairs plot (Figure 3(B)) displays individual studies together with paired 95% confidence intervals of sensitivity and specificity in a receiver-operator curves space, thereby facilitating the comparison of key diagnostic data between studies. Analoguous to forest plots, symbol sizes represent study weight, and summary estimates can be added.

The application of quality control charts has also been suggested (Kulinskaya and Koricheva, 2010). In $\bar{X}$ control charts (Figure 3(C)), the sample mean is plotted together with upper and lower control limits. The data are

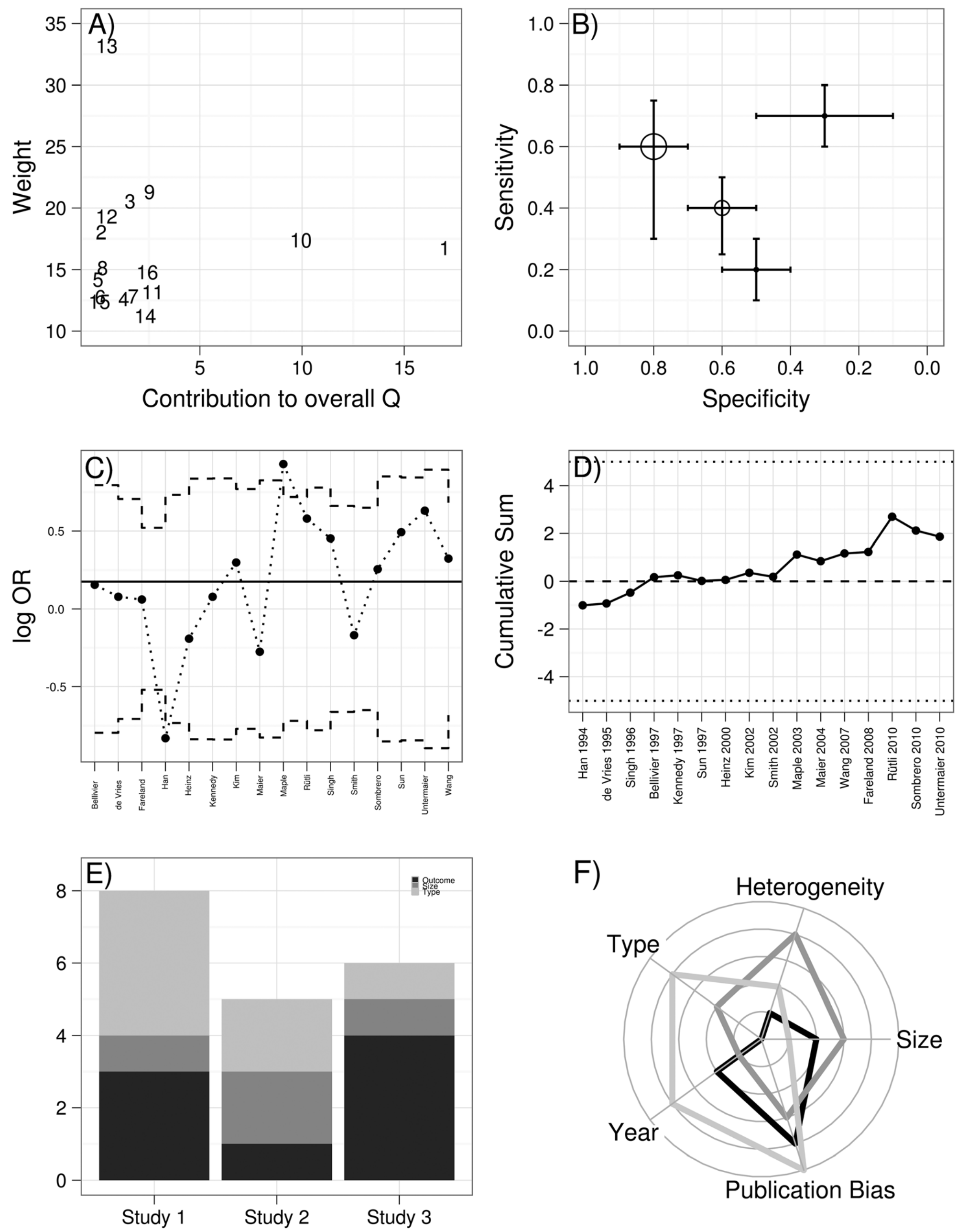

Figure 3. Examples of plots proposed for the visualization of meta-analyses. (A) Baujat plot, (B) crosshairs plot, (C) $\bar{X}$ control chart, (D) CUSUM chart, (E) harvest plot, and (F) veritas plot.

considered to be in control if the individual points fall within the limits and do not show a systematic pattern. In contrast to $\bar{X}$ control charts, in which the H_0 that the process is in control is tested for individual data points, CUSUM charts (Figure 3(D)) use cumulated sums of sample deviations from the target value, that is, they show accumulated sequential data. The authors recommend the implementation of these two methods to detect outliers and temporal trends.

The harvest plot (Figure 3(E)) allows for the simultaneous display of effects and study characteristics (Crowther *et al*., 2011). The plot is essentially a bar chart where the height of the bars corresponds to the effect size and different horizontally stacked fillings are used to visualize further information (e. g., study quality).

The veritas plot (Figure 3(F)) is intended to serve as a graphic assessment tool of meta-analysis quality and is meant to facilitate comparisons between multiple meta-analyses on the same topic. Relevant meta-analyses are coded on a number of different dimensions and ranked accordingly. For each meta-analysis, the scores for the individual dimensions are displayed in a radial plot, and the individual lines are connected with lines, that is, each line represents one meta-analysis. The veritas plot allows to compare both overall quality of meta-analyses and differences between individual meta-analyses regarding specific quality criteria. It is intended for use by practitioners and clinicians and may therefore be a vital tool in the implementation of evidence-based practice (Panesar *et al*., 2009).

Furthermore, a number of traditional statistical graphics may be utilized. This has the advantage of easy accessibility even for readers not familiar with more specialized graphs. Histograms, boxplots, or stem-and-leaf plots of effects are just three such possibilities.

The *Preferred Reporting Items for Systematic Reviews and Meta-Analyses* Statement (PRISMA Moher *et al*., 2009) and the *Meta-Analysis Reporting Standards* (MARS, APA Publications and Communications Board Working Group on Journal Article Reporting Standards, 2008) are comprehensive guidelines for the publication of meta-analyses. Unfortunately, these are largely limited to the textual reporting. Whereas PRISMA at least recommends the use of a forest plot in the presentation of results, there is no mention of graphical displays in MARS.

When flipping through an article reporting a meta-analysis, one might arrive at the conclusion that the paper is rather poor in figures. This subjective notion is strengthened by an informal review on graph use in meta-analyses in *Psychological Bulletin* between 1985 and 1991, which suggests that graph use is rather scarce (Light *et al*., 1994). An updated version of this review including the total of published articles between 2000 and 2005 in *Psychological Bulletin* and *Review of Educational Research* underlines this impression (Borman and Grigg, 2009). Although this updated review shows a notable increase in graph use in *Psychological Bulletin*, only slightly more than half of the examined articles included graphics at all. Of these, distributions of effect sizes played the most important role. Apart from the overall rather rare graph usage, it is surprising to note that only about a tenth of these made use of representations of uncertainty and, even more surprising, summary effects were shown in only 5% of the cases (Borman and Grigg, 2009). Overall, it seems that not only graph use was generally low but also the existing options were not used to their full potential. However, these results are hardly generalizable, as (i) they are limited to psychology journals, (ii) they are based on only two journals and cover only a limited space of time, and (iii) only the occurrence of graphs was recorded, neglecting other aspects (e.g., article length and number of included studies) completely.

The current review is stimulated by this concise review and the open questions it raises as well as by the subjective impression of the relative scarcity of graphs in meta-analyses. This work aims at evaluating actual graph use in the reporting of meta-analyses in top journals in three disciplines. Graph usage studies allow researchers to evaluate which methods are applied and how they are applied. Therefore, usage studies are vital for identifying design and usage errors, which enables the development of (i) usage and design guidelines, (ii) new graphical methods, and (iii) specialized software (Cleveland, 1984). The present paper aims at providing a detailed review of the development and current state of graph use in meta-analyses to lay the foundation for these three development goals.

2. Methods

2.1. Literature search

The search string *meta-anal* * or *meta anal* or *metaanal* * was entered into ISI Web of Knowledge (including Medline). The search was limited to paper titles, as the term 'meta-analysis' should be provided explicitly in the title of the article if a meta-analysis was conducted (APA Publications and Communications Board Working Group on Journal Article Reporting Standards, 2008). The search included all records available up until August 2, 2011, and yielded more than 30,000 hits.

Although the Cochrane library (http://www.thecochranelibrary.com) with more than 7000 systematic reviews constitutes the single largest resource of research synthesis, this database was excluded from the current review. Firstly, the Cochrane Collaboration publishes and enforces detailed guidelines on how to prepare suitable reviews (Higgins, 2008), and secondly, the reviews are dedicated solely to the field of healthcare. Hence, variation within this group should be minimal.

2.2. *Study selection*

From the complete list of records, the three subject areas (medicine, psychology, and business) with the highest number of records were selected using the *Subject Area* function in ISI Web of Knowledge. Subject areas were required to be broad and distinct in order to represent a general discipline rather than a specialization. For each of the three disciplines, the three journals with the highest number of records were chosen as target journals (medicine: *Lancet*, *British Medical Journal*, and *Journal of the American Medical Association*; psychology: *Psychological Bulletin*, *Clinical Psychology Review*, and *Journal of Applied Psychology*; business: *Academy of Management Journal*, *Journal of Organizational Behavior*, and *Journal of Marketing Research*).

2.3. *Data extraction*

The print versions of all selected articles were examined. All studies that reported meta-analyses were included (see Figure 4 for an overview of the process of literature search and study selection) in the analyses. References that did not report meta-analyses (e. g., letters and comments) were excluded upon manual examination ($N = 413$), and eight studies were excluded for other reasons (see Supplement 1 for a complete list of excluded references).

Data were abstracted from all studies that reported meta-analyses ($N = 993$, see Supplement 2 for a complete list of included references). Both general information (e.g., number of pages, number of graphs, and type of graphs) and graph specifications (e.g., inclusion of reference lines) were extracted. Information specific to the graphs was adapted from recommendations published in a recent overview on how to effectively use meta-analysis graphs (Anzures-Cabrera and Higgins, 2010).

Supplementary information published online was not considered, as, by definition, any material integral to the understanding and presentation of the research paper should be provided in the body of the manuscript (American Psychological Assocation, 2009). Furthermore, supplementary files typically receive less attention in the review process and are published in unedited format. Therefore, the degree of evidence of this additional information is not comparable to a published peer-reviewed paper. Moreover, the publication of online material is a rather recent phenomenon, whereas studies included in the complete set date back to 1981.

Only the number of figures per article was counted, that is, if one figure contained more than one graph (e. g., two forest plots in one figure), these were counted as one and the article was marked as containing multi-panel graphs. This coding variant was chosen to allow for more consistency in two cases. Firstly, especially individual forest plots can easily be combined into one graph, meaning that presenting two subgroup plots in one box conveys exactly the same information as presenting one summary plot of the two subgroups. Secondly, plots

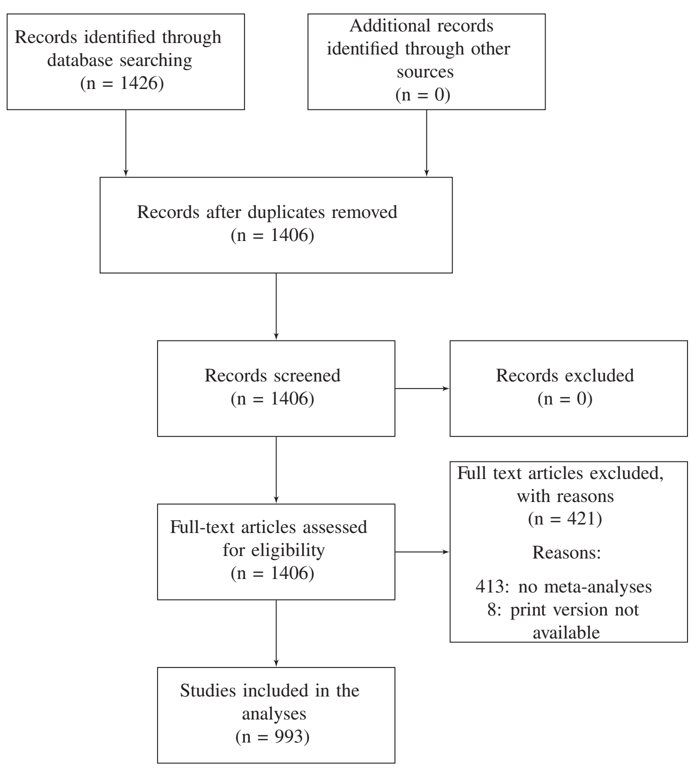

Figure 4. Flow chart of the process of identifying and selecting studies included in the analyses

may be superimposed onto each other (e.g., a random-effects forest plot superimposed onto a fixed effect forest plot) instead of being plotted separately.

In similar works, graph number and size were combined into a joint measure called fractional graph area (e.g., Cleveland, 1984; Smith *et al*., 2002). This approach was not followed, and graph size was not taken into account, as size largely depends on a number of formatting and journal constraints. Furthermore, graph size may be a direct function of the number of included samples, especially in forest plots.

2.4. Study characteristics

The analyses included 993 studies of which 523 belonged to the field of medicine, 415 were taken from psychology journals, and studies from business accounted for a total of 55. Median length of included meta-analyses was 11 printed pages ($IQR = 11$) and summarized data from a median of 30 studies ($IQR = 54$). Mean number of graphs ranged from 0.44 (*Journal of Organizational Behavior*) to 3.77 (*Lancet*) and differed widely between disciplines and journals (Table 1).

Disregarding between-discipline and between-journal differences leads to nearly 70% ($N = 686$) of meta-analyses that used visual aids. However, this percentage drops to as low as 51.6% ($N = 512$) when only graphs related to the meta-analysis are taken into account.

2.5. Quantitative data synthesis

All data were synthesized and analyzed using the open-source software R 2.15.0 (R Development Core Team, 2012). Only graphs referring to meta-analytical results were used in the analyses. Hence, flowcharts and graphs categorized as *other* were excluded.

3. Results

Bivariate correlations between all major variables (number of graphs, number of pages, number of included studies, and graph types) were mostly small (see Figure 5 for a graphical overview). The only larger correlations had little explanatory value (e.g., positive correlation between number of graphs and number of forest plots, which indicates only that forest plots accounted for the majority of graphs). The impact of potentially interesting variables (e. g., length in pages and number of included studies) is negligible and depends largely on the field and the specific journal. Therefore, potential differences between disciplines and journals are the strongest expectable factors.

Graphical representations of graph use per discipline (Figure 6(A)) show clear differences between the three areas. Graph use is by far the highest in medical meta-analyses and differs also between psychology and business, although those differences are much smaller in number. Figure 6(A) additionally shows the mean usage of multi-

Table 1. Number of graphs by type across and within disciplines.

Journal	N_{MA}	n_{graph}	*M* (*SD*)	Forest plot	Funnel plot	Flow chart	Frequency distribution	Other
Medicine								
BMJ	223	578	2.59 (2.04)	442	15	59	0	62
JAMA	174	449	2.58 (1.66)	290	18	73	3	65
Lancet	126	475	3.77 (2.16)	305	6	46	2	116
Total	523	1502	2.87 (2.01)	1037	39	178	5	243
Psychology								
Psychol Bull	199	250	1.25 (1.95)	26	25	5	38	156
J Appl Psychol	119	91	0.76 (1.24)	0	1	0	4	86
Clin Psychol Rev	97	120	1.24 (1.98)	35	7	11	12	55
Total	415	461	1.11 (1.79)	61	33	16	54	297
Business								
Acad Manage J	23	15	0.65 (1.23)	0	0	0	0	15
J Market Res	14	18	1.29 (0.99)	0	0	0	3	15
J Organ Behav	18	8	0.44 (0.86)	0	1	0	0	7
Total	55	41	0.75 (1.09)	0	1	0	3	37

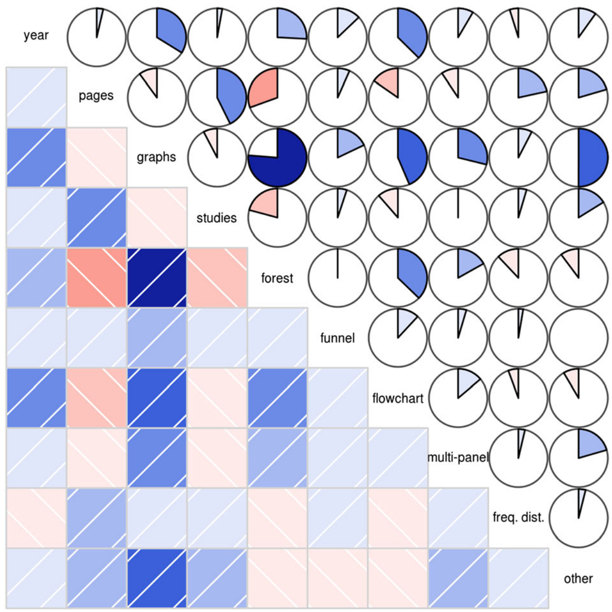

Figure 5. Correlogram. Color-coded correlation matrix with color saturation and pie area proportional to size and direction of correlations. Positive correlations are shown in blue shading and negative correlations in red shading.

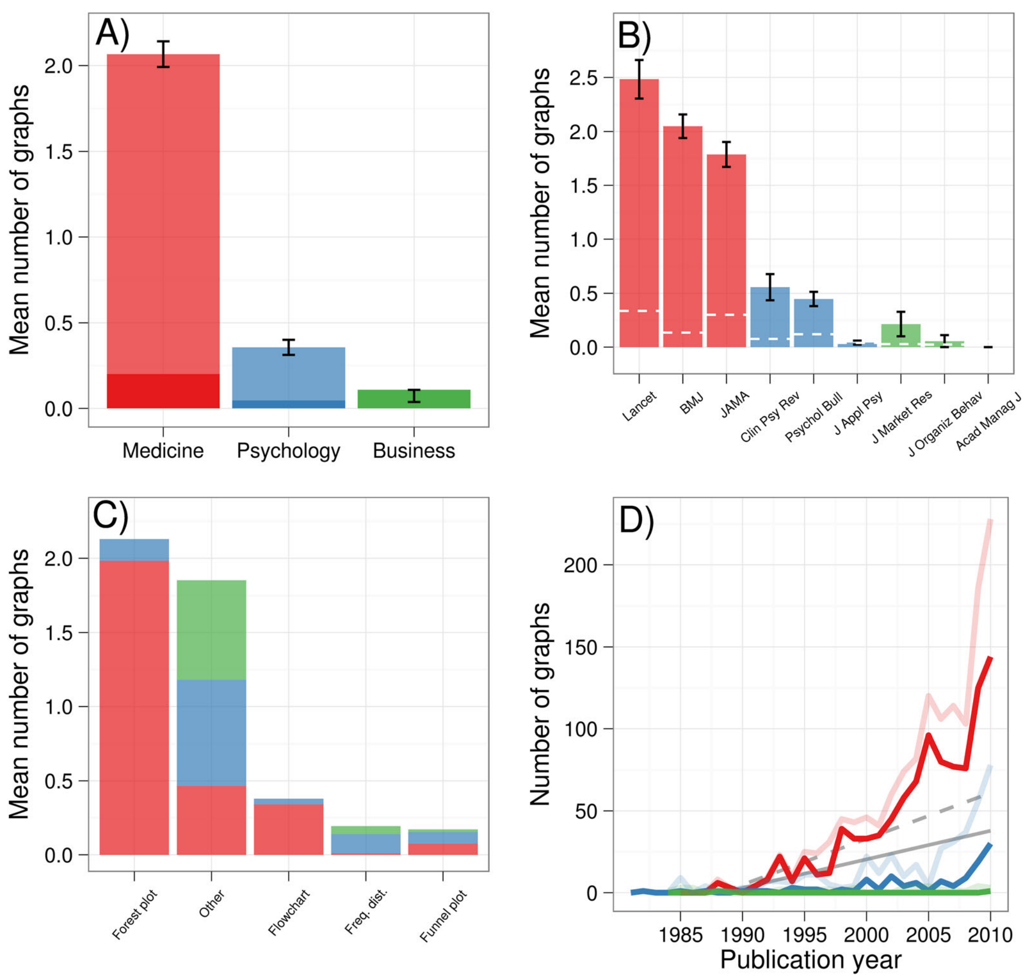

Figure 6. Results from systematic review. (A) Mean number of graphs per discipline. Darker shading indicates mean use of multi-panel graphs. (B) Mean number of graphs per journal. Dashed lines show journal impact factor divided by 100. (C) Mean number of graphs by type. (D) Increase in graph use over time by discipline. Gray dashed line shows increase in overall graph use, solid line refers to MA-specific graphs. Lines in light colors show overall graph use, lines in saturated colors show number of MA-specific graphs.

panel graphs (in darker shading). Apparently, such displays are rather uncommon in medicine, psychology, and business. However, the observed means differ between fields, and the proportion is highest in medicine. Previous research has shown a positive correlation between overall graph use and the use of multi-panel graphs (Best *et al*., 2001; Smith *et al*., 2002). The observed interdisciplinary differences support this finding.

Regarding mean graph use between journals (Figure 6(B)), the between-discipline differences are still apparent and hence not conditioned by discrepant journal requirements only. Moreover, Figure 6(B) further reveals that graph use in psychology and business is comparable. Considering the individual journal titles, it is clear that the three business journals are rather psychology-oriented (e. g., *Journal of Organizational Behavior*), and therefore, the observed lack of divergence between these two fields does not seem surprising.

One might argue that graphic use may be dependent on a journal's impact factor. Especially considering the field of medicine, it is obvious that the three chosen journals belong to a high-impact group. Therefore, the fact that papers published in these journals use significantly more graphs than papers from the other two disciplines may be rooted in the impact factor difference. This, however, does not seem to be the case, considering the pattern incoherence between mean graph use and journal's impact factors (white dashed lines).

Graph use in meta-analyses is limited not only by the actual number of graphs but also by their narrow range (Figure 6(C)). The front-runner among the graphs was the forest plot. It is interesting to note that this plot type is highly dominant in analyses pertaining to the field of medicine and virtually non-existent in the other two areas. It appears highly likely that this is rooted not only in journals' stylistic requirements but also in a copycat effect where potential authors stylistically reproduce recent articles published in their respective target journals. This is problematic in the sense that this phenomenon triggers a circular chain reaction in which high graph use presupposes itself.

Moreover, the variance in graph types used in medicine is much larger and spans the whole field of observed types, whereas plot type variation in psychology and business is much more limited. The only category distributed rather evenly over all three disciplines is the diffuse group *other*. Together with the observation that depiction of frequency distributions appears to be less common in medicine, one might argue that psychology and business show a preference toward more traditional displays. The application of more conventional graphs may be more acceptable in less visualization-friendly journals, and, furthermore, such journals might show an affinity for table-like graphs, which might explain the higher occurrence of stem-and-leaf plots in *Psychological Bulletin*.

A further fundamental limitation lies in the fact that diagnostic plots (e. g., funnel plots) are relatively rare. Although mentions of examination of funnel plots can be found in the text, the plots themselves are rarely given, a practice that is problematic given the low inter-rater agreement (Bax *et al*., 2009). Generally speaking, it appears that, if graphs are used at all, they are used very selectively. They may give the reader a chance to have a visual impression of effect sizes and general processes; however, diagnostics and inherent structures are presented in a text-oriented fashion and therefore less easily accessible.

Usage of meta-analysis graphs has increased in psychology and medicine over the last three decades (Figure 6 (D)). However, this phenomenon does not seem to be specific to graphs related to meta-analysis but is mediated by a general increase in graph use. Moreover, the growth in overall graph use has been slightly steeper ($r = 0.34$, $P < 0.001$) than the growth in usage of meta-analysis-related graphs ($r = 0.27$, $P < 0.001$).

4. Discussion

Although graphical displays should be an integral part in the application and reporting of meta-analysis, this, unfortunately, does not yet seem to be the case. It appears that their application largely depends on the scientific discipline. On a more positive note, one might argue that graph use in the reporting of meta-analyses may at least be a prominent feature in some fields. However, a more realistic (and unfortunately more discouraging) interpretation suggests that the observed interdisciplinary differences are not attributable to the method at all but depend solely on the field of study. Consequently, graph use could be considered an integral feature in the reporting within circumscribed fields but could be substantially detached from methodological considerations after all.

In a recent overview (Anzures-Cabrera and Higgins, 2010), which is the most comprehensive compilation of graphical applications in meta-analysis so far, particular prominence is given to four plot types: forest plots, funnel plots, Galbraith plots, and l'Abbé plots. This may give rise to the assumption that those four are not only commonly and prominently used in meta-analyses but also similarly frequent. The present results reveal, however, that both these subjective impressions are false. It becomes apparent that, if graphs are used at all, forest plots seem to be considered as a kind of 'gold standard' and other—possibly more adequate—options are neglected.

The claim that funnel plots are among the most widely used graphical displays in meta-analysis (Langan et al., 2012) appears highly unlikely in the light of the current results. It may well be that diagnostic usage is frequent, but generally they do not seem to appear in print. It has been argued that publication of diagnostic plots may be uninformative in most cases (Gelman, 2011). However, considering in tandem the implications of bias in meta-analyses and issues with inter-rater agreement (Bax *et al*., 2009), it might make sense to accompany analyses with funnel plots to permit the reader an individual assessment. Furthermore, in the light of a certain degree of subjectivity in the appraisal of funnel plots, their publication may be an important step in making the analyses more transparent and objective.

Although the results suggest an increase in the use of meta-analysis graphs, it is unlikely that this trend is the reflection of a change specific to the reporting of meta-analyses. Evidence has shown a general growth in graph use in scientific journals (Zacks *et al*., 2002), which is probably, for the most part, rooted in the greater availability of software tools for the production of statistical graphs.

The current review reveals a gaping void between existing methods and their actual application. Leaving aside considerations of discipline and journal, it appears that low graph use is foremost an issue, which is created not by a lack of options but by a failure to disseminate these and make them available to a broader audience.

It is highly likely that such discrepancies between availability and utilization of graphical methods are strongly dependent on statistical software. Table 2 gives an overview of specialized software packages for meta-analysis (adapted from Bax *et al*., 2007; Sterne, 2009; Wang and Bushman, 1999; Field and Gillett, 2010; Lipsey and Wilson, 2001) and the graphical methods implemented therein. Apparently, the most common graphics (cf. Anzures-Cabrera and Higgins, 2010) are more or less directly available. However, new developments (cf. Figure 3) have not yet found their way into software packages.

It is remarkable that most medical meta-analyses, which generally exhibited both higher graph use and frequency of forest plots, were in most cases prepared with RevMan (The Cochrane Collaboration, 2011). This software is intended for the preparation of Cochrane reviews, in which the use of forest plots is generally encouraged (Schuenemann *et al*., 2008; Higgins, 2008). It is only natural that a prominent implementation of graphic methods in a software package will facilitate the application of graphs, whereas a lack of availability will surely hinder the application of such methods. It goes without saying that a failure to implement novel methods and state-of-the-art graph design conditions low graph use.

Regarding reporting standards for meta-analyses, it becomes apparent that graphic use is an issue that has received little consideration as of yet. This shortcoming is reflected directly in published meta-analyses and in journals' stylistic requirements. More often than not, information on how to prepare meta-analyses for publication in specific journals is based strongly or entirely on either PRISMA (Moher *et al*., 2009) or MARS (APA Publications and Communications Board Working Group on Journal Article Reporting Standards, 2008). Although this is a highly positive development regarding the theoretical and textual preparation of meta-analyses, this unfortunately also entails minor attention to visualizations.

Although forest plots may be found in many meta-analyses, there is room for improvement regarding their design and the arrangement of single studies and subgroups within the plot. More often than not, studies are simply arranged in alphabetical order (Schriger *et al*., 2010), which is usually not ideal for conveying the maximum of the information in the complete set of studies. Optimally, authors should experiment with different arrangements. This way, firstly, the information content of the whole graphic may be enhanced, and, secondly,

Table 2. Available graphics in common software packages for meta-analysis.

	Forest[a]	Sub/Sum[b]	Cumul[c]	Funnel[d]	Cont[e]	Trim[f]	Galb[g]	l'Abbé[h]	Other
CMA[1]	✓	✓	✓	✓	✗	✓	✗	✗	
R[2,3,4]	✓	✓	✓	✓	✓	✓	✓	✓	
RevMan[5]	✓	✓	✗	✓	✓	✗	✗	✗	SROC curves, flow charts, risk of bias graphs
MIX[6]	✓	✓	✓	✓	✓	✓	✓	✓	Baujat plot, Bayesian triplots, sensitivity plots
Stata[7]	✓	✓	✓	✓	✓	✓	✓	✓	HSROC
SAS[8]	✓	✓	✓	✓	✗	✓	✓	✗	
SPSS[9]	✓	✗	✗	✓	✗	✗	✗	✗	

[a] Forest plot, [b] Summary forest plot and/or subgroup forest plot, [c] Cumulative forest plot, [d] Funnel plot, [e] Contour-enhanced funnel plot, [f] Trim-and-fill funnel plot, [g] Galbraith plot, [h] l'Abbé plot.
[1] Borenstein *et al.* (2005), [2] R: Viechtbauer (2010), [3] R: Schwarzer (2012), [4] R: Lumley (2009), [5] The Cochrane Collaboration (2011), [6] Bax (2011), [7] StataCorp (2011), [8] SAS Institute Inc. (2013), [9] IBM Corp. (2012).

the different versions may serve an exploratory function, that is, they may both trigger and accompany further analyses regarding for example time trends or subgroups.

In addition to the use of forest plots, funnel plots should find their way into published meta-analyses as well to enhance objectivity and transparency. The application of trim-and-fill methods and additional contours appear to be especially helpful in interpretation, and both these methods are readily available in a number of software packages (cf. Table 2).

Keeping in mind that meta-analyses most frequently present compact but comprehensive overviews and often summarize large amounts of information on extensive subject areas, graphical representations should be encouraged in all meta-analyses. Keeping in mind the interdisciplinary differences, more explicit recommendations and instructions for authors would be a helpful way to promote graphic use. Ideally, graphics should provide both a well-founded overview as a basis for decision making and should further present information relating to methodological and diagnostic considerations. In this regard, it is important to keep in mind that different audiences have different needs and that certain graph types may appear unintuitive to non-informed readers (e. g., Galbraith plots, Anzures-Cabrera and Higgins, 2010).

If visualizations are used to make your data talk, then they should be given the voice that best represents their particular strengths, weaknesses, and idiosyncrasies. Subsequently, the selection of graphs should, above all, be theoretically founded. Neither the choice of omitting graphs altogether nor the practice of simply adhering to tradition or convention fulfill this criterion. Using meta-analysis graphs to their full potential is a vital step in using meta-analysis to its full potential.

References

American Psychological Assocation. 2009. Publication Manual of the American Psychological Association. 6th edn. American Psychological Assocation, Washington, DC.

Anzures-Cabrera J, Higgins JPT. 2010. Graphical displays for meta-analysis: an overview with suggestions for practice. *Research Synthesis Methods* **1**: 66–80.

APA Publications and Communications Board Working Group on Journal Article Reporting Standards. 2008. Reporting standards for research in psychology: why do we need them? What might they be? *American Psychologist* **63**: 839–851.

Barrowman NJ, Myers RA. 2003. Raindrop plots: a new way to display collections of likelihoods and distributions. *The American Statistician* **57**: 1–6.

Baujat B, Mahe C, Pignon JP, Hill C. 2002. A graphical method for exploring heterogeneity in meta-analyses: application to a meta-analysis of 65 trials. *Statistics in Medicine* **21**: 2641–2652.

Bax L. 2011. MIX 2.0. Professional software for meta-analysis in Excel. Version 2.0.1.4. BiostatXL. Available at: http://www.meta-analysis-made-easy.com.

Bax L, Ikeda N, Fukui N, Yaju Y, Tsuruta H, Moons KGM. 2009. More than numbers: the power of graphs in meta-analysis. *American Journal of Epidemiology* **169**: 249–255.

Bax L, Yu LM, Ikeda N, Moons KGM. 2007. A systematic comparison of software dedicated to meta-analysis of causal studies. *BMC Medical Research Methodology* **7**: 40.

Best LA, Smith LD, Stubbs DA. 2001. Graph use in psychology and other sciences. *Behavioural Processes* **54**: 155–165.

Borenstein M, Hedges L, Higgins J, Rothstein H. 2005. Comprehensive Meta-analysis Version 2. Biostat, Englewood, NJ.

Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. 2009. Introduction to Meta-analysis. Wiley, West Sussex, England.

Borman GD, Grigg JA. 2009. Visual and narrative interpretation. In The handbook of research synthesis and meta-analysis, 2nd edn. Cooper H, Hedges LV, Valentine JC (eds.) Russell Sage, NY 495–519.

Cleveland WS. 1984. Graphs in scientific publications. *The American Statistician* **38**: 261–269.

Crowther M, Avenell A, MacLennan G, Mowatt G. 2011. A further use for the harvest plot: a novel method for the presentation of data synthesis. *Research Synthesis Methods* **2**: 79–83.

Egger AE, Carpi A. 2008. Data: using graphs and visual data. *Visionlearning* POS-1.

Few S. 2004. Show Me the Numbers: Designing Tables and Graphs to Enlighten. Analytics Press, Oakland, CA.

Field A, Gillett R. 2010. How to do a meta-analysis. *British Journal of Mathematical and Statistical Psychology* **63**: 665–694.

Gelman A. 2011. Why tables are really much better than graphs. *Journal of Computational and Graphical Statistics* **20**: 3–7.

Glass GV. 1976. Primary, secondary, and meta-analysis of research. *Educational Researcher* **5**: 3–8.

Higgins J. 2008. Considerations and recommendations for figures in Cochrane reviews: graphs of statistical data. Available at: www.cochrane.org/sites/default/files/style-guide/Graph_recom mendations9.pdf.

IBM Corp. 2012. IBM SPSS Statistics for Windows, Version 21.0. IBM Corp, Armonk, NY.

Jackson CH. 2008. Displaying uncertainty with shading. *The American Statistician* **62**: 1–8.

Khan K, Kunz R, Kleijnen J, Antes G. 2011. Systematic Reviews to Support Evidence-based Medicine. Marcel Dekker, New York.

Kulinskaya E, Koricheva J. 2010. Use of quality control charts for detection of outliers and temporal trends in cumulative meta-analysis. *Research Synthesis Methods* **1**: 297–307.

Langan D, Higgins JPT, Gregory W, Sutton AJ. 2012. Graphical augmentations to the funnel plot assess the impact of additional evidence on a meta-analysis. *Journal of Clinical Epidemiology* **5**: 511–519.

Lavis JN. 2005. Towards systematic reviews that inform health care management and policy-making. *Journal of Health Services Research & Policy* **10**: 35–48.

Lavis JN. 2009. How can we support the use of systematic reviews in policymaking? *PLoS Medicine* **6**: e1000141.

Lewis S, Clarke M. 2001. Forest plots: trying to see the wood and the trees. *British Medical Journal* **322**: 1479–1480.

Light RJ, Pillemer DB. 1984. Summing Up: The Science of Reviewing Research. Harvard University Press, Cambridge, MA.

Light RJ, Singer JD, Willet JB. 1994. The visual presentation and interpretation of meta-analyses. In The Handbook of Research Synthesis and Meta-analysis. 1st edn. Cooper H, Hedges LV (eds.) Russell Sage, NY 439–453.

Lipsey MW, Wilson DB. 2001. Practical Meta-analysis. Sage, Thousand Oaks, CA.

Lomas J. 2005. Using research to inform healthcare managers' and policy makers' questions: from summative to interpretative synthesis. *Healthcare Policy* **1**: 55–71.

Lumley T. 2009. rmeta: Meta-analysis. Available at: http://CRAN.R-project.org/package=rmeta.

Lundahl B, Yaffe J. 2007. Use of meta-analysis in social work and allied disciplines. *Journal of Social Service Research* **33**: 1–11.

Moher D, Liberati A, Tetzlaff J, Altman D. 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Medicine* **6**: e1000097.

Panesar SS, Rao C, Vecht JA, Mirza SB, Netuveli G, Morris R, Rosenthal J, Darzi A, Athanasiou T. 2009. Development of the veritas plot and its application in cardiac surgery: an evidence-synthesis graphic tool for the clinician to assess multiple meta-analyses reporting on a common outcome. *Canadian Journal of Surgery* **52**: E137–E145.

Phillips B, Stewart LA, Sutton AJ. 2010. 'Cross hairs' plots for diagnostic meta-analysis. *Research Synthesis Methods* **1**: 308–315.

R Development Core Team. 2012. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available at: http://www.R-project.org.

SAS Institute Inc. 2013. SAS Software. Cary, NC, SAS Institute Inc.

Schriger DL, Altmann DG, Vetter JA, Heafner T, Moher D. 2010. Forest plots in reports of systematic reviews: a cross-sectional study reviewing current practice. *International Journal of Clinical Epidemiology* **39**: 421–429.

Schuenemann HJ, Oxman AD, Higgins JPT, Vist GE, Guyatt GH. 2008. Presenting results and 'summary of findings' tables. In Cochrane Handbook for Systematic Reviews and Interventions. Wiley, Hoboken, NJ 335–358.

Schwarzer G. 2012. meta: Meta-Analysis with R. Available at: http://CRAN.R-project.org/package=meta.

Smith LD, Best LA, Stubbs DA, Archibald AB, Roberson-Nay R. 2002. Constructing knowledge: the role of graphs and tables in hard and soft psychology. *American Psychologist* **57**: 749–761.

Spiegelhalter D, Pearson M, Short I. 2011. Visualizing uncertainty about the future. *Science* **333**: 1393–400.

StataCorp. 2011. Stata Statistical Software: Release 12. StataCorp LP, College Station, TX.

Sterne JAC (ed.) 2009. Meta-Analysis in Stata: An Updated Collection from the Stata Journal. Stata Press, College Station, TX.

Stoto MA. 2000. Research synthesis for public health policy: experience of the institute of medicine. In Meta-analysis in Medicine and Health Policy. Stangl DK, Berry DA (eds.) Marcel Dekker, New York 321–357.

Terrin N, Schmid CH, Lau J. 2005. In an empirical evaluation of the funnel plot, researchers could not visually identify publication bias. *Journal of Clinical Epidemiology* **58**: 894–901.

The Cochrane Collaboration. 2011. Review Manager (RevMan) [Computer program]. The Nordic Cochrane Centre, The Cochrane Collaboration, Copenhagen.

Tufte ER. 1983. The Visual Display of Quantitative Information. Graphics Press, Cheshire, CT.

Tufte ER. 1990. Envisioning Information. Graphics Press, Cheshire, CT.

Tufte ER. 1997. Visual Explanations: Images and Quantities, Evidence and Narrative. Graphics Press, Cheshire, CT.

Tufte ER. 2006. Beautiful Evidence. Graphics Press, Cheshire, CT.

Tukey JW. 1977. Exploratory Data Analysis. Addison-Wesley, Reading, MA.

Viechtbauer W. 2010. Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software* **36**: 1–48.

Wang MC, Bushman BJ. 1999. Integrating results through meta-analytic review using SAS software. SAS Institue.

Zacks J, Levy E, Tversky B, Schiano D. 2002. Graphs in print. In Diagrammatic Representation and Reasoning. Anderson M, Meyer B, Olivier P (eds.). Springer, Berlin 187–206.