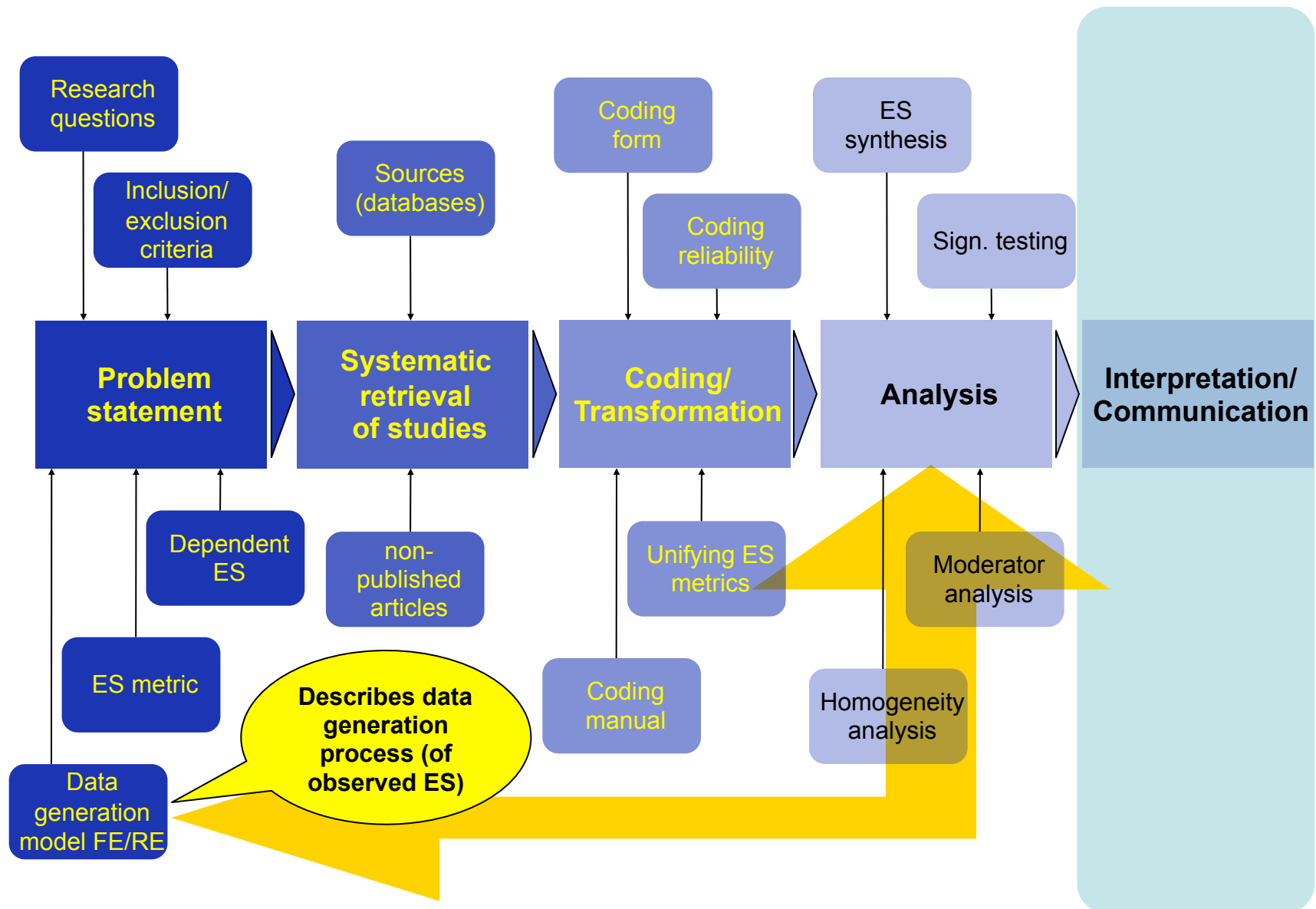# Meta-Analysis Workshop

Part 6: Reporting

Michael Bošnjak

Free University of Bozen-Bolzano, School of Economics and Management

michael.bosnjak@unibz.it, www.bosnjak.eu

Generic Procedure

Agenda

- **MARS: Meta-Analytic Reporting Standards (APA, 2008)**

- Reporting meta-analytic findings using …
  - Tables
  - Graphs

- Exercise: Compare MARS against a published meta-analysis, suggest at least three improvements.

MARS (APA, 2008)

- ´Checklist´of information to be included in a meta-analytic report (structure, content)

- High level of abstraction, no specific recommendations on tables, figures, formula notation, etc.

- Needs to be tailored towards specific audience(s) and content(s)

- Starting point for development: Reporting standards in other fields (QUORUM, PRISMA, MOOSE, etc.), refined and adapted to the needs in Psychology by members of the Society for Research Synthesis Methodology

- See APA (2008) in supplements folder, Table 4, for an overview

Agenda

- MARS: Meta-Analytic Reporting Standards (APA, 2008)

- **Reporting meta-analytic findings using …**
  - **Tables**
  - **Graphs**

- Exercise: Compare MARS against a published meta-analysis, suggest at least three improvements.

Tables and Figures/Graphs

- In general: Array of different approaches found in literature, *tendency to underuse figures*
- Tables/graphs *should communicate substantive messages*
- Tables are use to communicate
  - Raw data: Substantive difference in primary studies by listing (at least) ES and SEs (a) alphabetically or (b) arranged in meaningful subgroup(s)
  - Summary effect(s) and moderator analyses (mean ES and SEs, CIs, etc.)
- Figures / Graphs are used to communicate
  - Raw data and summary effects
  - Visualizing effect size distributions and systematic variations among mean ESs
  - Visualizing meta-regression(s)
  - Assessing publication bias

Example: Table: Raw Data Overview

Table 1 Summary of the 24 papers and 45 cases included in the meta-analysis

Case number	Reference	Web mode compared to:	Web mode: response rate in % (No. of eligible units contacted, no. of responses)		Other mode: response rate in % (No. of eligible units contacted, no. of responses)	
1	Bason (2000)	telephone	15.50	(742, 115)	23.98	(674, 161)
2	Bason (2000)	mail	15.50	(742, 115)	27.76	(735 204)
3	Bason (2000)	IVR	15.50	(742, 115)	17.39	(736, 128)
4	Bates (2001)	mail	55.70	(1571, 875)	44.23	(1569, 694)
5	Chatt & Dennis (2003)	telephone	82.13	(3627, 2979)	62.89	(477, 300)
6	Chisholm (1998)	email	24.00	(300, 72)	30.00	(300, 90)
7	Cobanoglu et al. (2001)	mail	44.21	(95, 42)	26.26	(99, 26)
8	Cobanoglu et al. (2001)	fax	44.21	(95, 42)	17.00	(100, 17)
9	Crawford et al. (2001)	mail	63.00	(3500, 2205)	52.00	(3500, 1820)
10	Elder & Incalcatera (2000)	mail	37.39	(690, 258)	54.26	(693, 376)
11	Fraze et al. (2002)	mail	43.16	(95, 41)	60.00	95, 57
12	Fraze et al. (2002)	email	43.16	(95, 41)	27.37	(95, 26)
13	Fricker et al. (2003)	telephone	51.61	(1058, 546)	97.43	(544, 530)
14	Grigorian et al. (2004)	telephone	61.88	(1941, 1201)	66.82	1941, 1297
15	Grigorian et al. (2004)	mail	61.88	(1941, 1201)	78.00	(27,982, 21,826)
16	Hayslett & Wildemuth (2004)	mail	28.00	(100, 28)	51.00	(100, 51)
17	Hayslett & Wildemuth (2004)	mail	39.00	(100, 39)	51.00	(100, 51)
18	Jones & Pitt (1999)	mail	18.50	(200, 37)	72.00	(100, 72)
19	Jones & Pitt (1999)	email	18.50	(200, 37)	34.00	(200, 68)
20	Kaplowitz et al. (2004)	mail	20.70	(4440, 919)	31.50	(2594, 817)
21	Kaplowitz et al. (2004)	mail	25.40	(4351, 1105)	31.50	(2594, 817)
22	Kaplowitz et al. (2004)	mail	29.70	(4327, 1285)	31.50	(2594, 817)
23	Kaplowitz et al. (2004)	mail	28.60	(4178, 1195)	31.50	(2594, 817)

(continued)

Table 1 (continued)

Case number	Reference	Web mode compared to:	Web mode: response rate in % (No. of eligible units contacted, no. of responses)		Other mode: response rate in % (No. of eligible units contacted, no. of responses)	
24	Kerwin et al. (2004)	mail	37.60	(359, 135)	27.69	(195, 54)
25	Knapp & Kirk (2003)	mail	15.88	(359, 57)	48.47	(359, 174)
26	Knapp & Kirk (2003)	touch-tone	15.88	(359, 57)	33.71	(359, 121)
27	Kwak & Radler (1999)	mail	27.36	(987, 270)	41.92	(990, 415)
28	Lesser & Newton (2001)	mail	18.87	(159, 30)	59.38	(389, 231)
29	Lesser & Newton (2001)	mail	21.89	(233, 51)	59.38	(389, 231)
30	Lesser & Newton (2001)	email	18.87	(159, 30)	39.26	(163, 64)
31	Lesser & Newton (2001)	email	21.89	(233, 51)	39.26	(163, 64)
32	Lesser & Newton (2001)	email	18.87	(159, 30)	52.98	(151, 80)
33	Lesser & Newton (2001)	email	21.89	(233, 51)	52.98	(151, 80)
34	Lozar Manfreda et al. (2001)	mail	77.00	(200, 154)	89.00	(200, 178)
35	Miller et al. (2002)	mail	14.30	(2805, 401)	37.00	(2811, 1040)
36	Miller et al. (2002)	mail	12.86	(2900, 373)	38.01	(2897, 1101)
37	Pötschke (2004)	mail	37.11	(380, 141)	50.75	(402, 204)
38	Sax et al. (2003)	mail	11.13	(737, 82)	10.28	(1478, 152)
39	Vehovar et al. (2001)	telephone	26.00	(300, 78)	51.94	(747, 388)
40	Vehovar et al. (2001)	mail	26.00	(300, 78)	39.19	(222, 87)
41	Vehovar et al. (2001)	fax	26.00	(300, 78)	31.58	(76, 24)
42	Weible & Wallace (1998)	mail	34.44	(151, 52)	35.71	(196, 70)
43	Weible & Wallace (1998)	fax	34.44	(151, 52)	30.86	(162, 50)
44	Weible & Wallace (1998)	email	34.44	(151, 52)	29.81	(161, 48)
45	Wygant & Lindorf (1999)	mail	49.53	(1270, 629)	31.56	(1299, 410)

From: Lozar Manfreda, Bosnjak, Berzelak, Haas & Vehovar (2008)

Example: Table: ES Overview

Table III

Atypical versus typical antipsychotics effects

Cognitive domain	k	N	g	95% CI	Z	p
Attention and vigilance	12	1574	0.06	-0.04 , 0.15	1.20	0.23
Automaticity and procedural learning	2	85	0.27	-0.13 , 0.67	1.31	0.19
General intellectual functioning	3	184	0.16	-0.21 , 0.52	0.84	0.40
Language and verbal comprehension	3	179	0.38	0.15 , 0.62	3.22	<0.01
Perceptual processing	2	136	0.04	-0.50 , 0.58	0.14	0.89
Psychomotricity	5	387	0.29	0.11 , 0.47	3.13	<0.01
Reasoning and problem solving	12	1569	0.07	-0.07 , 0.21	0.96	0.33
Speed of processing	11	909	0.26	0.13 , 0.39	3.99	<0.01
Verbal learning and memory	11	1615	0.19	-0.01 , 0.39	1.84	0.07
Visual learning and memory	6	725	0.16	-0.05 , 0.38	1.54	0.12
Working memory	8	1167	0.02	-0.10 , 0.15	0.33	0.74
Global cognitive index	18	1808	0.17	0.04 , 0.29	2.61	<0.01

k: number of studies; N: sample size; CI: confidence interval.

From: Guilera et al. (2009)

Example: Table: ES subgroup analysis I

Table 2 Summary of seven categorical moderator analyses predicting the response rate differences between web and other survey modes

Moderator variable	Categories (and number of cases)	Mean response difference estimate	95% CI	Q_B –test (Q for between categories)
Type of mode compared to	Mail (27)	−0.12	−0.17/−0.05	$Q_B = 4.52, df = 3, p = 0.21$
	Email (8)	−0.13	−0.27/0.00	
	Telephone (5)	−0.13	−0.32/0.06	
	Fax (3)	0.08	−0.32/0.48	
	Other (2)*			
Sample recruitment strategy	Panel/pre-recruited list (40)	−0.09	−0.14/−0.05	$Q_B = 7.18, df = 2, p = 0.01$
	One-time recruitment (4)	−0.28	−0.49/−0.07	
	Other (1)*			
Target population	Students (13)	−0.06	−0.14/0.02	$Q_B = 3.12, df = 2, p = 0.21$
	Employees/members of associations (20)	−0.12	−0.19/−0.06	
	General population (4)	−0.19	−0.40/0.03	
	Other (8)*			
Type of sponsorship	Academic (36)	−0.12	−0.17/−0.07	$Q_B = 1.68, df = 2, p = 0.43$
	Governmental (6)	−0.08	−0.24/0.07	
	Commercial (3)	−0.01	−0.39/0.36	
Solicitation mode	Mail (17)	−0.15	−0.21/−0.09	$Q_B = 6.69, df = 1, p = 0.01$
	Email (25)	−0.05	−0.10/0.00	
	Other (3)*			
Incentive	Yes (3)	−0.17	−0.55/0.21	$Q_B = 0.57, df = 1, p = 0.45$
	No (42)	−0.10	−0.15/−0.05	
Number of contacts	One–two (23)	−0.05	−0.11/0.01	$Q_B = 7.56, df = 1, p = 0.01$
	Three–five (22)	−0.16	−0.23/−0.10	

* Other categories dropped from the homogeneity analysis.

From: Lozar Manfreda, Bosnjak, Berzelak, Haas & Vehovar (2008)

Example: Table: ES subgroup analysis II

Table 3: Summary of seven categorical moderator analyses aimed at explaining in-homogeneity among the 211 self-congruity effect sizes

Moderator Variable	Categories (and number of effect sizes, k)	Fixed effects models			Random effects models (mixed effects)		
		Mean Z_r-score (r-Score)	95% CI of Z_r-score	Q_B	Mean Z_r-score (r-Score)	95% CI of Z_r-score	Q_B
Brand image facet (testing H$_1$)	Brand personality (129) Brand-user image (80) Other* (2)	.4146 (.3924) .3313 (.3197)	(.4023 / .4269) (.3162 / .3464)	$Q = 72.03$ $df = 1$ $p < .01$	.4175 (.3948) .3483 (.3349)	(.3816 / .4533) (.3033 / .3932)	$Q = 5.71$ $df = 1$ $p < .05$
Type of product stimulus (testing H$_2$)	Established brand name (122) New (fictitious) unknown brand (9) Product class name (3) Other* (2)	.3126 (.3028) .2521 (.2469) .4538 (.4250)	(.2983 / .3268) (.2000 / .3042) (.4403 / .4672)	$Q = 238.84$ $df = 2$ $p < .01$	.3120 (.3020) .2698 (.2634) .5177 (.4759)	(.2764 / .3475) (.1229 / .4167) (.4747 / .5607)	$Q = 57.50$ $df = 2$ $p < .01$
Degree of product involvement (testing H$_3$)	High (119) Low (87) Other* (5)	.3174 (.3072) .4366 (.4108)	(.3033 / .3315) (.4233 / .4499)	$Q = 148.61$ $df = 1$ $p < .01$	.3261 (.3150) .4751 (.4423)	(.2892 / .3631) (.4329 / .5173)	$Q = 27.81$ $df = 1$ $p < .01$
Self-congruity measure (testing H$_4$)	Indirect (164) Direct (47)	.3800 (.3527) .3767 (.3598)	(.3691 / .3910) (.3583 / .3951)	$Q = .10$ $df = 1$ $p = .75$	.3843 (.3664) .4085 (.3872)	(.3526 / .4161) (.3475 / .4696)	$Q = .50$ $df = 1$ $p = .48$
Type of dependent variable (testing H$_5$)	Attitudes (93) Intentions (82) Attitude/Intention composite (19) Behavior (17)	.3392 (.3268) .3972 (.3776) .4879 (.4525) .3334 (.3216)	(.3231 / .3553) (.3843 / .4101) (.4477 / .5281) (.2520 / .3749)	$Q = 69.84$ $df = 3$ $p < .01$	.3340 (.3221) .4130 (.3910) .5984 (.5359) .3419 (.3292)	(.2915 / .3765) (.3693 / .4567) (.4986 / .6982) (.2362 / .4475)	$Q = 28.14$ $df = 3$ $p < .01$
Type of independent variable (testing H$_6$)	Private-type facets (200) Public-type facets (7) Other* (4)	.3864 (.3683) .2777 (.2708)	(.3767 / .3962) (.2220 / .3333)	$Q = 21.82$ $df = 1$ $p < .01$	.3962 (.3767) .2641 (.2581)	(.3675 / .4250) (.0748 / .4542)	$Q = 2.79$ $df = 1$ $p = .09$

Note. * "Other" categories dropped from homogeneity analysis.

From: Rodriguez, Bosnjak & Sirgy (in press)

Example: Table: ES subgroup analysis III

Table II
Meta-analyses of studies comparing the effects of individual psychological treatments to group treatments for depression: Overall results and subgroup analyses

Study		N_{comp}	d	95% CI	Z	Q [a]	I^2 (%)	p
ALL STUDIES								
Overall effects								
All studies	REM [b]	19	0.20	0.05 0.35	2.60 **	16.46	0	
1 study excluded [c]	REM	18	0.23	0.08 0.39	2.92 **	13.83	0	
BDI [d]	REM	19	0.28	0.12 0.44	3.38 **	19.87	9.40	
Subgroup analyses								
CBT in both formats	Yes	11	0.15	-0.07 0.38	1.32	10.15	1.46	n.s.
	No	8	0.24	0.04 0.44	2.31 *	5.99	0	
Recruitment	Community	8	0.28	-0.02 0.58	1.82	5.29	0	n.s.
	Clinical	8	0.17	-0.10 0.44	1.25	8.58	18.43	
	Other	3	0.18	-0.10 0.45	1.27	2.23	10.37	
Depressive disorder [e]	Yes	14	0.17	0.01 0.34	2.02 *	11.80	0	n.s.
	No	5	0.34	-0.02 0.71	1.84 o	3.95	0	
Specific population [f]	Yes	4	0.20	-0.04 0.44	1.62	2.70	0	n.s
	No	15	0.20	0.01 0.39	2.03 *	13.77	0	
Analyses	Compl-only [g]	12	0.29	0.06 0.51	2.51 *	9.31	0	
	ITT	5	0.08	-0.16 0.32	0.67	4.39	8.95	n.s.
Effects at follow-up								
1-3 months follow-up	REM	7	0.05	-0.21 0.32	0.38	7.89	23.99	
6 months follow-up	REM	7	-0.17	-0.53 0.19	-0.92	9.89	39.31	
STUDIES WITH AN EQUIVALENT INTERVENTION IN BOTH FORMATS								
Overall effects								
All studies	REM	13	0.18	-0.01 0.37	1.82 o	10.31	0	
1 study excluded [c]	REM	12	0.23	0.03 0.42	1.22 *	7.83	0	
BDI [d]	REM	13	0.27	0.05 0.48	2.38 *	14.69	18.30	
Subgroup analyses								
CBT in both formats	Yes	11	0.15	-0.07 0.38	1.32	10.15	1.46	n.s.
	No	2	0.24	-0.12 0.59	1.31	0.01	0	
Recruitment	Community	7	0.30	0.01 0.62	1.88 o	5.05	0	n.s

From: Cuijpers et al. (2008)

Example: Graph: Stem-and-Leaf Plot

```
Stem Leaf                             #
  14 3                                1
  12 3                                1
  10 4                                1
   8 4494                             4
   6 16727                            5
   4 06937                            5
   2 014556913466                    12
   0 357115579                        9
  -0 966644331999888775332           21
  -2 7320008763300                   13
  -4 832188640                        9
  -6 82953                            5
  -8 7183                             4
 -10 70840                            5
 -12 1                                1
 -14 84                               2
     ----+----+----+----+-
Multiply Stem.Leaf by 10**-1
```

Figure 2. Stem-and-Leaf Plot of the Data Set

From: Zhao et al. (2005)

Example: Graph: Histogram and Box-Plot
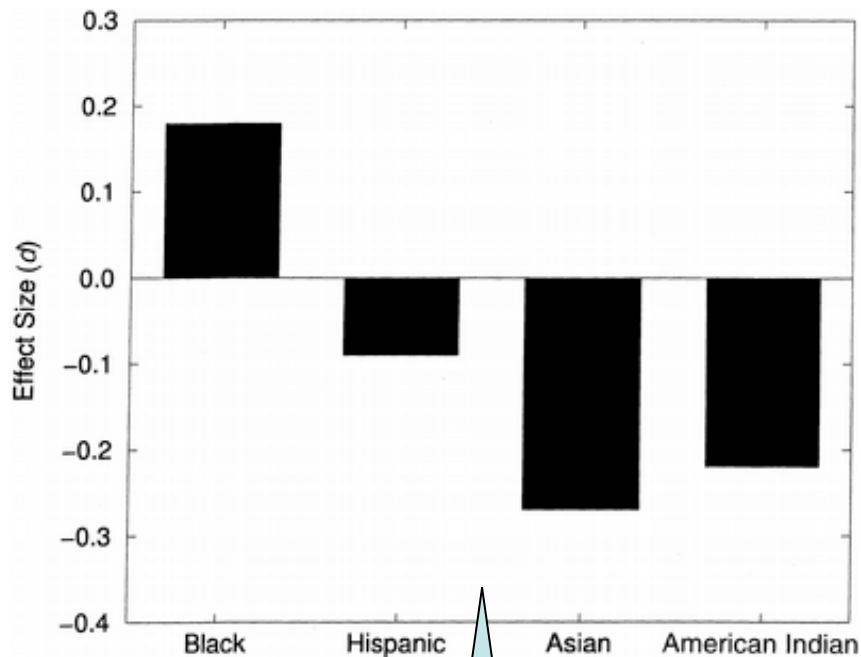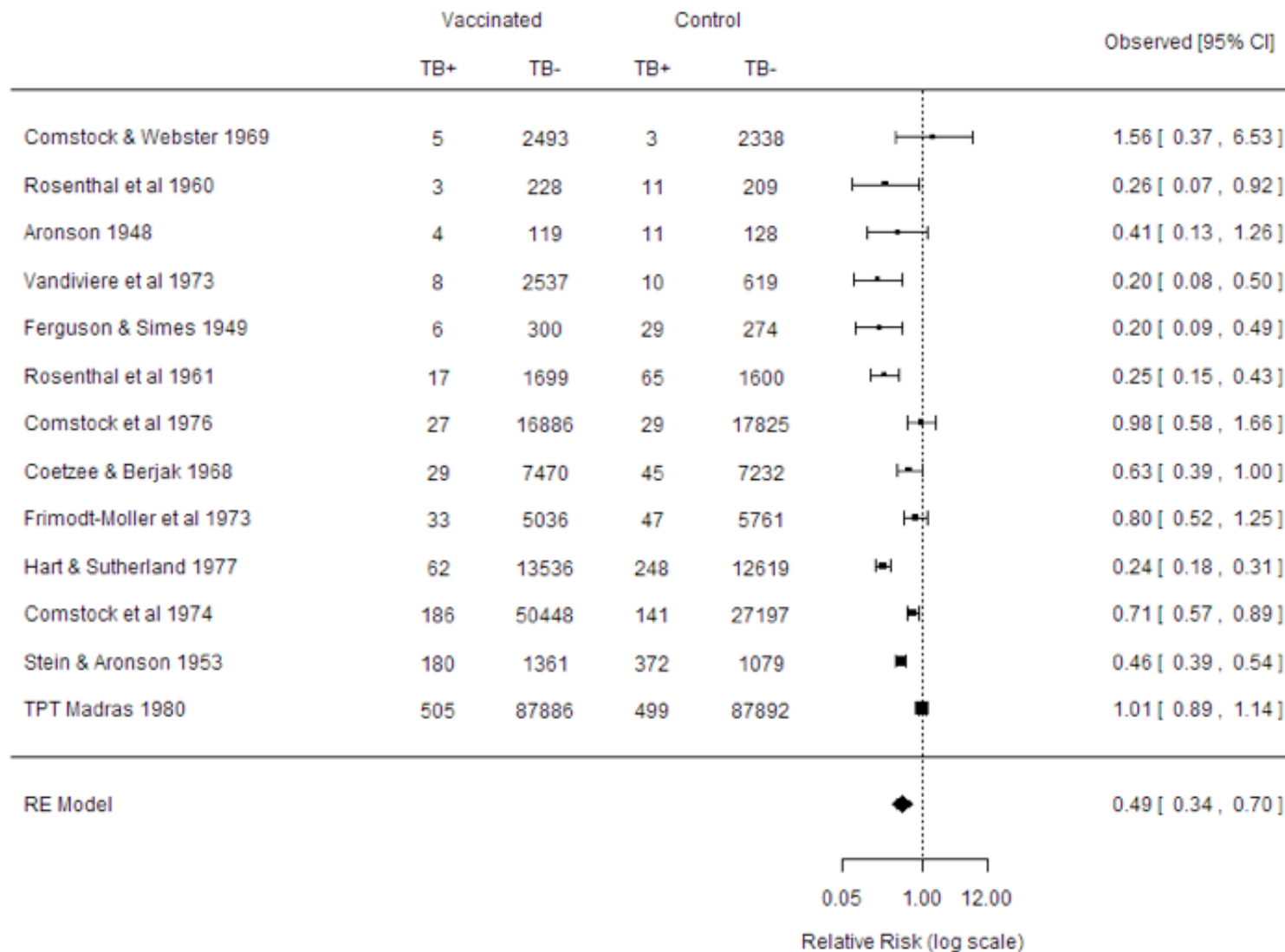

CIs missing!
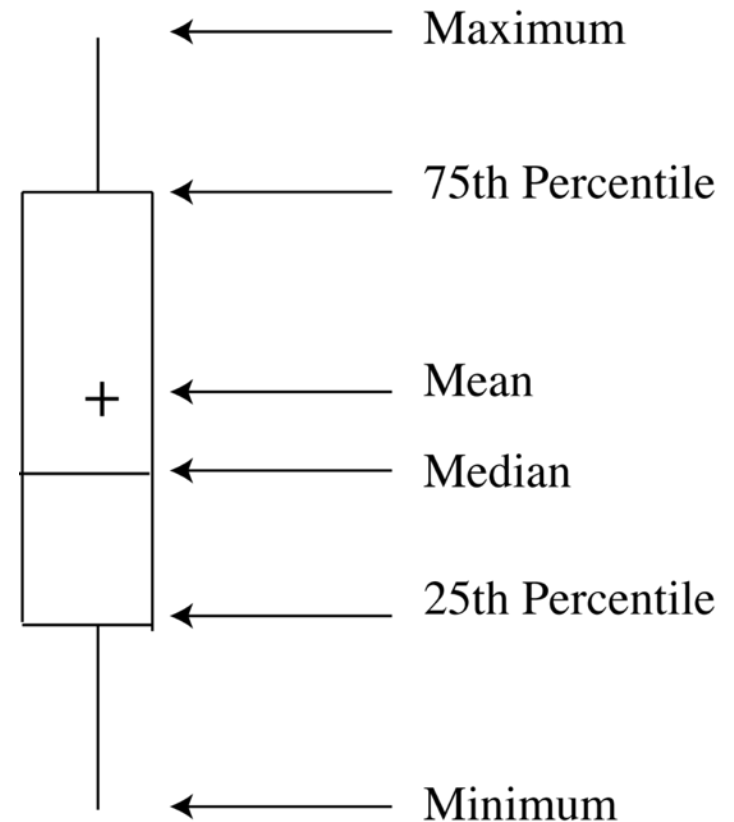
Example: Graph: Forest Plot
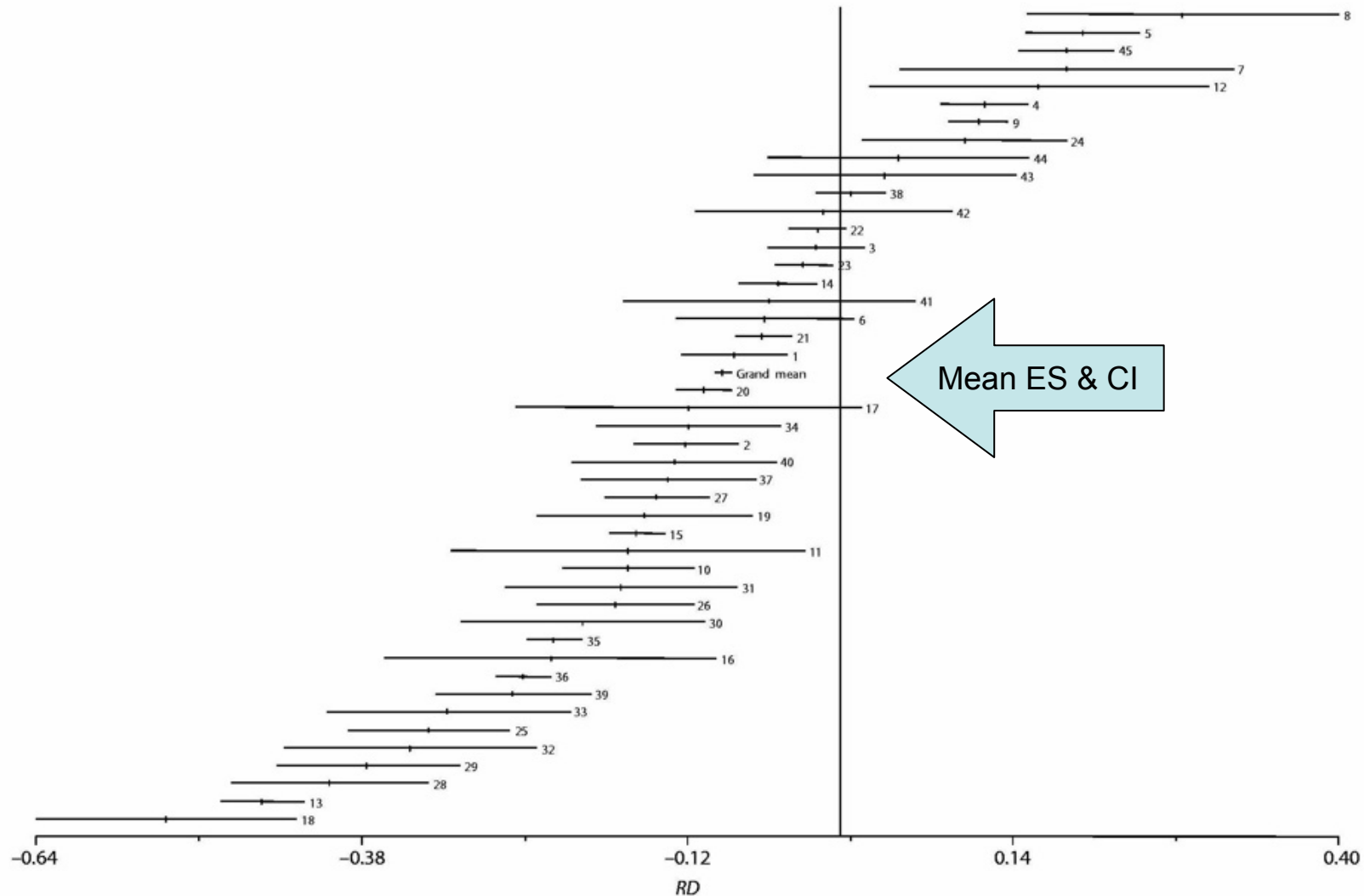
Figure 1 Distribution of effect sizes (response rate differences, *RD*) and their 95% confidence intervals (based on 45 comparisons between web survey and other survey modes; study numbers correspond to those reported in Table 1)

From: Lozar Manfreda, Bosnjak, Berzelak, Haas & Vehovar (2008)

Example: Graph: Forest Plot (annotated)

	Vaccinated		Control			Observed [95% CI]
	TB+	TB-	TB+	TB-		
Comstock & Webster 1969	5	2493	3	2338		1.56 [0.37 , 6.53]
Rosenthal et al 1960	3	228	11	209		0.26 [0.07 , 0.92]
Aronson 1948	4	119	11	128		0.41 [0.13 , 1.26]
Vandiviere et al 1973	8	2537	10	619		0.20 [0.08 , 0.50]
Ferguson & Simes 1949	6	300	29	274		0.20 [0.09 , 0.49]
Rosenthal et al 1961	17	1699	65	1600		0.25 [0.15 , 0.43]
Comstock et al 1976	27	16886	29	17825		0.98 [0.58 , 1.66]
Coetzee & Berjak 1968	29	7470	45	7232		0.63 [0.39 , 1.00]
Frimodt-Moller et al 1973	33	5036	47	5761		0.80 [0.52 , 1.25]
Hart & Sutherland 1977	62	13536	248	12619		0.24 [0.18 , 0.31]
Comstock et al 1974	186	50448	141	27197		0.71 [0.57 , 0.89]
Stein & Aronson 1953	180	1361	372	1079		0.46 [0.39 , 0.54]
TPT Madras 1980	505	87886	499	87892		1.01 [0.89 , 1.14]
RE Model						0.49 [0.34 , 0.70]

0.05   1.00   12.00

Relative Risk (log scale)

From: Wolfgang Viechtbauer, http://www.wvbauer.com/downloads.html

Example: Graph/Table: Forest Plot & Table

Study name	Statistics for each study					Std diff in means and 95% CI
	Std diff in means	Lower limit	Upper limit	Z-Value	p-Value	
Banken, 1993	0,05	-0,92	1,02	0,10	0,92	
Brown, 1984	-0,14	-0,81	0,53	-0,41	0,68	
Milgrom, 2005 A	0,00	-0,38	0,38	0,00	1,00	
Milgrom, 2005 B	0,24	-0,14	0,62	1,25	0,21	
Mohr, 2001	0,53	-0,11	1,17	1,63	0,10	
Ross, 1985	0,00	-0,56	0,56	0,00	1,00	
Rush, 1981	0,65	-0,17	1,47	1,55	0,12	
Schmidt, 1983 A	0,35	-0,47	1,17	0,83	0,41	
Schmidt, 1983 B	1,06	0,19	1,93	2,38	0,02	
Scott, 1990 A	-0,29	-0,90	0,32	-0,93	0,35	
Scott, 1990 B	-0,10	-0,80	0,60	-0,28	0,78	
Shapiro, 1982 A	0,13	-0,71	0,97	0,30	0,76	
Shapiro, 1982 B	-0,08	-0,86	0,70	-0,20	0,84	
Teri, 1986	0,67	0,13	1,21	2,41	0,02	
Wierzbicki, 1987	0,49	-0,45	1,43	1,02	0,31	
Wilson, 1990?	0,68	-0,59	1,95	1,05	0,30	
Wollersheim, 1991 A	0,20	-0,78	1,18	0,40	0,69	
Wollersheim, 1991 B	0,05	-0,93	1,03	0,10	0,92	
Zettle, 1992	0,36	-0,40	1,12	0,93	0,35	
	0,20	0,05	0,35	2,60	0,01	

-2,00   -1,00   0,00   1,00   2,00

Favours GRP   Favours IND

Figure 1. Standardized effect sizes indicating the difference in depressive symptoms between individual and group therapies for depression at post-test.

From: Cuijpers et al. (2008)

Example: Graph: Radial Plot

$z_I = \dfrac{y_I}{\sqrt{v_I}}$

0.45
0.10
-0.24
-0.59
-0.93
-1.28
-1.62

ES scale

$x_I = 1/\sqrt{v_I}$

ES precision

From: Wolfgang Viechtbauer, http://www.wvbauer.com/downloads.html
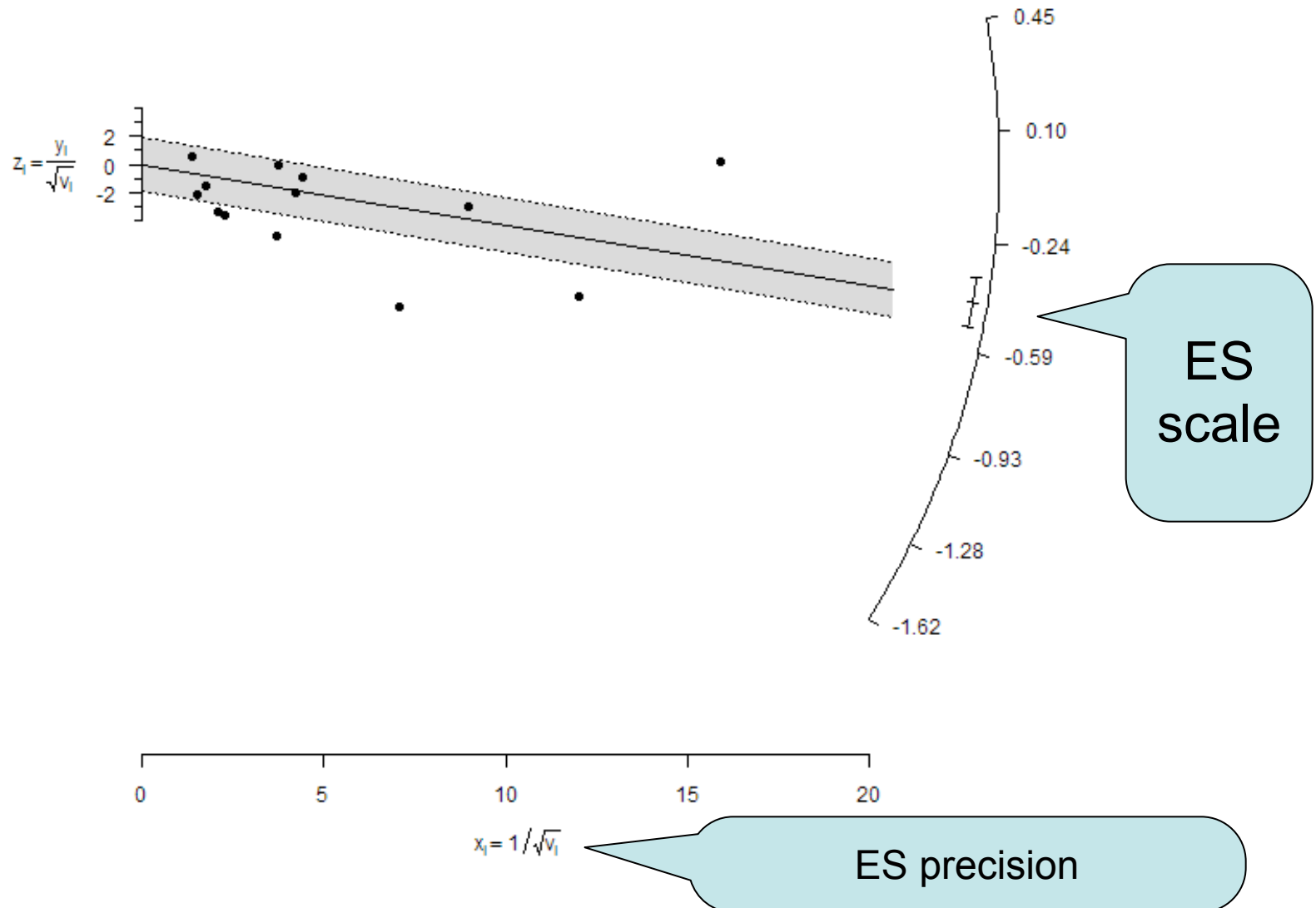
Example: Graph: Scatterplots (Meta-Reg)
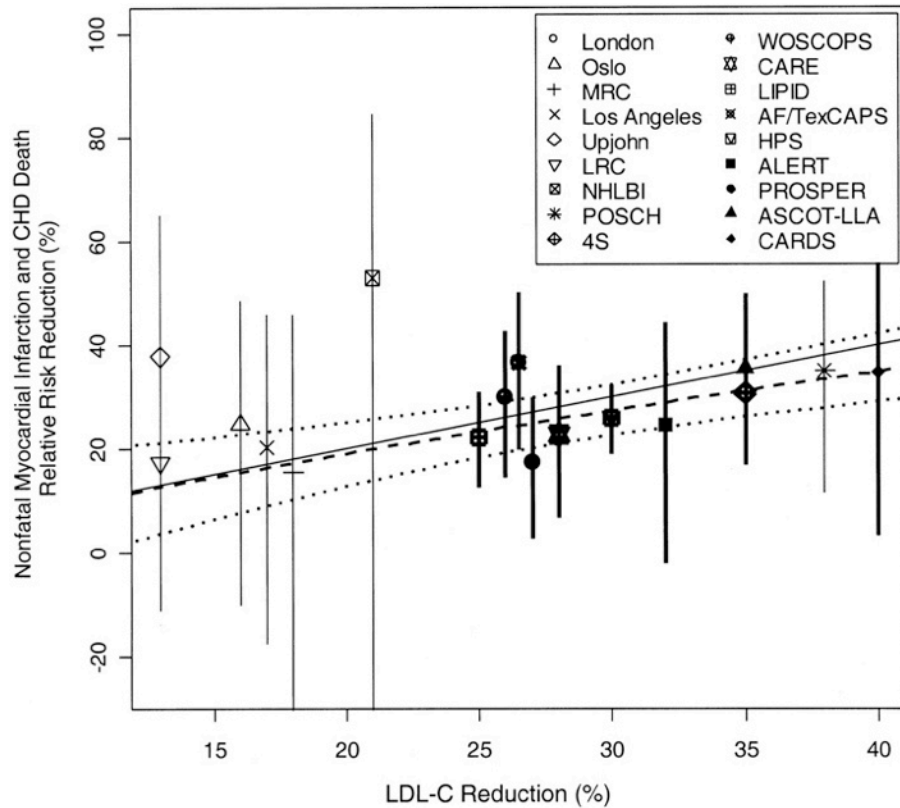
From: Robinson et al. (2005)
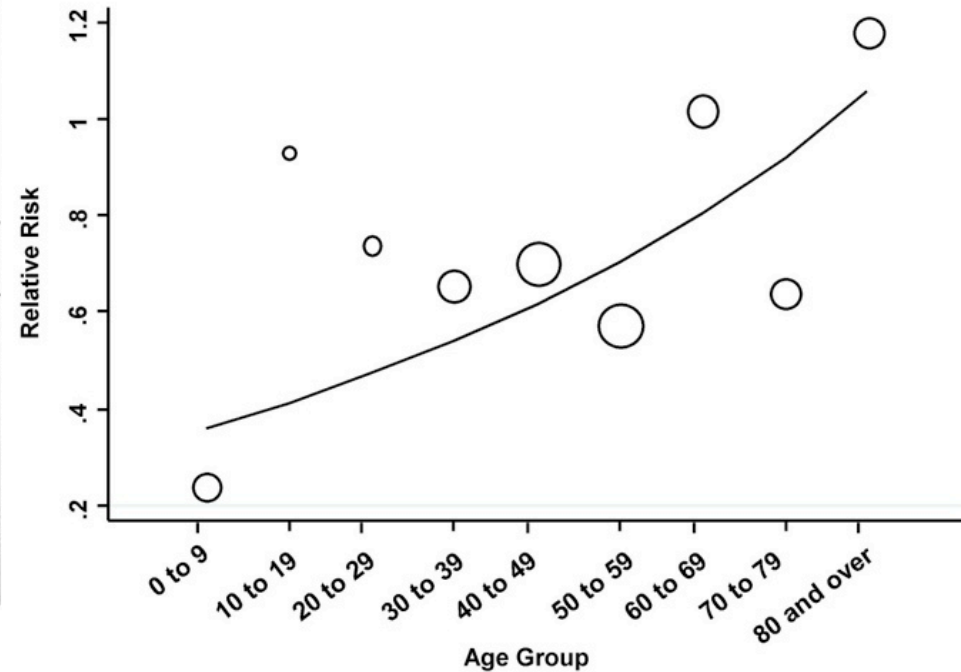
From: White et al. (2009)

Example: Graph: Funnel Plot(s) I
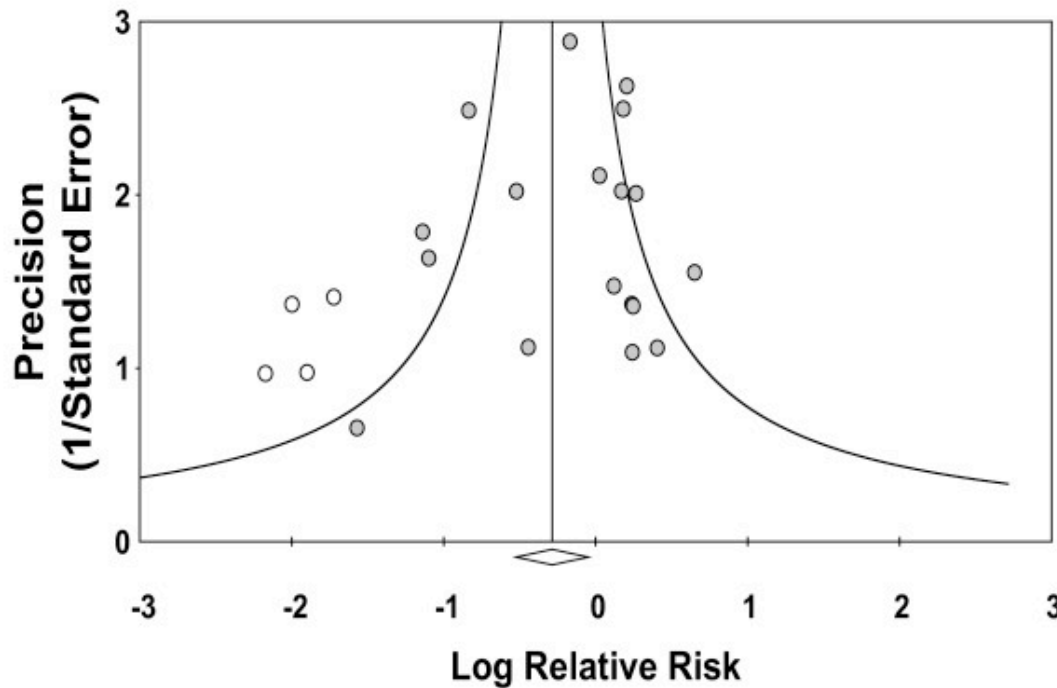
From: Shacher (2008)

From: Gonzales et al. (2007)
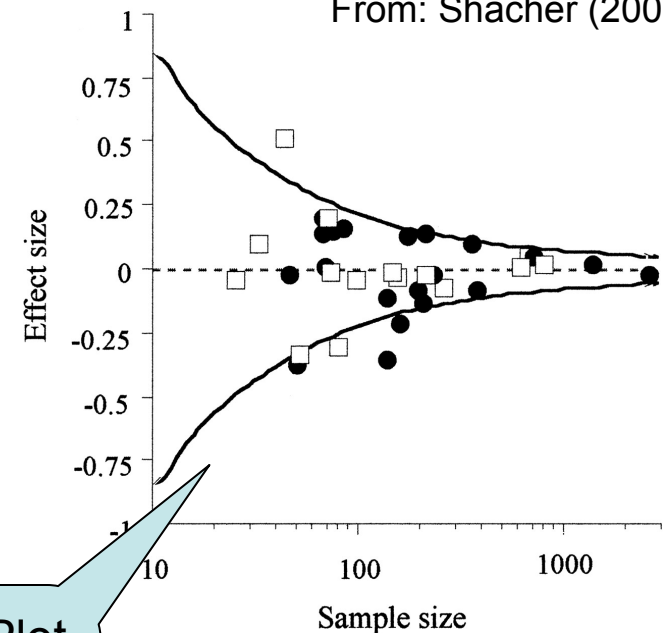
Galbraith Plot

From: Brown & Silk (2002)

Example: Graph: Funnel Plot II
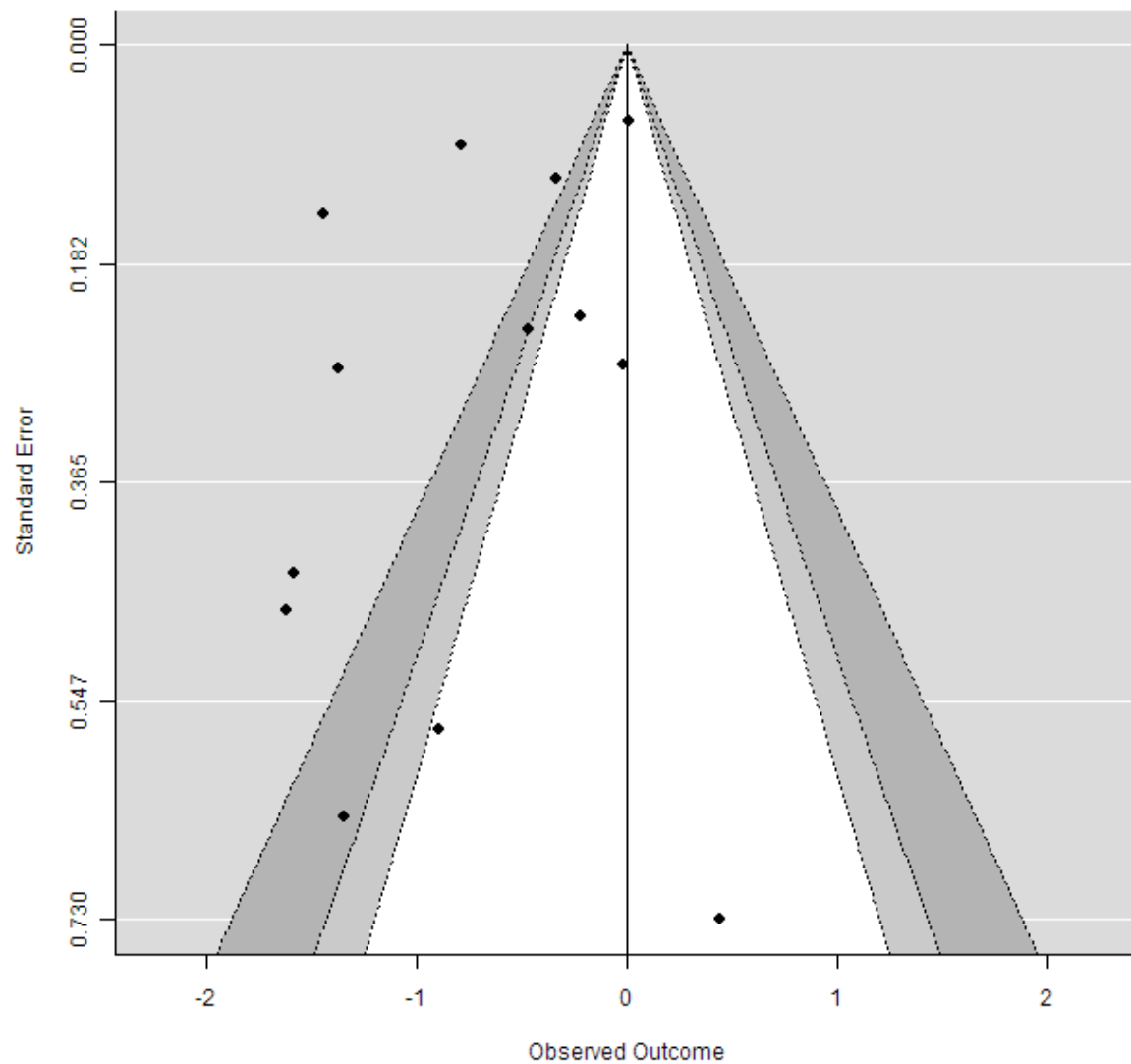
Example: Graph: Normal-Quantile-Plot I

From: Wolfgang Viechtbauer, http://www.wvbauer.com/downloads.html
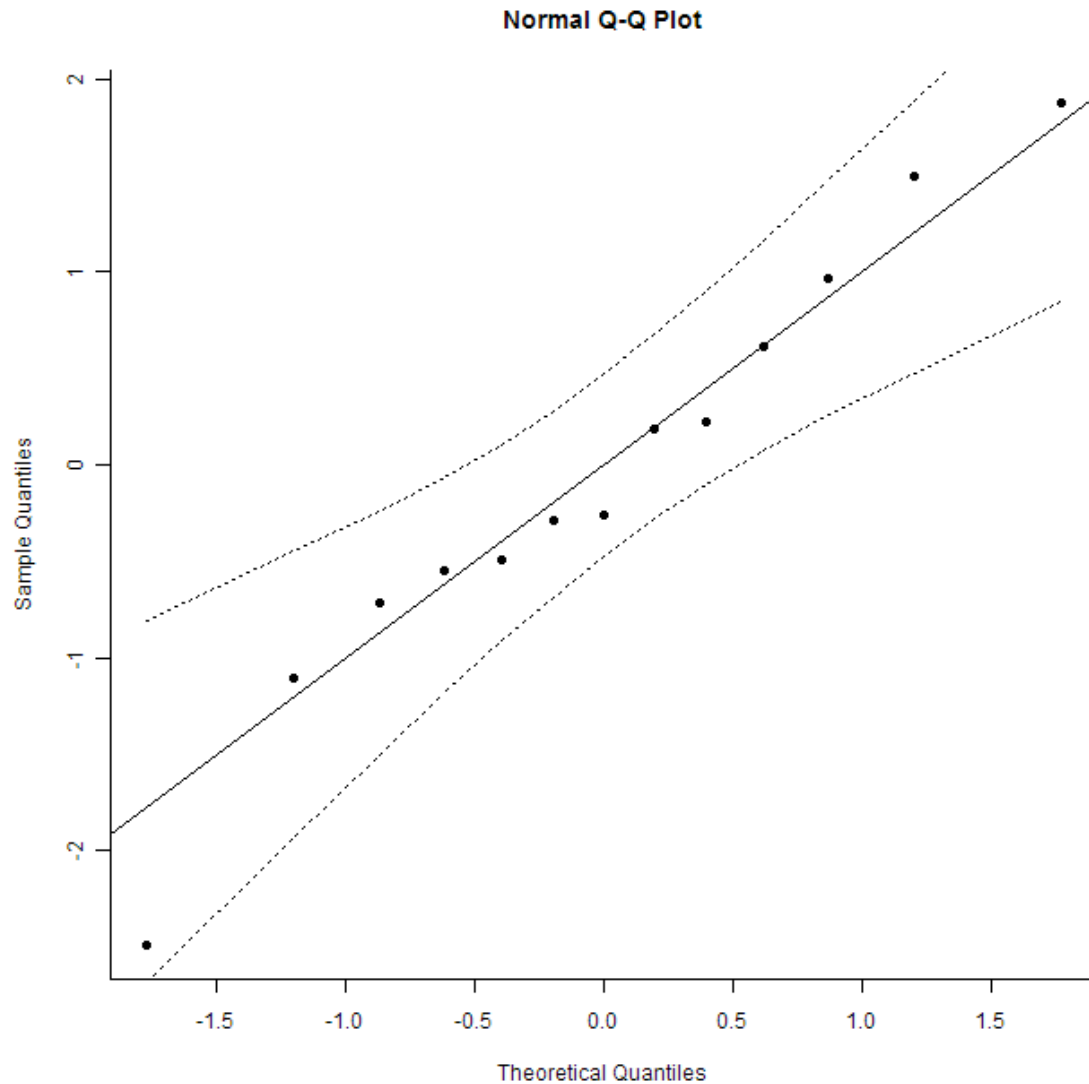
Example: Graph: Normal-Quantile-Plot II
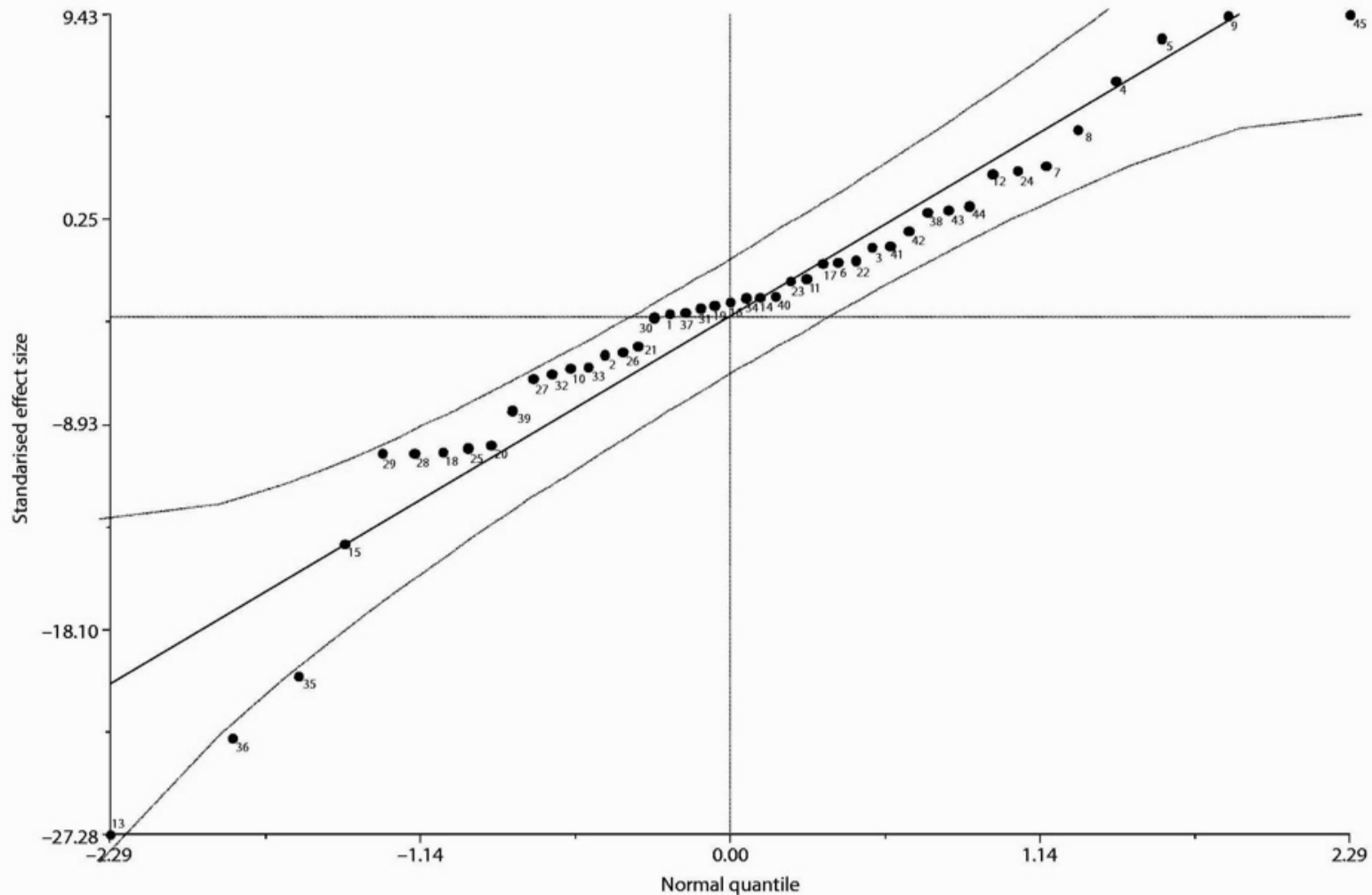
Figure 2 Normal quantile plot to detect any publication bias (based on 45 comparisons between web survey and other survey modes; study numbers correspond to those reported in Table 1)

From: Lozar Manfreda, Bosnjak, Berzelak, Haas & Vehovar (2008)

Example: Graph: L'Abbé plot
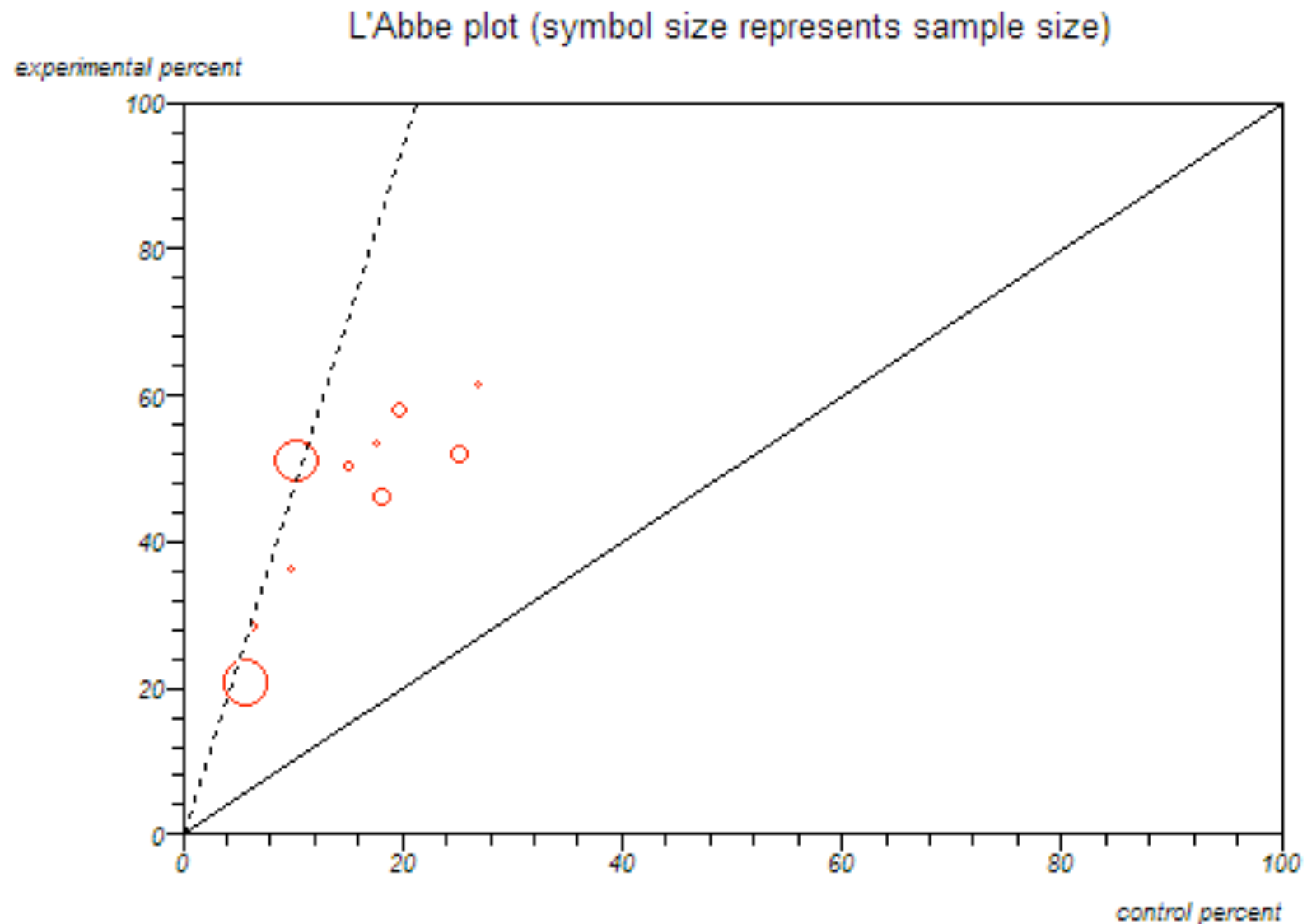
L'Abbe plot (symbol size represents sample size)

Agenda

- MARS: Meta-Analytic Reporting Standards (APA, 2008)

- Reporting meta-analytic findings using ...
  - Tables
  - Graphs

- **Exercise: Compare MARS against a published meta-analysis, suggest at least three improvements.**

Exercise: Reporting

- Please identify at least three aspects to be improved in the *abstract* and *results* section of a published meta-analysis (by using MARS as a guideline).

- Overall 22 meta-analyses in ´Literature´ folder:
  - Please check the overview first for topics of interest
  - Please read abstract and scan results section
  - Compare abstract and results against MARS guidelines
  - Suggest at least three improvements

Lipsey & Wilson (2001)
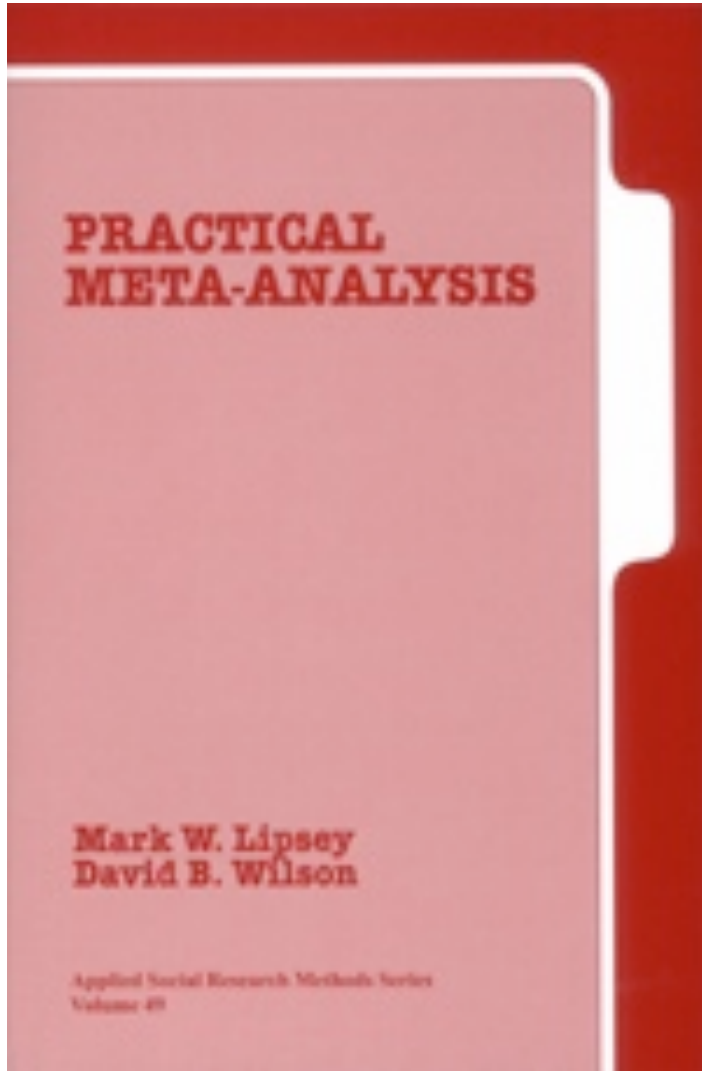
Lipsey, M.W., & Wilson, D.B.(2001). *Practical Meta-analysis.* Thousand Oaks: Sage.

Chapter 7: Computational Techniques for Meta-Analysis Data, Subchapter on Graphing Techniques

Cooper (2010)
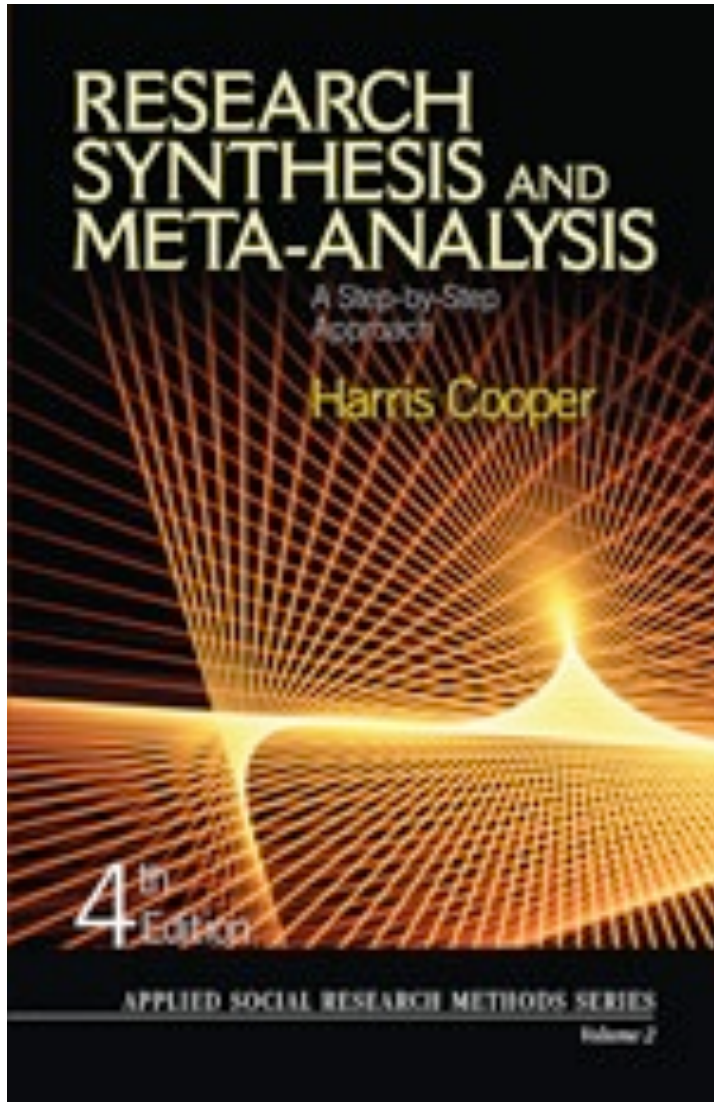
Cooper, H. (2010). *Research Synthesis and Meta-Analysis: A Step-by-Step Approach.* Thousand Oaks, CA: Sage.
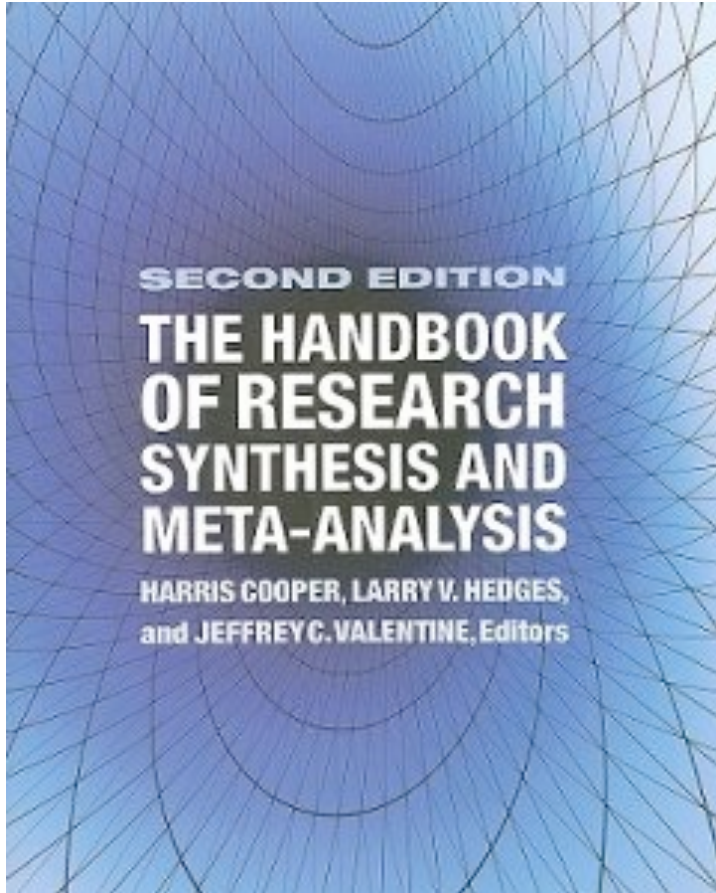
- Chapter 8: Step 7: Presenting the results
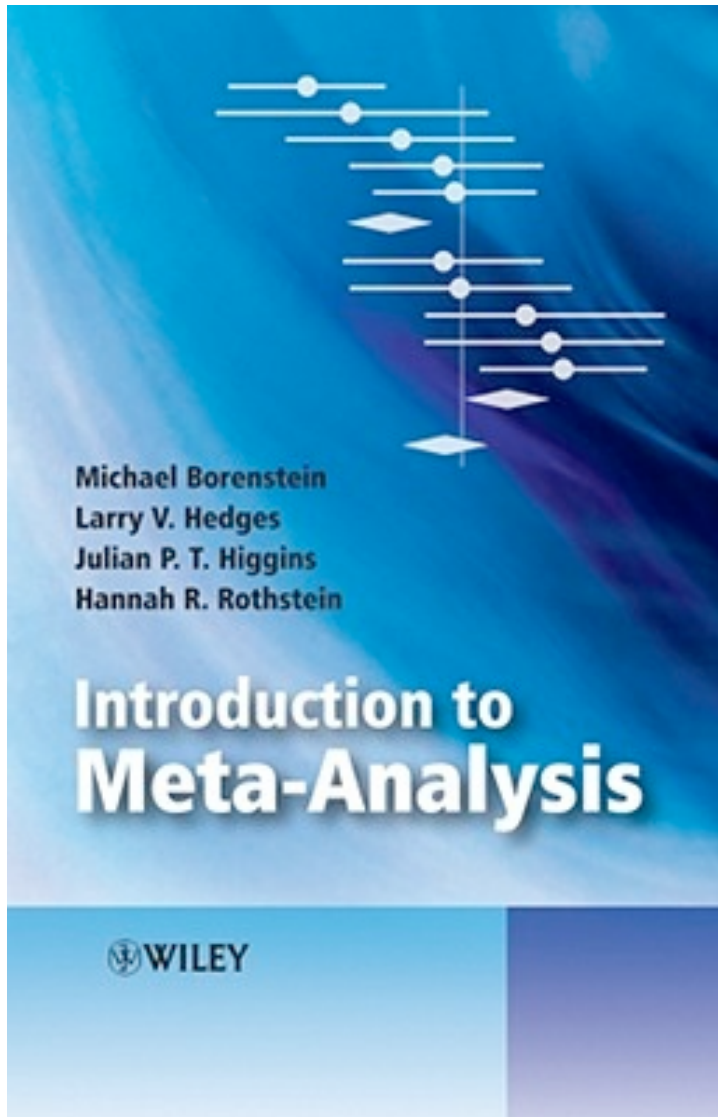
Borenstein, M., Hedges, L.V., Higgins, J.P.T, & Rothstein, H.R. (2009). *Introduction to Meta-Analysis*. Chichester, UK: Wiley.

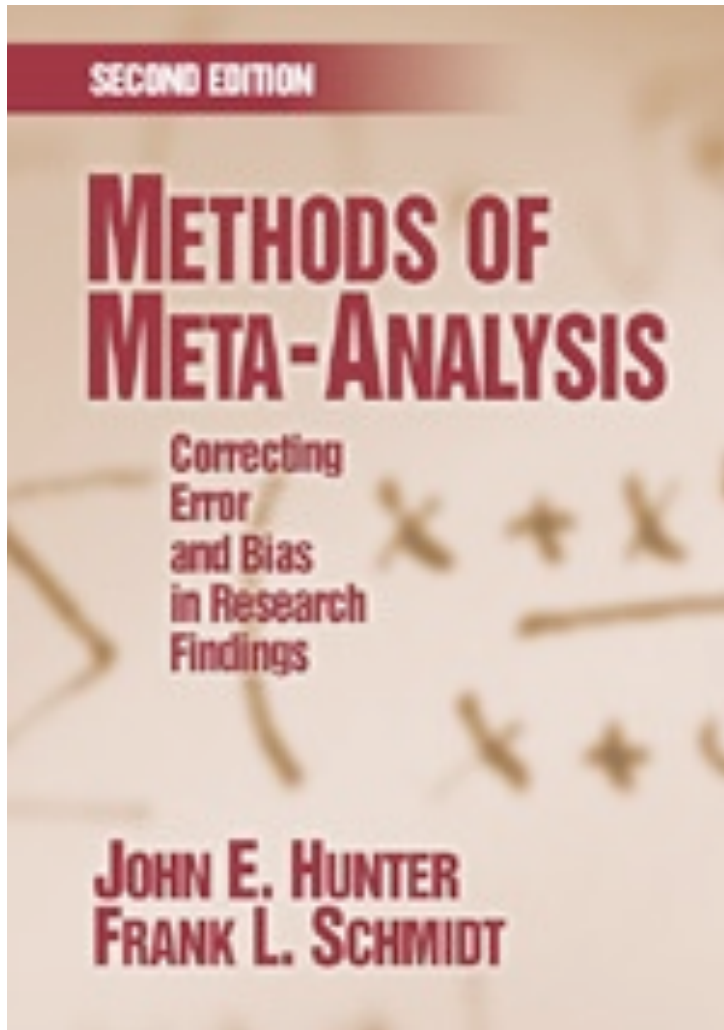● Chapter 41: Reporting the results of a meta-analysis

Cooper, Hedges & Valentine (2009)

Cooper, H., Hedges, L.V., & Valentine, J.C. (Eds.) (2009). *Handbook of Research Synthesis (2nd ed.)*. New York: Russell Sage Foundation.

● Chapter 26: Visual and narrative interpretation (by Geoffrey Borman and Jeffrey Grigg)
● Chapter 27: Reporting format (by Mike Clarke)

Hunter & Schmidt (2004)

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

● No specific chapter on reporting formats/standards.