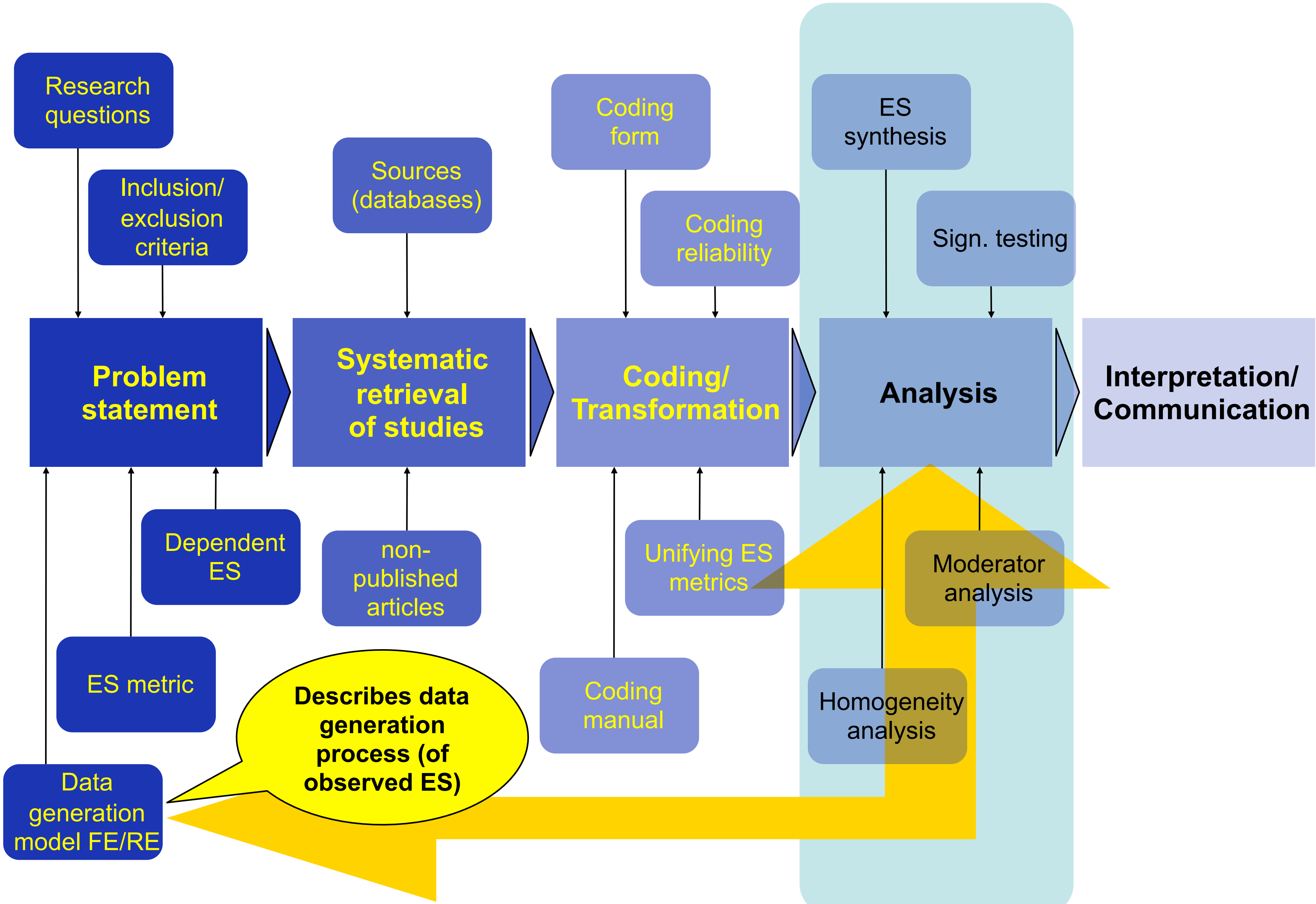
Meta-Analysis Workshop

Part 4: Analysis and Interpretation

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

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

Generic Procedure



Big Picture I: Meta-Analytic Models

Two broad data (= effect size) generation models and corresponding *inference goals* in (HO-based) meta-analysis can be distinguished (see also part 1):

- FIXED (FE) effects models:
  - *Data generation model:* One single (= fixed) true effect size does exist in the universe of studies -> Observed effect sizes are sample estimates of this single true score, confounded by subject-level sampling error.
  - *One-stage sampling process assumed* (only subject-level sampling error).
  - *Inference goal* under FE assumption is *conditional* upon the studies included in the meta-analysis (and to a set of studies identical to those included, except for sampling error).
- RANDOM (RE) effects model:
  - *Data generation model:* A (hyper-)distribution of related, but randomly distributed true effects does exist in the universe of studies -> Observed effect sizes are sample estimates of this (hyper-)distribution of true scores, confounded by subject-level sampling error AND study-level sampling error.
  - *Two-stage sampling process assumed* (subject-level *AND* randomly distributed between-study-level sampling error).
  - *Inference goal* under RE assumption is **un**conditional upon the studies included in the analysis, i.e. generalizable beyond the observed studies.

Big Picture II: Basic Steps

1. Decision about the inferential goal and corresponding meta-analytic model (FE *or* RE model)
2. Estimating the mean effect size and its standard error (for FE *or* RE model)
3. Significance testing of the estimated mean effect size (for FE *or* RE model)
4. Homogeneity testing: Identification and quantification of heterogeneity (for FE *or* RE model)
5. If *systematic* heterogeneity is *assumed*, i.e. sampling variation is not only due to random sampling error(s): Moderator analyses, *ideally guided by a-priori theoretical considerations*, are performed. (for FE *or* RE model)

Agenda

- Step-by-step analysis using MS Excel (following the HO tradition, exercise from Lipsey & Wilson, 2001)
  – FE Analysis
  – RE Analysis
- Special analysis issues:
  – Homogeneity / heterogeneity indicators
  – Sensitivity analyses
- Interpreting results
  – Theory testing/development (SICT)
  – Description of a research field (Web nonresponse)
  – Estimating the effectiveness of (here: Marketing) interventions (Fluency-effects)

Agenda

- **Step-by-step analysis using MS Excel (following the HO tradition, exercise from Lipsey & Wilson, 2001)**
  - **FE Analysis**
  - **RE Analysis**
- Special analysis issues:
  - Homogeneity / heterogeneity indicators
  - Sensitivity analyses
- Interpreting results
  - Theory testing/development (SICT)
  - Description of a research field (Web nonresponse)
  - Estimating the effectiveness of (here: Marketing) interventions (Fluency-effects)

Data Set to be Analyzed

- We have an independent set of effect sizes (ES) that have been transformed and/or adjusted, if needed.
- For each effect size we have an inverse variance weight (w).

Study	ES	w
1	-0.33	11.91
2	0.32	28.57
3	0.39	58.82
4	0.31	29.41
5	0.17	13.89
6	0.64	8.55
7	-0.33	9.80
8	0.15	10.75
9	-0.02	83.33
10	0.00	14.93

For instance, imagine Zr transformed correlations

$$ES_{Zr} = .5\ln\left[\frac{1+r}{1-r}\right]$$

$$se = \sqrt{\frac{1}{n-3}}$$

$$w = \frac{1}{SE^2} \qquad w = n-3$$

The Weighted Mean Effect Size

$$\overline{ES} = \frac{\sum (w \times ES)}{\sum w}$$

Study	ES	w
1	-0.33	11.91
2	0.32	28.57
3	0.39	58.82
4	0.31	29.41
5	0.17	13.89
6	0.64	8.55
7	-0.33	9.80
8	0.15	10.75
9	-0.02	83.33
10	0.00	14.93

- Start with the effect size (ES) and inverse variance weight (w) for 10 studies.

The Weighted Mean Effect Size

$$\overline{ES} = \frac{\sum (w \times ES)}{\sum w}$$

Study	ES	w	w*ES
1	-0.33	11.91	-3.93
2	0.32	28.57	
3	0.39	58.82	
4	0.31	29.41	
5	0.17	13.89	
6	0.64	8.55	
7	-0.33	9.80	
8	0.15	10.75	
9	-0.02	83.33	
10	0.00	14.93	

- Start with the effect size (ES) and inverse variance weight (w) for 10 studies.
- Next, multiply w by ES.

The Weighted Mean Effect Size

$$\overline{ES} = \frac{\sum (w \times ES)}{\sum w}$$

Study	ES	w	w*ES
1	-0.33	11.91	-3.93
2	0.32	28.57	9.14
3	0.39	58.82	22.94
4	0.31	29.41	9.12
5	0.17	13.89	2.36
6	0.64	8.55	5.47
7	-0.33	9.80	-3.24
8	0.15	10.75	1.61
9	-0.02	83.33	-1.67
10	0.00	14.93	0.00

- Start with the effect size (ES) and inverse variance weight (w) for 10 studies.
- Next, multiply w by ES.
- Repeat for all effect sizes.

The Weighted Mean Effect Size

$$\overline{ES} = \frac{\sum (w \times ES)}{\sum w}$$

Study	ES	w	w*ES
1	-0.33	11.91	-3.93
2	0.32	28.57	9.14
3	0.39	58.82	22.94
4	0.31	29.41	9.12
5	0.17	13.89	2.36
6	0.64	8.55	5.47
7	-0.33	9.80	-3.24
8	0.15	10.75	1.61
9	-0.02	83.33	-1.67
10	0.00	14.93	0.00
		269.96	41.82

- Start with the effect size (ES) and inverse variance weight (w) for 10 studies.
- Next, multiply w by ES.
- Repeat for all effect sizes.
- Sum the columns, w and ES.
- Divide the sum of (w*ES) by the sum of (w).

$$\overline{ES} = \frac{\sum (w \times ES)}{\sum w} = \frac{41.82}{269.96} = 0.15$$

The Standard Error of the Mean ES

Study	ES	w	w*ES
1	-0.33	11.91	-3.93
2	0.32	28.57	9.14
3	0.39	58.82	22.94
4	0.31	29.41	9.12
5	0.17	13.89	2.36
6	0.64	8.55	5.47
7	-0.33	9.80	-3.24
8	0.15	10.75	1.61
9	-0.02	83.33	-1.67
10	0.00	14.93	0.00
		269.96	41.82

- The standard error of the mean is the square root of 1 divided by the sum of the weights.

$$se_{\overline{ES}} = \sqrt{\frac{1}{\sum w}} = \sqrt{\frac{1}{269.96}} = 0.061$$

Mean, Standard Error, Z-test and Confidence Intervals

Mean ES

$$\overline{ES} = \frac{\sum (w \times ES)}{\sum w} =$$

SE of the Mean ES

$$se_{\overline{ES}} = \sqrt{\frac{1}{\sum w}} =$$

Z-value for the Mean ES

$$Z = \frac{\overline{ES}}{se_{\overline{ES}}} =$$

95% Confidence Interval

$$Lower = \overline{ES} - 1.96(se_{\overline{ES}}) =$$

$$Upper = \overline{ES} + 1.96(se_{\overline{ES}}) =$$

Mean, Standard Error, Z-test and Confidence Intervals

Mean ES

$$\overline{ES} = \frac{\sum (w \times ES)}{\sum w} = \frac{41.82}{269.96} = 0.15$$

SE of the Mean ES

$$se_{\overline{ES}} = \sqrt{\frac{1}{\sum w}} = \sqrt{\frac{1}{269.96}} = 0.061$$

Z-value for the Mean ES

$$Z = \frac{\overline{ES}}{se_{\overline{ES}}} = \frac{0.15}{0.061} = 2.46$$

95% Confidence Interval

$$Lower = \overline{ES} - 1.96(se_{\overline{ES}}) = 0.15 - 1.96(.061) = 0.03$$

$$Upper = \overline{ES} + 1.96(se_{\overline{ES}}) = 0.15 + 1.96(.061) = 0.27$$

Interpreting Effect Sizes

- Cohen (1992): Small, medium, large *descriptive (i.e., based on the magnitudes of past studies)* classification
  (see Literature folder)
- Ferguson (2009): Revised effect size interpretation, tried to resolve inconsistencies in Cohen (1992)
  (see Literature folder)
- NOT treated here (for good reasons):
  Binomial effect size display (BESD)

Homogeneity Analysis

- Homogeneity analysis tests whether the assumption that all of the effect sizes are estimating the same population mean is a reasonable assumption.

- If homogeneity is rejected, the distribution of effect sizes is assumed to be heterogeneous.
  - Single mean ES not a good descriptor of the distribution
  - There are real between study differences, that is, studies estimate different population mean effect sizes.
  - Two options:
    - model between study differences (> Moderator analysis)
    - fit a random effects model (if theoretically justifiable)
  - *CAVEAT(s): Low power of homogeneity tests (esp. Q) under certain conditions! Failing to reject homogeneity DOES NOT GUARANTEE the absence of systematic heterogeneity!*
    > see ´special analysis issues´ later

Q - The Homogeneity Statistic

$$Q_T = \sum_{i=1}^{k} \frac{(ES_i - \overline{ES})^2}{SE_i^2} = \sum_{i=1}^{k} w_i (ES_i - \overline{ES})^2$$

Study	ES	w	w*ES	w*ES^2
1	-0.33	11.91	-3.93	1.30
2	0.32	28.57	9.14	2.93
3	0.39	58.82	22.94	8.95
4	0.31	29.41	9.12	2.83
5	0.17	13.89	2.36	0.40
6	0.64	8.55	5.47	3.50
7	-0.33	9.80	-3.24	1.07
8	0.15	10.75	1.61	0.24
9	-0.02	83.33	-1.67	0.03
10	0.00	14.93	0.00	0.00
		269.96	41.82	21.24

- Calculate a new variable that is the ES squared multiplied by the weight.
- Sum new variable.

Calculating Q

We now have 3 sums:

$$\sum w = 269.96$$
$$\sum (w \times ES) = 41.82$$
$$\sum (w \times ES^2) = 21.24$$

Q is can be calculated using these 3 sums:

$$Q = \sum (w \times ES^2) - \frac{\left[\sum (w \times ES)\right]^2}{\sum w} = 21.24 - \frac{41.82^2}{269.96} = 21.24 - 6.48 = 14.76$$

Interpreting Q

- Q is distributed as a Chi-Square
- df = number of ESs - 1
- Running example has 10 ESs, therefore, df = 9
- Critical Value for a Chi-Square with df = 9 and p = .05 is:

  16.92
- Since our Calculated Q (14.76) is less than 16.92, we **fail to reject** the null hypothesis of homogeneity.
- Thus, the variability across effect sizes does not exceed what would be expected based on sampling error.
- BUT moderator analysis might still be required!

Heterogeneous Distributions: What Now?

- Moderator analysis:
  Analyze excess between study (ES) variability
  - categorical variables with the *analog* to the one-way ANOVA
  - continuous variables and/or multiple variables with weighted multiple regression

- IF *theoretically justifiable*, assume variability is random and fit a random effects model.
  - BUT heterogeneity might still be present (!)
    -> fit a *mixed effects model* (RE model after moderators have been introduced to explain systematic heterogeneity)

Analyzing Heterogeneous Distributions: The Analog to the ANOVA

Q partitioning: Q(total) = Q (between) + Q (within)

Q for all ESs

Pooled weighted sum-of-squares of the mean ES-es for each group around the grand mean

Pooled weighted sum-of-squares of the individual ES-es within each group around the group means

Study	Grp	ES	w	w*ES	w*ES^2
1	1	-0.33	11.91	-3.93	1.30
2	1	0.32	28.57	9.14	2.93
3	1	0.39	58.82	22.94	8.95
4	1	0.31	29.41	9.12	2.83
5	1	0.17	13.89	2.36	0.40
6	1	0.64	8.55	5.47	3.50
			151.15	45.10	19.90
7	2	-0.33	9.80	-3.24	1.07
8	2	0.15	10.75	1.61	0.24
9	2	-0.02	83.33	-1.67	0.03
10	2	0.00	14.93	0.00	0.00
			118.82	-3.29	1.34

- Calculate the 3 sums for each subgroup of effect sizes.

A grouping variable (e.g., random vs. nonrandom)

Analyzing Heterogeneous Distributions: The Analog to the ANOVA

Calculate a separate Q for each group:

$$Q_{GROUP\_1} = 19.90 - \frac{45.10^2}{151.15} = 6.44$$

$$Q_{GROUP\_2} = 1.34 - \frac{-3.29^2}{118.82} = 1.25$$

Analyzing Heterogeneous Distributions: The Analog to the ANOVA

The sum of the individual group Qs = Q within:

$$Q_W = Q_{GROUP\_1} + Q_{GROUP\_2} = 6.44 + 1.25 = 7.69$$

$$df = k - j = 10 - 2 = 8$$

Where k is the number of effect sizes and j is the number of groups.

The difference between the Q total and the Q within is the Q between:

$$Q_B = Q_T - Q_W = 14.76 - 7.69 = 7.07$$

$$df = j - 1 = 2 - 1 = 1$$

Where j is the number of groups.

Analyzing Heterogeneous Distributions: The Analog to the ANOVA

All we did was partition the overall Q(T) into two pieces, a within groups Q(W) and a between groups Q(B).

$Q_B = 7.69$	$df_B = 1$	$Q_{CV\_.05}(1) = 3.84$	$p_B < .05$
$Q_W = 7.07$	$df_W = 8$	$Q_{CV\_.05}(8) = 15.51$	$p_W > .05$
$Q_T = 14.76$	$df_T = 9$	$Q_{CV\_.05}(9) = 16.92$	$p_T > .05$

The grouping variable accounts for significant variability in effect sizes.

Mean ES for each Group

The mean ES, standard error and confidence intervals can be calculated for each group:

$$ES_{GROUP\_1} = \frac{\sum (w \times ES)}{\sum w} = \frac{45.10}{151.15} = 0.30$$

$$ES_{GROUP\_2} = \frac{\sum (w \times ES)}{\sum w} = \frac{-3.29}{118.82} = -0.03$$

Analyzing Heterogeneous Distributions: Multiple Regression Analysis

- Analog to the ANOVA is restricted to a single categorical between studies variable.
- What if you are interested in a continuous variable or multiple between study variables?

- Weighted Multiple Regression Analysis
  - as always, it is weighted analysis
  - can use "canned" programs (e.g., SPSS, SAS)
    - parameter estimates are correct (R-squared, B weights, etc.)
    - F-tests, t-tests, and associated probabilities are incorrect
    - Wilson/Lipsey SPSS macros which give correct parameters and probability values http://mason.gmu.edu/~dwilsonb/ma.html
  - use meta-analytic software, e.g.:
    - metafor: A Meta-Analysis Package for R (by Wolfgang Viechtbauer) http://www.wvbauer.com/downloads.html
    - Other software products listed at: http://www.meta-analysis.eu

Meta-Analytic Multiple Regression Results From the Wilson/Lipsey SPSS Macro (data set with 39 ESs)

Q partitioning: Q(total) = Q (model) + Q (residual/error)

Q for all ESs

Q explained by regression model. If significant, at least one predictor (moderator) is significant.

Variability unaccounted by the model. If significant, model is underspecified, fit a mixed effectsmodel.

```
*****  Meta-Analytic Generalized OLS Regression  *****

------- Homogeneity Analysis -------
                  Q            df             p
Model      104.9704        3.0000         .0000
Residual   424.6276       34.0000         .0000

------- Regression Coefficients -------
                B        SE   -95% CI   +95% CI         Z         P      Beta
Constant   -.7782     .0925    -.9595    -.5970   -8.4170     .0000     .0000
RANDOM      .0786     .0215     .0364     .1207    3.6548     .0003     .1696
TXVAR1      .5065     .0753     .3590     .6541    6.7285     .0000     .2933
TXVAR2      .1641     .0231     .1188     .2094    7.1036     .0000     .3298
```

Partition of total Q into variance explained by the regression "model" and the variance left over ("residual").

Interpretation is the same as will ordinal multiple regression analysis.

If residual Q is significant, fit a mixed effects model.

Review of Weighted Multiple Regression Analysis

- Analysis is weighted.
- Q for the model indicates if the regression model explains a significant portion of the variability across effect sizes.
- Q for the residual indicates if the remaining variability across effect sizes is homogeneous.
- If using a “canned” regression program, must correct the probability values (but this is done by freely available macros, too).

Random Effects Models

- Use if *unconditional inferences* are the goal of your meta-analysis!

- Three ´empirical´ reasons to use a random effects model
  - Total Q is significant and you assume that the excess variability across effect sizes derives from random differences across studies (sources you cannot identify or measure)
  - The Q within from an Analog to the ANOVA is significant
  - The Q residual from a Weighted Multiple Regression analysis is significant

Recap: The Logic of a Random Effects Model

- Fixed effects model assumes that all of the variability between effect sizes is due to sampling error
  - In other words, instability in an effect size is due simply to subject-level “noise”

- Random effects model assumes that the variability between effect sizes is due to sampling error ***plus*** variability in the population/universe of effects (unique differences in the set of true population effect sizes)
  - In other words, instability in an effect size is due to subject-level “noise” and true unmeasured differences across studies (that is, each study is estimating a slightly different population effect size)

The Basic Procedure of a Random Effects Model

- Fixed effects model weights each study by the inverse of the sampling variance.

$$w_i = \frac{1}{se_i^2}$$

- Random effects model weights each study by the inverse of the (subject-level) sampling variance ***plus*** a constant that represents the (study-level) variability across the population effects.

$$w_i = \frac{1}{se_i^2 + \underline{\hat{v}_\theta}}$$

This is the random effects variance component.

How To Estimate the Random Effects Variance Component

- The random effects variance component is based on Q.
- A formula (methods of moments estimate, i.e. the DerSimonian and Laird method, as *one* variant to estimate the random between-study variance component) is:

$$\hat{v}_{\theta} = \frac{Q_T - k - 1}{\sum w - \left(\frac{\sum w^2}{\sum w} \right)}$$

Important: Sometimes also termed ´T^2´, estimating ´tau-square´, the true between study variance in RE models.

Calculation of the Random Effects Variance Component

$$\hat{v}_\theta = \frac{Q_T - k - 1}{\sum w - \left(\dfrac{\sum w^2}{\sum w} \right)}$$

Study	ES	w	w*ES	w*ES^2	w^2
1	-0.33	11.91	-3.93	1.30	141.73
2	0.32	28.57	9.14	2.93	816.30
3	0.39	58.82	22.94	8.95	3460.26
4	0.31	29.41	9.12	2.83	865.07
5	0.17	13.89	2.36	0.40	192.90
6	0.64	8.55	5.47	3.50	73.05
7	-0.33	9.80	-3.24	1.07	96.12
8	0.15	10.75	1.61	0.24	115.63
9	-0.02	83.33	-1.67	0.03	6944.39
10	0.00	14.93	0.00	0.00	222.76
		269.96	41.82	21.24	12928.21

- Calculate a new variable that is the w squared.
- Sum new variable.

Calculation of the Random Effects Variance Component

- The total Q for this data was 14.76
- k is the number of effect sizes (10)
- The sum of w = 269.96
- The sum of w^2 = 12,928.21

$$\hat{v}_\theta = \frac{Q_T - k - 1}{\sum w - \left(\frac{\sum w^2}{\sum w} \right)} = \frac{14.76 - 10 - 1}{269.96 - \frac{12{,}928.21}{269.96}} = \frac{5.76}{269.96 - 47.89} = \underline{0.026}$$

Rerun Analysis with New Inverse Variance Weight

- Add the random effects variance component to the variance associated with each ES.
- Calculate a new weight.

$$w_i = \frac{1}{se_i^2 + \hat{v}_\theta}$$

- Rerun analysis.
- Congratulations! You have just performed a very complex statistical analysis.

SPSS Macro Output with Random Effects Variance Component

```
------- Homogeneity Analysis -------
                     Q            df             p
Model         104.9704        3.0000         .0000
Residual      424.6276       34.0000         .0000

------- Regression Coefficients -------
                 B        SE   -95% CI   +95% CI         Z         P      Beta
Constant    -.7782     .0925    -.9595    -.5970   -8.4170     .0000     .0000
RANDOM       .0786     .0215     .0364     .1207    3.6548     .0003     .1696
TXVAR1       .5065     .0753     .3590     .6541    6.7285     .0000     .2933
TXVAR2       .1641     .0231     .1188     .2094    7.1036     .0000     .3298

------- Estimated Random Effects Variance Component -------
v       =    .04715

Not included in above model which is a fixed effects model
```

Random effects variance component based on the residual Q.

Comparison of Random Effect with Fixed Effect Results

- Important conceptual differences:
  - FE: The *single (=´fixed´) true effect* size is estimated (by the mean ES)
  - RE: True distribution in the universe of studies is estimated, i.e. the mean effect size of a true distribution of effects along with the true distributions´ variance (tau square) is estimated.
- The biggest difference you will notice is in the significance levels and confidence intervals.
  - Confidence intervals will get bigger.
  - Effects that were significant under a fixed effect model may no longer be significant.
- Random effects models (compared to FE models)...
  - ... are more conservative
  - ... weights are more balanced (whenever T is nonzero)
    (see, e.g. Bornstein et. al, 2009, Chapter 12)

Agenda

- Step-by-step analysis using MS Excel (following the HO tradition, exercise from Lipsey & Wilson, 2001)
  - FE Analysis
  - RE Analysis
- **Special analysis issues:**
  - **Homogeneity / heterogeneity indicators**
  - Sensitivity analyses
- Interpreting results
  - Theory testing/development (SICT)
  - Description of a research field (Web nonresponse)
  - Estimating the effectiveness of (here: Marketing) interventions (Fluency-effects)

Homogeneity / Heterogeneity

Overview of measures to identify and quantify the heterogeneity in effect sizes:

- *Q* and its p-value (see formula above):
  - does only assess the viability of the null hypothesis (homogeneity), does not quantify (true) heterogeneity
  - Power of Q/p low for small sample sizes per study and small No of studies (e.g. true heterogeneity present despite n.s.)!
- *T-square* estimating *thau-square* (see formula above):
  - depends on scale, and does therefore quantify amount of true heterogeneity (RE model; T=0 in FE models, of course)
- *I-square (Higgins et al., 2003)*:
  - Ratio of true heterogeneity to total observed variation in % (0-100 range)
  - ´signal to noise´ratio: the larger, the more true heterogeneity

The I-square Statistic

- Equivalent expression of the statistic (in the notation of Borenstein et al., 2009, p. 177):

$$I^2 = \left(\frac{Variance_{betweenstudies}}{Variance_{total}} \right) * 100\% = \left(\frac{\tau^2}{\tau^2 + V_Y} \right) * 100\% = \left(\frac{Q - df}{Q} \right) * 100\%$$

- Interpretation:
  - Determines the proportion of the observed variance being ´real´ heterogeneity
  - The smaller, the more spurious (random) components in the observed variance
  - Tentative benchmarks to interpret I-square according to Higgins et al. (2003):
    - 25%: low, 50% moderate, 75% high
    - *DO NOT confuse* with 75% rule by Hunter/Schmidt!

Interpreting Q, T-square, I-square

- *Q*-Test and corresponding *p*-value:
  - serves as a homogeneity test of significance (only)
  - sensitive towards No of studies
  - not sensitive to the metric of the ES index
  - low power under certain conditions > do perform moderator analysis if theoretically sensible regardless of Q-result (!)
- *T-square:*
  - *quantifies true heterogeneity in the RE model*
  - *sensitive to the metric of the ES -> quantifies true variance (!)*
  - *not sensitive to the No of studies*
- *I-square:*
  - *´signal to noise ratio´*
  - *not sensitive to the No of studies*
  - not sensitive to the metric of the ES index
  - indicates if moderator analyses make sense

Agenda

- Step-by-step analysis using MS Excel (following the HO tradition, exercise from Lipsey & Wilson, 2001)
  - FE Analysis
  - RE Analysis
- Special analysis issues:
  - Homogeneity / heterogeneity indicators
  - **Sensitivity analyses**
- Interpreting results
  - Theory testing/development (SICT)
  - Description of a research field (Web nonresponse)
  - Estimating the effectiveness of (here: Marketing) interventions (Fluency-effects)

Sensitivity Analyses (Selection)

Overall: Give empirical answers on questions related to possible bias(es) introduced by certain meta-analytic design decisions.

- Vary inclusion/exclusion criteria
- Model subjective decisions during coding process (!)
- Model quality of coding information
  - for low-inference codes
  - for high inference codes
- Vary RE variance component estimator
- Vary multiple ES handling procedure(s)
- ... etc.

Agenda

- Step-by-step analysis using MS Excel (following the HO tradition, exercise from Lipsey & Wilson, 2001)
  - FE Analysis
  - RE Analysis
- Special analysis issues:
  - Homogeneity / heterogeneity indicators
  - Sensitivity analyses
- **Interpreting results**
  - **Theory testing/development (SICT)**
  - **Description of a research field (Web nonresponse)**
  - **Estimating the effectiveness of (here: Marketing) interventions (Fluency-effects)**

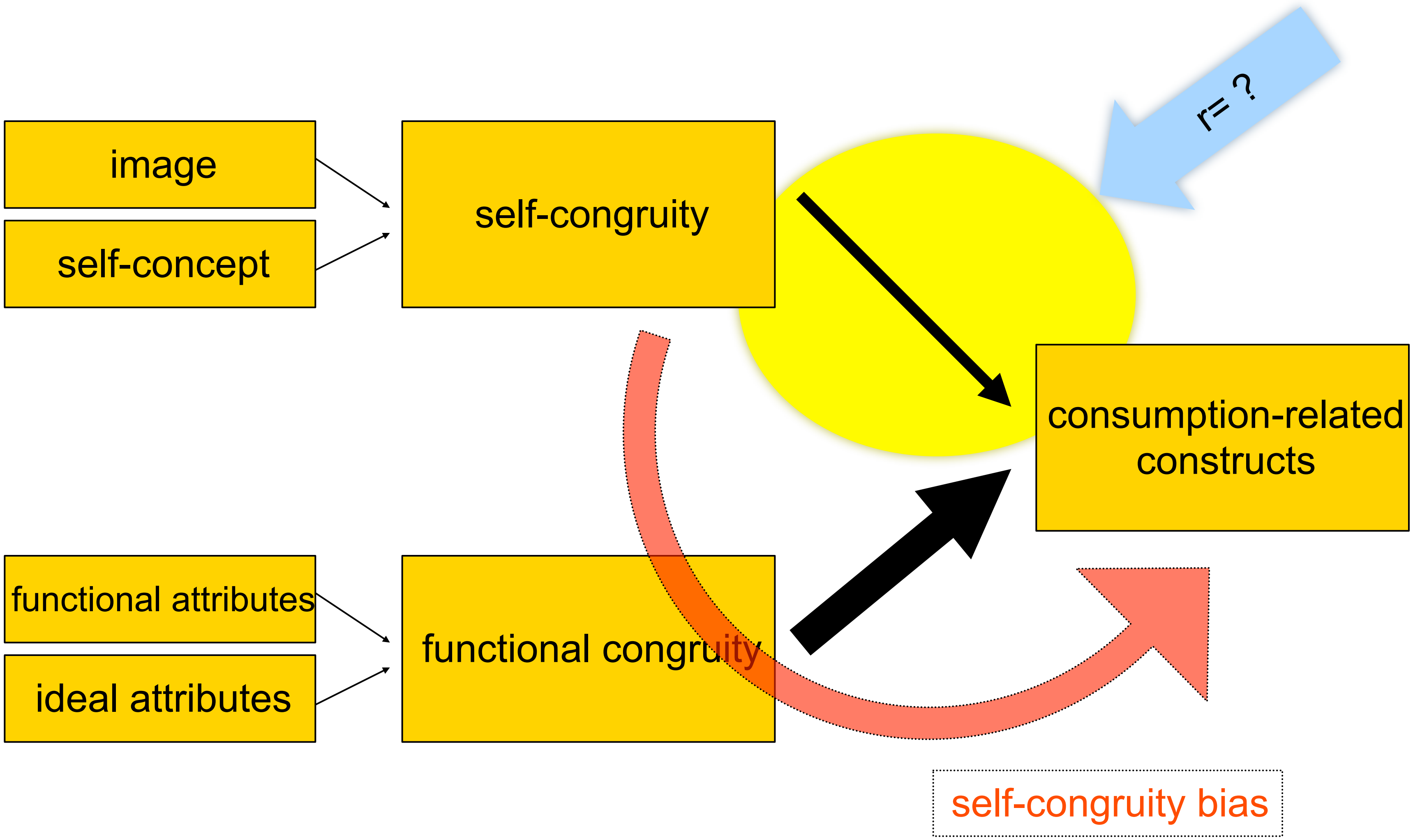
Selected Own Meta-Analyses

- **Example 1: Theory testing /development:** Self-image congruity meta-analysis (with M. Joseph Sirgy, VTech, and Alexandra Rodriguez, FSU; JBR in press)
- **Example 2: Description of a research field:** Response rate differences Web surveys versus other modes meta-analysis (with colleagues from U Ljubljana and U Mannheim, published in 2008 in IJMR)
- **Example 3: Estimating the effectiveness of (advertising) interventions:** Fluency-effects meta-analysis (with Norbert Schwarz, U Mich, and Marco Warth, U Mannheim, in progress)

Selected Own Studies in CB/MR

- **Example 1: Theory testing /development:** Self-image congruity meta-analysis (with M. Joseph Sirgy, VTech, and Alexandra Rodriguez, FSU; JBR revision under review)
- **Example 2: Description of a research field:** Response rate differences Web surveys versus other modes meta-analysis (with colleagues from U Ljubljana and U Mannheim, published in 2008 in IJMR)
- **Example 3: Estimating the effectiveness of (Marketing) interventions:** Fluency-effects meta-analysis (with Norbert Schwarz, U Mich, and Marco Warth, U Mannheim, in progress)

Self-Image Congruity Theory (M.J. Sirgy)



Example 1: Theory testing / development

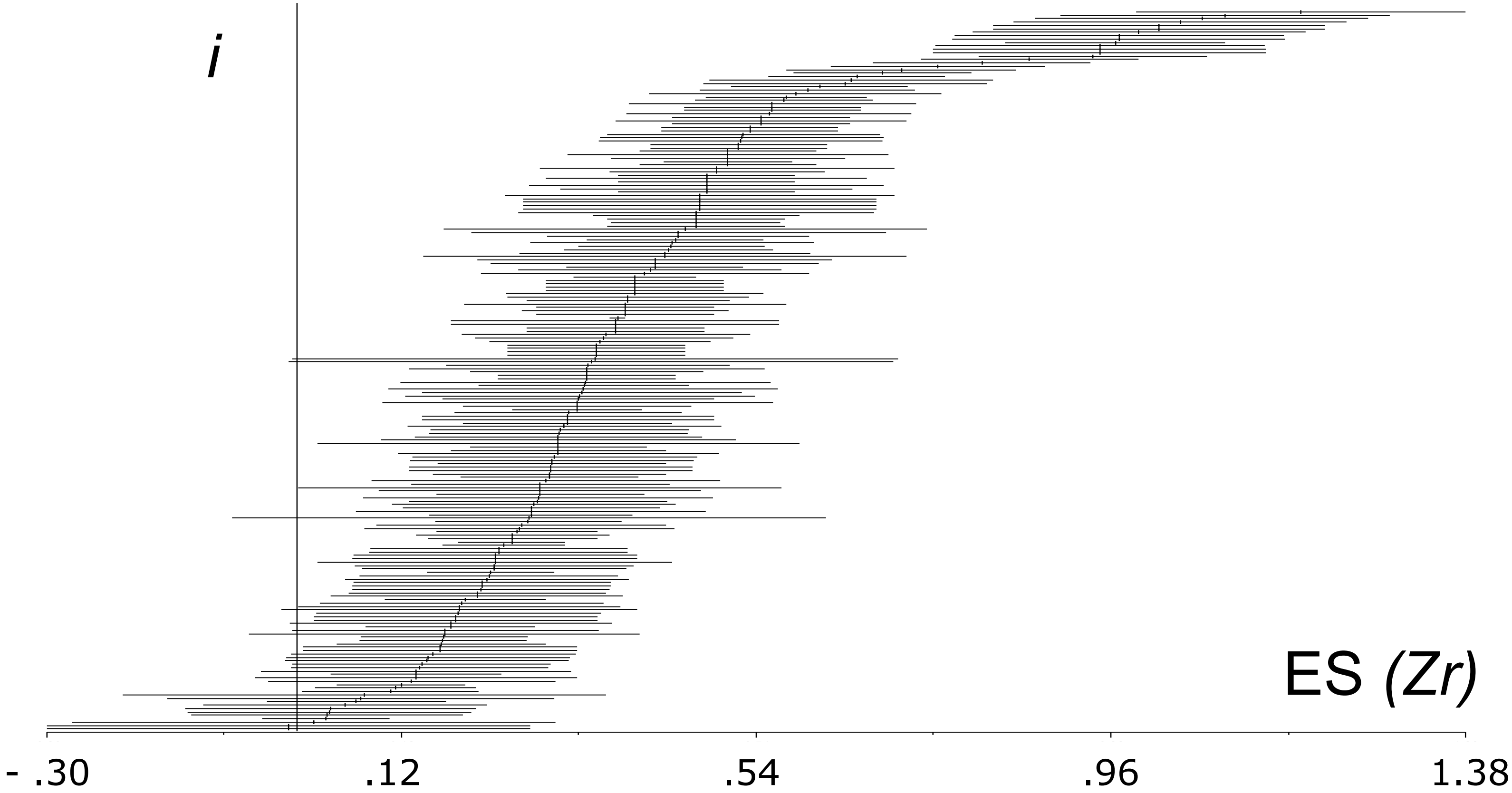
- Research questions:
  - Existence and strength of self-congruity effects?
  - Moderators of self-image congruity effects?
  - Avenues for future research?

- Sketch of methods:
  - 211 effect sizes out of 30 papers
  - ES measure: Correlation coefficients (*r*)
  - Moderators (selection):
    Type of self-congruity, dependent variable used
  - HO-type meta-analysis (random effects)

Example 1: Results (Mean Effect)

Z(r)= .39 (95% CI= .36, .42); *r*= .37

Q= 249.66, *df*= 210, *p*= .032



Example 1: Results (Moderators)

Moderator	Moderator Categories (and *No of ES*)	Mean *Zr*-Score	95% CI for *Zr*
Self-congruity type	Actual priv. congruity (***126***)	.39	.35/.42
	Ideal priv. congruity (***70***)	.42	.36/.47
	Actual social congruity (***5***)	.23	-.03/.49
	Ideal social congruity (***2***)	.36	-1.49/2.21
Dependent variable	Attitudes (***93***)	.33	.29/.38
	Intentions (***82***)	.41	.37/.46
	Behavior (***17***)	.34	.24/.45

Example 1: Results (Selection)

- Evidence for non-homogeneous self-image congruity effects (medium-size strength of effects according to Cohen´s classification)

- Despite a research tradition encompassing almost 30 years, clear gaps in research (selection):
  - *Herding tendency*: Predominant focus on private self-congruity, only a few studies on social self-congruity!
  - *Type III error tendency (right answers to ´wrong´ questions):* Only a few studies on actual (consumption-related behavior), but ample research on attitudes and intentions.
  - *Mono-method tendency*: No experimental studies systematically manipulating self-congruity!

Example 1: Selected follow-up projects

- Subsequent research projects, filling the gaps in research identified:
  - „Expanding the concept of self-congruity to other dimensions of destination image to explain and predict post-visit evaluations in South Tyrolean tourism", funded by FU Bozen (2008-2011)“
  - „Negative symbolic consumption“ (industry consortium financed, 2008-2009)
  - Various PhD and MA thesis (U Mannheim students): Exp. SICT, Beh. criteria, CMG ...
- Consecutive publications:
  - Bosnjak, M., Sirgy, M.J., Hellriegel, S., & Maurer, O. (in press). Post-visit destination loyalty judgments: Developing and testing a comprehensive congruity model. *Journal of Travel Research.*
  - Bosnjak, M. (in press). Negative symbolic aspects in destination branding: Exploring the role of the ´undesired self´ on Web-based vacation information search intentions among potential first-time visitors. *Journal of Vacation Marketing.*
  - Bosnjak, M. & Sirgy, M.J., Tidwell, J.B., & Kamra, K. (under revision). Global versus attribute-based measures of self-image congruence: An integrated mediation model. *Journal of Empirical Generalisations in Marketing Science.*
  - Sirgy, M.J., Lee, D.-J., Merunka, D., Bosnjak, M., Yu, G.B., Johar, J.S., & Mostert, P. (under revision). Expanding the concept of self-congruity to other dimensions of brand image to explain and predict post-purchase behavioral responses: In search of a comprehensive model. *Journal of the Academy of Marketing Science.*

Selected Own Studies in CB/MR

- **Example 1: Theory testing /development:** Self-image congruity meta-analysis (with M. Joseph Sirgy, VTech, and Alexandra Rodriguez, FSU; JBR revision under review)
- **Example 2: Description of a research field:** Response rate differences Web surveys versus other modes meta-analysis (with colleagues from U Ljubljana and U Mannheim, published in 2008 in IJMR)
- **Example 3: Estimating the effectiveness of (Marketing) interventions:** Fluency-effects meta-analysis (with Norbert Schwarz, U Mich, and Marco Warth, U Mannheim, in progress)

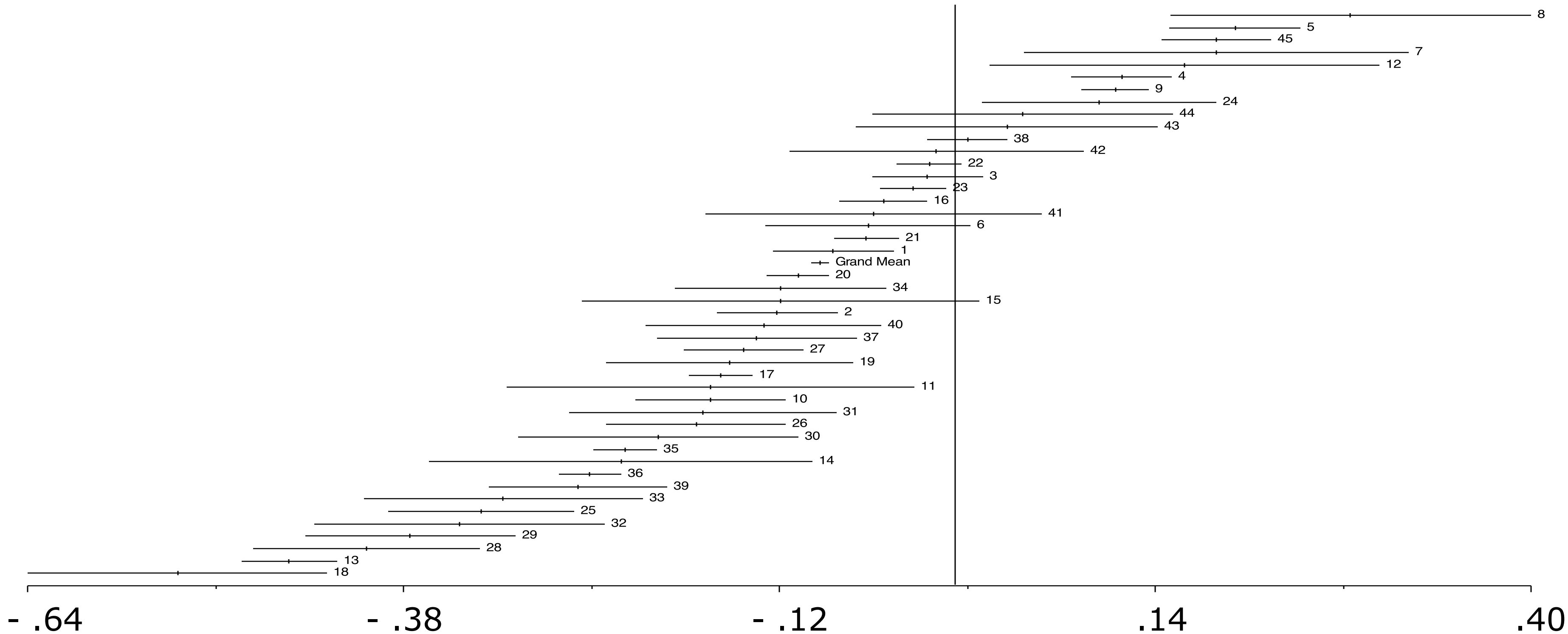
Example 2: Description of a research field

- Research question:
  Are Web surveys actually associated with lower response rates (in comparison to other survey modes)?

- Sketch of methods:
  - 45 experimental mode comparisons from 24 papers
  - ES: Response rate differences (RD)
  - Moderators: Type of mode compared to, sample recruitment strategy, target population, type of sponsorship, solicitation mode, incentives, number of contacts
  - HO-type meta-analysis (random effects)

Example 2: Results (mean effect)

Web: 33.6%, other modes: 44.4%, RD%= 10.8% (95%CI= 15%,6%)



LOWER
response rates in Web surveys

HIGHER
rersponse rates in Web surveys

Example 2: Results (Moderators)

Moderator	Categories (*and No of ES*)	RD	95%-CI
Sample recruitment strategy	Panel (*40*)	-.09	-.14 / -.05
	one-time recruitment (*4*)	-.28	-.49 / -.07
Solicitation/ contact mode	Postal mail (*17*)	-.15	-.21 / -.09
	E-Mail (*25*)	-.05	-.10 / .00
No of contacts	one – two (*23*)	-.05	-.11 / .01
	three – five (*22*)	-.16	-.23 / -.10

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.

Example 2: Selected follow-up projects

- Subsequent research projects:
  - "Increasing nonresponse error by reducing nonresponse rates? Investigating the biasing effect of methods and procedures aimed at increasing response rates in Web-based access panel surveys", DFG SPP project (2008-2010; with GESIS Mannheim, MPI Berlin, ISR Michigan).
  - Various PhD and MA theses (U Mannheim and FU Berlin students): Web survey usability, Web survey nonresponse.

- Consecutive publications:
  - Couper, M. & Bosnjak, M. (2010). Internet surveys (pp. 527-550). In J. D. Wright & P. V. Marsden (Eds.), *Handbook of Survey Research* (2nd edition). San Diego, CA: Elsevier.
  - Bosnjak, M., Neubarth, W., Couper, M., Bandilla, W., & Kaczmirek, L. (2008). Prenotification in Web surveys: The influence of mobile text messaging versus e-mail on response rates and sample composition. *Social Science Computer Review, 26*(2), 213-232.

Selected Own Studies in CB/MR

- **Example 1: Theory testing / development:** Self-image congruity meta-analysis (with M. Joseph Sirgy, VTech, and Alexandra Rodriguez, FSU; JBR revision under review)
- **Example 2: Description of a research field:** Response rate differences Web surveys versus other modes meta-analysis (with colleagues from U Ljubljana and U Mannheim, published in 2008 in IJMR)
- **Example 3: Estimating the effectiveness of (Marketing) interventions:** Incidental (adv.) stimuli exposure: Fluency-effects (with Norbert Schwarz, U Mich, and Marco Warth, U Mannheim, in progress)

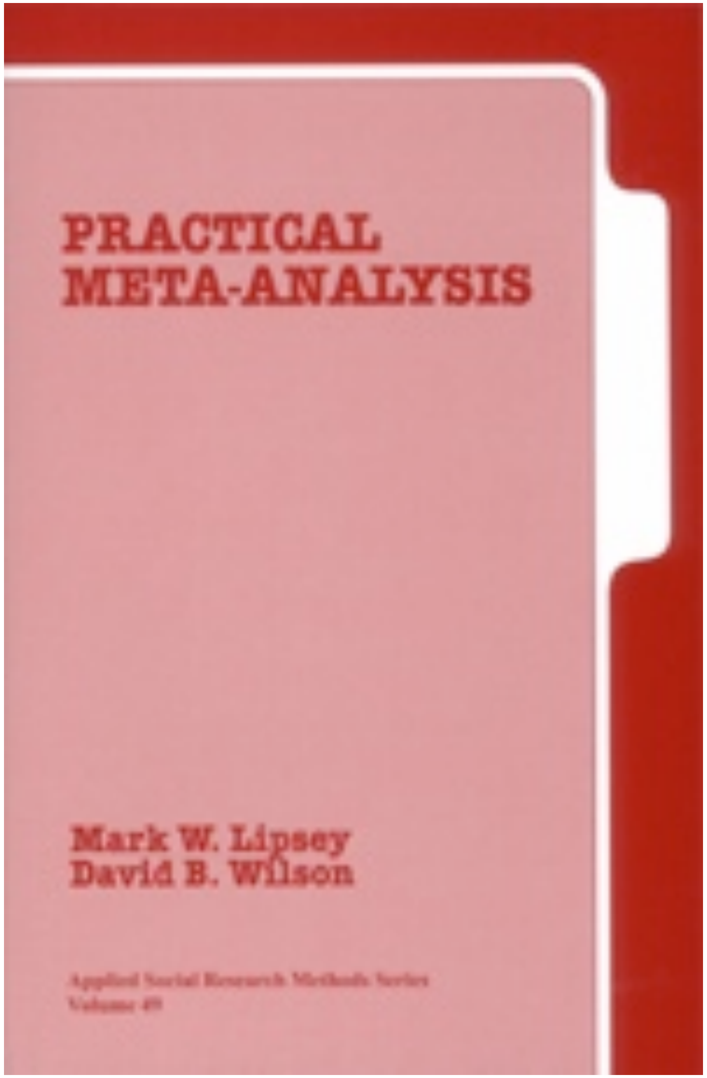
Example 3: Fluency-Effects Meta-Analysis

- Mean fluency effect: d= .50
  (non-homogeneous; 231 effect sizes, 90 studies, 51 papers)
- Moderators (selection):
  - Fluency effects only detectable when correction processes are unlikely (d= .75)
  - Strongest effects for choices and preferences (d= .73/.70), smaller for attitudes (d= .48) and intentions (d= .27)
  - Stimuli type:
    - More pronounced effects for abstract stimuli (d= .58) compared to concrete ones (d= .41)
    - No effects for negatively valenced material (d= 0) compared to positive and neutral (d= .53/.55)
- No differences (selection):
  - Fluency-Type (conceptual versus perceptual)
  - Measurement scale (negative, positive, bipolar scales)
  - Advertising versus non-advertising material

Example 3: Selected follow-up developments

- Subsequent research projects:
  - „Negative influences upon brand evaluations: The litter effect“ (with Manchester Business School): Exploring the influence of branded litter on brand reputation.
- Outlook:
  - Theoretical repositioning required? Selected considerations:
    - Stronger Fluency-effects on choices and preferences compared to attitudes/intentions: Automatic behavior instead of automatic judgmental processes influencing behavior?
    - No effect for negative valence material: Why not detected as theoretically expected?
    - Theoretical distinction between conceptual and perceptual fluency still justified in light of non-detectable empirical differences?
  - Practical implications?
    - Translating the empirically estimated mean fluency effect into an implicit/incidental advertising effectiveness measure

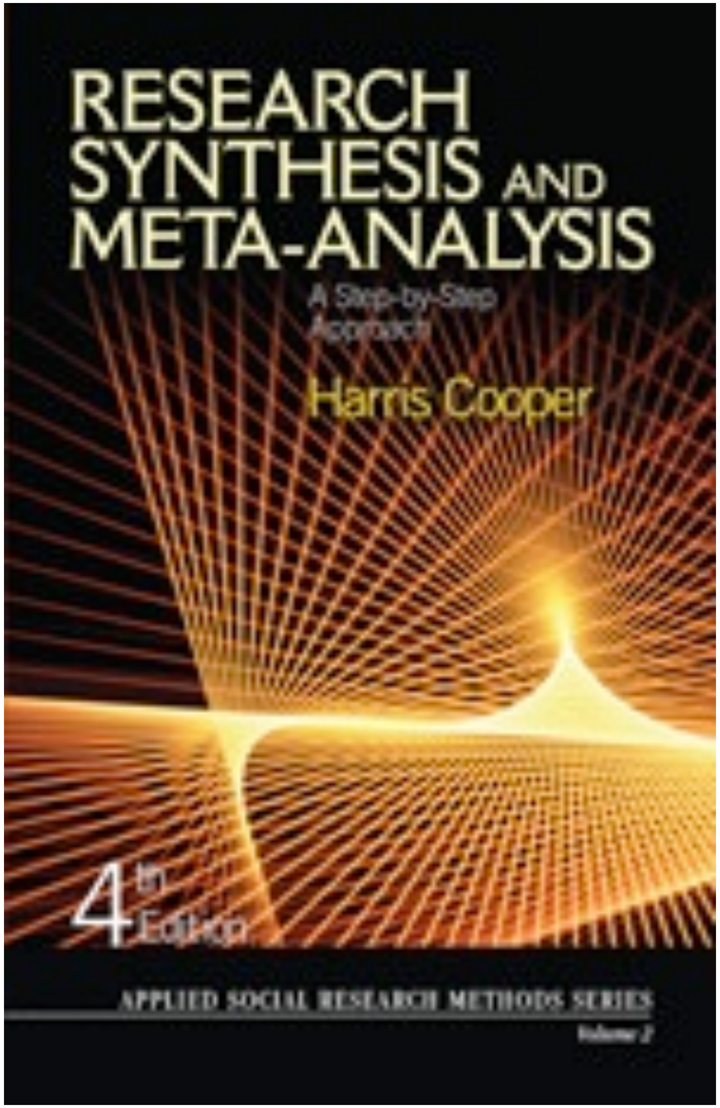
Lipsey & Wilson (2001)



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

- Chapter 6: Analysis issues and strategies
- Chapter 7: Computational techniques for meta-analysis data
- Appendix D: SPSS Macros

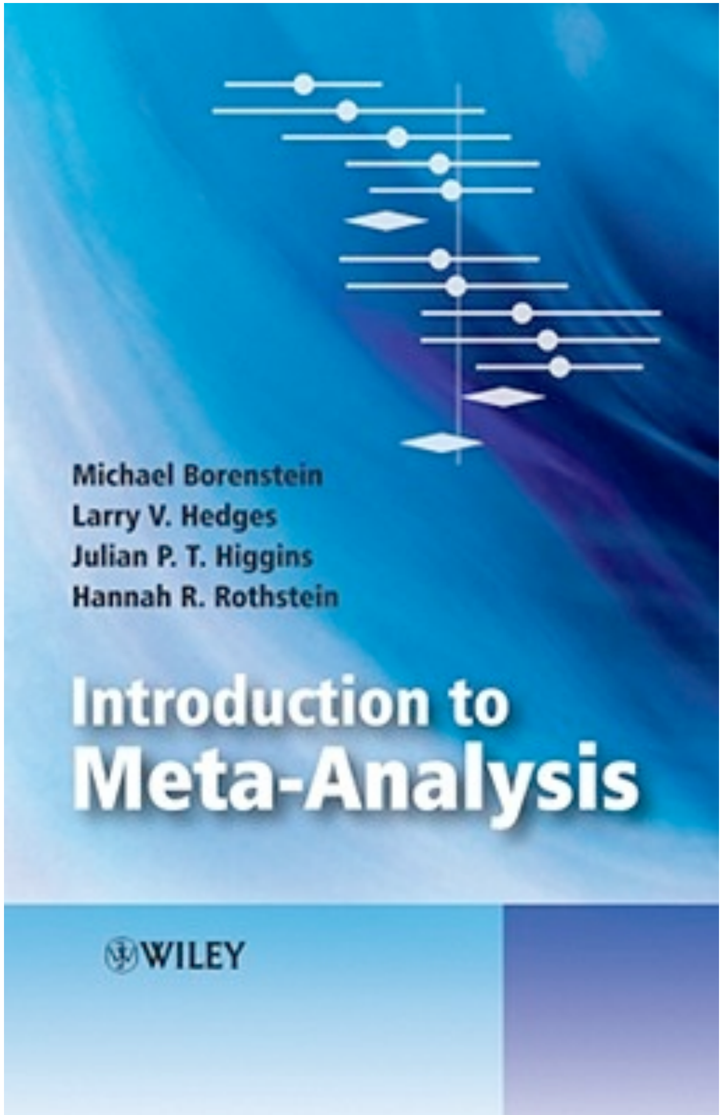
Cooper (2010)



Cooper, H. (2010). *Research Synthesis and Meta-Analysis: A Step-by-Step Approach.* Thousand Oaks, CA: Sage.

- Chapter 6: Step 5: Analyzing and integrating the outcomes of studies
- Chapter 7: Step 6: Interpreting the evidence

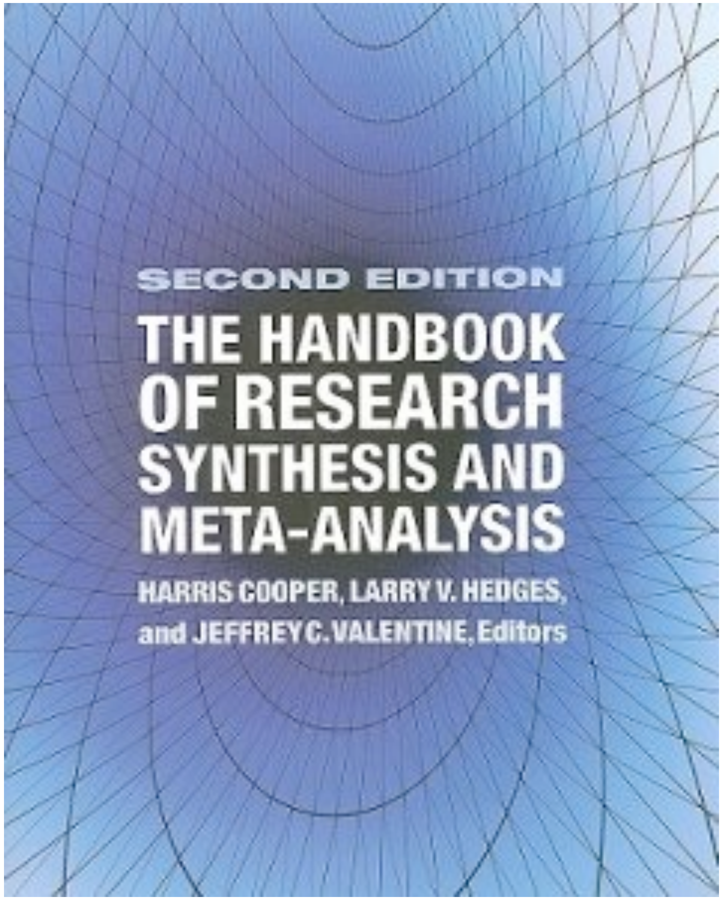
Borenstein et al. (2009)



Borenstein, M., Hedges, L.V., Higgins, J.P.T, & Rothstein, H.R. (2009). *Introduction to Meta-Analysis*. Chichester, UK: Wiley.

- Part 3: Fixed-effect and random effect models (Chapters 10-14)
- Part 4: Heterogeneity (Chapters 15-21)

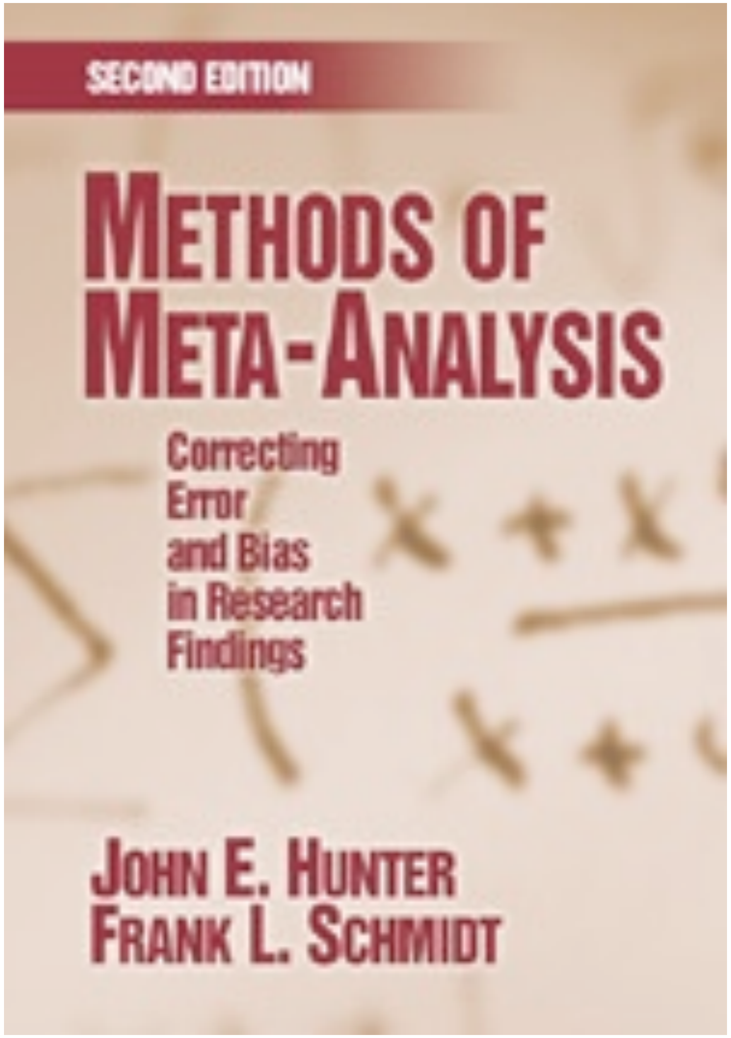
Cooper, Hedges & Valentine (2009)



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

- Part IV: Statistically combining effect sizes (encompassing Chapters 14-17)

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.

- Various pointers to non-HS approaches, but no comprehensive treatment of those
- Overview of non-HS approaches: Chapter 14