



Evidence-based Open Science Tools and Services & Selected Meta-Analytic Studies at ZPID

Landau, Nov. 20th, 2019

[Prof. Dr. Michael Bosnjak](#)

<http://dx.doi.org/10.23668/psycharchives.2640>



Evidence-based Open Science Tools and Services

&

Selected Meta-Analytic Studies at ZPID

- Part 1 -

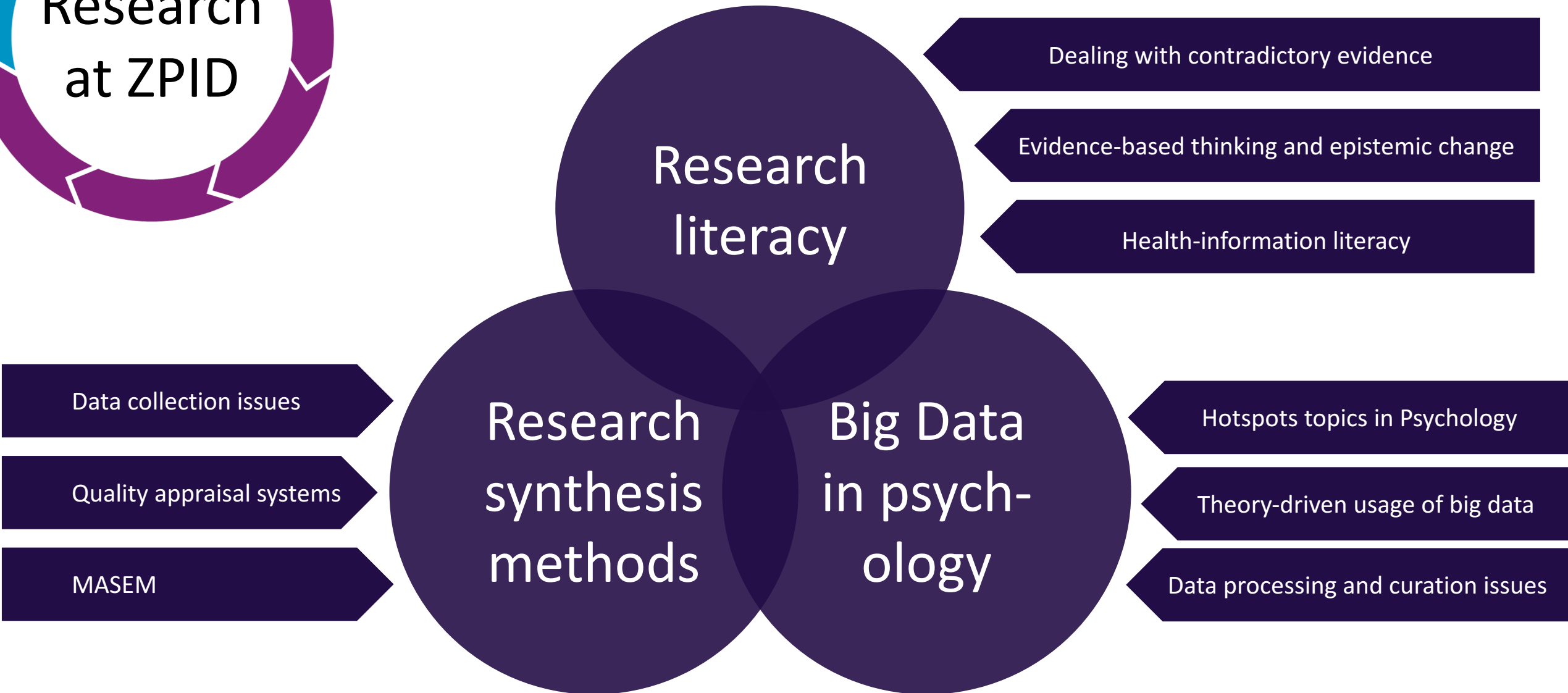
<http://bit.ly/ZPIDRoadshow19-20-mb>

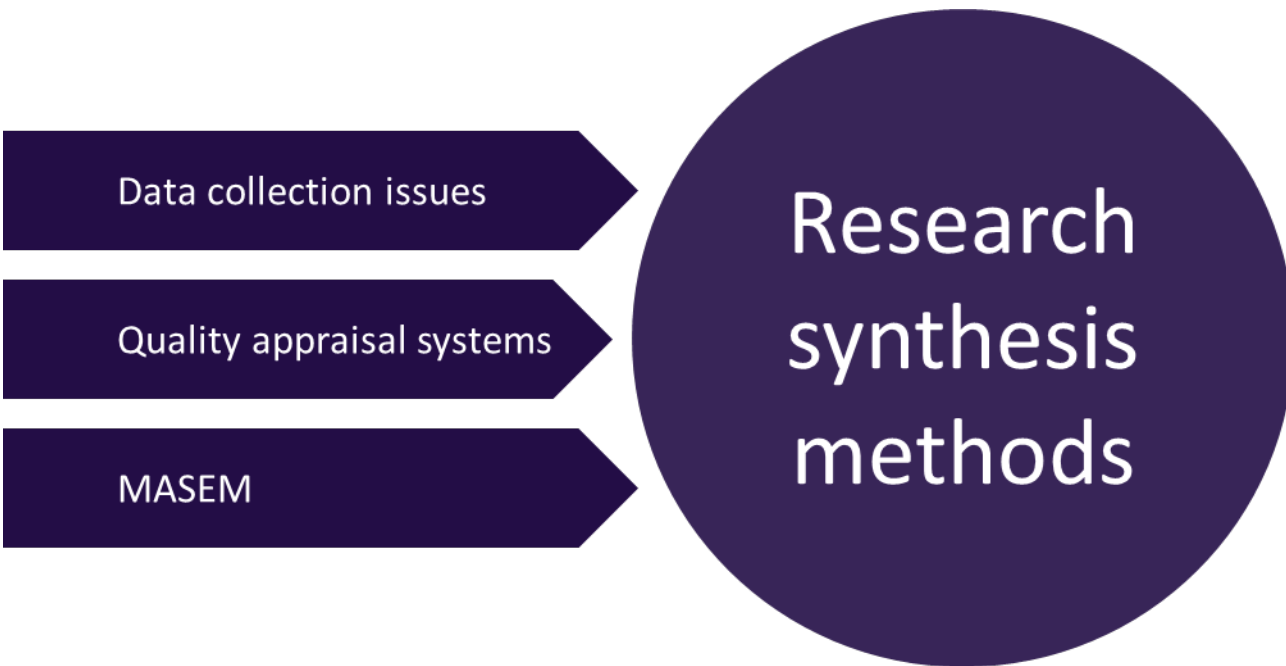


Evidence-based Open Science
Tools and Services
&
Selected Meta-Analytic Studies at
ZPID

- Part 2 -

<http://dx.doi.org/10.23668/psycharchives.2640>





Research
synthesis
methods

Data collection issues

Quality appraisal systems

MASEM



Michael Bosnjak
Director



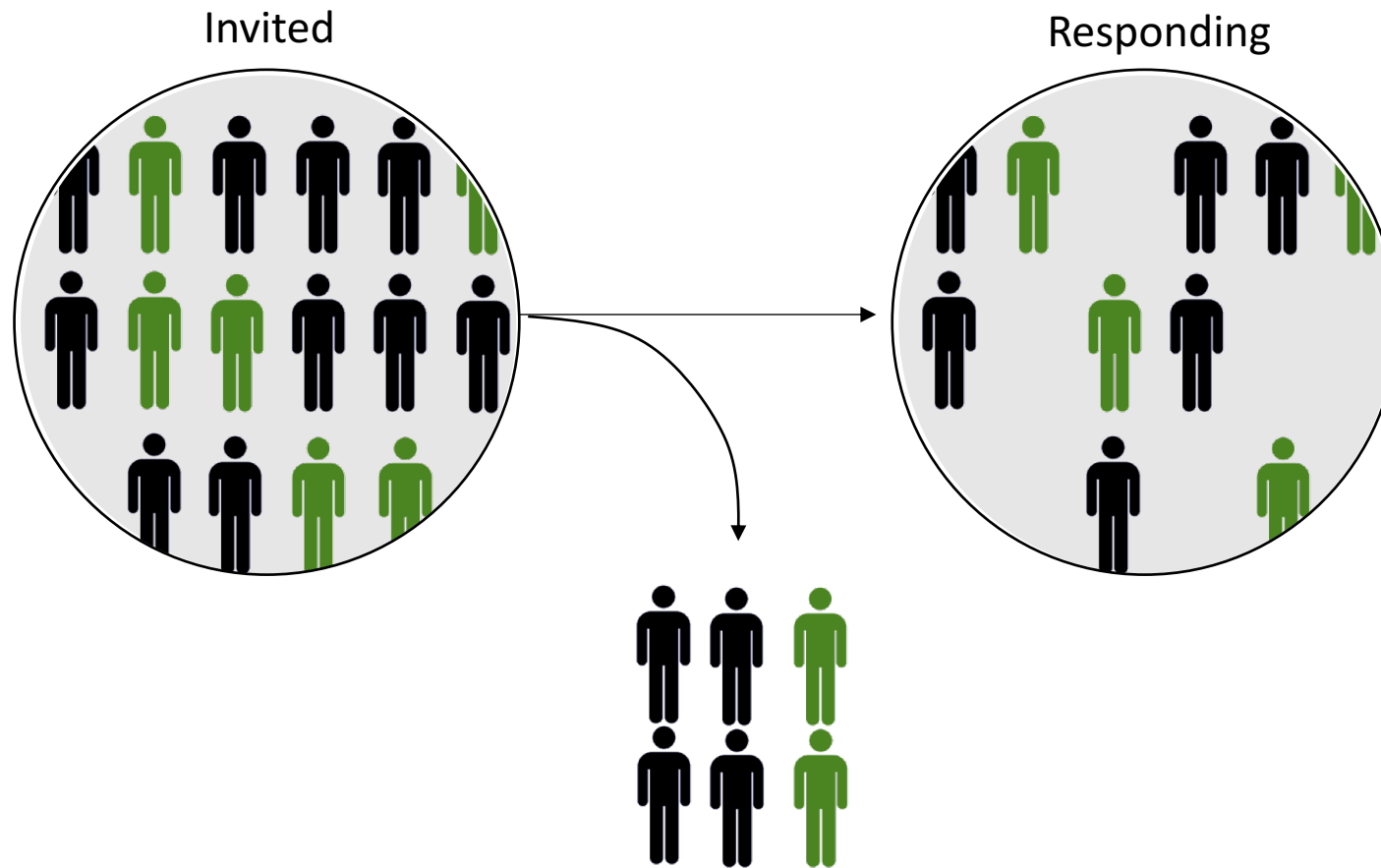
Nadine Wedderhoff
Postdoc Univ. Trier



Tanja Burgard
PhD Student

Nonresponse - Research Intention:

(Unit-) Nonresponse = participant are not willing or capable to respond to a survey

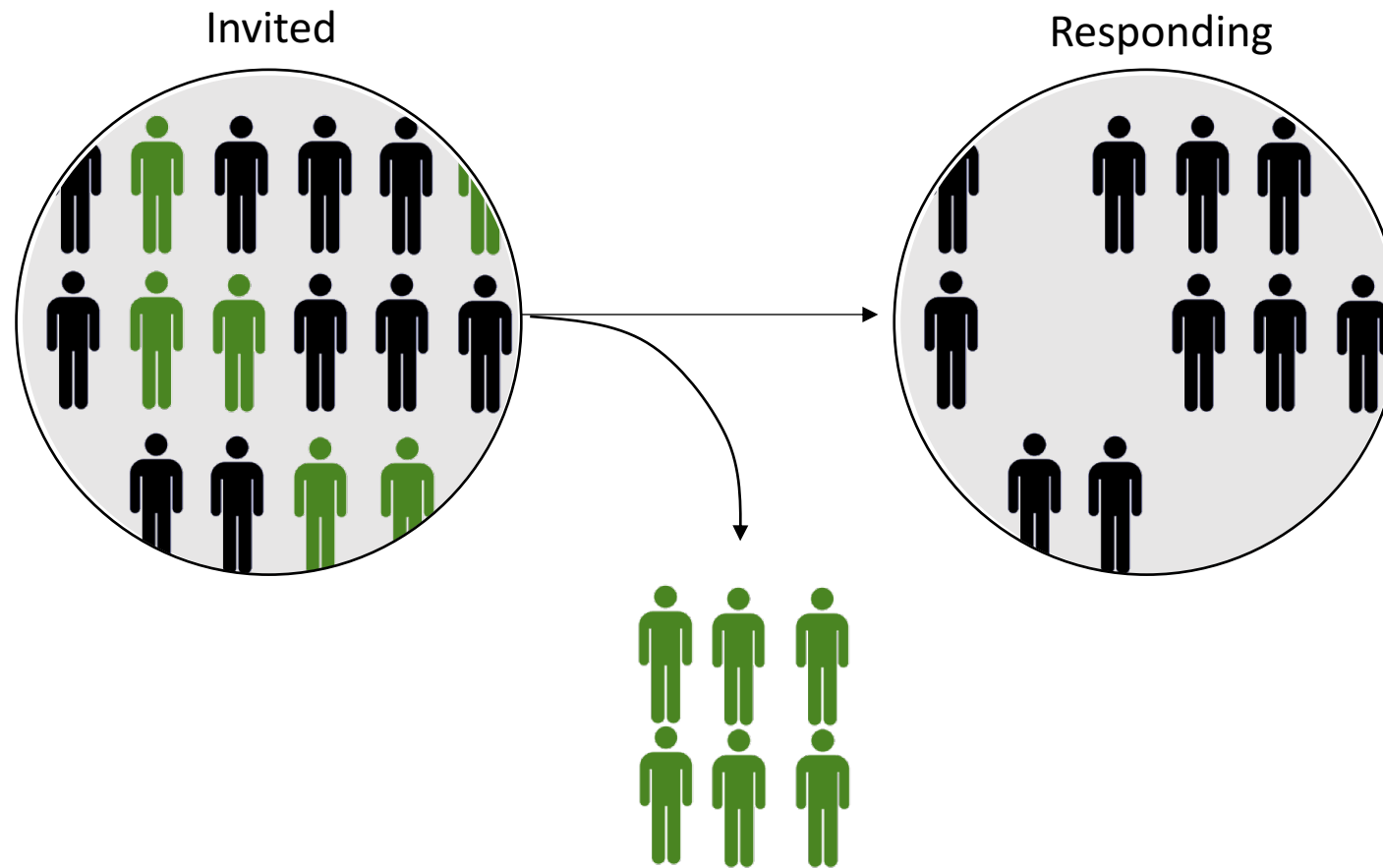


Severe consequences:

- Decreased N

Research Intention:

(Unit-) Nonresponse = participant are not willing or capable to respond to a survey



Severe consequences:

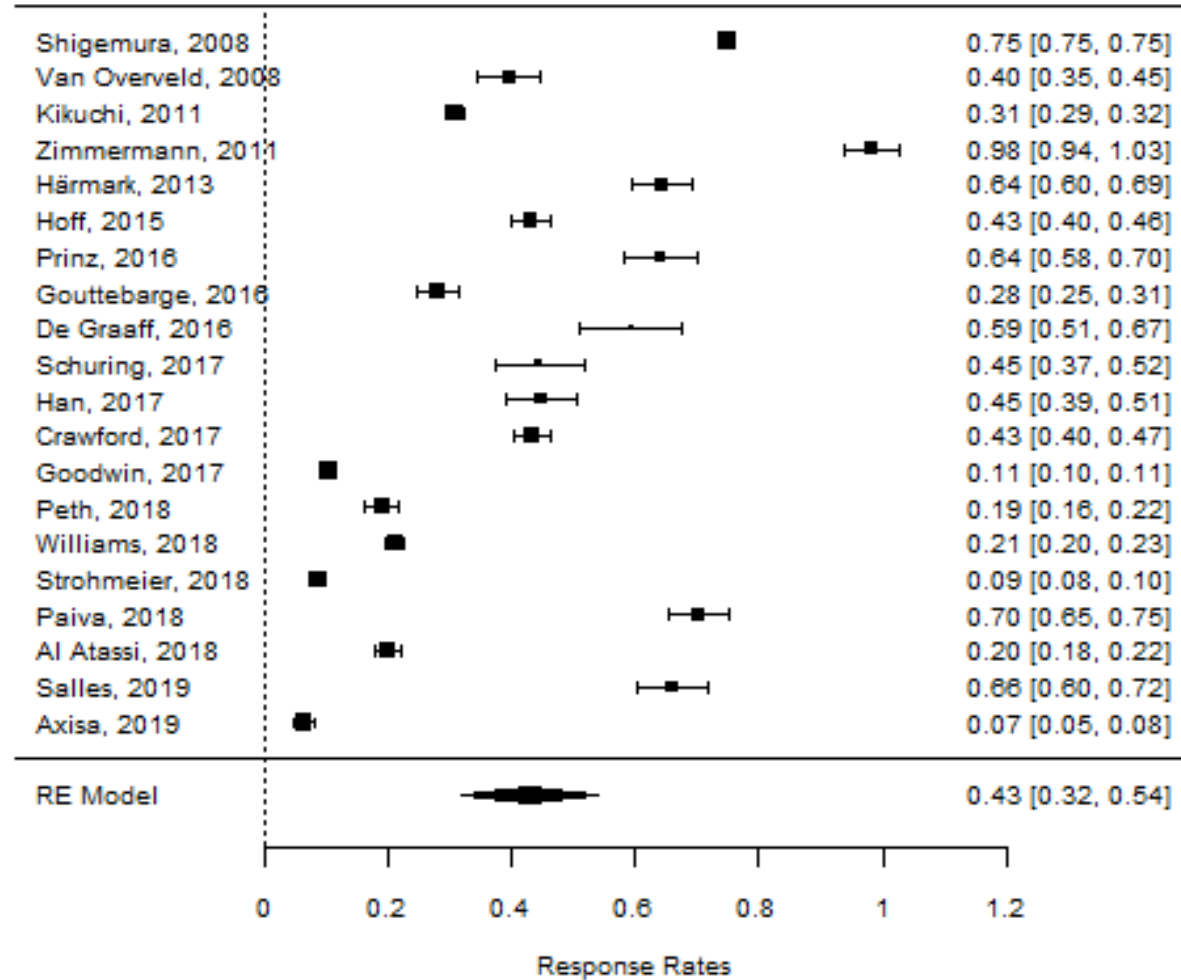
- Decreased N
- Nonresponse bias (if causes for missing are not random)

Aim and Method:

- Meta-analytic investigation on which factors influence the participation rate in psychological online surveys with specific populations
- The findings should guide researchers in how to optimally implement psychological online surveys to yield high participation rates
- First results from a subsample of studies including participants suffering from anxiety and depression $k = 20$.

Results:

Figure: Forest-Plot



Mean response rate: 0.428

 $Q(19) = 5904.967 ; p < .001$ $\tau = 0.251$ $I^2 = 99.92 \%$

Results:

Intercept	0.819 ***
Publication year (centered to 2016)	- 0.232 ***
Number of survey items	- 0.189 ***
Incentives (rc = no)	
yes	0.058
Aquisition procedure (rc = clinical list-based)	
Research panel	- 0.343 ***
Invitation medium (rc = other)	
E-Mail	- 0.393 **

 $R^2 = 0.541$

Note: rc = reference category; *** $p < .001$; ** $p < .01$; * $p < .05$

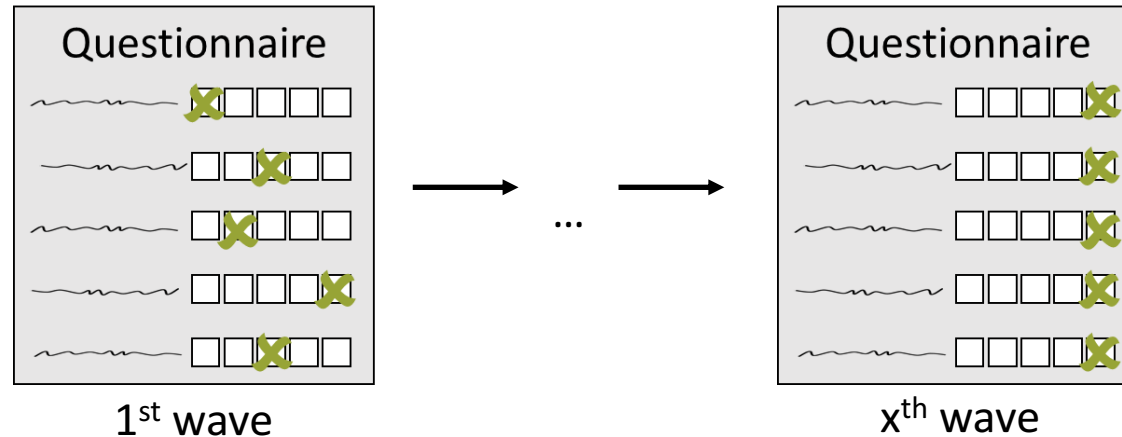
Implications:

- Mean response rate 43%. However: high variability between studies
- Many design specific characteristics influence the response rate and can be used as a guideline for future surveys (e.g. reduce burden of participation by decreasing number of items)
- Future Research: Testing the generalizability of our results across different populations and psychological research questions.

Panel Conditioning - Research Intention:

Panel Conditioning: Sample unit's responses are influenced by participation in earlier waves of a panel survey.

For example: „Straigthlining“ as a result of decreased motivation



Although not negative *per se*, panel conditioning is a potential threat to the validity of panel responses.

Research Intention:

Panel Conditioning is an umbrella term to describe diverse phenomena.
Accordingly: empirical evidence is mixed

Solution: Meta-analysis to integrate findings and to deduce recommendations regarding:

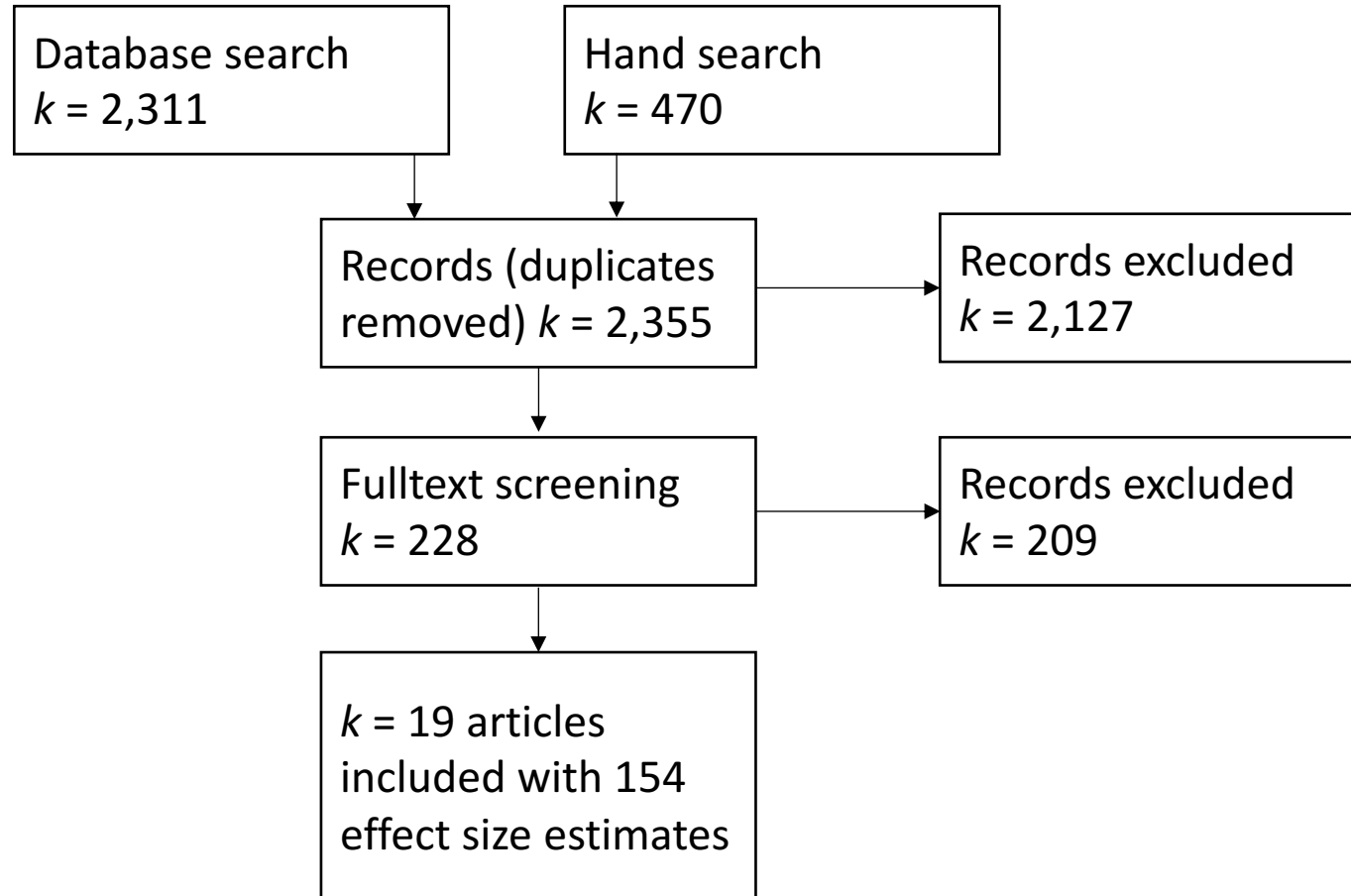
- participation frequency
- interval length between waves
- type of questions
- etc.

Fokus of this study:

- Panel conditioning for sensitive items
- Less social desirable answers expected for experiences vs. new respondents

Method:

Integration of panel-conditioning effects using a multilevel meta-analysis



Method:

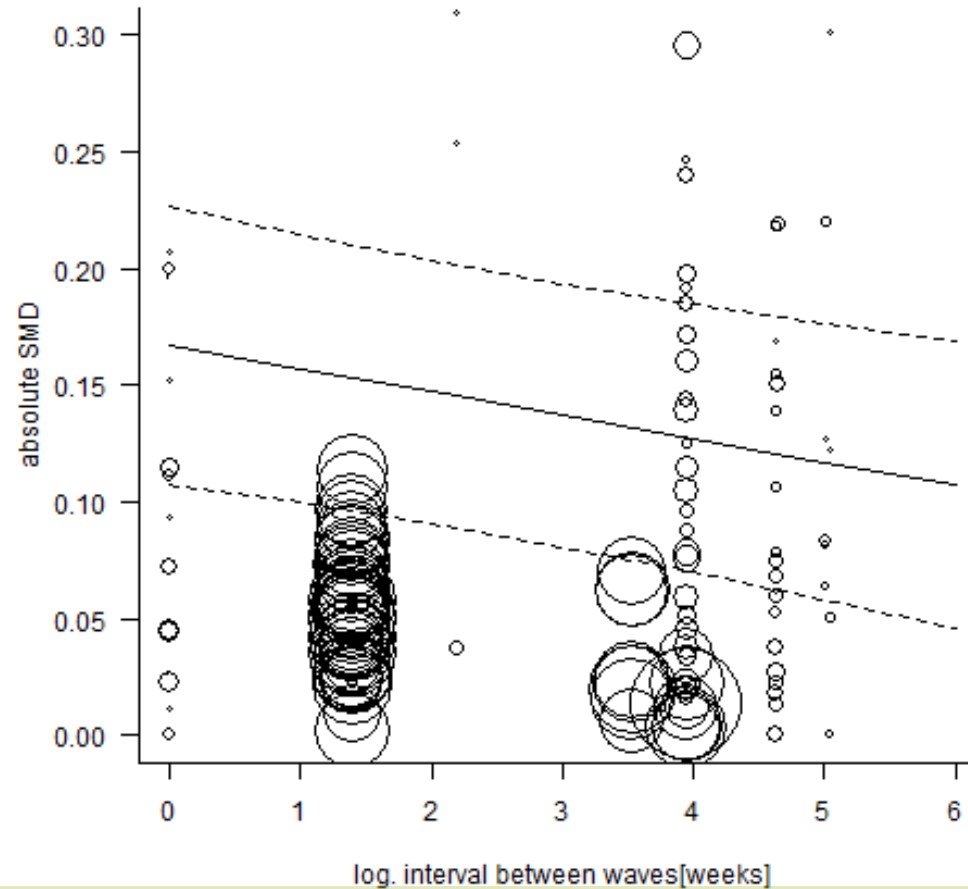
- The effect size is a standardized mean difference (SMDs) between first responders and experienced panel participants
- Since we are analyzing panel conditioning in the context of sensitive items, all SMDs are calculated that:
 - Negative SMDs = Experienced panelist exhibit more social desirable responding
 - Positive SMDs = Experienced panelist exhibit less social desirable responding
- Both are representing panel conditioning but differ in valence.

Results:

- The overall mean is small and not significant: $SMD = 0.040$
- high and significant variance components indicating substantial variability within ($\sigma_2^2 = 0.002$; $p < .001$) and between studies ($\sigma_3^2 = 0.029$; $p < .001$).
- Accordingly: Mean effect is not a good representative and moderator analyses are needed

Results:

- Moderator analyses: Only the interval between the waves has a significant main effect on the extent of panel conditioning:



Implications:

- Overall, panel conditioning seems to be less a problem than expected (at least for sensitive items)
- Regarding the interval between waves, longer intervals (i.e. more than one month) can be recommended
- Future research: Test the robustness of our results for a generalized item pool (i.e. not only sensitive content, and various type of items like knowledge or demographic questions)

Quality appraisal instruments - Research Intention:

Meta-analyses are subject to severe criticism, e.g.: „Garbage in – Garbage out“

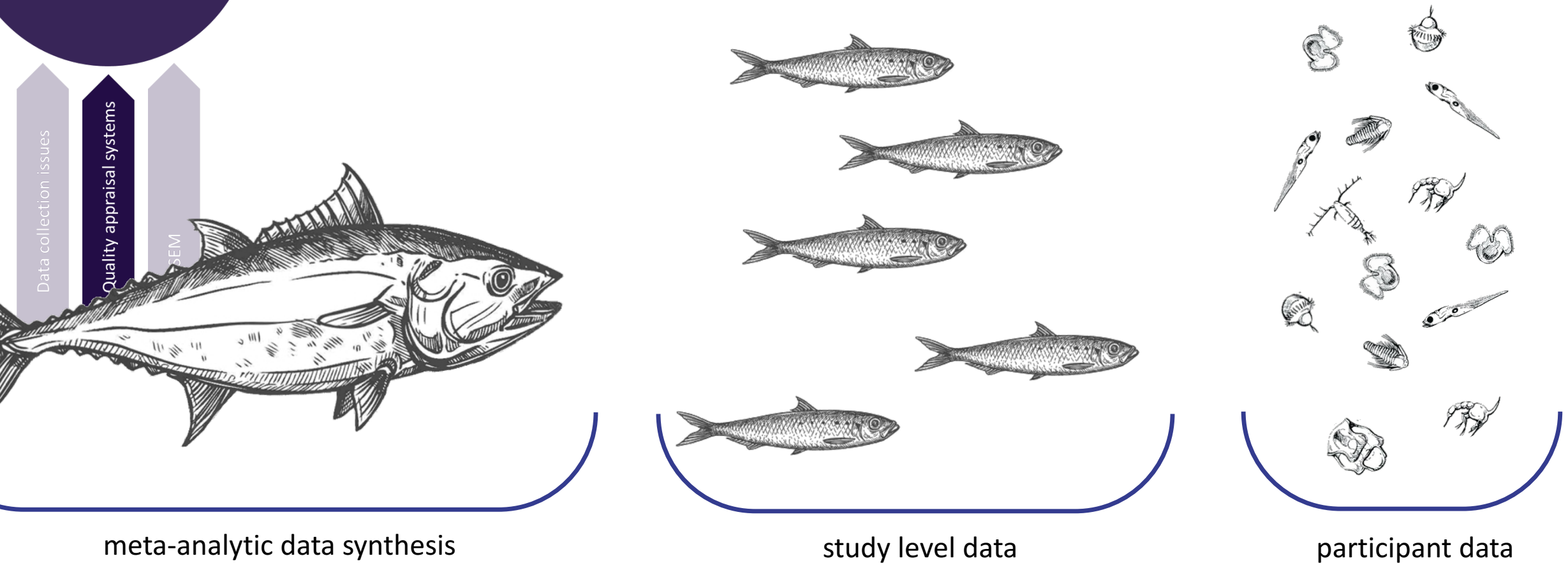
Data collection issues

Quality appraisal systems

MASEM

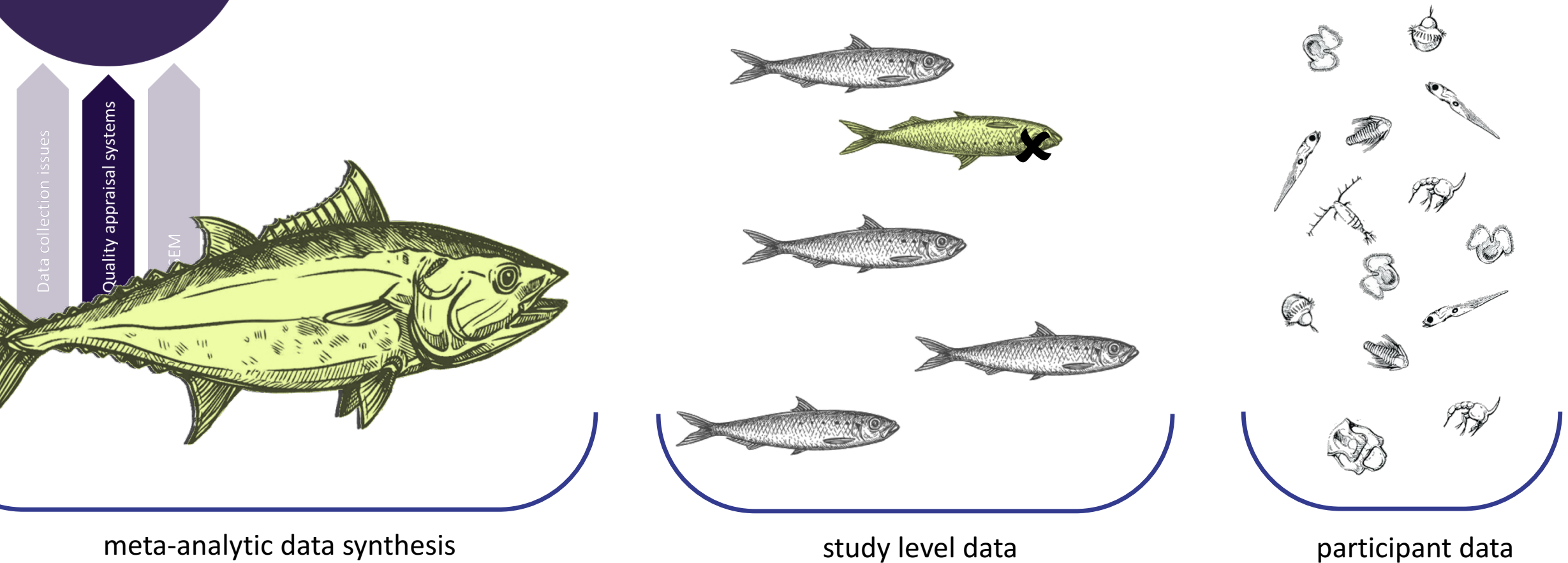
Research Intention:

Meta-analyses are subject to severe criticism, e.g.: „Garbage in – Garbage out“



Research Intention:

Meta-analyses are subject to severe criticism, e.g.: „Garbage in – Garbage out“



Research Intention:

Many empirical studies underline the influence of primary study quality and study results (e.g. Dechartres et al., 2016; Page et al., 2016; Shadish et al., 2008; Spinelli et al., 2015)

→ Meta-analyses should account for differences in primary study quality

Research question: How is the *status quo* in psychological meta-analyses

Method:

Systematic review of every meta-analysis published between 2009 and 2019 in *Psychological Bulletin*

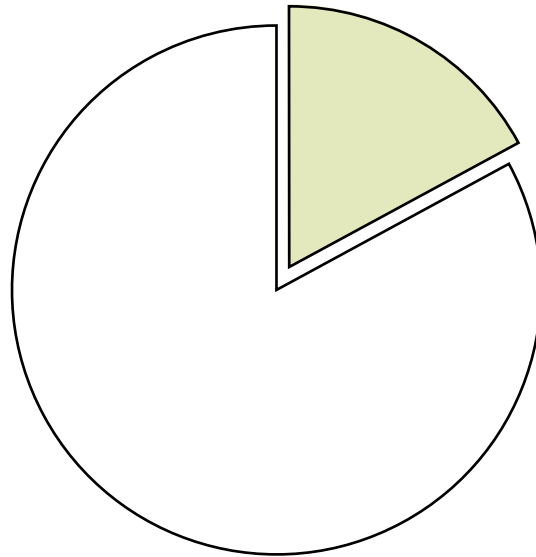
→ 222 Meta-analyses

Coding variables:

- Is primary study quality considered?
- How is it done?
- Which instruments are used to assess quality/risk-of-bias (validated vs. ad hoc)
- How many items are used to assess study quality?

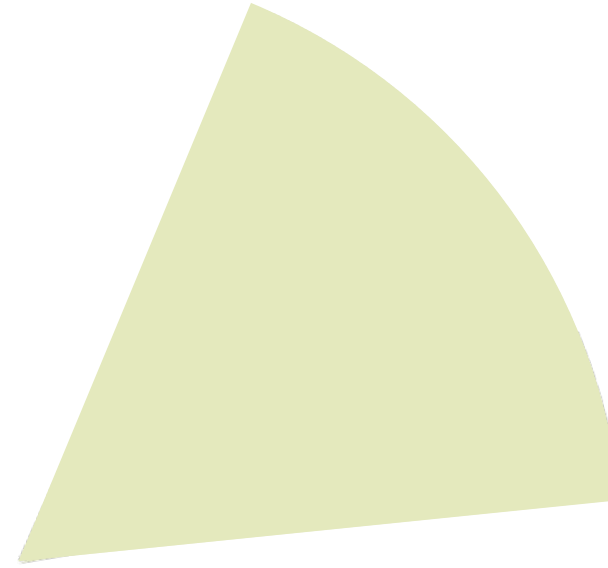
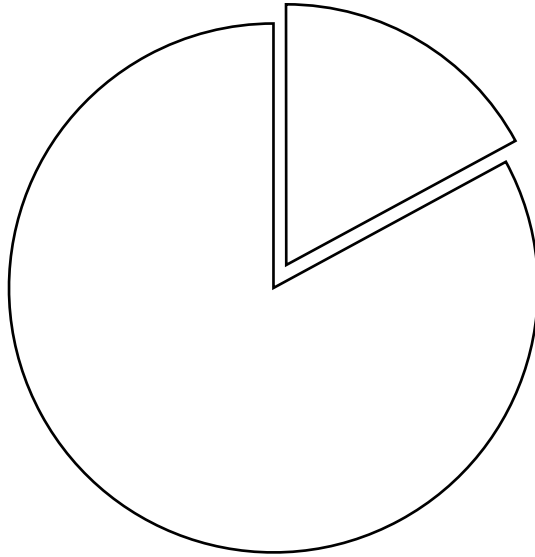
Results:

Only few meta-analyses ($k = 38$; 17 %) took primary study quality into account:



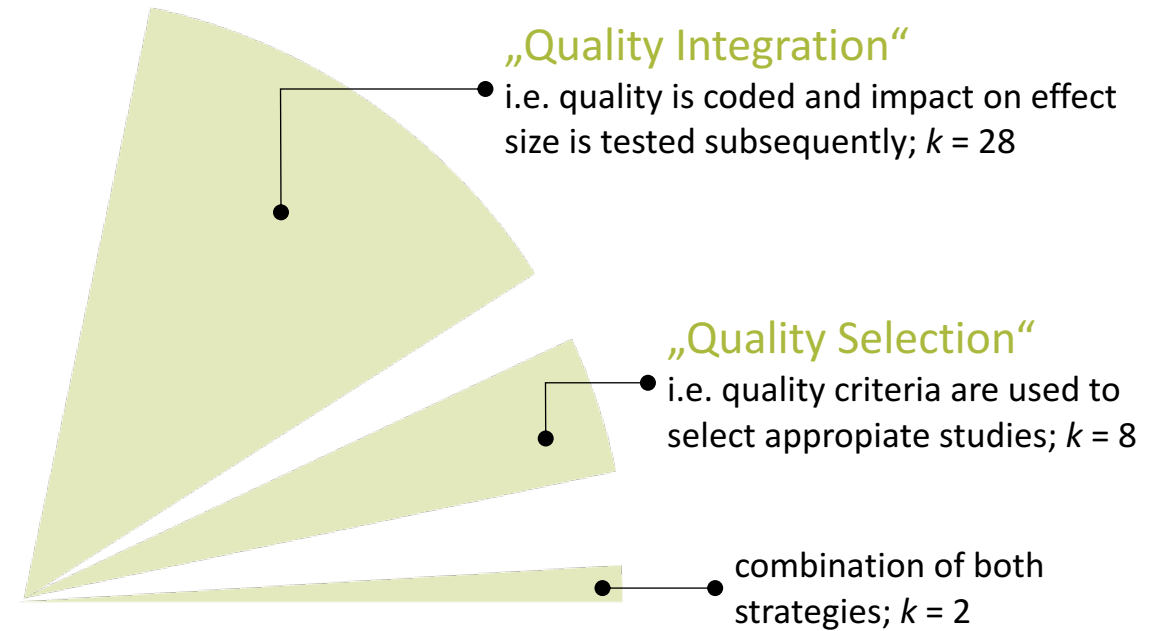
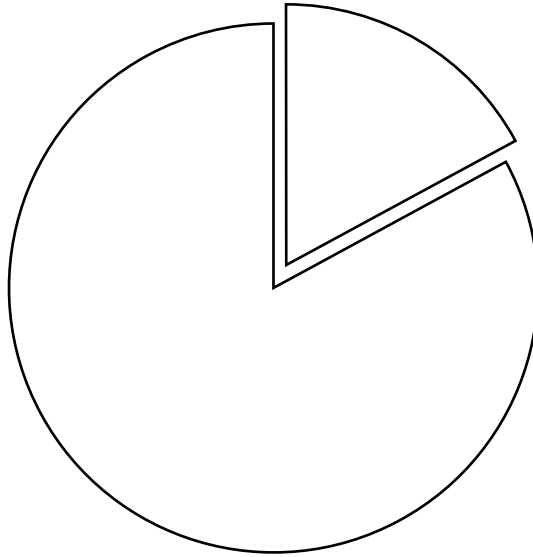
Results:

Only few meta-analyses ($k = 38$; 17 %) took primary study quality into account:



Results:

Only few meta-analyses ($k = 38$; 17 %) took primary study quality into account:



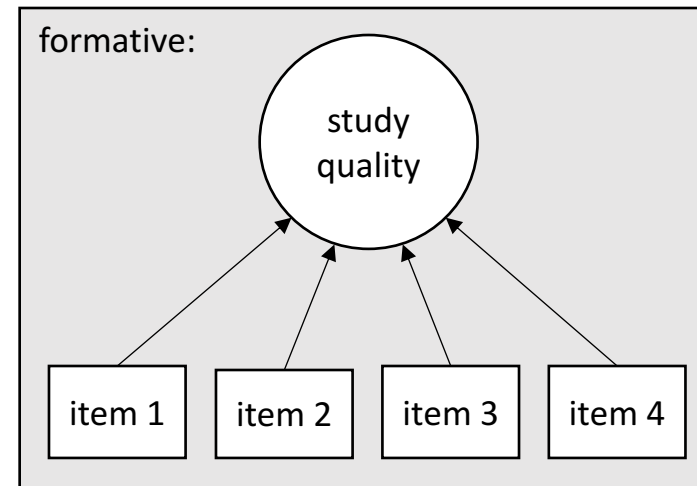
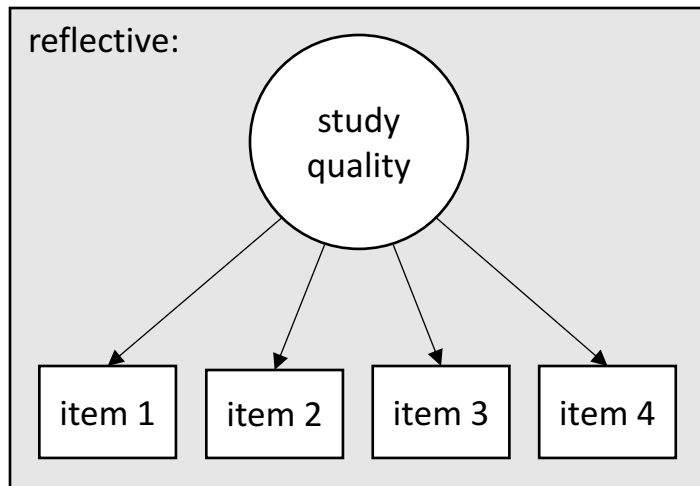
If quality was coded, most authors relied on self-defined standards and ad hoc developed items ($k = 20$).

Problematic: Results from different meta-analyses are not comparable and potentially depend on the choice of quality assessment

Results:

Further Problems:

- The items used differ to the extent they are associated with established validity concepts in Psychology (e.g., by Shadish, Cook, & Campbell, 2002)
- Although quality was typically assessed with multiple items, many authors used a total quality score (typicall one single score, which may not be appropriate, because quality items are *formative* rather than *reflective* indicators)



Implications:

The review underlies the need for

- a shared understanding on what indicators are critical
- a quality assessment tool specifically designed for psychology

The latter needs to meet some certain requirements:

- (a) Generalizable over a high range of study designs
- (b) Adoption of a psychological understanding of quality (i.e., validity)
- (c) easy to implement and reusable

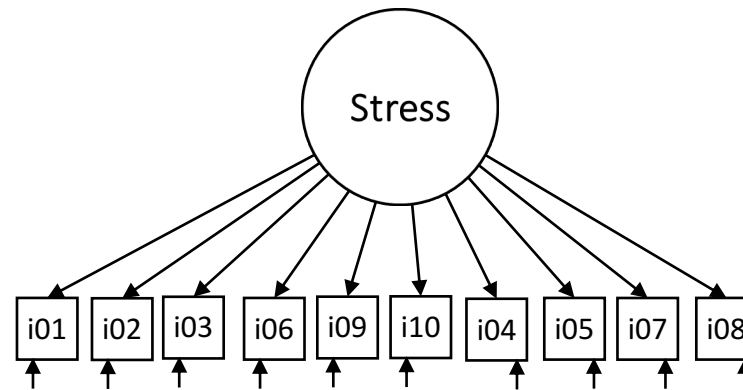
Factorial Structure MASEM(s) - Research Intention:

Using meta-analytic structural equation modeling (MASEM) to analyze the dimensionality of widely applied questionnaires, e.g. the Perceived Stress Scale, a ten item self-report measure to assess perceived stress

Problem: Different factor specifications are evaluated in the literature, e.g.:

Cohen et al. (1988)

Unidimensional model



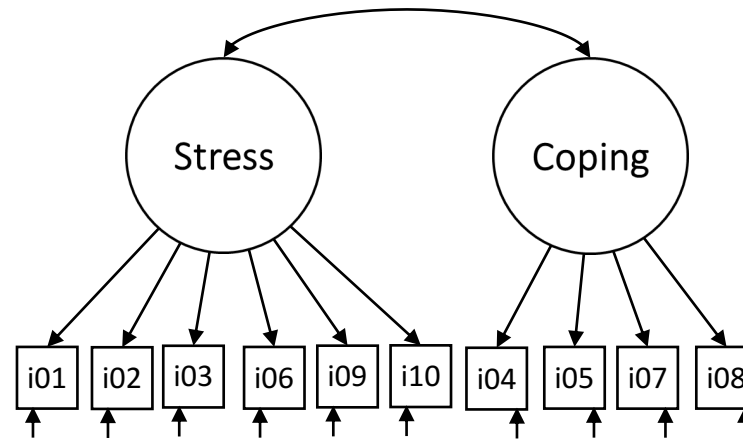
Research Intention:

Using meta-analytic structural equation modeling (MASEM) to analyze the dimensionality of widely applied questionnaires, e.g. the Perceived Stress Scale, a ten item self-report measure to assess perceived stress

Problem: Different factor specifications are evaluated in the literature, e.g.:

Cohen et al. (1988)

Correlated Two-factor model



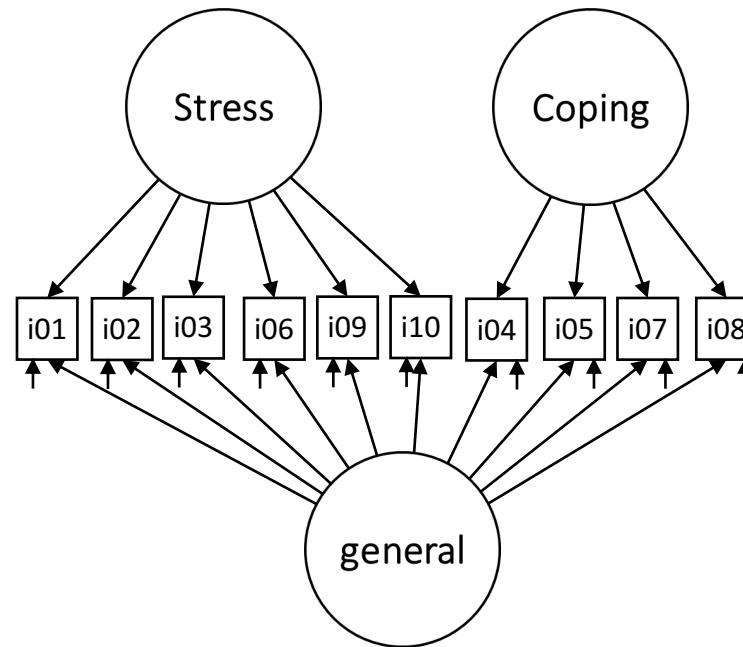
Research Intention:

Using meta-analytic structural equation modeling (MASEM) to analyze the dimensionality of widely applied questionnaires, e.g. the Perceived Stress Scale, a ten item self-report measure to assess perceived stress

Problem: Different factor specifications are evaluated in the literature, e.g.:

Denovan (2017)

Bifactor Specification



Methodological approach:

Systematic literature search to identify studies that report correlations between the 10 items or provide data sets to independently generate inter-item correlation matrices

Correlation matrices from
88 independent
samples

→ Input for random-effects meta-analytic structural equation modelling (MASEM) approach

Data collection issues

Quality appraisal systems

MASEM



Results:

MASEM is basically a two-step procedure:

Step 1:
Pool matrices across
samples

Step 2:
Fit CFA models to the
pooled matrix

Range: 0.150 - 0.543
Median: $r = 0.337$

	i01	i02	i03	i04	i05	i06	i07	i08	i09
i02	0.511								
i03	0.474	0.529							
i04	0.163	0.261	0.207						
i05	0.234	0.338	0.283	0.509					
i06	0.363	0.461	0.441	0.238	0.274				
i07	0.150	0.187	0.157	0.383	0.380	0.175			
i08	0.217	0.336	0.294	0.482	0.543	0.312	0.423		
i09	0.458	0.450	0.447	0.164	0.241	0.379	0.152	0.232	
i10	0.434	0.533	0.505	0.269	0.332	0.529	0.201	0.354	0.482

Wedderhoff, N., Wabner, L., Wunsch, K., & Bosnjak, M. (in prep.). The PSS structure revisited: A meta-analytic investigation of the dimensionality of the Perceived Stress Scale.

Results:

MASEM is basically a two-step procedure:

Step 1:
Pool matrices across
samples

Step 2:
Fit CFA models to the
pooled matrix

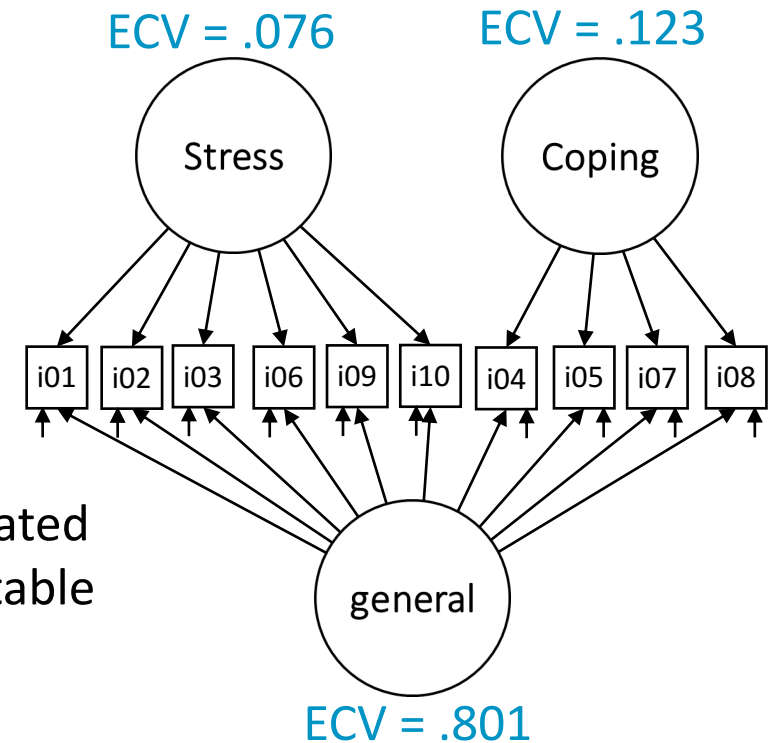
The Bifactor specification exhibited
best model fit

RMSEA: .037 [.035; .039]

SRMR: .021

CFI: .985

However: Explained common variance (ECV) indicated
that most (i.e., 80 %) of the variance was attributable
to the general factor



Conclusion:

The PSS is sufficiently unidimensional to warrant the use of total scores. The specific orthogonal factors only exhibited marginal additional contribution.

However: Substantial amount of heterogeneity in the pooled correlation matrix → Further analyses are needed to prove measurement invariance, e.g.

- (1) across different populations (e.g. clinical vs. non-clinical)
- (2) and application contexts (e.g. cross-cultural comparison)

Our research



➤ Research literacy

- Aims and scope
- Research projects
- Events
- Selected publications
- Members
- Visiting researchers
- Cooperation projects (selection)

➤ Research synthesis

- Aims and scope
- Research projects
- Events
- Edited volumes and series
- Selected publications
- Members
- Visiting researchers
- Cooperation partners

➤ Big data

- Aims and scope
- Events
- Members
- Visiting researchers
- Cooperation partners

