

REGEMA: Guidelines for Conducting and Reporting Reliability Generalization Meta-analyses

Julio Sánchez-Meca¹

José A. López-Pina¹

María Rubio-Aparicio¹

Fulgencio Marín-Martínez¹

Rosa M^a Núñez-Núñez²

Juan J. López-García¹

José A. López-López^{1,3}

¹University of Murcia, Spain

²University Miguel Hernández,
Spain

³University of Bristol, UK



This research was funded by a grant from the Ministerio de Economía y Competitividad of the Spanish Government and Fondo Europeo de Desarrollo Regional (FEDER) (Project No. PSI2016-77676-P)

Background

What is a reliability generalization (RG) meta-analysis?

- ➡ It is a type of **psychometric meta-analysis** that integrates reliability coefficients obtained across different applications of a test to different samples
- ➡ Its aim is to **estimate the average reliability** of the test scores, in order to determine whether reliability can be **generalizable** to different situations, contexts and populations or, otherwise, **identify characteristics** of the studies that are statistically associated with the reliability of the test scores
- ➡ **Examples** of characteristics: standard deviation of the test scores, mean of the scores, target population, test version...

Background

- ➡ Since 1998, when Vacha-Haase coined the term '**reliability generalization**' to refer to this type of meta-analysis, over 180 meta-analyses of this type have been published in Psychology and other fields
- ➡ An informal examination of published RG meta-analyses suggests **substantial heterogeneity** in the reporting of these studies
- ➡ To the best of our knowledge, **no guidelines have been proposed** to help researchers with the adequate reporting of RG meta-analyses

Background

Existing guidelines for the reporting of meta-analyses

- ☞ **PRISMA Statement** (**P**referred **R**eporting **I**tems for **S**ystematic Reviews and **M**eta-**A**nalyses; Moher et al., 2009)
 - ✓ MAs on efficacy of different interventions
 - ✓ MAs on individual participant data
 - ✓ Network meta-analyses

Background

Existing guidelines for the reporting of meta-analyses

- ☞ **AMSTAR guidelines** for the reporting of systematic reviews and meta-analyses on intervention efficacy (Grimshaw et al., 2007)
- ☞ **MOOSE guidelines** (**M**eta-analysis **O**bservational **S**tudies in **E**pidemiology) for the reporting of systematic reviews and meta-analyses of observational studies (Stroup et al., 2000)
- ☞ **MARS guidelines** (**M**eta-**A**nalysis **R**eporting **S**tandards) for the reporting of systematic reviews and meta-analyses on intervention efficacy (APA Publications and Communications Board Working Group on Journal Article Reporting Standards, 2008)

Background

What is the problem?

- ☞ The guidelines mentioned before **are not applicable** to RG meta-analyses

Aim of our study:

To develop guidelines for the adequate reporting of RG meta-analyses



REGEMA: RELiability GENeralization META-analysis

Stage 1

Developing the items

- ☞ We took existing guidelines (PRISMA, AMSTAR, MOOSE and MARS) as a starting point
- ☞ The **sections** of the new instrument were defined as follows:
 - ✓ Title/Abstract
 - ✓ Introduction
 - ✓ Method
 - ✓ Results
 - ✓ Discussion
 - ✓ Funding
- ☞ We suggest using a **flow diagram** to document the process of selection of studies

Some examples of REGEMA items

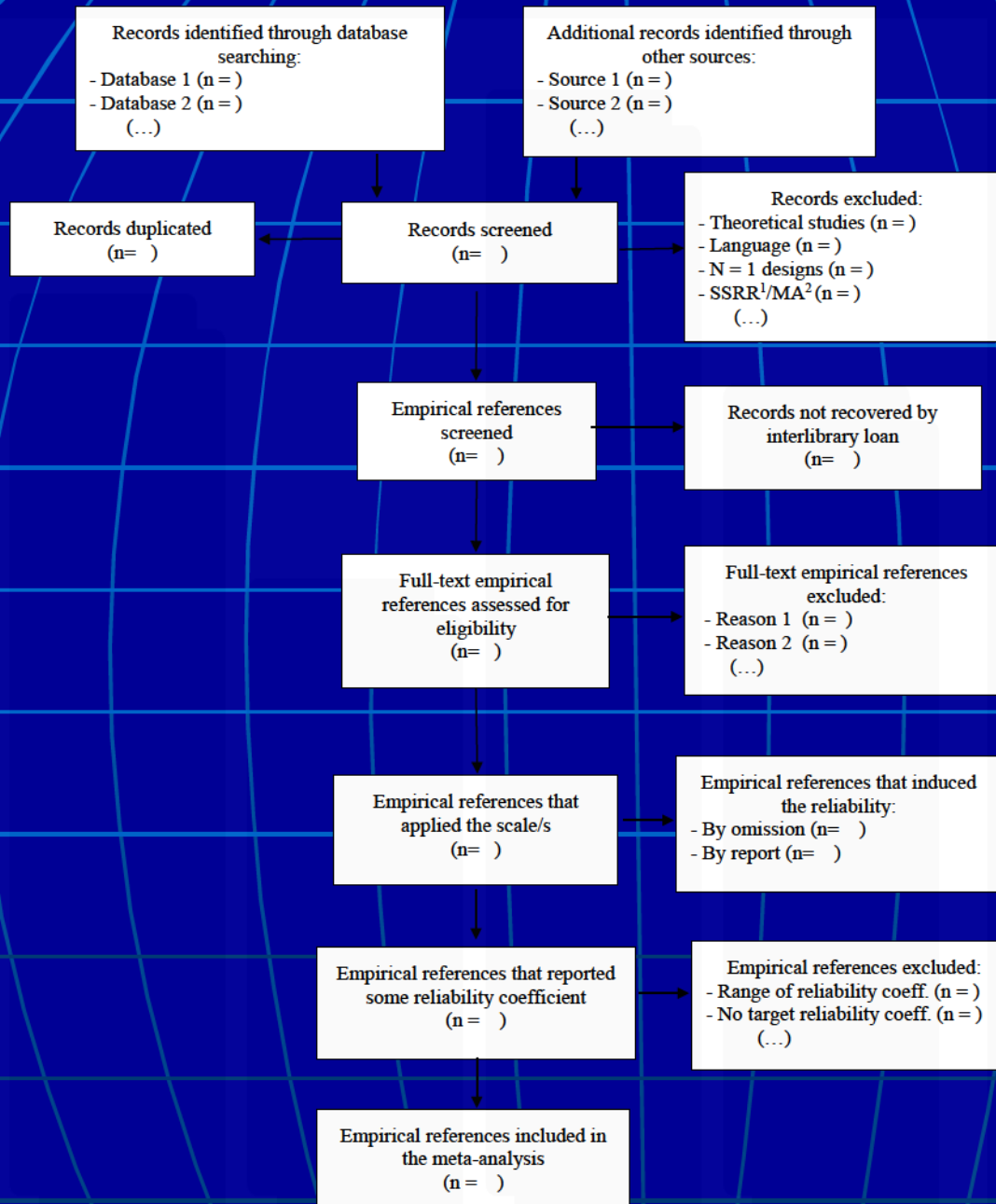
TITLE		Yes	No	Unclear	NA
1. <i>Title</i>	The title must include: (a) the term “reliability generalization” or “meta-analysis” together with some explicit indication to reliability (internal consistency, test-retest, inter- or intra- rater) and (b) the name of the scale or, if more than one scale, the attribute/outcome measure that the scales are assessing.				
ABSTRACT		Yes	No	Unclear	NA
2. <i>Abstract</i>	The abstract must state explicitly: (a) that the objective was to carry out a reliability generalization (RG) meta-analysis of one or several scales; (b) eligibility criteria of the studies; (c) data sources with the temporal range covered; (d) types of reliability coefficients analyzed ; (e) statistical model applied; (f) main results (e.g., pooled reliability coefficient and 95% CI, moderator variables related to reliability); and (g) main conclusions. In case of space limitation, (b) and (c) criteria can be omitted.				
INTRODUCTION		Yes	No	Unclear	NA
3. <i>Background</i>	The background must include: (a) a conceptual definition of the attribute/outcome measure assessed by the scale/s; (b) description of the target population/s to which the scale/s is/are applied and its/their purposes (e.g., screening, clinical diagnosis); (c) a complete description of the scale/s (length, number of categories), including the versions and adaptations to other languages/cultures; and (d) a brief presentation of reliability estimates obtained in previous psychometric studies of the scale/s. Optionally, a brief review of validation studies of the scale/s (e.g., exploratory/confirmatory factor analyses, concurrent/convergent/discriminant validity, responsiveness) could be included.				
4. <i>Objectives</i>	State whether the purpose of the meta-analysis was to obtain a more precise overall reliability coefficient estimate and/or investigate how reliability coefficients vary among different applications of the scales. Optionally, specify whether one objective of the meta-analysis is to estimate the reliability induction rates of the scale/s.				

Some examples of REGEMA items

METHOD		Yes	No	Unclear	NA
5. <i>Selection criteria</i>	Specify inclusion criteria: (a) name/s of the scale/s analyzed in the RG meta-analysis, as well as the versions and/or adaptations included; (b) years considered; (c) language restrictions; (d) publication status; (e) to report any reliability estimate based on the study-specific sample/s; (f) type/s of reliability considered (e.g., internal consistency, temporal stability, inter-/intra-rater reliability); and (g) target population/s (e.g., community, clinical, subclinical/analogue, university).				
6. <i>Search strategies</i>	Specify how the studies were located: (a) electronic databases consulted; (b) other formal search procedures (e.g., manual search in specific journals, backward search from references listed in selected studies); and (c) informal search procedures (e.g., internet searches, contacting study authors to identify additional studies). For electronic searches, describe the search strategy, including the keywords used and how they were combined, and the search limits (e.g., fields where the keywords were searched - title, abstract, full-text -, temporal range, language).				
7. <i>Data extraction</i>	Describe the characteristics extracted from the studies, including: (a) sample size/s, mean/s and standard deviation/s of total test scores and subscales (if applicable); (b) sample characteristics (e.g., target population, country, mean age, standard deviation of the age, gender distribution, ethnic distribution, disorder history –mean and SD in years); (c) test version (e.g., adaptation/version, number of items, reporting format –self-report, clinician); (d) methods (e.g., study design, purpose of the study –psychometric versus applied–, quality checklist); (e) extrinsic characteristics (e.g., publication status, researchers' affiliations, funding source).				
8. <i>Reported reliability</i>	Identify the types of reliability coefficients included in the RG meta-analysis: internal consistency (e.g., Cronbach's alpha, KR-21, parallel forms), temporal stability (test-retest), inter- and intra-rater reliability (e.g., intraclass correlation, kappa coefficient).				
9. <i>Estimating the reliability induction</i>	In case that the meta-analysis intends to estimate the reliability induction, identify the types of reliability induction: induction by omission (no mention of test reliability whatsoever) or reporting induction (vague or precise reporting).				

Stage 1

Title/Abstract	1. Title
	2. Abstract
Introduction	3. Background
	4. Objectives
Method	5. Selection criteria
	6. Search strategies
	7. Data extraction
	8. Reported reliability
	9. Type of reliability induction
	10. Data extraction of inducing studies
	11. Reliability of data extraction
	12. Transformation method
	13. Statistical model
	14. Weighting method
	15. Heterogeneity assessment
	16. Moderator analyses
	17. Graphical techniques
	18. Additional analyses
	19. Software
Results	20. Results of the study selection process
	21. Mean reliability and heterogeneity
	22. Moderator analyses
	23. Sensitivity analyses
	24. Comparison of inducing and reporting studies
	25. Data set
Discussion	26. Summary of results
	27. Limitations
	28. Implications for practice
	29. Implications for future research
Funding	30. Funding



Note. ¹Systematic Reviews. ²Meta-analyses

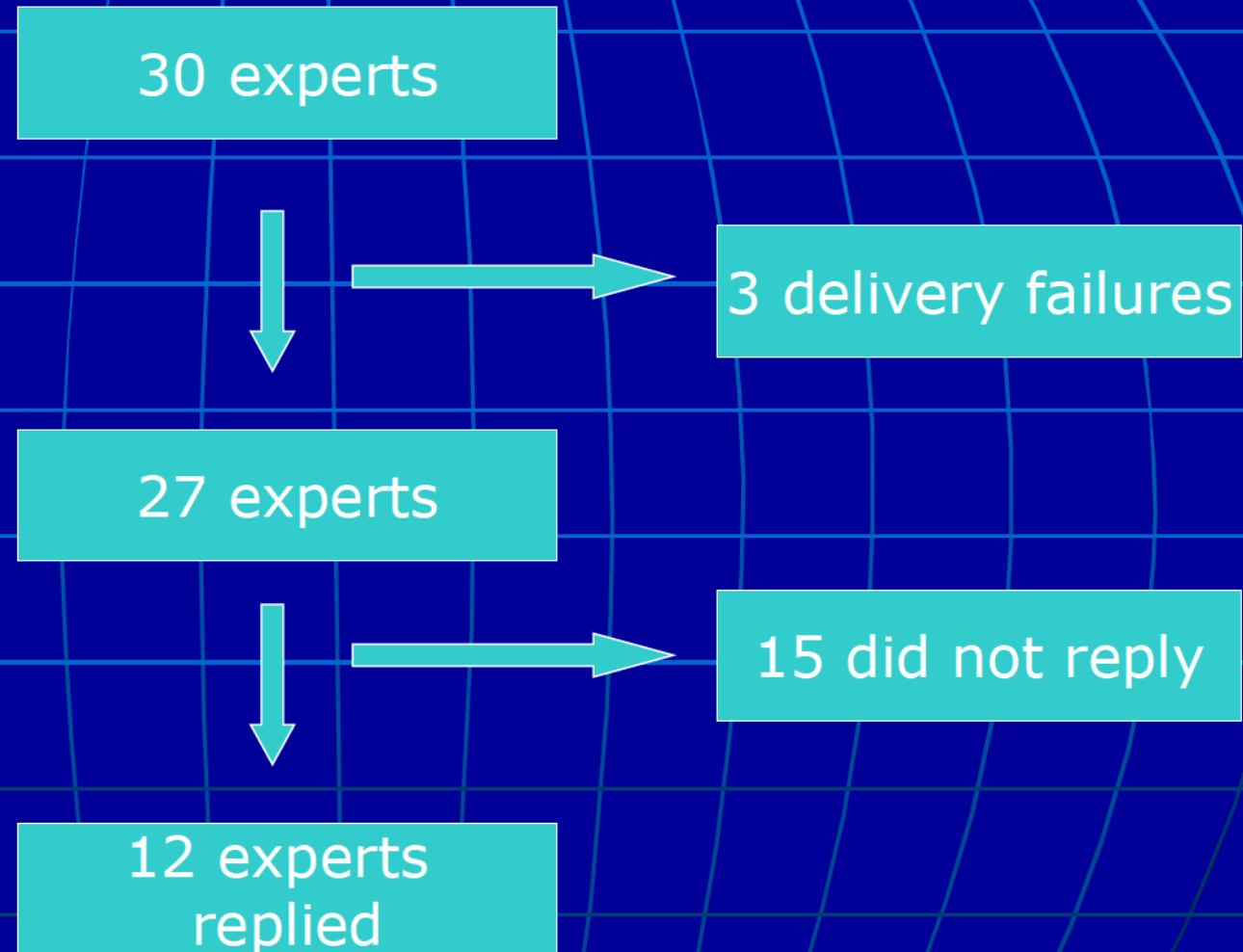
Stage 2

Construct validity

- ➡ We identified 30 prolific researchers in the fields of RG and/or meta-analysis
- ➡ We sent them the REGEMA checklist and asked them to appraise the adequacy of each item and to provide feedback on suggested changes
- ➡ Researchers were contacted via e-mail in order to facilitate a higher response rate

<i>Last name</i>	<i>Forename</i>	<i>Institution</i>
Aguayo	Raimundo	University of Granada, Spain
Becker	Betsy J.	Florida State University, Tallahassee, Florida, USA
Beretvas	Natasha	University of Texas at Austin, USA
Bonett	Douglas	University of California at Santa Barbara, USA
Botella	Juan	Autonomous University of Madrid, Spain
Brannick	Michael T.	University of South Florida, USA
Caruso	John C.	University of Montana, USA
Chacón	Salvador	University of Seville, Spain
Gnambs	Timo	Osnabrück University, Germany
Graham	James M.	Western Washington University, USA
Guillén	Alejandro	University of Granada, Spain
Hellman	Chan M.	University of Oklahoma, USA
Henson	Robin K.	University of North Texas, USA
Howell	Ryan T.	San Francisco State University, USA
Kieffer	Kevin M.	Saint Leo University, FL, USA
King	Chelsi	Alliant International University, Fresno, CA, USA
Kivisalu	Trisha M.	Alliant International University, Fresno, CA, USA
Kromrey	Jeffrey D.	University of South Florida, Tampa, FL, USA
Phillips	Colleen E.	Alliant International University, Fresno, CA, USA
Romano	Jeanine L.	American University of Sharjah, Sharjah, United Arab Emirates
Sanduvete	Susana	University of Seville, Spain
Shields	Alan L.	Adelphi Values Boston, MA, USA
Sun	Shuyan	University of Maryland, Baltimore, MD, USA
Thompson	Bruce	Texas A&M University, USA
Vacha-Haase	Tammi	Colorado State University, USA
Van den Noortgate	Wim	University of Louvain, Belgium
Vassar	Matt	Oklahoma State University, Tulsa, OK, USA
Viechtbauer	Wolfgang	Maastricht University, The Netherlands
Wheeler	Denna L.	University of Oklahoma, Tulsa, OK, USA
Worley	Jody A.	University of Oklahoma, Tulsa, OK, USA

Stage 2



Stage 3: analysis

Section	Item	%	Comments	Changes
Title/Abstract	1. Title	100	2	Yes
	2. Abstract	91,7	5	Yes
Introduction	3. Background	91,7	3	Yes
	4. Objectives	91,7	3	Yes
Method	5. Selection criteria	100	2	No
	6. Search strategies	100	3	No
	7. Data extraction	100	3	No
	8. Reported reliability	100	2	No
	9. Type of reliability induction	58,3	6	Yes
	10. Data extraction of inducing studies	66,7	5	No
	11. Reliability of data extraction	91,7	2	No
	12. Transformation method	100	1	No
	13. Statistical model	100	2	No
	14. Weighting method	100	3	Yes
	15. Heterogeneity assessment	100	1	Yes
	16. Moderator analyses	91,7	1	Yes
	17. Graphical techniques	66,7	4	Yes
	18. Additional analyses	100	1	No
	19. Software	100	2	Yes
Results	20. Results of the study selection process	100	3	No
	21. Mean reliability and heterogeneity	91,7	3	Yes
	22. Moderator analyses	83,3	4	No
	23. Sensitivity analyses	100	2	No
	24. Comparison of inducing and reporting studies	100	0	No
	25. Data set	100	2	Yes
Discussion	26. Summary of results	100	1	Yes
	27. Limitations	100	0	No
	28. Implications for practice	100	1	No
	29. Implications for future research	100	0	No
Funding	30. Funding	100	1	No

Removed

Stage 4: Final version of REGEMA

Section	Item
Title/Abstract	1. Title
	2. Abstract
Introduction	3. Background
	4. Objectives
Method	5. Selection criteria
	6. Search strategies
	7. Data extraction
	8. Reported reliability
	9. Type of reliability induction
	10. Data extraction of inducing studies
	11. Reliability of data extraction
	12. Transformation method
	13. Statistical model
	14. Weighting method
	15. Heterogeneity assessment
	16. Moderator analyses
	17. Additional analyses
	18. Software
Results	19. Results of the study selection process
	20. Mean reliability and heterogeneity
	21. Moderator analyses
	22. Sensitivity analyses
	23. Comparison of inducing and reporting studies
	24. Data set
Discussion	25. Summary of results
	26. Limitations
	27. Implications for practice
	28. Implications for future research
Funding	29. Funding

Conclusion

- ☞ We developed REGEMA, a checklist to guide researchers in the correct reporting of RG meta-analyses
- ☞ REGEMA offers an acceptable construct validity
- ☞ Limitation: only 44.4% of experts replied to our request

Conclusion

Who might benefit from REGEMA?

- ☞ Researchers interested in conducting an RG meta-analysis
- ☞ Researchers interested in a critical appraisal of RG meta-analyses undertaken by other teams
- ☞ Journal editors and reviewers, as REGEMA provides standards to assess the quality of an RG meta-analysis as part of the peer-reviewing process

Conclusion

What are our next steps?

- ➡ Publishing REGEMA
- ➡ Making it accessible to any researcher interested in using it, by adding a link to the website of the Meta-Analysis Unit at University of Murcia
<https://www.um.es/metaanalysis/>
- ➡ Applying REGEMA to every RG meta-analysis conducted to date, not only in Psychology but also in other related and/or Health Sciences

REGEMA: Guidelines for Conducting and Reporting Reliability Generalization Meta-analyses

Julio Sánchez-Meca¹

José A. López-Pina¹

María Rubio-Aparicio¹

Fulgencio Marín-Martínez¹

Rosa M^a Núñez-Núñez²

Juan J. López-García¹

José A. López-López^{1,3}

¹University of Murcia, Spain

²University Miguel Hernández,
Spain

³University of Bristol, UK



This research was funded by a grant from the Ministerio de Economía y Competitividad of the Spanish Government and Fondo Europeo de Desarrollo Regional (FEDER) (Project No. PSI2016-77676-P).