Transparency and Reporting Practices in Reliability Generalization Meta-analyses: Assessment with the REGEMA checklist

Julio Sánchez-Meca[1], Fulgencio Marín-Martínez[1], Rosa Mª Núñez-Núñez[2], María Rubio-Aparicio[3], José Antonio López-López[1], Desirée Mª Blázquez-Rincón[1], Carmen López-Ibáñez[1], Rubén López-Nicolás[1], José Antonio López-Pina[1] and Juan José López-García[1]

[1]University of Murcia (Spain)

[2]University Miguel Hernández at Elche (Spain)

[3]University of Alicante (Spain)

UNIVERSIDAD DE MURCIA

1. Fundaments
1.1 Focusing the problem

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

☞ It's a type of **psychometric meta-analysis** that aims to integrate reliability coefficients obtained when a given test is applied to different primary studies, to examine how reliability of a test scores varies from an application to the next

☞ Its purpose is **to estimate the average reliability** of a test scores, to investigate whether reliability **can be generalized** to different contexts, situations, and target populations, and in case of heterogeneity, **to identify study characteristics** that might be statistically associated to reliability coefficients (e.g., mean and SD of tests cores, target population, test version, etc.).

1. Fundaments
1.1 Focusing the problem

☞ Current checklists to guide the reporting of meta-analyses **are not adequate** for RG meta-analyses:

PRISMA checklist (**P**referred **R**eporting **I**tems for **S**ystematic Reviews and **M**eta-**A**nalyses; Moher et al., 2009)

AMSTAR checklist for reporting meta-analyses on intervention efficacy (Grimshaw, Wells, et al., 2007)

MOOSE checklist for reporting meta-analyses of observational studies (Stroup, Berlin, Morton et al., 2000)

MARS guidelines for reporting meta-analyses on intervention efficacy (APA Publications and Communications Board Working Group on Journal Article Reporting Standards, 2008)

1. Fundaments
1.1 Focusing the problem

☞ With this purpose in 2017 we developped the **REGEMA checklist** (**RE**liability **GE**neralization **M**eta-**A**nalysis)

Title	1 item
Abstract	1 item
Introduction	2 items
Method	14 items
Results	6 items
Discussion	4 items
Funding	1 item
Protocol	1 item
TOTAL	30 items

Research
Synthesis Methods WILEY

Improving the reporting quality of reliability generalization meta-analyses: The REGEMA checklist

Julio Sánchez-Meca[1] | Fulgencio Marín-Martínez[1] |
José Antonio López-López[1] | Rosa Maria Núñez-Núñez[2] |
María Rubio-Aparicio[3] | Juan José López-García[1] | José Antonio López-Pina[1] |
Desirée Mª. Blázquez-Rincón[1] | Carmen López-Ibáñez[1] | Rubén López-Nicolás[1]

[1]Department of Basic Psychology & Methodology, University of Murcia, Murcia, Spain

[2]Department of Behavioral & Health Sciences, Miguel Hernández University of Elche, Elche, Spain

[3]Department of Health Psychology, University of Alicante, Alicante, Spain

Correspondence

Abstract

Reliability generalization (RG) is a meta-analytic approach that aims to characterize how reliability estimates from the same test vary across different applications of the instrument. With this purpose RG meta-analyses typically focus on a particular test and intend to obtain an overall reliability of test scores and to investigate how the composition and variability of the samples affect reliability. Although several guidelines have been proposed in the meta-analytic liter-

The Meta-Analysis Unit

um.es

Reliability Generalization Meta-analysis: The REGEMA Checklist

Reliability generalization (RG) meta-analyses are a special kind of meta-analyses whose purpose is to investigate how r composition and variability of the test scores. In an RG meta-analysis, the outcomes from each primary study are reliab contextual characteristics of the studies are the potential moderators that can explain, at least, part of the variability typ

Several reporting guidelines have been developed in the meta-analytic arena to help researchers report meta-analyses applied on RG meta-analyses, but on meta-analyses on intervention effects or on the relationship between variables. Th proposed to date in the literature. The REGEMA checklist is a tool elaborated by the research team of the Meta-analysis

The REGEMA checklist is a tool composed by 30 items (1 item for the Title, 1 for the Abstract, 2 for the Introduction, 14 f is very easy to apply. REGEMA can be applied both by researchers that intend to write an RG meta-analysis and to make assess the adequacy of an RG meta-analysis.

A MS Word version of the REGEMA checklist can be freely downloaded from the following link:

REGEMA checklist

In addition, we have elaborated a flowchart, similar to that proposed by the PRISMA statement, but adapted to the spec searching and selecting the primary studies that fulfil the selection criteria of an RG meta-analysis.

A MS Word version of the REGEMA flowchart can be freely downloaded from the following link:

REGEMA flowchart

How to cite the REGEMA checklist:

Sánchez-Meca, J., Marín-Martínez, F., López-López, J. A., Núñez-Núñez, R. M., Rubio-Aparicio, M., López-García, J. J., Lóp the reporting quality of reliability generalization meta-analyses: The REGEMA checklist. *Research Synthesis Methods*

▪top

1. Fundaments
1.2 Objectives

1º To examine **the inter-rater reliability** of the REGEMA checklist

2º To investigate the capacity of REGEMA checklist to assess the **reporting quality** of RG meta-analyses

3º To assess the **degree of compliance** of REGEMA checklist by RG meta-analyses

2. Method
2.1 Selection criteria of the studies

(a) RG meta-analysis reporting reliability estimates of one or several tests to assess psychological constructs

(b) It had to be carried out between 1998 and 2019

(c) It had to include all primary studies that applied the test/s of interest (meta-analyses of only psychometric studies were excluded)

(d) In case of focusing on more than one test, it had to report separate reliability estimates of each one of them

(e) To be written in English or Spanish

2. Method
2.2 Searching for the studies

Electronic databases consulted:

Google Scholar

PsycInfo

Web of Science

Key-words:
Reliability generalization
Meta-analysis AND Reliability
Meta-analysis AND 'Internal consistency'
Meta-analysis AND 'Test-retest'
Meta-analysis AND ('interrater' OR 'intrarater')
Meta-analysis AND ('alpha coefficient' OR 'coefficient alpha')
Meta-analysis AND intraclass

Identification

Records identified through
database searching
(n = 221)

Additional records identified
through other sources
(n = 142)

Screening

Records after duplicates removed
(n = 350)

Records screened
(n = 350)

Records excluded
(n = 27)

Eligibility

Full-text articles assessed
for eligibility
(n = 323)

Full-text articles excluded,
with reasons
(n = 173)

Included

**Studies included in the
systematic review
(n = 150)**

19 unpublished studies

131 published studies

2. Method
2.3 Data extraction

(a) **REGEMA checklist** was applied to each RG meta-analysis

(b) **Other characteristics registered**: publication source, year of the meta-analysis, and idiom

(c) **Inter-rater reliability**: two coders independently extracted the data from the RG meta-analyses and applied the REGEMA checklist. Inconsistencies were resolved by consensus

2. Method
2.4 Statistical analysis

(a) To estimate the inter-rater reliability of REGEMA checklist, **Cohen's kappa coefficient and inter-rater agreement percentage** were calculated for each item and subitem of the checklist

(b) To assess the degree of compliance of each item of the REGEMA checklist, the **compliance percentage** from all the RG meta-analyses analyzed was calculated.

3. Results
3.1 Inter-rater reliability of REGEMA

Item/Subitem	κ	% agreement
1. Title: The title must include:		
1.1. The term "reliability generalization" or "meta-analysis"	.96	99.3
1.2. The name of the scale or, if more than one scale, the attribute/outcome measure that the scales are assessing	.39	98.0
2. Abstract. The abstract must state explicitly:		
2.1. That the objective was to carry out a reliability generalization (RG) meta-analysis of one or several scales	.69	95.3
2.4. Type/s of reliability analyzed (internal consistency, temporal stability, inter-rater / intra-rater agreement, etc.)	.87	94.7
2.6. Main results	.88	98.0
3. INTRODUCTION. Background. The background must include:		
3.1. A conceptual definition of the attribute/outcome measure assessed by the scale/s	.79	90.0
3.2. Description of the target population/s to which the scale/s is/are applied and/or its/their purposes	.73	90.6
3.3. A complete description of the scale/s (length, number of categories)	.78	90.7
4. INTRODUCTION. Objective. States that the purpose of the RG was to obtain a more precise overall reliability coefficient estimate and/or investigate how reliability coefficients vary among different applications of the scale/s	.69	96.7

In red color = Kappas under .61. In black color = kappas between .61 and .99. In blue color = perfect kappas

% of agreement under 90% = in red color

% of agreement of 100% = in blue color

3. Results
3.1 Inter-rater reliability of REGEMA

Item/Subitem	κ	% agreement
5. METHOD. Specify selection criteria of the studies	.76	88.0
6. METHOD. Search strategies. Specifies how the studies were located:		
6.1. Electronic databases consulted	.88	99.4
6.4. For electronic searches, describe the search strategy, including the keywords used and the search limits	.79	96.0
6.5. Temporal range	.89	94.7
6.6. Language restrictions	.86	93.3
7. METHOD. Data extraction. Describes the characteristics extracted from the studies, including:		
7.1. Sample size/s, mean/s and standard deviation/s of total test scores	.84	94.0
7.2. Sample characteristics (e.g., target population, country, mean age, standard deviation of the age, etc.)	.90	98.0
7.3. Test version (e.g., adaptation/version, number of items, reporting format –self-report, clinician)	.62	84.6
7.4. Methods (e.g., study design, purpose of the study –psychometric versus applied–, quality checklist)	.80	90.6
8. METHOD. Reported reliability. Identifies the types of reliability coefficients included in the RG meta-analysis	.64	94.6
9. METHOD. Estimating the reliability induction rates	.80	90.0
10. METHOD. Data extraction of inducing studies. Characteristics of inducing studies were also extracted	.59	97.3
11. METHOD. Reliability of data extraction. Describes how the reliability of data extraction process was appraised	.91	96.0
12. METHOD. Transformation method. States whether or not the reliability coefficients were transformed	.95	97.4
13. METHOD. Statistical model. Describes the statistical model(s) assumed in the meta-analytic integration	.85	92.7
14. METHOD. Weighting method. Specifies the weighting method applied in the meta-analytic integration	.83	89.3
15. METHOD. Heterogeneity assessment. Describes how heterogeneity among reliability coefficients was assessed	.82	92.0
16. METHOD. Moderator analyses. Describes how the influence of moderator variables was assessed	.86	93.4
17. METHOD. Additional analyses. Describes other additional analyses accomplished, such as sensitivity analyses	.83	92.6
18. METHOD. Software. Mentions the software and version used to carry out the statistical analyses	.97	98.7

In red color = Kappas under .61. In black color = kappas between .61 and .99. In blue color = perfect kappas

% of agreement under 90% = in red color
% of agreement of 100% = in blue color

3. Results
3.1 Inter-rater reliability of REGEMA

Item/Subitem	κ	% agreement
19. RESULTS. Results of the study selection process:		
19.1. Describes the selection process of the studies and/or present a flowchart.	.72	93.4
19.4. Presents a flowchart	1	100
20. RESULTS. Mean reliability and heterogeneity:		
20.1. Presents, at least, a pooled reliability coefficient and confidence/credibility intervals	.88	94.0
20.2. Presents the results in their original metric	.56	98.0
20.3. Illustrates the distribution of reliability coefficients with graphical techniques	.88	94.7
20.4. Describes the degree of heterogeneity by one or more heterogeneity measures	.71	92.0
20.5. Presents a forest plot of the individual reliability coefficients	.88	98.0
21. RESULTS. Moderator analyses: Presents the results of moderator analyses	.87	98.0
22. RESULTS. Sensitivity analyses: Reports the results of any sensitivity analyses	.77	88.6
23. RESULTS. Comparison of inducing and reporting studies. Comparing inducing and reporting studies	1	100
24. RESULTS. Data set:		
24.1. Tabulates the characteristics of the individual studies that reported reliability	.91	96.7
24.2. List of all studies included in the RG meta-analysis	.87	96.0

In red color = Kappas under .61. In black color = kappas between .61 and .99. In blue color = perfect kappas

% of agreement under 90% = in red color

% of agreement of 100% = in blue color

3. Results
3.1 Inter-rater reliability of REGEMA

Item/Subitem	κ	% agreement
25. DISCUSSION. Summary of results: Presents the main results	.59	97.3
26. DISCUSSION. Limitations: Discusses the limitations of the meta-analysis	.82	92.7
27. DISCUSSION. Implications for practice: Provides guidelines for professional practice	.64	82.0
28. DISCUSSION. Implications for future research: Includes recommendations for researchers	.65	93.0
29. FUNDING. State the financial sources of the meta-analysis	.91	96.6
30. PROTOCOL. State whether a protocol of the meta-analysis was previously published	--	100

In red color = Kappas under .61. In black color = kappas between .61 and .99. In blue color = perfect kappas

% of agreement under 90% = in red color

% of agreement of 100% = in blue color

3. Results
3.2 Compliance degree of REGEMA

Item	% compliance
1. Title	88.7
2. Abstract. The abstract must state explicitly	66.0
3. INTRODUCTION. Background	49.3
4. INTRODUCTION. Objective	96.0
5. METHOD. Specify selection criteria of the studies	49.3
6. METHOD. Search strategies	62.0
7. METHOD. Data extraction	73.3
8. METHOD. Reported reliability. Identifies the types of reliability coefficients included in the RG meta-analysis	92.7
9. METHOD. Estimating the reliability induction rates	56.0
10. METHOD. Data extraction of inducing studies. Characteristics of inducing studies were also extracted	2.0
11. METHOD. Reliability of data extraction. Describes how the reliability of data extraction process was appraised	36.7
12. METHOD. Transformation method. States whether or not the reliability coefficients were transformed	44.0
13. METHOD. Statistical model. Describes the statistical model(s) assumed in the meta-analytic integration	57.3
14. METHOD. Weighting method. Specifies the weighting method applied in the meta-analytic integration	60.0
15. METHOD. Heterogeneity assessment. Describes how heterogeneity among reliability coefficients was assessed	35.3
16. METHOD. Moderator analyses. Describes how the influence of moderator variables was assessed	57.3
17. METHOD. Additional analyses. Describes other additional analyses accomplished, such as sensitivity analyses	31.3
18. METHOD. Software. Mentions the software and version used to carry out the statistical analyses	31.3

% of Compliance	Degree	Color
< 50%	Deficient	Red
50 – 89%	Moderate	Black
≥ 90%	Excellent	Blue

3. Results
3.2 Compliance degree of REGEMA

Item	% Compliance
19. RESULTS. Results of the study selection process	87.3
20. RESULTS. Mean reliability and heterogeneity	52.0
21. RESULTS. Moderator analyses: Presents the results of moderator analyses	92.7
22. RESULTS. Sensitivity analyses: Reports the results of any sensitivity analyses	48.7
23. RESULTS. Comparison of inducing and reporting studies. Comparing inducing and reporting studies	2.0
24. RESULTS. Data set	82.0
25. DISCUSSION. Summary of results: Presents the main results	96.7
26. DISCUSSION. Limitations: Discusses the limitations of the meta-analysis	74.0
27. DISCUSSION. Implications for practice: Provides guidelines for professional practice	50.7
28. DISCUSSION. Implications for future research: Includes recommendations for researchers	88.7
29. FUNDING. State the financial sources of the meta-analysis	26.0
30. PROTOCOL. State whether a protocol of the meta-analysis was previously published	0

% of Compliance	Degree	Color
< 50%	Deficient	Red
50 – 89%	Moderate	Black
≥ 90%	Excellent	Blue

4. Conclusions

☞ REGEMA checklist exhibited a **very satisfactory degree of inter-rater reliability**

☞ REGEMA checklist is easy to be applied and enables assessing the reporting quality of RG meta-analyses

☞ The **degree of compliance** of REGEMA by RG meta-analyses is deficient in many of their items (< 50% of compliance), especially in:

- Background: defining the psychological construct assessed
- Determining the selection criteria of the studies
- The reliability analysis of data extraction data from the studies
- How heterogeneity among reliability coefficients was assessed
- Publishing a Protocol

4. Conclusions

☞ REGEMA checklist can be useful for:

● **Researchers** interested in conducting an RG meta-analysis

● Potential readers of RG meta-analyses, as a tool to make a **critical reading** of them

● **Editors** of scientific journals that publish RG meta-analyses, as a guide for an **adequate reporting** of this type of meta-analyses