Archives of Scientific Psychology Reporting Questions for Manuscripts Describing Meta-Analyses

MANUSCRIPT SECTION	Description
TITLE Meta-Analysis of the Discriminative Validity of Caregiver, Youth, and Teacher Rating Scales for Pediatric Bipolar Disorder	***Please answer "yes" or "no" to each question about the manuscript's Title, Author Note, and Scientific Abstract. If your answer is "No", please provide a brief explanation.*** • Does the Title make it clear that the manuscript describes a research synthesis and include the term "meta-analysis"? Yes ☒ No ☐ If no, please explain:
AUTHOR NOTE For a review of what should be included in the Author Note, see the *Publication Manual of the American Psychological Association*: http://www.apastyle.org/manual/	• Does the Author Note contain acknowledgment of special circumstances, for example: ○ use of data also appearing in previous publications, dissertations, conference papers? Yes ☐ No ☒ If yes, please explain: No special circumstances

	○ sources of funding or other support? Yes ■ No ☐ If yes, please explain: ○ relationships that may be perceived as conflicts of interest? Yes ■ No ☐ If yes, please explain: Yes, erring on the side of over-disclosure

SCIENTIFIC ABSTRACT	• Does the Scientific Abstract describe ○ the problem or relation(s) under investigation? Yes ■ No ☐ If no, please explain: ○ the study eligibility criteria? Yes ■ No ☐ If no, please explain: ○ the type(s) of participants included in primary studies? Yes ■ No ☐ If no, please explain:

- the meta-analysis methods (indicating whether a fixed or random model was used)?
  Yes ■ No ☐

  If no, please explain:

- the main results (including the more important effect sizes and any important moderators of these effect sizes)?
  Yes ■ No ☐

  If no, please explain:

- the conclusions (including limitations)?
  Yes ■ No ☐

  If no, please explain:

	○ implications for theory, policy and/or practice? Yes ■ No ☐ If no, please explain:
INTRODUCTION	***For the Introduction section, please indicate whether the requested information can be found in this section of the manuscript, in a supplemental file, or whether the information is not relevant to the study. If the information is not relevant, please provide a brief explanation.*** • Does the Introduction include ○ a clear statement of the question or relation(s) under investigation? In manuscript ■ In supplemental files ☐ Not relevant ☐ If not relevant, please explain: ○ the historical background? In manuscript ■ In supplemental files ☐ Not relevant ☐ If not relevant, please explain:

	○ the theoretical, policy and/or practical issues related to the question or relation(s) of interest? In manuscript ■ In supplemental files ☐ Not relevant ☐ If not relevant, please explain: ○ a rationale for the selection and coding of potential moderators and mediators of results? In manuscript ■ In supplemental files ☐ Not relevant ☐ If not relevant, please explain: ○ the types of study designs used in the primary research, their strengths, and their weaknesses? In manuscript ■ In supplemental files ☐ Not relevant ☐ If not relevant, please explain:

- the types of predictor and outcome measures used, their psychometric characteristics?
  In manuscript ◘ In supplemental files ☐ Not relevant ☐

  If not relevant, please explain:

- the populations to whom the question or relation is relevant?
  In manuscript ◘ In supplemental files ☐ Not relevant ☐

  If not relevant, please explain:

- hypotheses, if any?
  In manuscript ◘ In supplemental files ☐ Not relevant ☐

  If not relevant, please explain:

METHOD	***For the Method section, please provide the information requested below, regardless of whether it also appears in the manuscript or supplemental files.***
Inclusion and exclusion criteria:	• What were the o operational characteristics of independent (predictor) and dependent (outcome) variable(s)? **Informant: caregiver, teacher, or youth** **Type of scale: did the scale contain symptoms specific to mania or comprise of items/subscales originally designed to measure other pathology** **Distilled sample design: were healthy controls included as part of comparison group?** **Interview strategy: were the criterion dx derived from interviews soley w/ the primary caregiver or did it also involve direct interview of the youth?** o eligible participant populations? A) cases w/ a dx of bp made via a structured/semi-structured interview b) as well as a comparison grou c) w/ both groups completing the same checklists assessing manic/hypomanic/externalizing symptoms d) data reported for participants 18 yrs or younger o eligible research design features (e.g., random assignment only, minimal sample size)? *drawn from clinical or community samples * had to have 10 or more bp dx *had to include a rating scale * be published in English format * data from both bp and comparison grou * published after 1993 * participants under the age of 18

	○ time periods in which studies needed to be conducted?
	1993 or later (to insure the availablity/use of DSM-IV criteria
	○ geographical and/or cultural restrictions, if any?
	No restrictions
----------	***----------***
Moderator and mediator analyses:	• Please provide a definition of all coding categories used to test moderators or mediators of the relation(s) of interest

Variable	Definition
Informant	who completed the checklist or scale
Type of scale	If the scale comprised of items/subscales desigened to measure other pathology or contained symptoms specific to mania
Distilled	if the original study included healthy controls as part of the comparison group

---

Search strategies:

- Please identify
    - in a table the reference and citation databases and registries searched (including prospective registries), along with
        - keywords used to enter databases and registries
        - search software used and version (e.g., Ovid)
        - time period in which studies needed to be conducted, if applicable

Database Searched	Years Included	Keywords Used	Search Software
Pubmed, PsycINFO, ERIC, and Google Scholar	**1993 – Sept. 1, 2014**	(Pediatric OR juvenile OR child* OR adolescen*) AND ("bipolar disorder" OR man* OR cyclothymi*) AND [(sensitivity AND specificity) OR comparison]	

    - any other efforts to retrieve all available studies (e.g., listservs queried, contacts made with authors and how authors were chosen, reference lists of reports examined):

- the method of addressing reports in languages other than English, if used:

did not locate any primary reports published in languages other than English (although some reports published in English language journals gathered data using translated versions of measures into Korean/Hangul, French, and Dutch)

- the process for determining study eligibility, including
  - the aspects of reports that were examined (i.e., title, abstract, and/or full text)
  - the number and qualifications of individuals who judged the relevance of studies
  - any indication of judge agreement
  - how disagreements were resolved

*Reviewed titles and abstracts
* Four relavance judges completed the search training (included a review of guidelines and a session of orientation and consultation w/ a reference lib about search optimization) and then conducted/reviewed the searches
* Content expert (EAY) reviewed ambiguous cases and instances of disagreement

- how unpublished studies were treated, and, if they were not included, why:

Unpublished studies were included when raw data were available

Coding procedures:	• Please describe o the number and qualifications of coders (e.g., level of expertise in the area, training): Undergrad psych majors, doctoral students, & senior invesitgator Training: reading methodology papers (QUADAS, PRISMA, STARD), sample meta-analyses focusing on PBD, orientation to diagnostic efficiency states, coding two articles & comparing the scores to content expert to resolve discrepancies/clarify concepts o the intercoder reliability or agreement: ICC >.87 for demographics & moderator variables >.95 for effect size metrics >.80 for quality ratings o whether each report was coded by more than one coder, and, if so, how disagreements were resolved: double coded all studies for effect sizes, moderator variables, and qualtiy content expert reviewed all discrepancies and assigned final code after perusing the source material

- how study quality was assessed
  - if a quality scale was employed, a description of criteria and the procedures for application:

Kowatch system: assigned points for adequate sample size, interviewing both caregiver and youth, using formal concensus process, following DSM criteria, including spectrum dx, recording comorbid dx, asking about lifetime episodes

QUADAS-2: operationally defines coding criteria for STARD guidelines

  - if study design features were coded, what these were:

Kowatch system and QUADAS-2

- how missing data were handled:

multivariate meta-regression made it possible to include all studies despite differences in informants used

- Please describe
  - the effect size metric(s)
    - the effect sizes calculating formulas (e.g., means and SDs, use of univariate F to r transform, etc.):

Hedges' g

Converted reported stats into g
AUC --> Cohen's d --> g
Sens&spec --> auc --> d --> g
means&st dev & n --> g

    - any corrections made to effect sizes (e.g., small sample bias, correction for unequal ns, etc.):

Hedges' g corrects Cohen's d for slight upward bias

- the effect size averaging and/or weighting method(s):

Inverse variance weighting

- how the effect size confidence intervals (or standard errors) were calculated:

standard methods -- 95% confidence intervals based on critical z score times SE

- how effect size credibility intervals were calculated, if used:

METAFOR provides these; available as supplemental material

- how studies with more than one effect size were handled:

Nesting
Multivariate meta-regression in METAFOR

- whether fixed and/or random effects models were used and the model choice justification:

Used mixed model approach b/c we had several hypothesis-driven moderators of interest but also wanted to preserve generalizability

- how heterogeneity in effect sizes was assessed or estimated:

Cochran's Q tested homogeneity of effect sizes, along with graphical methods (e.g., forest plots)

- the means and SDs for measurement artifacts, if construct-level relationships were the focus:

N/A - informant is measured categorically, not as a psychometric construct

- tests and any adjustments for data censoring (e.g., publication bias, selective reporting):

Mixed model extension of Egger's test for publication bias

- tests for statistical outliers:

Examined standardized residuals from fitted models to test for influental outliers while accounting for the nested structure of the data

- the statistical power of the meta-analysis, if calculated:

power exceeded 99.9% to reject null hypothesis of g~0
power to detect moderate heterogeneity was between .64-.91 (27 independent samples vs. 63 disaggregated effects)
power for large heterogeneity .86-.99

- the statistical programs or software packages used to conduct statistical analyses:

Metafor

RESULTS	***For the Results section, please provide the information requested in the questionnaire or in the text box provide the page number, table, or supplemental file in which the information can be found.*** **If your manuscript is accepted for publication, you will need to deposit your data set in an approved data repository. Please see Instructions to Authors for more information: www.apa.org/pubs/journals/arc** • Please provide o the number of citations examined for relevance: 4094 hits in Pubmed, 1325 hits in PsycINFO (fig 1) o a list of citations included in the synthesis 17 artiicles selected for inclusion & coded (figure 1) o the number of citations relevant on many but not all inclusion criteria that ultimately were excluded from the meta-analysis: 52 (figure 1)

	▪ Number of exclusions for each exclusion criteria (e.g., effect size could not be calculated), with examples: 4025 articles excluded due to: not studying bp, no original data reported, no bp case reported, adult sample, repeat of previously included sample 52 articles excluded due to: did not separate child and adult data, did not provide sufficient statistics to estimate a standardized effect size (figure 1) ○ a table with descriptive information for each included study, including effect size and sample size: table 2: effect size level characteristics and moderators

	○ the results of the assessment of study quality, if any: table 1: summary of sample-level characteristics of studies included in meta-analysis ○ tables and/or graphic summaries of ▪ overall characteristics of the database (e.g., number of studies with different research designs) ▪ overall effect size estimates, including measures of uncertainty (e.g., confidence and/or credibility intervals) table 1: summary of sample-level characteristics of studies included in meta-analysis table 2: effect size level characteristics and moderators table 4: file drawer estimates for disaggregated effect size estimates grouped by informant table 3: tests of homogeneity and estimates of random effects variances between effect sizes (level 1) and between samples (level 2) for multivariate meta-regression models using maximum likelihood estimation table 5: multivariate meta-regression estimates of the effects of moderators entered together in the model

- the results of moderator and mediator analyses (analyses of subsets of studies), including
  - the number of studies and total sample sizes for each moderator analysis
  - any assessment of interrelations among variables used for moderator and mediator analyses

table 3: tests of homogeneity and estimates of random effects variance between effect sizes (level 1) and between sample (level 2) for multivariate meta-regression models using maximum likelihood estimation

- any assessments of bias, including possible data censoring:

table 4: file drawer estimates for disaggregated effect size estimates grouped by informant

DISCUSSION	***For the Discussion section, please indicate whether the requested information can be found in this section of the manuscript, in a supplemental file, or whether the information is not relevant to the study. If not relevant, please provide a brief explanation.***

- Does the Discussion contain
  - a statement of major findings?
    In manuscript ■ In supplemental files ☐ Not relevant ☐

    If not relevant, please explain:

  - consideration of alternative explanations for observed results, including the impact of data censoring?
    In manuscript ■ In supplemental files ☐ Not relevant ☐

    If not relevant, please explain:

- an assessment of the generalizability of conclusions, e.g.,
  - relevant populations;
  - treatment variations;
  - dependent (outcome) variables;
  - research designs, etc.?

In manuscript ■ In supplemental files ☐ Not relevant ☐

If not relevant, please explain:

- an assessment of general limitations (including assessment of the quality of studies reviewed)?

In manuscript ■ In supplemental files ☐ Not relevant ☐

If not relevant, please explain:

	○ implications and interpretation for theory, policy, or practice? In manuscript ■ In supplemental files ☐ Not relevant ☐ If not relevant, please explain: ○ guidelines for future research In manuscript ■ In supplemental files ☐ Not relevant ☐ If not relevant, please explain: