

# Grundlagen von Systematic Reviews und Meta-Analysen

Hanna Mütze  
Doktorandenworkshop der 39. DGM-Tagung DGM  
Hannover, den 07.09.2023

# Vorstellung



Hintergrund



Methoden



Zusammenfassung



Diskussion

# Hintergrund

Methoden

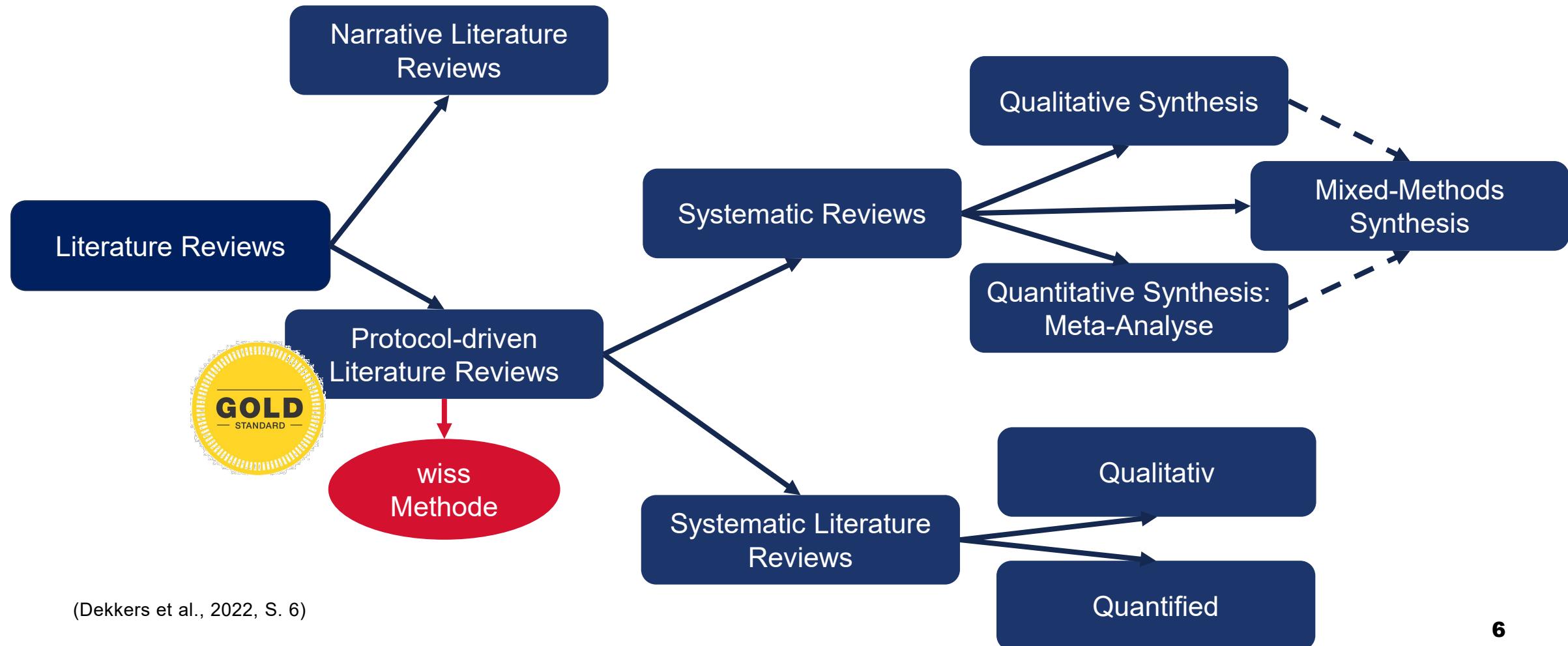
Zusammenfassung

Diskussion

# Wovon sprechen wir?

-  Kennt jemand eine Meta-Analyse oder ein Systematic Review aus dem Bereich Musikpsychologie oder Musikpädagogik?
  
- Definition: „a systematic, explicit, reproducible method for identifying, evaluating, and synthesising the existing body of completed and recorded work produced by researchers, scholars, and practitioners“ . Fink (2019, S. 6)

# Taxonomie der Literatur-Reviews



# Funktionen der verschiedenen Reviews

- Narrative reviews: detaillierte Analyse existenter Literatur, Diskussion verschiedener Argumente / Widersprüche, Literatur zugänglich machen, Forschungslücken
- Systematic literature reviews: Alle relevante Literatur berücksichtigt durch Protokollansatz bei Suche und Analyse, beinhaltet auch Sekundärliteratur & theoretische Arbeiten
- Systematic reviews: nur Primärstudien, engere Fragestellung (Effekte, Interventionen, Treatments, Policies und Praktiken), Methodologie wie bei empirischen Studien: Objektivität, Reliabilität, Validität -> Maßnahmen, um Bias gering zu halten

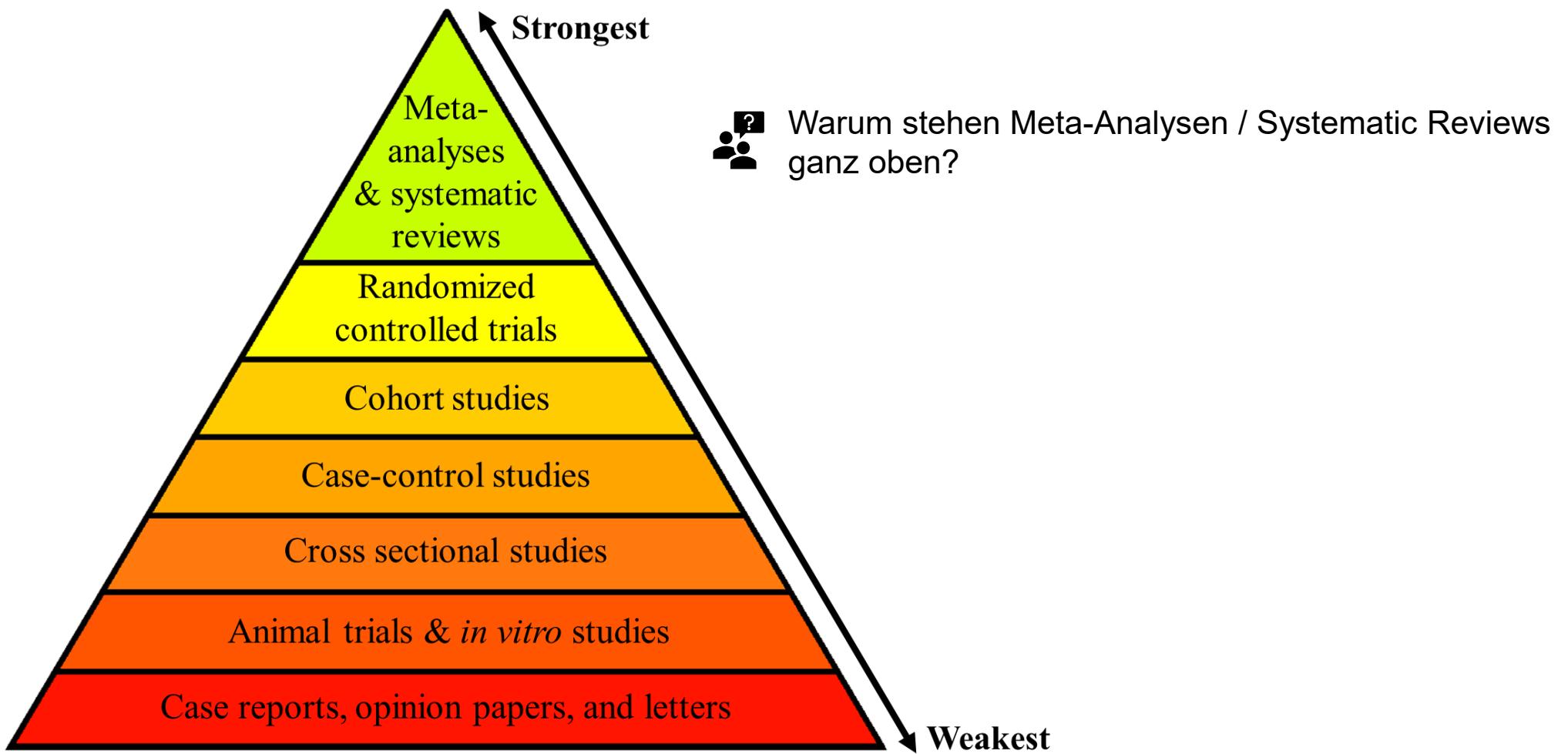
# Funktionen



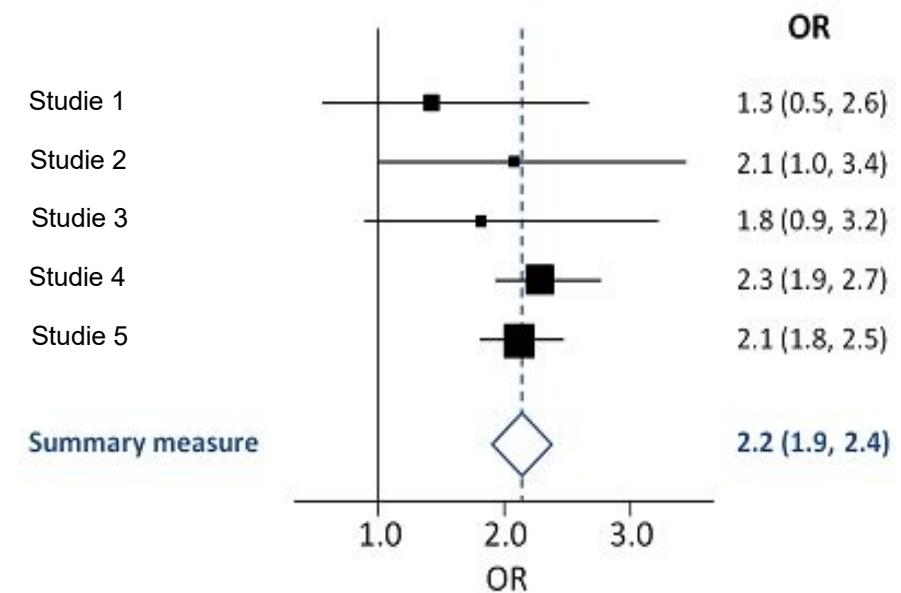
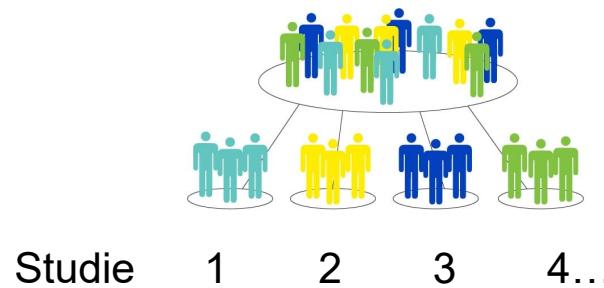
Welche Funktionen können Meta-Analysen / Systematic Reviews haben?

- Identifikation aktueller Forschung:  
Vermeidung von Dopplungen  
Erkennen von Forschungslücken & Limitationen  
Kontextualisieren der eigenen Studie
  - Theoriebildung & -evaluation:  
kritische Bewertung, Verbinden & Kontextualisieren von Einzelbefunden
  - Wissenschaftskommunikation: Handlungsbasis für Entscheidungsträger
- relevant in allen Phasen des Forschungsprozesses
- 
- Narrative
- Protocol-driven

## Hierarchy of Scientific Evidence



# Exkurs: Stichprobenfehler



- ABER: Auch Meta-Analysen ziehen quasi eine Stichprobe (nur aus Studien und nicht Probanden)!
- Methoden zur (allgemeinen) Bias-Reduktion, hier: „selection bias“

Hintergrund

Methoden

Zusammenfassung

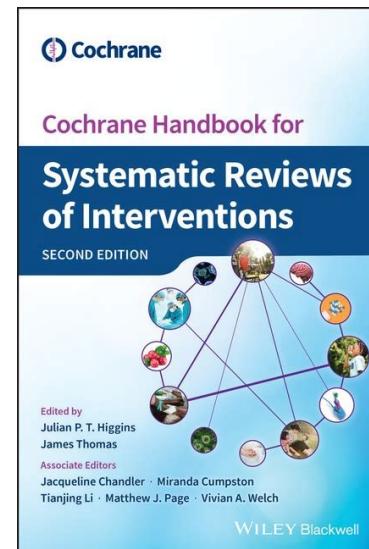
Diskussion



# Wie gehe ich vor?



PRISMA

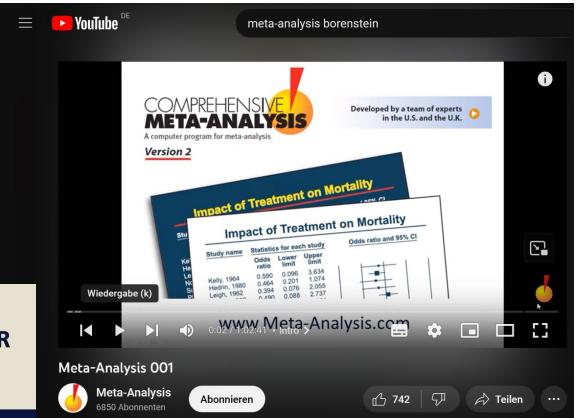
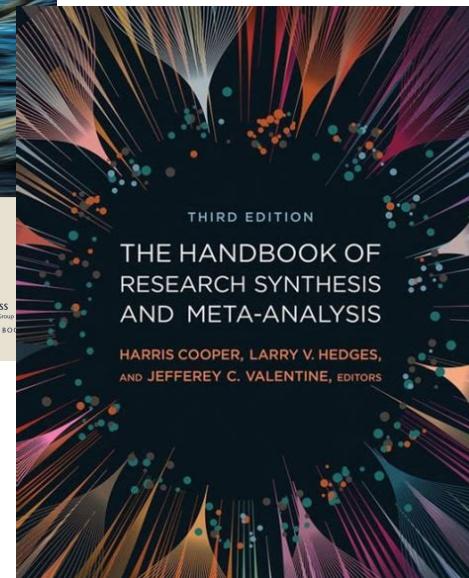


Doing Meta-Analysis with R  
A Hands-On Guide

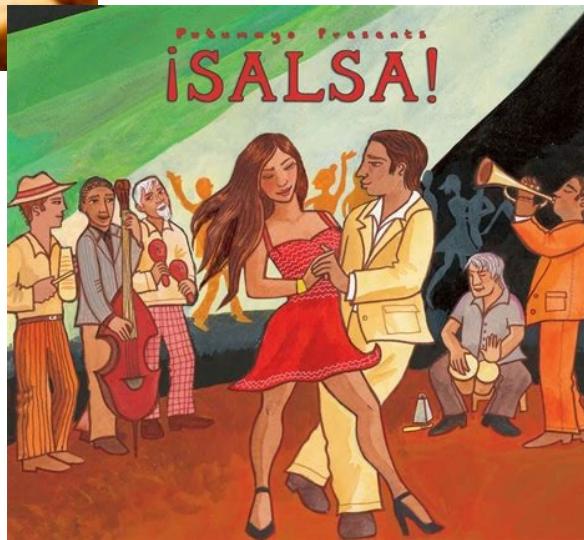


Mathias Harrer  
Pim Cuijpers  
Toshi Furukawa  
David Ebert

CRC Press  
Taylor & Francis Group  
A CHAPMAN & HALL BOOK



# A SALSA?



**Ask**  
**Search**  
**Appraisal**  
**Synthesis**  
**Analysis**



# Ask

1. Festlegen des Forschungsgegenstandes → Entwicklung von Tonalität
2. Festlegen der Funktion des Literature Reviews → Meta-Analyse
3. Festlegen der Forschungsfrage des Reviews
  - a. Frage nach Effekt
  - b. Frage nach Methode
  - c. Frage nach Konzepten

→ Entwicklung = Alter  
(→ Optional auch hierzu)  
(→ Optional auch hierzu)



Paarweise: Formuliert schriftlich eine Forschungsfrage mit o. g. Randbedingungen!

Wie stark beeinflusst das Alter die Entwicklung interner Repräsentationen tonaler Hierarchie  
(=> das Konzept tonaler Hierarchie)?



# In- and Exclusion Criteria

- "Population" - Population und Phänomen
- "Intervention" - Behandlung / Maßnahme
- "Comparison" - Alternativmaßnahme oder keine Behandlung
- "Outcome" - Abhängige Variable (Leistung, latente Variable...)
- („Context“) - Welches Land? Welche Tonalität?



Kleingruppen: Findet zu jedem PICOC-Aspekt mind. je zwei Ein- und Ausschlusskriterium!

Wie stark beeinflusst das Alter die Entwicklung interner Repräsentationen tonaler Hierarchie  
(=> das Konzept tonaler Hierarchie)?



# In- and Exclusion Criteria

Topic	In-/Exclusion Criteria
Methodology	Primary data, hypothesis-testing design
Publication Date	1982 – April 2022
Participants	Healthy, normal developed pupils in elementary and secondary school
Outcome	Measurement of IRT; exclusion of psychoacoustics, neuroimaging and musical long-term memory
Design	Cross-sectional (group comparisons) or longitudinal

# In- and Exclusion Criteria

Zwei Meta-Analysen zum Einfluss von Deliberate Practice auf die Leistung:

- Mancnamara, B. N., Hambrick, D. Z., & Oswald, F. L. (2014). Deliberate practice and performance in music, games, sports, education, and professions: A meta-analysis. *Psychological Science*, 25(8), 1608–1618.
- Platz, F., Kopiez, R., Lehmann, A. C., & Wolf, A. (2014). The influence of deliberate practice on musical achievement: A meta-analysis. *Frontiers in Psychology*, 5(646).

Ergebnisse:

- Mancnamara et al. (2014):  $r_c = 0.35, 95\% CI [0.30, 0.39]$
- Platz et al. (2014):  $r_c = 0.61, 95\% CI [0.54, 0.67]$

- Statistische Methoden ersetzen keine inhaltliche Validität!
- Die Möglichkeiten / Grenzen der inhaltlichen Interpretation werden bereits hier abgesteckt!



# Search: Was, wo und wie?

## Was?

1. Keywords aus PICOC ableiten:

- "Population" → *Children*
- "Intervention" → *Instrumental lessons*
- "Comparison" → *Development*
- "Outcome" → *feeling for tonality*
- („Context“) → *Western classical music*

2. Thesaurus → Synonyme zu Keywords finden



4 Gruppen: Findet je mindestens zwei Synonyme für die bereits genannten Keywords.



# Search: Was, wo und wie?

## Wo?

### 1. Datenbanken



Welche Datenbanken kennt ihr, die für unser Fachgebiet relevant sind?

PubMed, PsycINFO, Psyndex, ProQuest Dissertations Abstracts, Web of Science, ERIC, pedocs, RILM Abstracts, ...

### 2. Handsuche z. B. in Elektronische Fachzeitschriften (high impact journals)

### 3. Grey Literature: Bibliothekskataloge, Grey-Net International, Konferenz-Proceedings, Experten kontaktieren



# Search: Was, wo und wie?

## Wie?

1. Datenbanken: keywords & in-/ exclusion criteria mit Filtern
2. Boolesche Operatoren (Volltext / Abstract / Titel):

Operator	Bedeutung	Wirkung auf Trefferzahl
*	Platzhalter Wortmitte oder –ende	>
AND (+)	Beide Ausdrücke müssen zutreffen	<
OR	Mindestens ein Ausdruck trifft zu	>
XOR	Es darf nur einer der Ausdrücke zutreffen	<
NOT (-)	Der erste Ausdruck muss, der zweite darf nicht zutreffen	<

3. Grey Literature: Snowball Sampling
4. Dokumentation mit Literaturverwaltungsprogramm

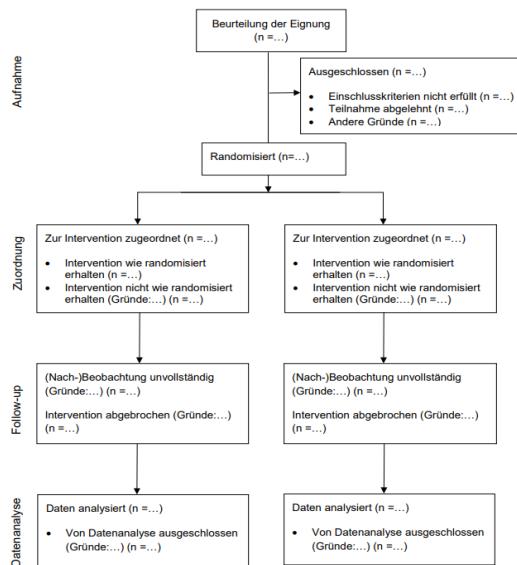




# Appraisal



- Vielzahl von Checklisten für unterschiedliche Studiendesigns
- Übersicht: <https://guides.lib.unc.edu/systematic-reviews/assess-quality>
- Bsp.: CONSORT Statement for Randomized Controlled Studies (Butcher et al., 2022)



## Ziele:

- Reduktion
- keine Exklusion relevanter Forschung
- qualitativ hochwertige Primärstudien
- vergleichbare Fragestellungen / Methoden

# Synthesize

Vorgehensweise: Mindestens zwei Kodierer

1. Volltext für alle Studien?
2. Welche Informationen sind wichtig? (PICO, Moderatorvariablen / Hintergrundinformationen,...)
3. Methode zur Datensammlung? (Software: Covidence, Revman; Manuell)
4. Datentablle erstellen mit Kodierhandbuch
5. Sammeln & extrahieren der Daten
6. Interrater-Reliabilität berechnen & Fehlerkontrolle

# Beispiele und Übung

## Items für die Datenextraktion (Bsp.)

- 1) Artikelinformationen: Autor, Erscheinungsjahr, Titel, DOI
- 2) Studieninformationen: Studientyp, Probandenrekrutierung, Studienqualität
- 3) Teilnehmerinformationen: Alter, Geschlecht, musikalische Expertise
- 4) Intervention: Qualität & Quantität, Dauer, Setting
- 5) Ergebnis: Qualität & Quantität: Maß und Größe des Effekts



Stellt relevante Items zur Datenextraktion für die Meta-Analyse zur Tonalitätsentwicklung tabellarisch zusammen!



# Hintergrundvariablen für Meta-Analyse: Paper

Kodieranweisung: Was heißt: „formal musical training“?



Study ID	Author (Year)	DV	n		Age range		Relative frequency of children with formal musical training	
			treatment	control	treatment	control	treatment	control
1	Krumhansl & Keil (1982)	probe-tone	14	14	grades 3-4	grades 1-2	.71	.43



Kodieranweisung: Alter oder Klassenstufe?



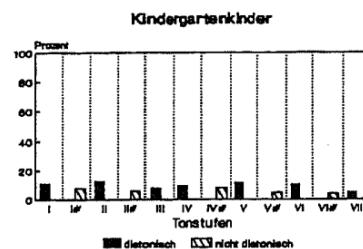
# Collecting Effect sizes

Ziel (Voraussetzung für statistische Aggregation):

Author (Year)	ID	N	Fisher'z-transformed $r$	Fisher'z-transformed $v$
Krumhansl & Keil (1982)	1	28	1.020	0.060
Cuddy & Badertscher (1987)	2	110	0.701	0.010
Schwarzer et al. (1997)	3	54	0.000	0.030



Realität:



$$\beta \quad \eta_p^2 \quad d$$

$$r \quad OR$$

“[...] the profiles reflected a similar trend at each age level...” Cuddy & Badertscher (1987)





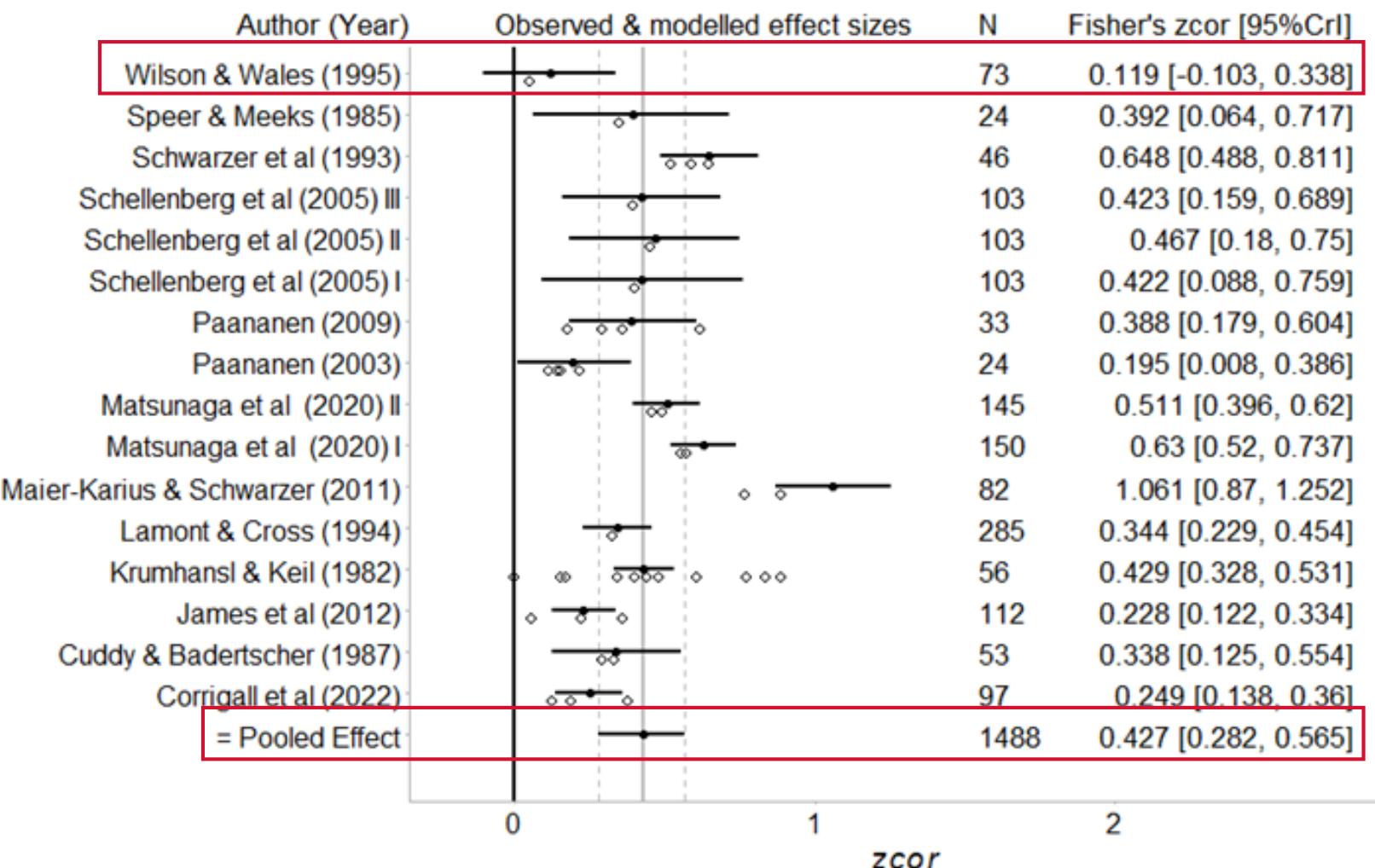
# Analyze

- 1) Inhaltliche Analyse: Gruppierung, Identifikation von Themen und Forschungslücken
- 2) Statistische Analyse
  - Berechnung eines aggregierten, gewichteten mittleren Effekts  
→ genauere Aussage als Primärstudien
  - Aufklärung von Heterogenität





# Analyze: gewichteter mittlerer Effekt





# Analyse: Heterogenität

Sind die Studien überhaupt (inhaltlich) homogen genug für eine Zusammenfassung?

- keine Pauschalantwort, abhängig von vielen Faktoren, z. B.: Fragestellung

Subgruppenanalyse:

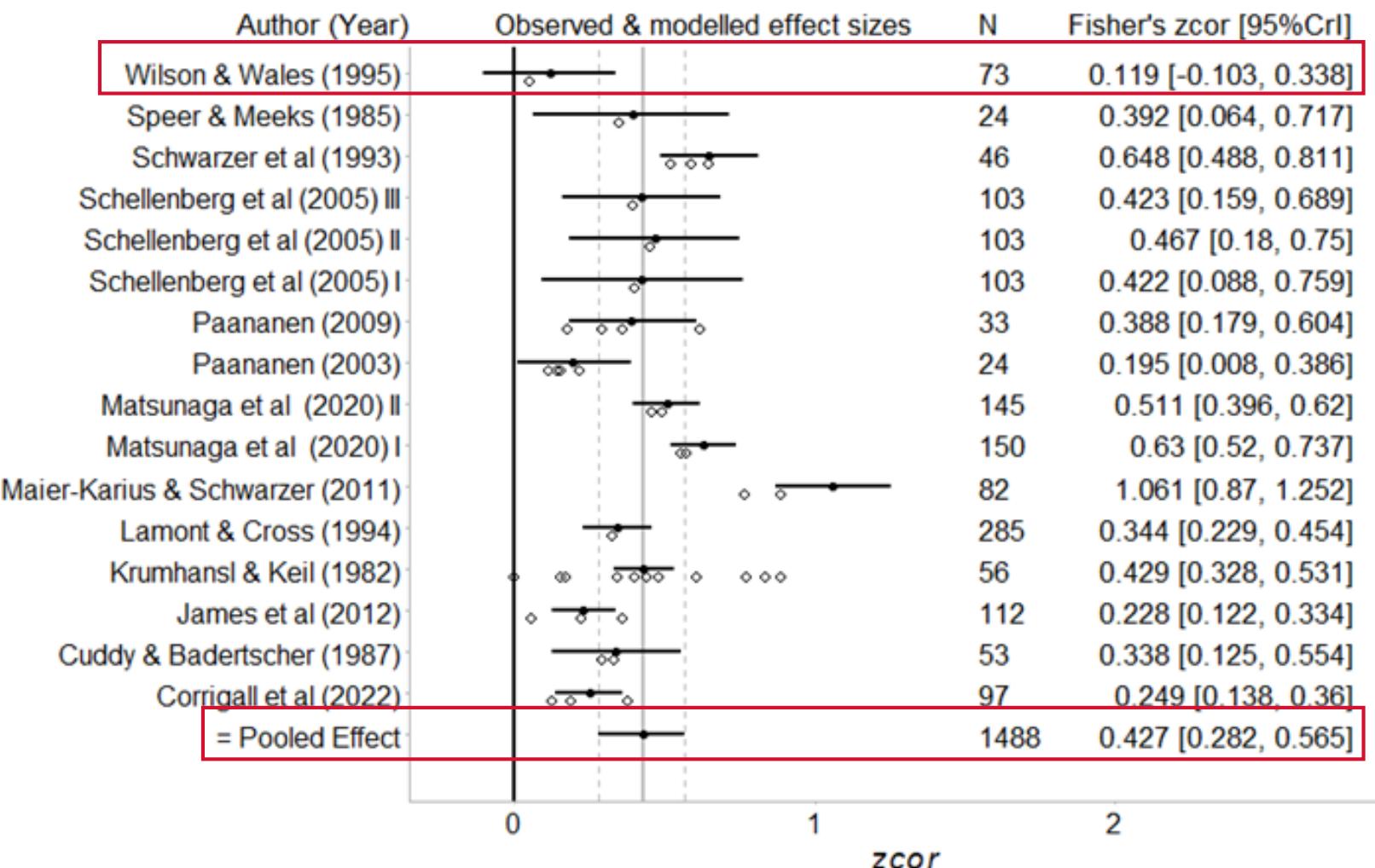
- z. B.: Altersgruppen, Operationalisierung der abhängigen Variable (bei mir kein Effekt)

Mehrebenenmodell:

- auf welcher Ebene entsteht Heterogenität (innerhalb einzelner Studien oder zwischen Studien?)

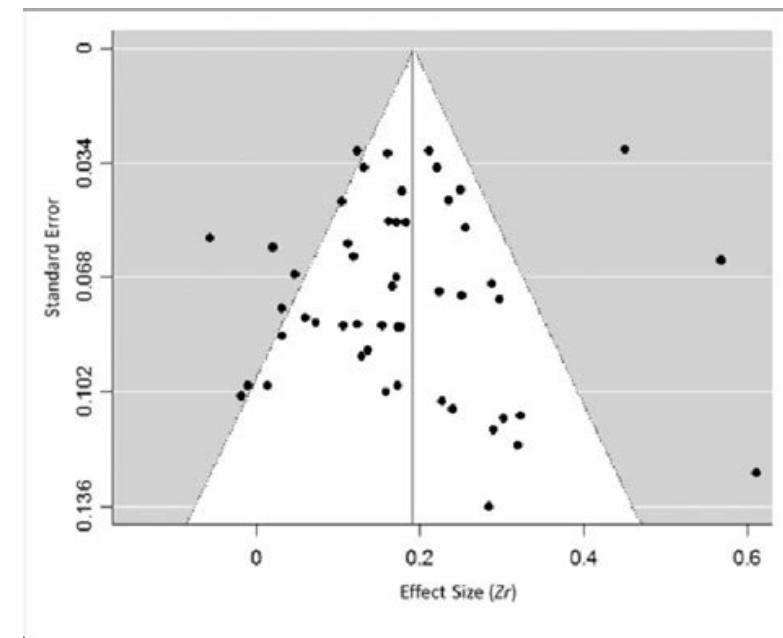


# Analyze: gewichteter mittlerer Effekt



# Diskussion / Reduktion von möglichen Biasarten

- Conflict of interests
- Publicationbias
- Languagebias
- Outlier detection and influential case diagnostics
- ...
- Verschiedene Korrekturverfahren, z. B. Reliabilität des Messinstruments (Hunter & Schmidt, 2014)
- Schwierig: z. B. bei aggregierten Daten wie Altersgruppen



Hunter, J. E., & Schmidt, F. L. (Hrsg.). (2014). *Methods for Meta-Analysis: Correcting error and bias in research findings* (3. Aufl.). Sage Publications.

# Interpretation

- Mögliche Erklärungen der Ergebnisse diskutieren
- Praktische Relevanz
- Einbettung in andere Forschung, Theorien, ...
- Limitationen
- Forschungslücken
- ...



"We believe numbers don't lie, especially when interpreted advantageously."

Hintergrund

Methoden

Zusammenfassung

Diskussion

# Zusammenfassung

- 1) Methoden generell ähnlich zu frequentistischen / Bayes-Statistik
- 2) Standards: What Works Clearinghouse / Cochrane Collaboration / PRISMA
- 3) Für's Lesen / Bewerten einer Meta-Analyse: kürzere Checklisten (z. B. Hancock, Stapleton & Mueller, 2019)
- 4) Planen → Protokoll-Registrierung
- 5) ALLES dokumentieren
- 6) Durchführung: ‚A SALSA‘
- 7) ‚PICOC‘

## TAKE-HOME:

- Meta-Analysen / Systematic Reviews stehen an der Spitze der wissenschaftlichen Evidenz
- relevant an den meisten Stellen im Forschungsprozess



## ABER:

- importieren Qualität der Primärstudien (z. B.: WEIRD-participants)
  - Qualitätsmanagement: Reduktion von Heterogenität und Bias
  - Viele inhaltliche und methodische Freiheitsgrade
- Meta-Analysen und Systematic-Reviews decken keine ‚Wahrheit‘ auf

Hintergrund

Methoden

Zusammenfassung

Diskussion

# Diskussion

- Welche Aspekte nützen euch in eurer Abschlussarbeit?
- Welche Möglichkeiten / Grenzen seht ihr in der Stellung von Meta-Analysen / Systematic Reviews in der Hierarchie der Evidenz?
- Kann das nicht alles Chat-GPT viel besser?

# Typische Fallen

1. does not clearly relate the findings of the literature review to the researcher's own study;
2. does not take sufficient time to define the best descriptors and identify the best sources to use in review literature related to one's topic;
3. relies on secondary sources rather than on primary sources in reviewing the literature;
4. uncritically accepts another researcher's findings and interpretations as valid, rather than examining critically all aspects of the research design and analysis;
5. does not report the search procedures that were used in the literature review;
6. reports isolated statistical results rather than synthesizing them by chi-square or meta-analytic methods; and
7. does not consider contrary findings and alternative interpretations in synthesizing quantitative literature.

(pp. 161-162)

# Literatur

- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2021). *Introduction to Meta-Analysis*. John Wiley & Sons.
- Cooper, H., Hedges, L. V., & Valentine, J. C. (2009). *The Handbook of Research Synthesis and Meta-Analysis: Bd. 2nd ed.* Russell Sage Foundation; eBook Collection (EBSCOhost).
- Dekkers, R., Carey, L., & Langhorne, P. (2022). *Making Literature Reviews work: A multidisciplinary guide to systematic approaches*. Springer International Publishing.
- Fink, A. (2019). *Conducting research literature reviews*. UCLA.
- Hancock, G. R., Stapleton, L. M., & Mueller, R. O. (2019). *The reviewer's guide to quantitative methods in the social sciences* (2. Aufl.). Routledge/Taylor & Francis Group.
- Harrer, M., Cuijpers, P., Furukawa, T. A., & Ebert, D. D. (2021). *Doing meta-analysis in R: A hands-on guide*. Chapman and Hall/CRC.
- PRISMA-P Group, Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., & Stewart, L. A. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Reviews*, 4(1), 1. What Works Clearinghouse. (2020). *What Works Clearinghouse™ procedures handbook, Version 4.1*. U.S. Department of Education, Institute of Education Science, National Center for Education Evaluation and Regional Assistance.