

Bernd Weiß und Michael Wagner

### 54.1 Meta-Analysen und systematische Reviews<sup>1</sup>

Der Begriff „Meta-Analyse“ wurde im Jahr 1976 von Gene V. Glass eingeführt. Er verstand darunter die „Analyse von Analysen“ („analysis of analyses“) (Glass 1976), die neben Primär- und Sekundäranalysen einen dritten Forschungstyp darstellt. Im Gegensatz zu Primär- und Sekundäranalysen werden hier nicht Originaldaten erstmalig oder wiederholt ausgewertet. Vielmehr richtet sich die Meta-Analyse in der von Glass vorgeschlagenen Form auf die *statistische Zusammenfassung publizierter empirischer („aggregierter“) Befunde aus möglichst allen für die jeweilige Fragestellung relevanten Studien*. Dabei geht es nicht nur darum, empirische Befunde wie etwa Mittelwertdifferenzen, Korrelations-/Assoziationsmaße oder Regressionskoeffizienten zusammenzufassen, sondern auch darum, die Verteilung der Forschungsbefunde zu beschreiben und zu erklären. Diese Forschungsbefunde entstammen vor allem Fachartikeln. Man sollte jedoch auch „graue Literatur“ (z. B. Konferenzbeiträge, Arbeitspapiere oder Examensarbeiten) in eine Meta-Analyse einbeziehen. Mit einer Meta-Analyse lässt sich dann untersuchen, ob sich Befunde aus verschiedenen Quellen mit unterschiedlichen Qualitätsstandards systematisch voneinander unterscheiden.

Während die von Glass propagierte „klassische“ Meta-Analyse auch Aggregierte-Personen-Daten Meta-Analyse (engl. *aggregate person data (APD) meta-analysis*) genannt wird, wird ein zweiter Typ von Meta-Analysen, die Individual-Personen-Daten Meta-Analyse (engl. *individual person data (IPD) meta-analysis*) (Riley et al. 2021) immer populärer. IPD Meta-Analysen zielen nicht auf eine Synthese publizierter empirischer Befunde, sondern auf die *Reanalyse der Originaldaten*, die den jeweiligen Studien zugrunde liegen. Diese Datensätze können je einzeln in einheitlicher Weise

1 Wir danken Jessica Daikeler, Anna-Carolina Haensch, Dastan Jasim, Sonja Schulz, und Sebastian Sterl für wertvolle Hinweise und Kürzungsvorschläge.

oder aber in zusammengefasster („pooled“) Form analysiert werden. IPD Meta-Analysen sind aufwendig, haben aber mehrere Vorteile: die Vergleichbarkeit der Variablen kann durch Ex-post-Harmonisierung erhöht werden, das Spektrum möglicher Analyseverfahren ist größer und es lassen sich Hypothesen untersuchen, die bislang nicht untersucht wurden. Nachfolgend befassen wir uns hauptsächlich mit APD Meta-Analysen und verweisen mit Blick auf das Thema IPD Meta-Analyse auf Riley et al. (2021) für eine aktuelle und gute Übersicht.

Im Zusammenhang mit Meta-Analysen ist häufig von „*systematischen Reviews*“ die Rede (Higgins/Green 2008, Petticrew/Roberts 2006). Ein systematisches Review beabsichtigt *alle* empirischen Befunde zu vorab definierten Suchkriterien zusammenzustellen, um auf Grundlage dieser empirischen Evidenz eine bestimmte Fragestellung zu beantworten. Solche Reviews zeichnet aus, dass sie systematische und transparente Verfahren verwenden, um Verzerrungen (etwa den „Publication Bias“, vgl. Abschnitt 54.5.4) zu vermeiden und damit zu zuverlässigen Schlussfolgerungen kommen. Je nach disziplinärer Zugehörigkeit werden die Begriffe „Meta-Analyse“ und „systematisches Review“ entweder synonym verwendet oder eine Meta-Analyse ist Teil eines quantitativen systematischen Reviews (Higgins/Green 2008: 6).

Meta-Analysen sind in vielen wissenschaftlichen Disziplinen fest etabliert. Dies zeigen nicht zuletzt die zahlreichen Lehr- und Handbücher zur Theorie und ihrer praktischen Durchführung (etwa Borenstein et al. 2021, Cooper et al. 2019). In der Medizin und Epidemiologie hat sich die Meta-Analyse bereits in Form der *Evidenzbasierten Medizin* institutionalisiert, am bekanntesten ist die *Cochrane Collaboration*. Die *Cochrane Collaboration* ist eine internationale Non-Profit-Organisation, die im Feld der medizinischen Versorgung systematische Reviews motiviert, begutachtet und veröffentlicht. Für die Sozialwissenschaften ist die *Campbell Collaboration* zu nennen – ein Forschungsnetzwerk, das systematische Reviews und Meta-Analysen über die Effekte sozialer Interventionen erstellt.

---

## 54.2 Ziele

APD sowie IPD Meta-Analysen haben zum Ziel, den empirischen Forschungsstand zu einer Fragestellung zu bestimmen und zu genaueren und sichereren Schlussfolgerungen zu gelangen, als es mit einer oder nur wenigen Studien oder mit einem einzelnen Datensatz möglich wäre. Solche Forschungssynthesen werden immer dringlicher, weil die Menge der Forschungsbefunde in den letzten Jahrzehnten in allen wissenschaftlichen Feldern explosionsartig gestiegen ist (Weiß/Wagner 2008).

Meta-Analysen sind auch deshalb geboten, weil empirische Ergebnisse zu demselben Variablenzusammenhang, die aber aus verschiedenen Studien stammen, sehr heterogen und inkonsistent sein können. Das mag an differierenden Operationalisierungen (Stein, Kapitel 8 in diesem Band) oder dem Stichprobendesign (Häder/Häder, Kapitel 27 in diesem Band) liegen. Mit Hilfe von Meta-Analysen lässt sich ermitteln,

wie groß die Heterogenität der Befunde ist. Ferner besteht die Möglichkeit, diejenigen Faktoren zu identifizieren, die möglicherweise für diese Befundheterogenität verantwortlich sind. Meta-Analysen tragen dazu bei, drei bedeutende wissenschaftliche Ziele besser zu erreichen: (1) *Empirische Forschung soll sich kumulativ entwickeln*. Die Erklärungskraft von Theorien und ihr praktischer Nutzen nehmen zu, wenn bekannt ist, welche Hypothesen widerlegt wurden und welche nicht. Meta-Analysen können überdies Forschungslücken identifizieren und diese gezielt zum Gegenstand künftiger Forschung machen. (2) Die Anwendung von Maßnahmen und Interventionen kann nur dann optimiert werden, wenn zuverlässiges Wissen um ihre Effizienz vorhanden ist. Die Aufgabe von Meta-Analysen ist es, *Forschungsergebnisse zu synthetisieren*, um die besten Methoden herauszufinden, mit denen bestimmte wissenschaftliche oder praktische Probleme gelöst werden können. (3) Häufig besteht ein Forschungsziel darin, den *Informationsgehalt von Theorien durch eine Suche nach universalen Gesetzen zu verbessern*. Lassen sich Erkenntnisse zu empirischen Fragestellungen mit mehreren unabhängigen Stichproben replizieren, dann erhöht sich der Grad an Generalisierbarkeit.

---

### 54.3 Typische Forschungsprobleme in den Sozialwissenschaften

Drei prototypische sozialwissenschaftliche Forschungsprobleme lassen sich mit Meta-Analysen untersuchen: (1) *Deskription und Exploration sozialer Sachverhalte*: In den Sozialwissenschaften kommt es häufig vor, dass man die Verbreitung eines sozialen Phänomens in der Gesellschaft nicht genau kennt, weil die amtliche Statistik (Hartmann/Lengerer, Kapitel 106 in diesem Band) darüber keine Auskunft gibt oder keine entsprechend große Primäruntersuchung vorliegt. So ist unbekannt, wie viele Jugendliche in Deutschland die Schule schwänzen. Allerdings gibt es in Deutschland zahlreiche und teilweise lokal begrenzte Schüler- und Jugendstudien, in denen das Schulschwänzen (bzw. Schulabsentismus) erhoben wird. Weiß (2008) hat eine IPD Meta-Analyse von 15 Studien vorgelegt und eine Hochrechnung zur prozentualen Verbreitung des Schulabsentismus vorgenommen. Ein weiteres, umfragemethodisch motiviertes Beispiel ist eine Meta-Analyse von Daikler et al. (2020) in der sie der Frage nachgehen, ob web-basierte Umfragen niedrigere Ausschöpfungsquoten erzielen als andere Umfragemodi. (2) *Testen von Hypothesen*: Um herauszufinden, ob und warum Scheidungsrisiken in Europa variieren, führten Wagner und Weiß (2006) eine Meta-Analyse auf der Basis von 120 empirischen Studien aus 19 europäischen Ländern durch. Getestet wurden Hypothesen zum Zusammenhang von Makrovariablen und individuellen Scheidungsrisiken. Es zeigt sich also, dass Meta-Analysen auch in der international vergleichenden Forschung (Braun, Kapitel 79 in diesem Band) gewinnbringend eingesetzt werden können. (3) *Evaluation von Maßnahmen*: Meta-Analysen werden auch oft zur Evaluation von Maßnahmen eingesetzt, sei es im therapeutisch-medizinischen Bereich oder in der Sozialpolitik, die vor allem mit Hilfe von

Experimentalstudien untersucht werden (Eifler/Leitgöb, Kapitel 13 in diesem Band). In der Präventionsforschung zum abweichenden Verhalten Jugendlicher publizierten Wilson et al. (2001) eine Meta-Analyse zur Frage, welche Interventionen von Schulen das abweichende Verhalten von Jugendlichen reduzieren können.

---

## 54.4 Durchführung von Meta-Analysen

Eine Meta-Analyse umfasst alle Elemente eines sozialwissenschaftlichen Forschungsprozesses (Stein, Kapitel 8 in diesem Band): Problemspezifikation und Hypothesenbildung, Datenerhebung, Datenaufbereitung, Datenanalyse und die Interpretation der Befunde. Der erste Schritt, die *Problemspezifikation* und gegebenenfalls die *Hypothesenbildung*, verläuft analog zu anderen empirischen Studien. Häufig setzt eine APD Meta-Analyse voraus, dass zu einer Forschungsfrage schon publizierte Befunde vorliegen. In einem zweiten Schritt sollte bei einer publikationsbasierten APD Meta-Analyse eine möglichst vollständige und systematische *Literaturrecherche* durchgeführt werden (White 2019). Es lassen sich drei Quellen ausmachen, um an relevante Literatur zu gelangen: (a) Befragungen von Expertinnen/Experten (per E-Mail, auf Konferenzen etc.) (Helfferich, Kapitel 55 in diesem Band), (b) (Fach)Datenbanken (Ohly/Stahn/Weber, Kapitel 3 in diesem Band) sowie (c) das Durchsuchen der durch (a) und (b) gefundenen Literatur nach weiteren Studien („Schneeballsystem“). Ähnlich einer Inhaltsanalyse (Mayring/Fenzl, Kapitel 43 in diesem Band) wird nach der Literaturrecherche das *Material vercodet und elektronisch gespeichert; in einer Primärstudie entspricht dieser Schritt der Dateneingabe* (Lück/Landrock, Kapitel 33 in diesem Band). Relevant sind die eigentlichen Befundstatistiken. In vielen Fällen müssen die benötigten Befundstatistiken erst selbsttätig ermittelt werden, etwa wenn Studienbefunde nur in tabellarischer Form präsentiert werden und für die Meta-Analyse Assoziationsmaße wie beispielsweise Odds Ratios oder relative Risiken benötigt werden. Um die Heterogenität der Befunde aufzuklären, sollten aber auch Publikationsmerkmale (Publikationstyp, -jahr etc.), Studien- bzw. Stichprobenmerkmale (Erhebungsjahr, Stichprobendesign etc.) oder auch Qualitätsmerkmale berücksichtigt werden.

Im vierten Schritt findet die statistische Analyse statt. Die *Datenanalyse* einer APD Meta-Analyse besteht mindestens aus drei Teilen (Weiß/Wagner 2008): (1) *Befundintegration*: Hier geht es darum, die einzelnen Befundstatistiken zu einer Gesamtstatistik zusammenzufassen (Abschnitt 54.5.1). (2) *Heterogenitätsanalyse*: Bei der Heterogenitätsanalyse wird die Variation der Befunde analysiert und gegebenenfalls erklärt (Abschnitt 54.5.3). (3) *Fehlende Werte*: Bei der Analyse fehlender Werte (Engel/Schmidt, Kapitel 29 in diesem Band) wird unter anderem überprüft, ob ein Publikationsbias vorliegt. Dieser liegt dann vor, wenn die Vorbereitung, Einreichung oder Publikation von Forschungsbefunden von bestimmten Merkmalen dieser Befunde – wie etwa der statistischen Signifikanz oder der Richtung der Effekte – abhängt (Abschnitt 54.5.4).

Bei der Publikation einer Meta-Analyse wird empfohlen, sich an den zahlreichen, disziplinspezifischen Berichtsstandards zu orientieren, die sich leicht mit Hilfe einschlägiger Suchmaschinen und der Suchbegriffe „reporting standards“ und „meta-analysis“ finden lassen.

---

## 54.5 Statistische Verfahren

### 54.5.1 Verfahren und Modelle der Befundintegration

Es existieren zahlreiche statistische Verfahren zur Befundintegration im Rahmen von publikationsbasierten APD Meta-Analysen, die sich vor allem nach Art und Informationsgehalt der zugrundeliegenden Befundstatistiken unterscheiden (Borenstein et al. 2021). Typische Analyseeinheiten einer Meta-Analyse sind etwa Korrelationskoeffizienten, Mittelwertdifferenzen, Odds Ratios oder Regressionskoeffizienten (Blasius/Baur, Kapitel 45 in diesem Band). Univariate Anteils- oder Mittelwerte, in einigen Fällen auch Teststatistiken (etwa  $t$ -,  $F$ -,  $\chi^2$ -Werte) oder  $p$ -Werte, können ebenfalls genutzt werden. Im Folgenden wird angenommen, dass die Befundstatistiken unabhängig voneinander sind, d. h. pro Publikation wird nur eine einzige Statistik genutzt. Für fortgeschrittene Verfahren der Befundsynthese, die auch mit abhängigen Statistiken umgehen können, wird auf Abschnitt 54.5.3 verwiesen.

In der meta-analytischen Literatur findet sich, wenn von Befundstatistiken die Rede ist, überwiegend der Begriff Effektstärke („effect size“), worunter ein „scale-free index of effect magnitude“ (Hedges/Olkin 1985: 7) zu verstehen ist. Da in diesem Beitrag nicht nur Effektstärken thematisiert werden, verwenden wir die allgemeineren Begriffe Forschungs- oder Studienbefunde.

Wenngleich es naheliegend erscheint, für die Befundsynthese ein einfaches arithmetisches Mittel zu verwenden, verbietet sich dieser Ansatz bei einer Meta-Analyse aus zwei Gründen: (1) Die Forschungsbefunde aus verschiedenen Studien basieren für gewöhnlich auf unterschiedlichen Stichproben mit unterschiedlichen Stichprobengrößen. Da der Standardfehler bzw. die Varianz von der Stichprobengröße abhängt, verletzen meta-analytische Daten die Annahme der Homoskedastizität. Die Kleinst-Quadrat-Schätzfunktion (engl. *ordinary least squares*, OLS) liefert in diesem Fall verzerrte Schätzer; im Fall ohne Prädiktoren ist dies das einfache arithmetische Mittel. Ein angemessenes Schätzverfahren ist dagegen die Methode der gewichteten kleinsten Quadrate (engl. *weighted least squares*, WLS). Im univariaten Fall (ohne Prädiktoren) erfolgt die meta-analytische Befundsynthese damit durch die Verwendung eines gewichteten arithmetischen Mittels. Die Gewichte berechnen sich aus den inversen Varianzen der Forschungsbefunde. (2) Sofern die Heterogenität der Befundstatistiken relativ gering ist, erhalten Befundstatistiken, die auf großen Fallzahlen basieren – und damit eine höhere Reliabilität besitzen –, einen entsprechend großen Einfluss bei der Parameterschätzung. Die Anwendung dieser statistischen Verfahren setzt also voraus,

dass sowohl ausreichend Informationen über die Befundstatistiken als auch über deren Standardfehler bzw. Varianz vorliegen.

Nachdem nun die Wahl des Schätzverfahrens in Form des gewichteten arithmetischen Mittels (bzw. WLS-Schätzers) begründet wurde, geht es um die Berechnung der Gewichte. Es wurde oben bereits erklärt, dass sich die Gewichte aus der inversen Varianz der Forschungsbefunde ergeben. Die Gewichtung kann sich an zwei Modellen orientieren: Die Befundstatistiken werden bereits aufgrund des Stichprobenfehlers eine gewisse Variation aufweisen. Eine zentrale Frage bei einer Meta-Analyse ist, ob diese Befundstreuung nur auf den Stichprobenfehler zurückzuführen ist (Modell 1) oder ob die Streuung der Effektstärken das allein durch den Stichprobenfehler erwartete Niveau übersteigt und sich diese zusätzliche Heterogenität auf weitere (möglicherweise unbekannte) Einflussgrößen zurückführen lässt (Modell 2). Die Beantwortung dieser Frage hat sowohl Konsequenzen für die Parameterschätzung als auch für die statistische Inferenz. Modell 2 ist als Modell mit zufälligen Effekten (*random-effects model*, REM) bekannt. Modell 1 ist ein Spezialfall des REM und wird als Modell mit festem Effekt (*fixed-effect model*, FEM) bezeichnet. Konkret wird beim REM ein zusätzlicher Fehlerterm berücksichtigt, so dass sich die Varianz der Forschungsbefunde erhöht (vgl. ausführlich Borenstein et al. 2021).

Im Hinblick auf die statistische Inferenz sprechen Hedges und Vevea (1998) beim FEM auch vom Modell der eingeschränkten Inferenz (engl. *conditional inference*), während das REM eine unbeschränkte Inferenz (engl. *unconditional inference*) ermöglicht. Das FEM ist in seiner Generalisierbarkeit auf die vorliegende Verteilung der Befundstatistiken beschränkt, während im Fall des REM eine solche Beschränkung nicht angenommen wird. Besteht das Ziel der Meta-Analyse darin, über die vorliegende Stichprobe von Befundstatistiken hinaus zu verallgemeinern oder sind die zugrundeliegenden Studien nicht unmittelbar vergleichbar, so ist das REM die angemessene Wahl.

Vor allem bei der Befundintegration und Bewertung von Experimentalbefunden mit mehr als einer Interventionsmaßnahme lassen sich sogenannte Netzwerk Meta-Analysen (*network meta-analysis*, *multiple-treatment meta-analysis*) durchführen. Über gemeinsame Vergleichsgruppen sind auch indirekte Vergleiche zwischen unterschiedlichen Interventionen möglich, die bislang noch nicht direkt miteinander verglichen wurden (Kiefer et al. 2015).

### 54.5.2 Homogenitätstests

Die Entscheidung, inwieweit eine zusätzliche Varianzkomponente bei der Berechnung der zusammengefassten Befundstatistik berücksichtigt werden sollte (also REM vs. FEM), kann einerseits von statistischen Homogenitätstests abhängen, andererseits aber auch methodische oder inhaltliche Gründe haben (Borenstein et al. 2021: 99–117). Solche Gründe liegen etwa dann vor, wenn die Literaturrecherche nicht

vollständig war oder die Befunde aus Studien mit einem unterschiedlichen Stichprobendesign stammen. Es ist dann nicht angebracht, von einem gemeinsamen Effektstärkeparameter (FEM) auszugehen. Wenn mindestens eine der beiden genannten Bedingungen zutrifft, sollte mit Hilfe des REM diese zusätzliche Befundvariation sowohl im Rahmen der statistischen Inferenz als auch bei der Schätzung einer mittleren Befundstatistik berücksichtigt werden. Es gibt verschiedene graphische und statistische Verfahren, mit denen sich die Homogenität der Verteilung der Befundstatistiken evaluieren lässt. In den Sozialwissenschaften ist in den allermeisten Fällen aus methodischen oder inhaltlichen Gründen das REM angemessen.

### 54.5.3 Verfahren der Heterogenitätsaufklärung und Umgang mit abhängig Befundstatistiken

Neben der univariaten Befundsynthese ist bei einer Meta-Analyse die Heterogenitätsaufklärung von zentraler Bedeutung. Werden kategoriale Merkmale zur Aufklärung der Heterogenität herangezogen, kommen *varianzanalytische Modelle* zum Einsatz. Im Fall metrischer Merkmale können *Meta-Regressionen* geschätzt werden (Borenstein et al. 2021). So können sich Effektstärken etwa danach unterscheiden, ob sie bivariate Zusammenhänge widerspiegeln oder ob sie mit multiplen Regressionsmodellen geschätzt wurden, in denen für mögliche Konfounder kontrolliert wurde. Die Effektstärken lassen sich also in zwei Gruppen unterteilen und es lässt sich statistisch testen, ob die mittleren Effektstärken zwischen den beiden Gruppen signifikant differieren.

Die bisherigen Ausführungen haben das Vorgehen für statistisch unabhängige Befundstatistiken beschrieben. Sofern diese voneinander abhängig sind – etwa, wenn sich mehrere Effektstärken auf denselben Datensatz beziehen – gibt es unterschiedlich anspruchsvolle Verfahren. Dazu gehören die einfache Befundsynthese innerhalb identischer Datensätze, die Schätzung robuster Standardfehler (Hedges et al. 2010) sowie Mehrebenen-Meta-Analysen (Van den Noortgate et al. 2013).

### 54.5.4 Publikationsbias

Die Validität einer publikationsbasierten Meta-Analyse ist gefährdet, wenn die Publikationschancen unter anderem von den Befunden selbst abhängen, also ein Publikationsbias vorliegt (Rothstein et al. 2005). In vielen Disziplinen zeigt sich, dass Studien mit statistisch signifikanten Befunden eine größere Publikationswahrscheinlichkeit haben (Berning/Weiß 2016). Dahinter verbirgt sich die irrtümliche Vorstellung, dass statistische Signifikanz (Häder/Häder, Krebs/Menold, Kapitel 27 und 35 in diesem Band) die Gültigkeit einer Theorie „beweist“. Angenommen, es werden 50 Interventionsstudien zur Reduzierung der Schulverweigerung durchgeführt. Allerdings zei-



gen sich nur in fünf Studien statistisch signifikante Befunde (im Sinne der Intervention). Auf Grund des Publikationsbias werden aber nur diese fünf Studien publiziert und können damit Gegenstand einer Meta-Analyse werden. Diese Meta-Analyse würde anschließend „belegen“, dass die Intervention funktioniert und unter Umständen käme es dann im Weiteren zu Fehlallokationen von finanziellen Mitteln. Die Folgen in der Medizin können sogar noch dramatischer sein. Es ist daher unumgänglich, im Rahmen einer Meta-Analyse das Vorliegen eines Publikationsbias zu testen. Es existiert hierfür inzwischen ein umfangreiches statistisches Instrumentarium (vgl. Weiß/Wagner 2011 oder Rothstein et al. 2005). Eine begrüßenswerte Entwicklung der letzten Jahre zur Prävention von Publikationsbias ist die Zunahme an Präregistrierungen von empirischen Studien (Warren 2018), in denen das Studiendesign, aber noch nicht die eigentlichen empirischen Analysen dargestellt werden.

#### 54.5.5 Software

Die Cochrane Collaboration bietet etwa unter <https://community.cochrane.org/help/tools-and-software> (abgerufen am 20.03.2022) eine Reihe von Programmen an (Literatur- und Datenverwaltung), die bei der Review-Erstellung eine große Hilfe sein können. Für die statistischen Analysen verweisen wir insbesondere auf das freie Statistikpaket R. Für R gibt es zahlreiche Pakete sowie inzwischen auch mehrere Bücher, die eine Einführung in die Nutzung von R für Meta-Analysen geben. Eine Übersicht der verschiedenen R-Pakete bietet der „CRAN Task View: Meta-Analysis“ (Version: 2022-03-07, <https://cran.r-project.org/web/views/MetaAnalysis.html>, abgerufen am 20.03.2022). Für Stata werden Informationen auf der folgenden Webseite dargestellt: <https://www.stata.com/features/meta-analysis/> (abgerufen am 20.03.2022). Inzwischen bietet auch SPSS Funktionen für Meta-Analysen an: <https://www.ibm.com/docs/en/spss-statistics/SaaS?topic=features-meta-analysis> (abgerufen am 20.03.2022).

---

## 54.6 Probleme und offene Fragen

Gegen Meta-Analysen können mehrere Argumente ins Feld geführt werden. Häufig wird behauptet, man vergleiche „Äpfel mit Birnen“, d. h. Meta-Analysen würden Unvergleichbares vergleichen. Diese Heterogenität kann sich auf die abhängigen und unabhängigen Variablen oder auf die Stichprobe (z. B. unterschiedliche Untersuchungseinheiten oder geographische Räume) beziehen. Es werden zwei Lösungen für das *Heterogenitätsproblem* vorgeschlagen. Die theoretische Lösung richtet sich auf eine angemessene Generalisierung der Kategorien (statt „Äpfel“ und „Birnen“ wird „Obst“ zur Bezugskategorie). Die methodische Lösung besteht darin, im Rahmen regressionsanalytischer Verfahren (Blasius/Baur, Kapitel 45 in diesem Band) zu über-



prüfen, ob unterschiedliche Kategorisierungen zu signifikant unterschiedlichen Effektstärken führen.

Eine weitere Kritik richtet sich auf die *Berücksichtigung qualitativ „schlechter“ Studien* („garbage in, garbage out“). Es wird die Auffassung vertreten, dass bei einer Kombination methodisch „guter“ und „schlechter“ Studien Meta-Analysen irreführende Ergebnisse erbringen würden. Meistens wird diesem Problem durch Rating-Skalen begegnet, welche die Qualität der Studien bewerten. Die Skalen werden als Gewichtungsfaktoren in der Meta-Analyse verwendet oder bei der Heterogenitätsaufklärung als Prädiktoren genutzt (siehe ausführlich Valentine 2019).

Kritisiert werden auch *publikationsbasierte Meta-Analysen, die auf nicht-experimentellen Beobachtungs- oder Umfragestudien basieren*. Bei einem *nicht-experimentellen Studiendesign* lassen sich vergleichbare Gruppen erst nachträglich (ex-post-facto) durch statistische Verfahren – üblicherweise multiple Regressionsmodelle (Blasius/Baur, Kapitel 45 in diesem Band) – konstruieren. Es wird auch argumentiert, dass sich Koeffizienten, die aus unterschiedlich spezifizierten multiplen Regressionsmodellen stammen, nicht sinnvoll zusammenfassen lassen. Abgesehen davon, dass unklar ist, was der „eigentliche“ Effekt jeweils ist, kann man ermitteln, ob die Befunde signifikant variieren, je nachdem, welche Drittvariablen kontrolliert werden. Es sollte sich jedenfalls eine Publikationspraxis durchsetzen, bei der immer auch bivariate Befunde, Standardfehler und Fallzahlen präsentiert werden (Friedrichs/Leßke, Kapitel 19 in diesem Band). Dann können Befunde mit und ohne Drittvariablenkontrolle in die Meta-Analyse einfließen und angemessen berücksichtigt werden. Eine bedenkenwerte Alternative zu APD Meta-Analysen sind IPD Meta-Analysen, die den Vorteil haben, dass mit den Originaldaten vergleichbare Regressionsmodelle geschätzt werden können (Weiß/Wagner 2008). Eine familiensoziologische Anwendung einer IPD Meta-Analyse ist das von der DFG geförderte Forschungsprojekt „Harmonisierung und Synthese von paarbiografischen Daten: Ein Modellprojekt zur Verknüpfung von Forschungsdaten aus verschiedenen Infrastrukturen“ (HaSpaD) (<https://www.gesis.org/en/services/processing-and-analyzing-data/data-harmonization/harmonizing-and-synthesizing-partnership-histories>, abgerufen am 17. 9. 2021).

## Literatur

- Berning, Carl C./Weiß, Bernd (2016): Publication Bias in the German Social Sciences: An Application of the Caliper Test to Three Top-Tier German Social Science Journals. In: *Quality & Quantity* 50 (2): 901–917. doi: 10.1007/s11135-015-0182-4
- Borenstein, Michael/Hedges, Larry V./Higgins, Julian P. T./Rothstein, Hannah R. (2021): *Introduction to Meta-Analysis*. Chichester: John Wiley
- Cooper, Harris/Hedges, Larry V./Valentine, Jeffrey C. (Hg.) (2019): *The Handbook of Research Synthesis and Meta-Analysis*. New York: Russel Sage Foundation
- Daikeler, Jessica/Bošnjak, Michael/Lozar Manfreda, Katja (2020): Web Versus Other Survey Modes: An Updated and Extended Meta-Analysis Comparing Response Rates. In: *Journal of Survey Statistics and Methodology* 8(3): p. 513–539. doi: 10.1093/jssam/smz008
- Glass, Gene V. (1976): Primary, Secondary and Meta-Analysis of Research. In: *Educational Researcher* 5 (10): 3–8. doi: 10.2307/1174772
- Hedges, Larry V./Olkin, Ingram (1985): *Statistical Methods for Meta-Analysis*. Orlando: Academic Press
- Hedges, Larry V./Vevea, Jack L. (1998): Fixed- and Random-Effects Models in Meta-Analysis. In: *Psychological Methods* 3 (4): 486–504. doi: 10.1037/1082-989X.3.4.486
- Hedges, Larry V./Tipton, Elizabeth/Johnson, Matthew C. (2010): Robust Variance Estimation in Meta-Regression with dependent Effect Size Estimates. In: *Research Synthesis Methods* 1 (1): 39–65. doi: 10.1002/jrsm.5
- Higgins, Julian P. T./Green, Sally (2008): *Cochrane Handbook for Systematic Reviews of Interventions*. Chichester: John Wiley & Sons
- Kiefer, Corinna/Sturtz, Sibylle/Bender, Ralf (2015): Indirekte Vergleiche und Netzwerk-Metaanalysen. In: *Deutsches Ärzteblatt* 112(47): 803–8. doi: 10.3238/arztebl.2015.0803.
- Petticrew, M./Roberts, H. (2006): *Systematic Reviews in the Social Sciences. A Practical Guide*. Malden: Blackwell Publishing
- Riley, Richard D./Tierney, Jayne F./Stewart, Lesley A. (Hg.) (2021). *Individual Participant Data Meta-Analysis: A Handbook for Healthcare Research*. Hoboken, NJ: Wiley.
- Rothstein, Hannah R./Sutton, Alexander J./Borenstein, Michael (Hg.) (2005): *Publication Bias in Meta-Analysis. Prevention, Assessment and Adjustments*. Chichester: John Wiley
- Valentine, Jeffrey C. (2019): Incorporating Judgments About Study Quality into Research Syntheses. In: Cooper, Harris/Hedges, Larry V./Valentine, Jeffrey C. (Hg.): *The Handbook of Research Synthesis and Meta-Analysis*. New York: Russel Sage Foundation, 129–140

- Van den Noortgate, Wim/López-López, José Antonio/Marín-Martínez, Fulgencio/Sánchez-Meca, Julio (2013): Three-level Meta-analysis of Dependent Effect Sizes. In: *Behavior Research Methods* 45 (2): 576–594. doi: 10.3758/s13428-012-0261-6
- Wagner, Michael/Weiß, Bernd (2006): On the Variation of Divorce Risks in Europe: Findings from a Meta-Analysis of European Longitudinal Studies. In: *European Sociological Review* 22 (5): 483–500. doi: 10.1093/esr/jcl014
- Warren, Matthew (2018): First Analysis of ‚Pre-Registered‘ Studies Shows Sharp Rise in Null Findings. In: *Nature* d41586-018-07118-1. doi: 10.1038/d41586-018-07118-1
- Weiß, Bernd (2008): Meta-Analyse als Verfahren der Forschungssynthese in der Soziologie. Dargestellt anhand zweier Fallbeispiele zu Schulabsentismus. Dissertation, Wirtschafts- und Sozialwissenschaftliche Fakultät: Universität zu Köln
- Weiß, Bernd/Wagner, Michael (2008): Potentiale und Probleme von Meta-Analysen in der Soziologie. In: *Sozialer Fortschritt* 10 (11): 250–256. doi: 10.3790/sfo.57.10-11.250
- Weiß, Bernd/Wagner, Michael (2011): The Identification and Prevention of Publication Bias in the Social Sciences and Economics. In: *The Journal of Economics and Statistics (Jahrbücher für Nationalökonomie und Statistik)* 231 (5/6): 661–684. doi: 10.1515/jbnst-2011-5-608
- White, Howard D. (2019): Scientific Communication and Literature Retrieval. In: Cooper, Harris/Hedges, Larry V./Valentine, Jeffrey C. (Hg.): *The Handbook of Research Synthesis and Meta-Analysis*. New York: Russel Sage Foundation, 51–72
- Wilson, David B./Gottfredson, Denise C./Najaka, Stacy S. (2001): School-Based Prevention of Problem Behaviors: A Meta-Analysis. In: *Journal of Quantitative Criminology* 17 (3): 247–272. doi: 10.1023/A:1011050217296

**Bernd Weiß** arbeitet bei GESIS – Leibniz-Institut für Sozialwissenschaften in Mannheim, ist stellvertretender Abteilungsleiter der Abteilung Survey Design and Methodology und leitet das GESIS Panel, ein probabilistisches Mixed-Mode Access Panel. *Ausgewählte Publikationen:* Predicting Nonresponse in Future Waves of a Probability-based Mixed-mode Panel with Machine Learning, in: Journal of Survey Statistics and Methodology (zusammen mit Christoph Kern und Jan-Philipp Kolb, 2021); Exploring increasing divorce rates in West Germany: Can we explain the iron law of increasing marriage instability?, in: European Sociological Review 31, 2 (zusammen mit Michael Wagner und Lisa Schmid, 2015); The identification and prevention of publication bias in the social sciences and economics, in: Journal of Economics and Statistics 231, 5–6 (zusammen mit Michael Wagner, 2011). *Webseite:* <https://berndweiss.net>. *Kontaktadresse:* bernd.weiss@gesis.org.

**Michael Wagner** ist Professor für Soziologie am Institut für Soziologie und Sozialpsychologie der Universität zu Köln und geschäftsführender Herausgeber der Kölner Zeitschrift für Soziologie und Sozialpsychologie. *Ausgewählte Publikationen:* The transition from living apart together to a coresidential partnership. In: Advances in Life Course Research 39 (zusammen mit Clara H. Mulder und Sandra Krapf, 2019); Union Dissolution and Mobility: Who Moves From the Family Home After Separation? In: Journal of Marriage and Family 72, 5 (zusammen mit Clara H. Mulder, 2010); Bilanz der deutschen Scheidungsforschung. Versuch einer Meta-Analyse, in: Zeitschrift für Soziologie 32, 1 (zusammen mit Bernd Weiß, 2003). *Webseite:* [www.iss-wiso.uni-koeln.de](http://www.iss-wiso.uni-koeln.de). *Kontaktadresse:* mwagner@wiso.uni-koeln.de.