

Meta-Analyse

Weiß, Bernd; Wagner, Michael

Veröffentlichungsversion / Published Version

Sammelwerksbeitrag / collection article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:

Weiß, B., & Wagner, M. (2019). Meta-Analyse. In N. Baur, & J. Blasius (Hrsg.), *Handbuch Methoden der empirischen Sozialforschung* (S. 1511-1522). Wiesbaden: Springer VS. https://doi.org/10.1007/978-3-658-21308-4_112

Nutzungsbedingungen:

Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use:

This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.



Bernd Weiß und Michael Wagner

112.1 Meta-Analysen und systematische Reviews¹

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 die 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. Die Forschungsbefunde, die bei einer Meta-Analyse zusammengefasst werden, entstammen vor allem Fachartikeln. Es wird jedoch empfohlen, auch „graue Literatur“, beispielsweise Konferenzbeiträge, Arbeitspapiere oder Examensarbeiten in eine Meta-Analyse einzubeziehen. Im Rahmen einer Meta-Analyse lässt sich dann untersuchen, ob sich Befunde aus verschiedenen Quellen 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*) (Pigott 2012) 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 oder aber in zusammengefasster

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

(„pooled“) Form analysiert werden. IPD Meta-Analysen sind aufwendig, haben aber mehrere Vorteile: die Vergleichbarkeit der Variablen kann erhöht werden, (Ex-post-Harmonisierung) das Spektrum möglicher Analyseverfahren ist größer und es lassen sich Hypothesen untersuchen, die bislang nicht untersucht wurden. Nachfolgend beschränken wir uns auf die Darstellung von APD Meta-Analysen.

In der Literatur ist im Zusammenhang mit Meta-Analysen 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 112.5.4) 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. 2009, Cooper et al. 2009). In der Medizin und Epidemiologie hat sich die Meta-Analyse bereits in Form der *Evidenzbasierten Medizin* institutionalisiert, deren sichtbarste Ausprägung die *Cochrane Collaboration* ist. 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.

112.2 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 beispielsweise an differierenden Operationalisierungen (Stein, Kapitel 8 in diesem Band) oder dem Stichprobendesign (Häder/Häder, Kapitel 22 in diesem Band) liegen. Mit Hilfe von Meta-Analysen lässt sich nicht nur bestimmen, *wie groß die Heterogenität der Befunde ist*. Vielmehr besteht auch die Möglichkeit, diejenigen *Faktoren zu identifizieren, die 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 Forschungsbemühungen 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) Auf vielen Gebieten besteht ein zentrales 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.

112.3 Typische Forschungsprobleme für Meta-Analysen in den Sozialwissenschaften

Die folgenden drei prototypischen sozialwissenschaftlichen 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 89 in diesem Band) darüber keine Auskunft gibt oder keine entsprechend große Primäruntersuchung vorliegt. So ist beispielsweise unbekannt, wie viele Jugendliche in Deutschland die Schule schwänzen. Allerdings gibt es in Deutschland zahlreiche lokal begrenzte Schüler- und Jugendstudien, in denen das Schulschwänzen erhoben wird. Weiß (2008) hat eine IPD Meta-Analyse von 15 Studien vorgelegt und eine Hochrechnung zur prozentualen Verbreitung des Schulschwänzens vorgenommen. Mit Hilfe dieser Meta-Analyse konnte hochgerechnet werden, dass der Anteil der Jugendlichen, die überhaupt schon einmal geschwänzt haben, bei 17 % liegt; für das häufige Schulschwänzen ergibt sich ein deutlich niedrigerer Anteilswert von knapp 2 %. (2) *Test 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 konnte zunächst nachgewiesen werden, dass es Faktoren gibt, die in allen Ländern und Studien signifikant mit dem Scheidungsrisiko assoziiert sind. Ferner wurde belegt, dass die Übertragung des Scheidungsrisikos zwischen der Eltern- und der Kindergeneration umso geringer ist, je niedriger die Scheidungsbarrieren in einem Land sind. Es zeigt sich also, dass Meta-Analysen auch in der international vergleichenden Forschung (Braun, Kapitel 63 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. 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. Ein Ergebnis der Meta-Analyse ist, dass Interventionen zur Erhöhung der Selbstkontrolle oder der sozialen Kompetenz besonders erfolgreich sind.

112.4 Der Forschungsprozess einer Meta-Analyse

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 (Cooper 2010). 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. Allerdings ist es bei komparativen APD Meta-Analysen möglich, auch neue Fragestellungen zu untersuchen, wenn diese beispielsweise Ländervergleiche betreffen (Wagner/Weiß 2006).

In einem zweiten Schritt sollte bei einer publikationsbasierten APD Meta-Analyse eine möglichst vollständige und systematische *Literaturrecherche* durchgeführt werden (White 2009). Es lassen sich drei Quellen ausmachen, um an relevante Literatur zu gelangen: (a) Expertenbefragungen (per E-Mail, auf Konferenzen etc.) (Helfferich, Kapitel 44 in diesem Band), (b) (Fach)Datenbanken (Ohly/Weber, Kapitel 6 in diesem Band) sowie (c), dass Durchsuchen der durch (a) und (b) gefundenen Literatur nach weiteren Studien („Schneeballsystem“). Ähnlich einer Inhaltsanalyse (Mayring/Fenzl, Kapitel 42 in diesem Band) wird nach der Literaturrecherche das gefundene *Material vercodet und elektronisch gespeichert; in einer Primärstudie entspricht dieser Schritt der Dateneingabe* (Lück/Landrock, Kapitel 32 in diesem Band). Relevant sind vor allem 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 aufklären zu können, 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*: Bei der Befundintegration geht es darum, die einzelnen Befundstatistiken zu einer Gesamtstatistik zusammenzufassen (Abschnitt 112.5.1). (2) *Heterogenitätsanalyse*: Bei der Heterogenitätsanalyse wird die Variation der Befunde analysiert und gegebenenfalls erklärt (Abschnitt 112.5.3). (3) *Analyse fehlender Werte*: Bei der Analyse fehlender Werte (Engel/Schmidt, Kapitel 27 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 112.5.4).

Bei der Publikation der Meta-Analyse wird empfohlen, sich an den zahlreichen, disziplinspezifischen Berichtsstandards zu orientieren. Aus Platzgründen verzichten wir auf die Darstellung dieser Standards und verweisen stattdessen auf einschlägige Suchmaschinen, wo sich diese leicht mit Hilfe der Suchbegriffe „reporting standards“ und „meta-analysis“ finden lassen.

112.5 Statistische Verfahren

In diesem Abschnitt geben wir eine knappe Einführung in die statistischen Grundlagen von publikationsbasierten APD Meta-Analysen vor. Dabei konzentrieren wir uns darauf, die beiden zentralen Modelle der Befundintegration vorzustellen sowie einige wichtige Konzepte einzuführen.

112.5.1 Verfahren und Modelle der Befundintegration

Es existieren zahlreiche statistische Verfahren zur Befundintegration, die sich vor allem nach Art und Informationsgehalt der zugrundeliegenden Befundstatistiken unterscheiden (Borenstein et al. 2009). Typische Analyseeinheiten einer Meta-Analyse sind etwa Korrelationskoeffizienten, Mittelwertdifferenzen, Odds Ratios oder Regressionskoeffizienten (Blasius/Baur, Kapitel 103 in diesem Band). Univariate Anteils- oder Mittelwerte, in einigen Fällen auch Teststatistiken (etwa t -, F -, χ^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 112.5.3 verwiesen.

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

Wenngleich es naheliegend erscheint, für die eigentliche Befundsynthese ein einfaches arithmetisches Mittel zu verwenden, verbietet sich dieser Ansatz im Rahmen einer Meta-Analyse aus zwei Gründen: (1) Die Forschungsbefunde aus verschiedenen Studien basieren für gewöhnlich auf unterschiedlichen Stichproben mit unterschiedlichen Stichprobengrößen. Da der Standardfehler beziehungsweise die Varianz von der Stichprobengröße abhängt, verletzen meta-analytische Daten die Annahme der Homoskedastizität. Die Kleinste-Quadrate-Schätzfunktion (engl. *ordinary least squares*, OLS) liefert 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 (d.h. den inversen quadrierten Standardfehlern) 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 (siehe Jackson/Turner 2017). 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 (beziehungweise 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 im Rahmen 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 unbekannt) 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. 2009).

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.

112.5.2 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. 2009: 105-125). Solche Gründe liegen beispielsweise vor, wenn vermutet wird, dass die Literaturrecherche nicht

vollständig war oder die Befunde aus Studien mit einem unterschiedlichen Stichprobendesign stammen. Es ist dann nicht angebracht, von einem gemeinsamen Effektstärkeparameter 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 schon aus methodischen oder inhaltlichen Gründen das REM angemessen.

112.5.3 Verfahren der Heterogenitätsaufklärung und Umgang mit abhängigen 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. 2009). Beispielsweise können sich Effektstärken 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. Ebenso lässt sich überprüfen, ob die Effektstärken vom Jahr der Datenerhebung, vom Stichprobendesign oder vom geografischen Kontext abhängen.

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).

112.5.4 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 22 und 34 in diesem Band) die Gültigkeit einer Theorie „beweist“. Angenommen, es werden 50 Interventionsstudien zur Reduzierung der

Schulverweigerung durchgeführt. Allerdings zeigen 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).

112.5.5 Software

Die Durchführung eines systematischen Reviews sowie der statistischen Analysen im Rahmen einer Meta-Analyse lässt sich ohne Computer nicht mehr leisten: angefangen von der Web-basierten Literatur- bzw. Studienrecherche, über die Archivierung und Verwaltung der Rechercheergebnisse, der Vercodung und Speicherung der Forschungsbefunde bis hin zu umfangreichen statistischen Analysen und Visualisierungen. Die Cochrane Collaboration bietet etwa unter <http://community.cochrane.org/tools> (abgerufen am 30.11.2017) eine Reihe von Programmen an (Literatur- und Datenverwaltung), die bei der Review-Erstellung eine große Hilfe sein können. Für die statistischen Analysen verweisen wir 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 (Schwarzer et al. 2015). Eine Übersicht der verschiedenen R-Pakete geben Polanin et al. (2017) sowie Dewey (2017). Für SPSS gibt es Makros für basale meta-analytische Verfahren von David B. Wilson, die aber in ihrer Funktionalität sehr begrenzt sind (<http://mason.gmu.edu/~dwilsonb/ma.html>). Für Stata ist das Buch von Palmer und Sterne (2016) die einschlägige Referenz.

112.6 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 103 in diesem Band) zu überprüfen, 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. Gewöhnlich 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 2009).

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 103 in diesem Band) – konstruieren. Es wird auch argumentiert, dass sich Koeffizienten, die aus unterschiedlich spezifizierten multiplen Regressionsmodellen stammen, nicht sinnvoll zusammenfassen lassen (Brüderl 2004). 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, Kapitel 19 in diesem Band). Dann können Befunde mit und ohne Drittvariablenkontrolle in die Meta-Analyse einfließen und angemessen berücksichtigt werden. Eine bedenkenswerte Alternative zu APD Meta-Analysen sind IPD Meta-Analysen, die den Vorteil haben, dass direkt die Originaldaten analysiert werden und vergleichbare Regressionsmodelle geschätzt werden können (Weiß/Wagner 2008, für ein weiteres Beispiel siehe Klein et al. 2013).

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
- Borenstein, Michael/Hedges, Larry V./Higgins, Julian P.T./Rothstein, Hannah R. (2009): *Introduction to Meta-Analysis*. Chichester: John Wiley
- Brüderl, Josef (2004): Meta-Analyse in der Soziologie: Bilanz der deutschen Scheidungsforschung oder „statistischer Fruchtsalat“? In: *Zeitschrift für Soziologie* 33 (1): 84-86
- Cooper, Harris/Hedges, Larry V./Valentine, Jeffrey C. (Hg.) (2009): *The Handbook of Research Synthesis and Meta-Analysis*. New York: Russel Sage Foundation
- Cooper, Harris M. (2010): *Research Synthesis and Meta-Analysis. A Step-by-Step Approach*. Thousand Oaks/California: Sage Publications
- Dewey, Michael (2017): CRAN Task View: Meta-Analysis. URL: <https://cran.r-project.org/web/views/MetaAnalysis.html>
- Glass, Gene V. (1976): Primary, Secondary and Meta-Analysis of Research. In: *Educational Researcher* 5 (10): 3-8
- 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
- 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 (2): 39-65
- Higgins, Julian P.T./Green, Sally (2008): *Cochrane Handbook for Systematic Reviews of Interventions*. Chichester: John Wiley & Sons
- Jackson, Dan/Turner, Rebecca (2017): Power Analysis for Random-Effects Meta-Analysis: Power Analysis for Meta-Analysis. In: *Research Synthesis Methods* 8 (3): 290-302
- Klein, Thomas/Johannes Kopp/Rapp, Ingmar (2013): Metaanalyse mit Originaldaten. Ein Vorschlag zur Forschungssynthese in der Soziologie. In: *Zeitschrift für Soziologie* 42: 222-38
- Palmer, Tom M./Sterne, Jonathan A. C. (Hg.) (2016): *Meta-Analysis in Stata: An Updated Collection from the StatJournal*. College Station/Texas: StataCorp LP
- Petticrew, M./Roberts, H. (2006): *Systematic Reviews in the Social Sciences. A Practical Guide*. Malden: Blackwell Publishing
- Pigott, Therese D. (2012): *Advances in Meta-Analysis*. New York: Springer
- Polanin, Joshua R./Hennessy, Emily A./Tanner-Smith, Emily E. (2017): A Review of Meta-Analysis Packages in R. In: *Journal of Educational and Behavioral Statistics* 41 (2): 206-242

- Rothstein, Hannah R./Sutton, Alexander J./Borenstein, Michael (Hg.) (2005): *Publication Bias in Meta-Analysis. Prevention, Assessment and Adjustments*. Chichester: John Wiley
- Schwarzer, Guido/Carpenter, James R./Rücker, Gerta (2015): *Meta-Analysis with R*. Heidelberg: Springer
- Valentine, Jeffrey C. (2009): Judging the Quality of Primary Research. In: Cooper, Harris/Hedges, Larry V./Valentine, Jeffrey C. (Hg.): *The Handbook of Research Synthesis and Meta-Analysis*. New York: Russel Sage Foundation, 129-146
- 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
- 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
- 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
- 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
- White, Howard D. (2009): 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, 51-71
- 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

Michael Wagner ist Professor für Soziologie am Institut für Soziologie und Sozialpsychologie der Universität zu Köln. *Ausgewählte Publikationen:* 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 Lisa Schmid und Bernd Weiß, 2015); 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. *Kontaktadresse:* mwagner@wiso.uni-koeln.de.

Bernd Weiß arbeitet bei GESIS – Leibniz-Institut für Sozialwissenschaften in Mannheim und leitet das GESIS Panel, ein probabilistisches Mixed-Mode Access Panel. *Ausgewählte Publikationen:* Effectiveness of a standardized patient education program on therapy-related side effects and unplanned therapy interruptions in oral cancer therapy: A cluster-randomized controlled trial, in: *Supportive Care in Cancer* 25, 11 (zusammen mit Christoph Riese et al., 2017); 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:* <http://berndweiss.net>. *Kontaktadresse:* Bernd.Weiss@gesis.org.