

Metaanalyse mit Originaldaten

Ein Vorschlag zur Forschungssynthese in der Soziologie

Meta-Analysis of Original Data

A Proposal for Research Synthesis in Sociology

Thomas Klein*, Johannes Kopp** & Ingmar Rapp*

* Max Weber-Institut für Soziologie, Universität Heidelberg, Bergheimer Straße 58,
69115 Heidelberg, Germany
thomas.klein@soziologie.uni-heidelberg.de; ingmar.rapp@soziologie.uni-heidelberg.de

** Institut für Soziologie, TU Chemnitz, Thüringer Weg 9, 09107 Chemnitz, Germany
johannes.kopp@soziologie.tu-chemnitz.de

Zusammenfassung: Die Metaanalyse ist eine Methode der quantitativen Zusammenfassung von Forschungsergebnissen verschiedener Studien und wird bislang vorwiegend in experimentell orientierten Wissenschaften eingesetzt. Ein Hauptproblem der Metaanalyse in der Soziologie liegt darin, dass soziologische Analysen selten ein experimentelles Design mit Randomisierung verfolgen, das den Einfluss nicht interessierender Faktoren unterbinden würde. Der Beitrag diskutiert die Anwendbarkeit von Metaanalysen in der Soziologie und wirbt für die Metaanalyse mit Originaldaten. Es werden drei Anwendungsbeispiele dieses Verfahrens aus dem Bereich der Scheidungsforschung gegeben. Metaanalysen mit Originaldaten bieten insbesondere den Vorteil, für den Vergleich verschiedener Studien einheitliche Modellspezifikationen und Auswertungskonzepte verwenden zu können, und eine Reihe weiterer Vorzüge. Nachteile des Ansatzes liegen im Aufwand der Harmonisierung von Datensätzen sowie in der Beschränkung verwendbarer Variablen und der theoretisch abbildbaren Prozesse.

Schlagerworte: Scheidung; Metaanalyse; Review; Methodologie.

Summary: Meta-analysis is a method of quantitatively summarizing findings from different studies and has mainly been employed in experimentally oriented disciplines. A major problem of meta-analysis in sociology is that studies rarely employ experimental designs with randomization which preclude distorting influences. This article discusses the applicability of meta-analyses in sociology and advocates conducting meta-analyses of original data. Three examples from divorce research are given. Meta-analyses of original data offer numerous advantages, most notably allowing the use of model specifications and evaluation concepts across different sets of data. The limitations of this method are that it requires a laborious harmonization of data and that it restricts the variables that can be used and thus the theoretical processes that can be examined.

Keywords: Divorce; Meta Analysis; Review; Methodology.

1. Einleitung

Wenn in einem Forschungsfeld viele – womöglich widersprüchliche – Ergebnisse vorliegen, ist es naturgemäß schwierig, einen Überblick zu bekommen. Einige Zeitschriften informieren in mehr oder weniger regelmäßigen großen Abständen mit Überblicksbeiträgen, so genannten Reviewartikeln über den Stand der Forschung in einem bestimmten Forschungsfeld. So geben beispielsweise die Beiträge im Heft 3 des Jahres 2010 des *Journal of Marriage and Family* für die verschiedensten Felder der Familienforschung einen Überblick über die Entwicklun-

gen im letzten Jahrzehnt. Solche rein sprachliche Beschreibungen des Forschungsstands beruhen auf der Kompetenz und dem Überblick ihrer Autoren, sind aber natürlich nicht frei von subjektiven Einschätzungen.

Der Begriff der Metaanalyse bezeichnet demgegenüber eine quantitative Zusammenfassung publizierter Forschungsergebnisse aus verschiedenen Studien zur selben Fragestellung nach festgesetzten Regeln. Im Vergleich mit einem qualitativen Review von Forschungsergebnissen ist mit dieser Quantifizierung eine Standardisierung und Systematisierung des Forschungsüberblicks verbunden. Bei einer hinrei-

chend großen Anzahl von Einzelstudien lässt sich mit der Metaanalyse u. U. auch die Unterschiedlichkeit von Forschungsergebnissen quantitativ aufklären – beispielsweise wenn diese vom Jahr der Untersuchung abhängen (z. B. Amato 2001).

Beginnend mit dem Beitrag von Glass (1976) hat sich die Metaanalyse mittlerweile als Verfahren zur Systematisierung von Forschungsergebnissen fest etabliert und vielfach bewährt. So berichten beispielsweise Borenstein et al. (2009: XXI), dass durch den rechtzeitigen Einsatz metaanalytischer Verfahren in der Medizin tausende Fälle des so genannten plötzlichen Kindstods vermieden worden wären, weil der dafür verantwortliche Faktor (die Bauchlage des Säuglings) erst bei einer Zusammenschau zahlreicher Studien sichtbar wurde. Hunt veröffentlichte bereits im Jahr 1997 eine Erfolgsgeschichte der Metaanalyse, in der er feststellt (Hunt 1997: 19): „Whether one sees meta-analysis as a set of recondite techniques for getting precise answers to irrelevant questions or as an epochal advance in scientific methodology, it unquestionably has come to occupy a major place in contemporary scientific research.“ Während jedoch in der Medizin und der Pharmazie metaanalytische Ergebnisse unter dem Stichwort „evidence based practice“ zum Standard geworden sind und beispielsweise auch in der Psychologie eine weite Verbreitung gefunden haben (Rosenthal & DiMatteo 2001), sind Metaanalysen in der Soziologie noch erstaunlich selten zu finden. Ziel dieses Beitrages ist es aufzuzeigen, dass es zum einen gute Gründe für diese auf den ersten Blick unverständliche Vernachlässigung gibt und wie zum anderen alternative Verfahren aussehen könnten, die die Vorteile der Metaanalyse nutzen, ohne die Schwächen aufzuweisen, die klassische Metaanalysen aus forschungslogischen Gründen in der Soziologie haben müssen.

Hierzu rekapitulieren wir in einem ersten Schritt kurz, was Metaanalysen eigentlich sind (2.1). Anschließend werden Metaanalysen vorgestellt, die in der Soziologie bislang durchgeführt wurden, und die mit ihnen verbundenen forschungslogischen Schwierigkeiten diskutiert (2.2). Diese Schwierigkeiten bringen uns dazu, die Metaanalyse mit Originaldaten als alternatives Verfahren vorzustellen (2.3). Dessen praktische Umsetzung am Beispiel der Scheidungsforschung bildet den Schwerpunkt des darauffolgenden Abschnitts (3.). Wie zu zeigen sein wird, besitzt auch die Metaanalyse mit Originaldaten bestimmte Einschränkungen und Probleme, die abschließend diskutiert werden (4.).

2. Metaanalyse in der Soziologie: Bestandsaufnahme, Probleme und Alternativen

2.1 Metaanalysen als Methode der empirischen Forschungssynthese

Seit den 1970er Jahren hat sich die Metaanalyse als ein Verfahren etabliert, um vielfältige, aber oftmals uneinheitliche Ergebnisse empirischer Forschungen zusammenfassend zu betrachten und zu bewerten (z. B. Glass 1976; Mann 1990, 1994; Eisend 2004; Borenstein et al. 2009). Betrachtet man die betreffenden Einführungen und Lehrbücher, so hat sich dieses Verfahren vor allem in der medizinischen und pharmazeutischen Forschung, aber auch in der Psychologie und vereinzelt in der Pädagogik etabliert. Ausgangspunkt der Metaanalyse in all diesen Forschungsfeldern ist die Unübersichtlichkeit einer Vielzahl empirischer Ergebnisse aufgrund eines rasch wachsenden Feldes an wissenschaftlichen Untersuchungen, Forschungsergebnissen und Publikationen (de Solla-Price 1974; Kölbel 2001).

Die klassische, aber letztlich unangemessene Reaktion (Hunt 1997: 6) auf diese Problematik bestand in der Erstellung von Überblicksartikeln, die in Gestalt entsprechender Abschnitte mit der Überschrift „Zum Stand der Forschung“ bis heute in nahezu jeder Dissertation, aber auch in einer Vielzahl von Artikeln bis heute weit verbreitet sind. Der Überblick, den sie bieten, ist häufig sehr hilfreich, doch es fehlt an Kriterien, um die Güte dieser Beiträge zu bewerten und damit ein methodologisch gesichertes Fazit zu einem gewissen Forschungsgebiet zu ziehen. Dies zeigt sich unter anderem daran, dass Überblicksartikel immer wieder zu diametral entgegengesetzten Schlussfolgerungen kommen (Hunt 1997: 6ff.). Neben dem meist nicht hinreichend genau geschilderten Zugang zur Literatur spielen dabei persönliche Überzeugungen und individuelle Urteilsheuristiken eine große Rolle. So werden z. B. neuere Studien stärker beachtet, auch wenn ihre Erklärungskraft eventuell geringer ist. Alternativ werden ab und an die empirischen Effekte einfach abgezählt. Dieses auch als „vote counting“ bezeichnete Verfahren zeichnet sich in der Regel durch eine systematische Darstellung der Literaturrecherche und somit durch einen breiteren Blick auf das Forschungsfeld aus, vernachlässigt jedoch methodische Unterschiede. Darüber hinaus wird die Effektgröße bei diesem Vorgehen nicht berücksichtigt (vgl. Borenstein 2009: 251ff.; Mosteller & Colditz 1996: 12f.).

Mit der Metaanalyse steht seit den 1970er Jahren ein Verfahren zur Verfügung, das die genannten Probleme der Forschungssynthese lösen kann und in der Lage ist, auch quantitativ statistisch abgesicherte Aussagen über den kumulierten Stand der Forschung und somit eine fundierte Aussage über den Stand des Wissens zu liefern. Ausgangspunkt dieser Forschungstradition ist eine Arbeit von Glass aus dem Jahr 1976, in der er neben der Analyse von Primärdaten und der Sekundärdatenanalyse als Verfahren der Reanalyse bereits vorhandener Daten die Metaanalyse als „analysis of analyses“ (Glass 1976: 3) vorschlägt. Während in den Anfangsjahren der Metaanalyse die möglichst breite Suche nach einschlägiger Literatur zu einem vorab klar definierten Forschungsproblem ein großes Problem darstellte (Hunt 1997), liegen heute zu vielen Forschungsfragen zahlreiche Untersuchungen vor.

Das wichtigste Grundkonzept der Metaanalyse ist die Effektgröße. Aufgrund unterschiedlicher Operationalisierungen und Messungen sind die Einzeleffekte unterschiedlicher Studien in der Regel nicht direkt miteinander vergleichbar. Zur Lösung dieser Problematik werden die einzelnen Effektstärken der interessierenden Größen standardisiert: „Most important was the effect size (...): the mean difference on the outcome variable between treated and untreated subjects divided by the within group standard deviation“, so Glass (1976: 6).¹ Auf diese Art und Weise ist es möglich, auch die Ergebnisse kleinerer Stichproben angemessen zu bewerten und in die Gesamtschau einfließen zu lassen.

Die erste Metaanalyse beschäftigte sich inhaltlich mit der Frage, inwieweit Psychotherapien einen Einfluss auf das Wohlbefinden haben, wobei sowohl unterschiedliche Outcome-Variablen wie verschiedene Therapieformen Berücksichtigung fanden. Seit 1976 hat die Zahl der Metaanalysen stetig zugenommen, insbesondere in der Medizin, der Pharmakologie, der Public-Health-Forschung, der empirischen Pädagogik sowie in der Psychologie, dort bezeichnet als das Verfahren „for the systematic review of randomized controlled trials“ (Mosteller & Colditz 1996: 2).

Gegen die Metaanalyse werden beinahe schon routinemäßig einige Einwände vorgebracht, die sich unter drei Stichworten zusammenfassen lassen: (i) „garbage in – garbage out“, (ii) „oranges and apples“ sowie (iii) „publication bias“ (Rosenthal & DiMatteo 2001: 66ff.; vgl. Eisend 2004: 19ff.). Das erste Argument bezieht sich darauf, dass die methodische Qualität der in einer Metaanalyse berücksichtigten Studien sehr unterschiedlich ausfallen kann. Wenn Studien sich in der Qualität der Messungen oder des Designs deutlich unterscheiden, sollten sie nicht gemeinsam berücksichtigt werden. Schon früh wurde deshalb vorgeschlagen (vgl. Rosenthal & DiMatteo 2001: 67), die Qualität der Studien zu erfassen und als Kontrollvariable in der Metaanalyse zu berücksichtigen. Das zweite Argument zielt auf die Unvergleichbarkeit der verwendeten Operationalisierungen und Messungen. Auch hier kann man diese Unterschiede als Moderatorvariablen in einer Metaanalyse berücksichtigen, und zudem gilt (Rosenthal & DiMatteo 2001: 68): „It is a good thing to mix apples and oranges, particularly if one wants to generalize about fruits.“ Ein dritter Kritikpunkt bezieht sich auf die Veröffentlichungsstrategien von Autoren, aber auch von Zeitschriften (Hunt 1997: 118ff.). Statistisch signifikante Ergebnisse haben sowohl eine erhöhte Chance eingereicht als auch publiziert zu werden. Der Einfluss der betreffenden Faktoren wird überschätzt, wenn Studien, die kein signifikantes Ergebnis gebracht haben, seltener publiziert werden. Die Antwort auf dieses so genannte „file-drawer problem“ liegt in der Berechnung der Anzahl von nichtsignifikanten Studien, die notwendig wären um die Ergebnisse einer Metaanalyse inhaltlich zu ändern. In der Regel liegt diese Zahl in einem Bereich, die es unwahrscheinlich macht, dass inhaltliche Interpretationen geändert werden müssten.

Auch wenn mit den drei genannten Schwierigkeiten wichtige Probleme angesprochen werden, so sind damit keine prinzipiellen Hindernisse gegen Metaanalysen als Verfahren der Forschungssynthese geltend gemacht. Im Gegenteil: Eine Vielzahl von Analysen besonders in den so genannten ‚life sciences‘ zeigen den enormen Vorteil von Metaanalysen gegenüber Verfahren wie beispielsweise qualitativen Forschungsreviews oder dem bereits erwähnten „vote counting“. Metaanalysen bieten die Möglichkeit, statistisch abgesicherte Aussagen über den Forschungsstand in den unterschiedlichsten Bereichen zu treffen und somit Beiträge zum Erkenntnisfortschritt zu liefern, der Kosten sparen und beispielsweise in der Medizin Leben retten kann. Auch wenn die Konsequenzen soziologischer Forschung

¹ An dieser Stelle findet sich in vielen Arbeiten (vgl. exemplarisch Wagner & Weiß 2003: 39ff.) die berechnete Effektgröße, die die berechnete Effektgröße darstellt, aber auch der empirischen Ergebnisse nicht hinreichend genau ist. Für Metaanalysen ist es eigentlich Voraussetzung, dass die Standardfehler der Koeffizienten angegeben sind; vgl. Borenstein et al. 2009 für einige Hinweise, wie diese Standardfehler teilweise zu schätzen sind.

selten derart dramatisch erscheinen, kann man sich die Frage stellen, warum Metaanalysen in der Soziologie nur selten Anwendung finden.

2.2 Metaanalyse in der Soziologie: wenige Großstudien und die Logik soziologischer Umfrageforschung²

Eine der Ursachen für die weite Verbreitung von Metaanalysen in den oben genannten Wissenschaften ist die große Anzahl von eher kleinen Studien zu derselben Fragestellung. Soziologische Untersuchungen setzen hingegen oft große Fallzahlen voraus. Dementsprechend konzentriert sich die betreffende Forschung auf wenige Großstudien, mit denen sich die Datengrundlage der empirischen Sozialforschung in den letzten Jahrzehnten zunehmend verbessert hat.

Zu den wenigen Metaanalysen in der Soziologie, die auf eine große Zahl vorliegender Studien Bezug nehmen können, gehören die Metaanalysen von Jose et al. (2010) zum Einfluss vorehelicher Lebensgemeinschaften auf die Qualität und die Stabilität von Ehen sowie die von Shor et al. (2012) zum Einfluss der Verwitwung auf das Mortalitätsrisiko des überlebenden Ehepartners. In Deutschland haben Dölling und Hermann (Hermann 2010; Dölling & Hermann 2012; Dölling et al. 2006, 2009) eine Metaanalyse zur Abschreckungswirkung des Strafrechts durchgeführt, die auf mehr als 700 Einzelstudien beruht.

² In diesem Abschnitt sollen nur Untersuchungen vorgestellt werden, die den oben genannten wesentlichen Merkmalen einer Metaanalyse genügen, so dass etliche Beiträge, die unter dem Schlagwort Metaanalyse zu finden sind, hier nicht weiter diskutiert werden können – und müssen. So fordern Ruddat & Schulz (2010) eine quantitativ-qualitative Metaanalyse oder diskutieren Formen der partizipativen Metaanalyse, meinen damit jedoch eher Experteninterviews beziehungsweise so genannte Delphi-Befragungen. Baur (2008) stellt eine klassische Literaturübersicht über die Parallelen und Unterschiede zwischen sozialwissenschaftlichen und historischen Arbeiten vor, bezeichnet aber die Art und Weise der Aufbereitung als Metaanalyse (Baur 2008: 221). Kalmijn & Kraaykamp (2007) untersuchen selbst die verschiedenen Länder des European Value Surveys und führen dann eine zusammenfassende Analyse der einzelnen Effekte und somit letztlich eine Mehrebenenanalyse durch, bezeichnen ihr Vorgehen jedoch als „meta-analytical design“ (Kalmijn & Kraaykamp 2007: 547). Fulkerson & Thompson (2008) untersuchen die verschiedenen Definitionen des Sozialkapitalbegriffes, bezeichnen dieses Vorgehen aber auch als Metaanalyse.

Metaanalysen können jedoch auch sinnvoll sein, wenn die Zahl verfügbarer Studien wesentlich kleiner ist, insbesondere wenn aus diesen wenigen Studien viele, zum Teil unterschiedliche Ergebnisse publiziert wurden. Dies zeigen Wagner & Weiß (2003) anhand einer Metaanalyse der deutschen Scheidungsforschung, für die nur (aber immerhin) 7 Studien – darunter 6 Großstudien – vorliegen: die Mannheimer Scheidungsstudie, das Sozioökonomische Panel, der Familiensurvey, die Lebensverlaufsstudie, der Family and Fertility Survey, der ALLBUS sowie die Kölner Gymnasiastenstudie (Wagner & Weiß 2003: 31). Ausgehend von 42 Publikationen zum Ehescheidungsrisiko aus diesen Studien werden in der Metaanalyse von Wagner und Weiß die Ergebnisse der einzelnen Publikationen für die jeweiligen Studien gemittelt und dann mit Hilfe metaanalytischer Verfahren untersucht. Auf ähnliche Weise wurde eine weitere Metaanalyse der Scheidungsforschung in Europa (Wagner & Weiß 2006b) und eine Metaanalyse zur Stadt-Umland-Wanderung (Bleck & Wagner 2006) durchgeführt.

Die geringe Verbreitung von Metaanalysen in der Soziologie erklärt sich jedoch nicht nur mit der eher kleinen Zahl von wenigen Großstudien in vielen Forschungsfeldern, sondern es gibt auch forschungslogische Gründe, die Zweifel an der Einsatzfähigkeit der Metaanalyse in der Soziologie nahelegen, wenn nicht aufdrängen (vgl. Brüderl 2004). Diese betreffen die Verrechnung von Effektstärken, die auf ganz unterschiedliche Weise zustande kamen und nicht auf einem experimentellen Design beruhen, das sich durch die Konstanzhaltung einer Vielzahl von Faktoren auszeichnet. Hintergrund des Problems ist, dass ein untersuchter Einflussfaktor – z. B. der des Heiratsalters auf das Scheidungsrisiko – oft eng mit anderen Faktoren (Drittvariablen) zusammenhängt, die das Scheidungsrisiko ebenfalls beeinflussen: Ein hohes Heiratsalter geht beispielsweise tendenziell damit einher, dass es sich eher um eine Zweitehe handelt als bei jüngeren Heiratsalter. Während das Scheidungsrisiko durch ein höheres Heiratsalter zwar reduziert wird, sind Zweitehen hingegen im Durchschnitt instabiler. Geht man nun bei der Untersuchung des Einflusses, den das Heiratsalter auf das Scheidungsrisiko ausübt, nicht einheitlich entweder von Ersten oder von Zweitehen aus, führt das dazu, dass der stabilisierende Effekt des höheren Heiratsalters unterschätzt wird, weil bei höherem Heiratsalter Zweitehen häufiger vorkommen. Je nachdem, ob Heiratsalter und Scheidungsrisiko bivariat in Zusammenhang gebracht werden oder welche und wie viele Drittvariablen kontrolliert

werden, ergeben sich unterschiedliche Effektstärken, die nicht miteinander vergleichbar sind und nicht sinnvoll miteinander verrechnet werden können.

Im Unterschied zur Soziologie ist die empirische Forschung insbesondere in der Medizin und in der Psychologie stark von randomisierten, kontrollierten Studien (randomized controlled trial, RCT) geprägt, und eine zufällige Aufteilung der Versuchspersonen auf die Experimental- und die Kontrollgruppe (die Randomisierung) sorgt dafür, dass Drittvariablen keinen systematischen Unterschied zwischen beiden Gruppen erzeugen, so dass Unterschiede bei der untersuchten Variablen ausschließlich der Experimentalvariable zugeschrieben werden können. Die in randomisierten Experimenten gefundenen Ergebnisse lassen sich deshalb relativ problemlos in einer Metaanalyse verrechnen.³

In den von Wagner & Weiß (2003) berücksichtigten soziologischen Großstudien kann hingegen die Heterogenität der Population nicht durch eine einfache Randomisierung kontrolliert werden. In der soziologischen Umfrageforschung, insbesondere in den wichtigen Großstudien, findet man deshalb in der Regel nur ein so genanntes Ex-post-facto-Design (Schnell et al. 2011: 223ff.), bei dem man alle relevanten Drittvariablen statistisch (statt schon durch das Untersuchungsdesign) kontrollieren (und dafür insbesondere überhaupt erst kennen) muss, um den tatsächlichen Einfluss eines Faktors zu ermitteln. Um inhaltlich interessierende Effekte abzubilden, sind daher immer multivariate Verfahren notwendig, die sich durch eine grundlegend eigenständige Logik der Datenanalyse auszeichnen (Kopp & Lois 2012). Die grundlegende Vorgehensweise ist dabei, mit Hilfe so genannter Mediations- und Modera-

tionsanalysen (MacKinnon 2008; Kopp & Lois 2012) einen bivariaten Zusammenhang mit Hilfe weiterer soziologischer, sozialstruktureller oder einstellungstheoretischer Variablen zu erklären. In der Regel finden sich dabei in soziologischen Publikationen unterschiedliche Berechnungen, die sich vor allem durch unterschiedliche Variablen unterscheiden, die konstant gehalten werden. Oft finden sich unterschiedliche Berechnungen auch in derselben Publikation. Auf diese Weise wird schrittweise zur Erklärung des Ausgangsproblems beigetragen. In einer Metaanalyse werden jedoch diese inhaltlich sehr unterschiedlichen Modelle als gleichgewichtig angesehen und gehen alle in die Metaanalyse ein.

Dies ist ein grundlegendes Problem von Metaanalysen in der Soziologie, das auch in der Literatur gesehen wird (Wagner & Weiß 2006b: 488): „Coefficients from bivariate and multivariate methods differ according to their magnitude and standard errors. (...) meta-analysis misses adequate procedures of multivariate result integration.“ Lipsey & Wilson (2001: 67) stellen weiter fest: „Multivariate relationships present special challenges to meta-analysis. (...) the varying sets of independent variables across regression equations complicates their synthesis.“ Und wieder Wagner & Weiß (2003: 35) „Insbesondere die Integration von Partialkoeffizienten ist ein Problem, das noch nicht befriedigend gelöst ist.“ Gerade die oben geschilderten Anwendungen der Metaanalyse in der Soziologie verwenden jedoch relativ unhinterfragt trotz allem derartige Ergebnisse. „We do not know of any meta-analysis that ignores an effect size because coefficients were estimated in multivariate models. Because it is common to aggregate effect sizes which are related to different subgroups (...) and which estimate different parameters, it is reasonable to use this method“, so Wagner & Weiß (2006b: 488).

Diese Schlussfolgerung ist jedoch nicht haltbar. Allenfalls ließe sich darüber nachdenken, die jeweils konstant gehaltenen Drittvariablen – so wie andere Eigenschaften einzelner Studien – in eine Metaanalyse mit aufzunehmen. In der Praxis wird dies jedoch an der Vielfalt und Unterschiedlichkeit der Modellierungen scheitern und geschieht eben auch nicht. Der Umstand, dass soziologische Forschung häufig diesen Rahmenbedingungen der Umfrageforschung unterliegt, ist – neben der kleineren Anzahl empirischer Studien in der Soziologie – ein zweiter Grund dafür, warum Metaanalysen in der Soziologie wenig verbreitet sind.

Hinzu kommen weitere Aspekte, die die Möglichkeiten der Metaanalyse besonders in der Soziologie weiter einschränken. Drittens spielt im Rahmen der

³ Selbstverständlich unterscheiden sich auch Experimente in ihrer Güte, beispielsweise hinsichtlich der Qualität der Verblindung der Untersuchung oder ähnlicher Faktoren. Hier besteht jedoch die Möglichkeit, diese Gütekriterien in eine Metaanalyse aufzunehmen und die Ergebnisse unter Kontrolle der methodischen Güte der Untersuchung zu interpretieren (vgl. Borenstein et al. 2009). Diesbezüglich wird auch zwischen „fixed-effect models“ und „random-effect models“ in einer Metaanalyse unterschieden, je nachdem, ob die statistischen Modelle der Metaanalyse davon ausgehen, dass die eigentlich interessierende Effektgröße in allen Studien gleich oder unterschiedlich ist (vgl. Borenstein et al. 2009: 61ff.). Die wenigen Metaanalysen in der Soziologie verzichten aus guten Gründen auf derartige Modellierungsstrategien, da die entsprechenden Verfahren bei kleinen Fallzahlen – und hiermit sind die Zahl der zugrundeliegenden Studien gemeint – nur sehr eingeschränkt einsetzbar sind (vgl. Borenstein et al. 2009: 84).

Modellspezifikation nicht nur eine Rolle, welche Drittvariablen kontrolliert, sondern ebenso, welche Berechnungen im Einzelfall durchgeführt werden (multiplikative versus additive Regressionsmodelle, Einzelfall- versus logistische Regression usw.). Und viertens sind durch die limitierte Anzahl weniger Großstudien bei vielen soziologischen Fragestellungen der quantitativen Aufklärung unterschiedlicher Ergebnisse enge Grenzen gesetzt. Auch diese Einschränkungen haben in anderen Wissenschaften, die auf Metaanalysen zurückgreifen, teilweise eine geringere Bedeutung, z. B. weil die Auswertungsmodalitäten in klinischen Studien stärker standardisiert oder sogar reglementiert sind (geringere Unterschiede der Modellspezifikation), oder weil viele Einzelstudien zur quantitativen Aufklärung unterschiedlicher Ergebnisse vorliegen.

2.3 Metaanalysen mit Originaldaten – ein Vorschlag

Vor dem Hintergrund der beschriebenen Probleme, mit denen die Metaanalyse speziell in der Soziologie konfrontiert ist, schlägt der vorliegende Beitrag vor, verschiedene Studien nicht auf der Basis publizierter Ergebnisse, sondern auf der Basis der Originaldaten zusammenzufassen und zu vergleichen. Diese Form der Metaanalyse bezieht sich auf die in den Originaldaten enthaltenen Untersuchungseinheiten statt auf Regressionskoeffizienten, in denen sich publizierte Ergebnisse oft kristallisieren. Der Begriff der Metaanalyse bezieht sich dabei nicht auf die Untersuchungseinheiten – Originaldaten versus Regressionskoeffizienten –, sondern auf die übergreifende Zusammenfassung und den übergreifenden Vergleich verschiedener Studien auf der Basis der originären Untersuchungseinheiten, wie sie den Primär- und Sekundäranalysen zugrunde liegen. Auch in der Literatur zur Metaanalyse wird in den letzten Jahren ein entsprechendes Verfahren diskutiert: „When the meta-analyst has access to all of the original data from each study, the meta-analysis may be referred to as an individual participant data (or individual patient data) meta-analysis.“ (Borenstein et al. 2009: 316; vgl. Sutton et al. 2008; Berlin et al. 2002; Sutton & Higgins 2008). In der methodischen Literatur zu Metaanalysen werden derartige Verfahren immer bedeutsamer. Bereits Stewart und Parmar (1993) untersuchen diese auf Individualdaten zurückgreifenden Metaanalysen (meta-analysis of individual patient data = MAP) und vergleichen sie mit den herkömmlichen, auf publizierte Ergebnisse zurückgreifenden Verfahren (meta-analysis of the literature = MAL). Sie können deutliche

Unterschiede in den Schlussfolgerungen beider Verfahren zeigen und fordern deshalb (Stewart & Parmar 1993: 418): „The results of a meta-analysis of the literature alone may be misleading. Whenever possible, a meta-analysis of updated individual patient data should be done.“ Riley et al. (2010) geben einen Überblick über die Vorteile dieser Metaanalysen mit Originaldaten. Sie betonen die Möglichkeit einer konsistenten Vorgehensweise bei der Datenanalyse und insbesondere bei der statistischen Modellierung und kommen zu folgendem Fazit, das man – wie unten deutlich werden wird – sicherlich nicht in seiner ganzen Konsequenz teilen muss (Riley et al. 2010: 4): „It is thus very clear why individual participant data are deemed the ‚gold standard‘ for meta-analysis.“⁴ Andererseits machen sie aber auch deutlich, dass trotz der zunehmenden Beliebtheit dieses Verfahrens – so hat die Anzahl entsprechender Analysen deutlich zugenommen, und seit 2005 lassen sich rund 50 Metaanalysen mit Originaldaten im Jahr finden – die absolute Zahl im Vergleich zu herkömmlichen Metaanalysen immer noch relativ gering ist. In der sozialwissenschaftlichen Diskussion zu Metaanalysen werden diese Ideen bislang vollkommen ignoriert. Wie die folgenden Ausführungen deutlich machen, könnten jedoch durch die Metaanalyse mit Originaldaten vielfältige Probleme der wenigen bislang vorliegenden soziologischen Metaanalysen vermieden und konsistentere und aussagekräftigere Ergebnisse erzielt werden.

Gerade die Analysen zum Stand der Scheidungsforschung von Wagner & Weiß (2003, 2006a, 2006b) liefern den Beleg, dass möglicherweise gar keine Notwendigkeit besteht, sich in die Probleme der Metaanalyse in der soziologischen Umfrageforschung zu begeben – zumindest nicht in der soziologischen Scheidungsforschung. Die Metaanalyse der deutschen Scheidungsforschung beruht auf insgesamt nur sieben Datensätzen, die wissenschaftsöffentlich (i. d. R. über das Datenarchiv für Sozialwissenschaften, das ehemalige Zentralarchiv für Empirische Sozialforschung ZA) zugänglich sind. Damit stellt sich die Situation in der soziologischen Forschung deutlich anders dar als beispielsweise in der Medizin oder Pharmakologie, wo die Zahl der Datenbestände leicht dreistellig sein kann und diese Daten zudem oft nicht für Reanalysen zur Verfügung stehen.

Die Metaanalyse mit Originaldaten weist einen Ausweg aus den geschilderten Problemen der tradi-

⁴ Eine ähnliche Einschätzung findet sich in dem Überblick von Broeze et al. (2010).

tionellen Metaanalyse, und diese Vorgehensweise bietet zudem erweiterte Analysemöglichkeiten. Wir sehen zumindest vier Vorteile.

Ein erster Vorteil ergibt sich aus der Möglichkeit, einheitliche Modellspezifikationen und Auswertungskonzepte zu verwenden. Es ist unmittelbar einsichtig, dass sich in Bezug auf Modellspezifikation und Auswertungskonzept auf der Grundlage von Originaldaten einiges vereinheitlichen lässt.

Ein zweiter Vorteil betrifft die Aufklärung von Heterogenität zwischen verschiedenen Studien. Auf der Basis von Originaldaten lässt sich beispielsweise zur Aufklärung von Heterogenität zwischen verschiedenen Studien eine gemeinsame Schnittmenge verschiedener Stichproben vergleichen, etwa ein gemeinsamer Alters-Range oder bei Längsschnittstudien auch ein gemeinsamer Untersuchungszeitraum. In Betracht kommt auch eine statistische Konstanzhaltung der Variablen, hinsichtlich derer sich verschiedene Stichproben unterscheiden. Beides verlangt obendrein keine große Zahl empirischer Studien, sondern kann schon zur Aufklärung von Heterogenität zwischen nur zwei verschiedenen Studien beitragen.

Ein dritter Vorteil ist die Ausschaltung eines möglichen Publikationsbias. Ob statistisch signifikante Ergebnisse häufiger publiziert werden, spielt für eine Metaanalyse mit Originaldaten keine Rolle. Voraussetzung ist allerdings, dass bei der Auswahl der Datensätze auch Studien berücksichtigt werden, zu denen keine publizierten Befunde vorliegen.

Viertens eröffnen gepoolte Daten die Möglichkeit vertiefter Analysen. Dank der vergrößerten Fallzahlen lassen sich auch Zusammenhänge beleuchten, die bisher aus Fallzahlgründen nicht untersucht werden konnten.

Der vorliegende Beitrag demonstriert diese vier Vorteile der Metaanalyse mit Originaldaten anhand einer Analyse der Ehestabilität. Hierbei werden allerdings auch Beschränkungen und Probleme einer Metaanalyse mit Originaldaten deutlich, die abschließend diskutiert werden.

3. Metaanalyse mit Originaldaten: Ein Anwendungsbeispiel aus der Scheidungsforschung

3.1 Auswahl und Harmonisierung der Studien

Am Beginn einer Metaanalyse mit Originaldaten steht die Sichtung und Auswahl geeigneter Studien, die in die Metaanalyse einbezogen werden. Als

Kriterien für die Eignung von Studien lassen sich einerseits eine hinreichende Entsprechung der Stichproben benennen, insbesondere in Bezug auf Grundgesamtheit und Auswahlverfahren, sowie andererseits eine hinreichende Entsprechung der benötigten Variablen hinsichtlich Verfügbarkeit und Operationalisierung.

Grundlage der nachfolgend dargestellten Metaanalyse zur Ehestabilität sind Originaldaten aus fünf sozialwissenschaftlichen Großstudien. Diese wurden ursprünglich im Rahmen einer Analyse zur „(In-)Stabilität von Partnerschaften im mittleren und höheren Erwachsenenalter“ harmonisiert und kumuliert.⁵ Bei den fünf Studien handelt es sich um die Allgemeine Bevölkerungsumfrage der Sozialwissenschaften (ALLBUS), den Generations and Gender Survey für Deutschland (GGG), die Lebensverlaufsstudie, die Mannheimer Scheidungsstudie und das Sozioökonomische Panel (SOEP).⁶

Bei der Sichtung und Auswahl der Studien wurde zunächst darauf geachtet, dass neben sämtlichen Studien, die bereits für Scheidungsanalysen für Deutschland herangezogen wurden, auch weitere potentiell für Trennungs- und Scheidungsanalysen geeignete Studien berücksichtigt werden.⁷ Alle identifizierten Datensätze wurden anschließend auf eine hinreichende Entsprechung in Bezug auf Grundgesamtheit, Auswahlverfahren, Erhebungszeiten und Erhebungsmethode geprüft. Ausgewählt wurden nur Studien, die die deutsche oder westdeutsche Bevölkerung repräsentieren. Außerdem wurden nur

⁵ Die Arbeiten wurden unter dem genannten Titel als DFG-Projekt gefördert (Antragsteller Thomas Klein, Geschäftszeichen KL 946/11-1).

⁶ Mit Ausnahme des später durchgeführten GGS von 2005 sind diese Studien auch in die Metaanalyse von Wagner und Weiß (2003) eingegangen, der ALLBUS und das SOEP natürlich nur mit den bis dahin ausgewerteten Daten. Demgegenüber bleiben in der vorliegenden Untersuchung mit dem Familiensurvey, dem Family and Fertility Survey und der Kölner Gymnasiastenstudie drei Studien unberücksichtigt, die Eingang in die Metaanalyse von Wagner und Weiß fanden, die aber ausschließlich jüngere und mittlere Altersbereiche abdecken und die deshalb nicht in die Datenkumulation einbezogen worden sind. Aus diesem Grund finden auch nur die beiden Teilstichproben LV1 und LV2 der Lebensverlaufsstudie Berücksichtigung, die sich auf ältere Geburtskohorten der 1919–21, 1929–31, 1939–41 und 1949–51 Geborenen beziehen. Die Datengrundlage der vorliegenden Untersuchung überschneidet sich somit nur zu einem – allerdings großen – Teil mit den Studien, die in die Metaanalyse von Wagner und Weiß eingegangen sind.

⁷ Vergleiche für eine ausführliche Beschreibung der Sichtung und Auswahl der Studien Rapp (2013).

Zufallsstichproben berücksichtigt, da nur diese garantieren, dass die Stichprobe nicht systematisch verzerrt ist.⁸ Die fünf oben genannten Studien erfüllen beide Kriterien. Zwar erstrecken sich die Erhebungen über den Zeitraum von 1980, dem Jahr der ersten ALLBUS-Erhebung, bis zum Jahr 2009, in dem die jüngste hier einbezogene Erhebungswelle aus dem SOEP stattfand. Dennoch überlappen sich die Längsschnittdaten der zumeist retrospektiv erhobenen Ehebiografien zu einem großen Teil. Als Erhebungsmethode verwenden alle fünf Studien die standardisierte Befragung.

Neben einer hinreichenden Entsprechung der Stichproben setzt eine Harmonisierung und Kumulation verschiedener Studien eine hinreichende Entsprechung der benötigten Variablen voraus. Dies ist bei der hier getroffenen Auswahl der Fall. So liegen in allen fünf Studien Informationen zum Beginn und zum Ende aller erfassten Ehen vor, aus denen sich die Ehedauer berechnen lässt.⁹ In allen Studien außer dem SOEP beruhen diese Informationen durchweg auf Retrospektivangaben. Im SOEP ist die Ehebiografie sowohl retrospektiv (für die Zeit vor Aufnahme in das Panel) als auch innerhalb des Panelzeitraums erfasst. Da an dieser Stelle nur diejenigen Ehen aus dem SOEP berücksichtigt werden, die im Panelzeitraum aufscheinen, beruhen die Informationen zum Ende der Ehen allesamt auf prospektiv erhobenen Daten. Bei den Angaben zum Beginn der Ehe handelt es sich entweder um retrospektiv oder um prospektiv erhobene Daten, je nachdem, ob die Ehe vor oder während des Panelzeitraums geschlossen wurde. In die Berechnung der Ehedauer fließen erstens Informationen zum Heiratsjahr ein, das in allen fünf Surveys in vergleichbarer Weise erhoben wurde. Zweitens fließen Informationen zum Endzeitpunkt der Ehe ein. Diese kann entweder mit dem Tod oder durch Trennung und Scheidung beendet werden. Wann immer möglich, wird bei der Berechnung der Ehedauer auf die Trennung und damit auf die faktische Auflösung der Partnerschaft Bezug genommen. Im GGS, in der Teilstichprobe LV2 der Lebensverlaufsstudie und im SOEP ist das Trennungsjahr für alle Ehen einer Person erfragt.

Dies gilt auch für die Mannheimer Scheidungsstudie, die allerdings nur Erstehen enthält. Im ALLBUS ist das Trennungsjahr nicht erfasst, in der Teilstichprobe LV1 der Lebensverlaufsstudie nur in bestimmten Fällen. Für die Ehen aus dem ALLBUS und aus LV1 wird das Ende der Ehe deshalb aus dem Scheidungsjahr minus 1 erschlossen, da Scheidungen in der Regel um ein Jahr verzögert auf Trennungen folgen (vgl. hierzu ausführlich Rapp 2013: 74 f.). Schließlich liegen in allen Studien außer dem GGS Informationen zum Verwitwungsjahr vor, das als Zensierungszeitpunkt behandelt wurde.¹⁰

Als unabhängige Variable werden das Heiratsjahr (metrisch), die Ordnungsnummer der Ehe als erste Ehe oder als Folgeehe (als Dummy-Variable), das Heiratsalter (metrisch) und der Altersabstand der Ehepartner (über ein Set von Dummy-Variablen) berücksichtigt. Das Heiratsjahr wurde in allen fünf Studien erfragt. Für alle Ehen liegen ebenfalls Informationen darüber vor, ob es sich um eine Erstehe handelt oder um eine Folgeehe, bei der ein oder beide Partner früher schon einmal verheiratet waren. Bei den Ehen aus dem ALLBUS, aus dem GGS und aus der Lebensverlaufsstudie bezieht sich diese Information nur auf die Befragungsperson. Für den anderen Ehepartner ist hier unbekannt, ob dieser früher schon einmal verheiratet war. Bei den Ehen aus der Mannheimer Scheidungsstudie handelt es sich generell um die erste Ehe der Befragungsperson. Im Unterschied zu den zuvor genannten Studien ist aber bekannt, ob der Ehepartner früher schon einmal verheiratet war. Ist dies der Fall, wird die Ehe als Folgeehe gekennzeichnet. Für die Ehen aus dem SOEP liegen in der Regel für beide Partner Informationen darüber vor, ob sie früher schon einmal verheiratet waren. Schließlich ist in allen fünf Studien das Heiratsalter der Befragungsperson bekannt, das sich in allen Studien aus dem Heiratsjahr und aus dem Geburtsjahr der Befragungsperson erschließt. Zum Altersabstand liegen hingegen nur für die Ehen aus dem GGS, aus der Teilstichprobe LV2 der Lebensverlaufsstudie, aus der Mannheimer Scheidungsstudie und aus dem SOEP verwertbare Informationen vor.

Die berücksichtigten Studien erfassen insgesamt Daten über 44.894 Ehen: 18.401 Ehen im ALLBUS, 7.079 Ehen im GGS, 3.629 Ehen in der Lebensverlaufsstudie, 5.020 Ehen in der Mannheimer

⁸ Dieses Auswahlkriterium legen auch Dinkel & Milenovic (1992) bei der Kumulation verschiedener Datensätze zum Zweck der Untersuchung der Kohortenfertilität zugrunde.

⁹ Siehe Rapp (2013) für eine ausführliche Dokumentation der Operationalisierung dieser und nachfolgend genannter Variablen in den Originaldatensätzen und der vorgenommenen Variablen-Harmonisierung im kumulierten Datensatz.

¹⁰ Die durch Verwitwung beendeten Ehen aus dem GGS sind, da der Zeitpunkt der Verwitwung und somit auch die Ehedauer unbekannt sind, aus den Analysen ausgeschlossen.

Scheidungsstudie und 10.765 Ehen im SOEP.¹¹ Von dieser Menge stehen jene Ehen nicht für Trennungsanalysen zur Verfügung, bei denen keine Informationen zum Beginn (d. h. zum Heiratsjahr) oder Ende der Ehe (d. h. zum Trennungs- oder Verwitwungsjahr) vorliegen. Des Weiteren werden jene Ehen aus den Analysen ausgeschlossen, bei denen bekannt ist, dass die beiden Ehepartner nicht bereits seit der Eheschließung in Deutschland leben. Nach Ausschluss dieser Fälle verbleiben 17.686 Ehen aus dem ALLBUS, 6.068 Ehen aus dem GGS, 3.602 Ehen aus der Lebensverlaufsstudie, 4.633 Ehen aus der Mannheimer Scheidungsstudie und 10.371 Ehen aus dem SOEP. Für die nachfolgenden Analysen stehen somit 42.360 Ehen mit 6.603 Trennungsergebnissen zur Verfügung.¹²

3.2 Auswertungsverfahren

Ein zentraler Vorteil der Metaanalyse mit Originaldaten ergibt sich aus der Möglichkeit, einheitliche Modellspezifikationen und Auswertungskonzepte zu verwenden. Zur Analyse ehelicher Stabilität stellen ereignisanalytische Verfahren (Allison 1995; Blossfeld & Rohwer 2002) das geeignete Instrumentarium dar. Im Folgenden werden ereignisanalytische Parameterschätzungen dargestellt, die sich auf ein erstmals von Klein (1995, 2003) vorgeschlagenes Verlaufmodell stützen. Es handelt sich dabei um eine Generalisierung des von Diekmann & Mitter (1983) vorgeschlagenen Sichelmodells, das die

¹¹ Beim ALLBUS bezogen auf die Erhebungen 1980, 1982, 1984, 1986, 1988, 1991 und 2000, in denen die Ehebiografie erfasst wurde, bei der Lebensverlaufsstudie bezogen auf die beiden Teilstichproben LV1 und LV2 und beim SOEP bezogen auf die Teilstichproben A, C, E, F und H sowie bis zum Jahr 2009. Aus dem SOEP unberücksichtigt bleiben die Samples B (ausländische Haushalte seit 1984), D (Zuwanderer-Privathaushalte seit 1994/95) und G (Hocheinkommens-Privathaushalte seit 2002). Von den im SOEP erfassten Ehen sind außerdem nur Ehen berücksichtigt, die im Panelzeitraum aufscheinen, bei denen beide Ehepartner eindeutig identifizierbar sind und die nicht erstmals im letzten Befragungsjahr bestehen.

¹² Abzüglich der Fälle mit fehlenden Werten bei den unabhängigen Variablen stehen für die Analysen zum Einfluss des Heiratsjahres, der Ordnungsnummer der Ehe und des Heiratsalters auf das Trennungsrisiko 42.256 Ehen zur Verfügung und 22.285 Ehen für die Analysen zum Einfluss des Altersabstands auf das Trennungsrisiko. Für eine Beschreibung der Mittelwerte bzw. Anteilswerte der unabhängigen Variablen in den Originaldatensätzen und im kumulierten Datensatz siehe Rapp (2013: 110ff.).

Trennungsrate in Abhängigkeit von der Ehedauer in der folgenden Form spezifiziert:

$$h(t) = \exp(a + bt + clnt + \beta_1 x_1 + \dots + \beta_k x_k).$$

Dabei repräsentieren die Koeffizienten die Einflüsse von erklärenden Variablen, zum Beispiel des Heiratsalters, auf die Trennungsrate. Unter Verwendung der Methode des Episodensplittings (Blossfeld & Rohwer 2002) werden alle Ehen in einjährige Subepisoden aufgesplittet. Durch die Aufnahme eines linearen sowie eines logarithmierten Terms für die Ehedauer wird der erst ansteigende und dann wieder abfallende Zeitverlauf der Trennungsrate modelliert. Das Modell erlaubt eine sehr flexible Anpassung an den tatsächlichen Verlauf des Trennungsrisikos und wurde bereits mehrfach zur Analyse ehelicher Stabilität herangezogen (Klein 1995, 1999; Klein & Rapp 2010; Klein & Stauder 1999; Lois 2008; Rapp 2008, 2012, 2013; Stauder 2002). Schließlich wird für die Ehen aus der Mannheimer Scheidungsstudie eine Design-Gewichtung vorgenommen (vgl. Hartmann 1997), da es sich um eine disproportional (nach verheiratet und bereits geschieden) geschichtete Zufallsstichprobe handelt.

3.3 Ergebnisse

Modellspezifikation und Drittvariablenkontrolle: Wie groß ist der Einfluss von Heiratsalter, Heiratsjahr und Folgeehe auf das Trennungsrisiko?

Aus Tabelle 1 geht hervor, welche erheblichen Einfluss unterschiedliche Modellspezifikationen auf die Einzeleffekte haben können. In Spalte 1 von Tabelle 1 ist neben den Variablen zur Ehedauer, welche die sichelförmige Prozesszeitabhängigkeit der Trennungsrate modellieren, nur der Effekt des Heiratsalters enthalten. Nach den Ergebnissen unserer Berechnung fällt die Trennungsrate mit jedem Altersjahr, um das die Eheschließung verschoben wird, um den Faktor 0,988 bzw. um $(0,988 * 100 - 100 =)$ 1,2 Prozent niedriger aus (vgl. Tabelle 1). Der Effekt ist aber nicht signifikant und deutlich geringer als das Ergebnis der Metaanalyse von Wagner und Weiß, das $-4,3$ % beträgt.

Aus Spalte 2 von Tabelle 1 geht hervor, dass Folgeehen ein höheres Trennungsrisiko haben als Erstehen.¹³ Berücksichtigt man zusätzlich (in Spalte 3

¹³ In unserer bivariaten Betrachtung haben Zweitehen im Vergleich zu Erstehen ein um 93,6 % höheres – d. h. fast doppelt so hohes – Trennungsrisiko. Die Metaanalyse von Wagner und Weiß zeigt einen wesentlich geringeren Wert, vermutlich deshalb, weil in den zugrundeliegenden Koeffizienten einige vermittelnde Faktoren wie zum Beispiel die

Tabelle 1 Effekte des Heiratsalters, der Ordnungsnummer der Ehe und des Heiratsjahres auf das Trennungsrisiko (relative Risiken, Standardfehler in Klammern)

Parameter	Modell 1	Modell 2	Modell 3	Modell 4	Modell 5	Modell 6
Ehedauer (b-Parameter)	0,902** (0,004)	0,903** (0,004)	0,903** (0,004)	0,920** (0,003)	0,920** (0,003)	0,921** (0,003)
<i>ln</i> Ehedauer (c-Parameter)	1,862** (0,140)	1,863** (0,138)	1,859** (0,134)	1,837** (0,140)	1,823** (0,136)	1,817** (0,131)
Heiratsalter ¹	0,988 (0,013)		0,962** (0,011)		0,971** (0,005)	0,950** (0,005)
Folgeehe ²		1,936** (0,154)	2,881** (0,136)			2,470** (0,265)
Heiratsjahr ¹				1,032** (0,003)	1,034** (0,004)	1,033** (0,003)
a-Parameter	-5,1**	-5,2**	-5,2**	-5,5**	-5,5**	-5,6**
Ereignisse	6579	6579	6579	6579	6579	6579
Episoden	775705	775705	775705	775705	775705	775705
Log-Likelihood	-28780,5	-28694,7	-28599,1	-28231,7	-28161,7	-28025,9

Signifikanzlimits: ** $p < 0,01$; * $p < 0,05$; + $p < 0,10$

Anmerkung: Es wurden robuste Standardfehler berechnet, die ein mögliches Clustering auf Ebene der Einzelstudien berücksichtigen.

¹ Mittelwertzentriert. Die Mittelwerte betragen 24,4 Jahre (Heiratsalter) bzw. 1961,3 (Heiratsjahr).

² Dummy-Variable, die bei Vorliegen der genannten Ausprägung mit 1, ansonsten mit 0 kodiert ist.

Quelle: kumulierter Datensatz, eigene Berechnung

von Tabelle 1), dass das Heiratsalter und die Wahrscheinlichkeit einer Zweitehe eng zusammenhängen, und hält man diesen Einfluss konstant, dann erlangt der Effekt des Heiratsalters statistische Signifikanz und vergrößert sich auf $(0,962 \cdot 100 - 100) = -3,8\%$, fällt also mehr als dreimal so groß aus!¹⁴

Die Vernachlässigung eines weiteren Faktors ist ebenfalls geeignet, den Einfluss des Heiratsalters zu unterschätzen. Zahlreiche Studien zeigen einen Trend, demzufolge nachfolgende Eheschließungsjahrgänge instabiler sind als die Vorausgegangenen.¹⁵ Der stabilisierende Einfluss des heutzutage höheren Heiratsalters wird ebenfalls unterschätzt, wenn man vernachlässigt, dass diese Ehen durch den genannten Trend stärker belastet werden (Spalte 5 von Tabelle 1).

Werden gar beide Faktoren – Erst- versus Zweitehe und Eheschließungsjahrgang – konstant gehalten,

dann vergrößert sich der Effekt des Heiratsalters auf 5 %. Während also der bivariate Zusammenhang zwischen Heiratsalter und Trennungsrisiko eher moderat ausfällt, unterscheidet sich der „wahre“ Kausaleffekt schon bei der Einführung zweier zusätzlicher Konstrukte beträchtlich, wobei jedoch zu bedenken ist, dass bei Kontrolle weiterer Variablen dieser Effekt auch wieder sinken kann. Dieses Beispiel illustriert jedoch in eindrucksvoller Weise, dass die Effekte aus verschiedenen spezifizierten Modellen nicht ohne Weiteres miteinander verrechnet werden können, sondern dass hier eine alternative Vorgehensweise angemessen ist.

Zur Heterogenität zwischen verschiedenen Studien

Die Analysen von Wagner & Weiß (2003) enthalten vor dem Hintergrund der eingangs diskutierten Rahmenbedingungen für eine traditionelle Metaanalyse in der Soziologie keine Aussagen zur Heterogenität zwischen den verschiedenen Studien. Es ist jedoch ein zentrales Anliegen der empirischen Sozialforschung zu erfahren, ob und inwieweit verschiedene Großstudien in Bezug auf elementare Grundaussagen zu möglicherweise recht unterschiedlichen Resultaten kommen. Der Heterogenitätsfrage wird hier in Bezug auf die oben betrachteten Einflussgrößen Heiratsalter, Heiratsjahr und Folgeehe nachgegangen. Tabelle 2 gibt die Ergebnisse der entsprechenden Analysen wieder.

in Zweitehen wesentlich höhere Anzahl an Stiefelternschaftskonstellationen konstant gehalten sind.

¹⁴ Best & Wolf (2012) zeigen allerdings in einer Simulationsstudie für Logit- und Probit-Regressionen, dass ein Vergleich von Koeffizienten zwischen verschachtelten Modellen problematisch sein kann. Möglicherweise besteht diese Problematik auch bei Übergangsratenmodellen.

¹⁵ Mitverantwortlich hierfür ist eine Eigendynamik, die man auch als Scheidungsspirale bezeichnet; vgl. zum Überblick Hill & Kopp 2006: 292ff.; Klein 2005: 217f.

Tabelle 2 Effekte des Heiratsalters, der Ordnungsnummer der Ehe, des Heiratsjahres und der Datenquelle auf das Trennungsrisiko (relative Risiken, Standardfehler in Klammern)

Parameter	Modell 1	Modell 2	Modell 3
Ehedauer (b-Parameter)	0,919** (0,001)	0,918** (0,004)	0,918** (0,004)
<i>ln</i> Ehedauer (c-Parameter)	1,815** (0,130)	1,832** (0,127)	1,841** (0,139)
Heiratsalter ¹	0,948** (0,004)	0,956** (0,000)	0,940** (0,011)
Folgeehe ²	2,416** (0,291)	1,847** (0,013)	2,254** (0,149)
Heiratsjahr ¹	1,032** (0,004)	1,032** (0,005)	1,033** (0,005)
ALLBUS ²	0,827+ (0,088)	0,789 (0,118)	0,799+ (0,107)
Generations and Gender Survey ²	0,843** (0,004)	0,869 (0,095)	0,859+ (0,078)
Lebensverlaufsstudie ²	0,941 (0,121)	0,700* (0,111)	0,719* (0,102)
Mannheimer Scheidungsstudie ²	0,682** (0,055)	0,553** (0,071)	0,550** (0,061)
Referenzkategorie: SOEP	1	1	1
Heiratsalter* ALLBUS		0,985** (0,000)	0,999 (0,009)
Heiratsalter* Generations and Gender Survey		0,996** (0,000)	0,999 (0,002)
Heiratsalter* Lebensverlaufsstudie		0,989** (0,001)	1,009 (0,012)
Heiratsalter* Mannheimer Scheidungsstudie		0,998** (0,001)	1,006 (0,006)
Folgeehe* ALLBUS		1,300** (0,005)	1,104* (0,055)
Folgeehe* Generations and Gender Survey		1,704** (0,004)	1,636** (0,017)
Folgeehe* Lebensverlaufsstudie		1,855** (0,002)	1,445** (0,138)
Folgeehe* Mannheimer Scheidungsstudie		1,340** (0,004)	1,214** (0,040)
Heiratsjahr* ALLBUS		1,001 (0,003)	1,001 (0,003)
Heiratsjahr* Generations and Gender Survey		0,994* (0,003)	0,995* (0,002)
Heiratsjahr* Lebensverlaufsstudie		0,966** (0,003)	0,967** (0,002)
Heiratsjahr* Mannheimer Scheidungsstudie		1,016** (0,003)	1,017** (0,002)
Folgeehe* Heiratsjahr			0,993** (0,003)
Heiratsalter* Heiratsjahr			1,001 (0,000)
a-Parameter	-5,4**	-5,4**	-5,4**
Ereignisse	6579	6579	6579
Episoden	775705	775705	775705
Log-Likelihood	-28000,3	-27952,8	-27949,5

Signifikanzlimits: ** $p < 0,01$; * $p < 0,05$; + $p < 0,10$

Anmerkung: Es wurden robuste Standardfehler berechnet, die ein mögliches Clustering auf Ebene der Einzelstudien berücksichtigen.

¹ Mittelwertzentriert. Die Mittelwerte betragen 24,4 Jahre (Heiratsalter) bzw. 1961,3 (Heiratsjahr).

² Dummy-Variable, die bei Vorliegen der genannten Ausprägung mit 1, ansonsten mit 0 kodiert ist.

Quelle: kumulierter Datensatz, eigene Berechnung

Im Modell 1 in Tabelle 2 ist deutlich zu sehen, dass sich – auch unter Kontrolle des Heiratsalters, der Frage, ob es sich um die erste Ehe handelt, und vor allem des Heiratsjahres und der damit verbundenen historischen Effekte – deutliche Unterschiede zwischen den einzelnen Studien zeigen. Vor allem in der Mannheimer Scheidungsstudie, aber auch im GGS sowie tendenziell im ALLBUS findet sich eine geringere Trennungsneigung als im Sozioökonomischen Panel. Ob dies eventuell auf die unterschiedliche Art der Datenerhebung – retrospektive

Befragungen einerseits und zumindest teilweise Erhebungen im Rahmen einer Panelbefragung andererseits – zurückzuführen ist, kann hier nicht geprüft werden.

Bedeutsamer als diese eher deskriptiven Befunde sind jedoch Unterschiede zwischen den einzelnen Untersuchungen hinsichtlich der hier untersuchten Wirkmechanismen. Hierzu wurden im Modell 2 in Tabelle 2 Interaktionseffekte zwischen den Studien und dem Heiratsalter, der Frage nach einer Folgeehe sowie dem Heiratsjahr aufgenommen. Dabei

sind die negativen Effekte des Heiratsalters auf das Trennungsrisiko in allen Studien signifikant stärker als im Sozioökonomischen Panel. Aber auch im SOEP ist der Effekt des Heiratsalters auf das Trennungsrisiko signifikant negativ, d. h. die Grundaussage ist in allen Surveys dieselbe. Der (positive) Effekt von Zweitehen ist ebenfalls in allen Studien stärker als im Sozioökonomischen Panel, wobei aber auch in diesem Fall alle fünf betrachteten Studien zu derselben Grundaussage gelangen. Der Kohorteneffekt, also der Einfluss des Heiratsjahres, differenziert sich ebenfalls deutlich zwischen den Studien. So findet sich beispielsweise in der Lebensverlaufsstudie kein Effekt des Heiratsalters (der Haupt- und der Interaktionseffekt nivellieren sich). Tabelle 2 zeigt somit, dass sich inhaltliche Interpretationen je nach verwendeter Datengrundlage systematisch unterscheiden können. Die gleichberechtigte Verrechnung aller Effekte in einer Metaanalyse kann diese Problematik nicht berücksichtigen und liefert somit ein eventuell verzerrtes Bild der wirklichen Kausalmuster.

Zur Aufklärung von Heterogenität zwischen den Studien wird in Modell 3 von Tabelle 2 untersucht, ob unterschiedliche Effekte in den betrachteten Studien damit zusammenhängen, dass die Studien teilweise unterschiedliche Kohorten repräsentieren, während sich gleichzeitig die Wirkung der hier untersuchten Einflussfaktoren mit der Zeit verändert hat. Zu diesem Zweck sind in Modell 3 zusätzlich Interaktionseffekte zwischen Heiratsalter bzw. Folgehe und dem Heiratsjahr einbezogen. Als Ergebnis zeigt sich, dass der destabilisierende Effekt von Folgeehen in jüngeren Heiratskohorten schwächer geworden ist. Gleichzeitig fallen bei Berücksichtigung dieser Interaktion die Unterschiede des Folgeheeffekts zwischen den Studien geringer aus. Dies bedeutet, dass der schwächere Folgeheeffekt im Sozioökonomischen Panel teilweise darauf zurückzuführen ist, dass das SOEP vergleichsweise jüngere Heiratskohorten abbildet, bei denen Folgeehen das Trennungsrisiko nicht mehr im selben Maße erhöhen wie dies noch bei den älteren Heiratskohorten der Fall war.

Vertiefte Analyse: Hat die Altershomogamie wirklich keinerlei Einfluss?

Eines der überraschenden Ergebnisse der Metaanalyse von Wagner & Weiß (2003: 41ff.) ist, dass die theoretisch erwarteten stabilisierenden Effekte der Altershomogamie statistisch nicht bedeutsam sind: Keine der Variablen zur Erfassung sozialer Homogamie „steht mit dem Ehescheidungsrisiko in einem statistisch signifikanten Zusammenhang“

(Wagner & Weiß 2003: 44). Nun ist es gerade der Reiz empirischer Analysen, auch Ergebnisse zutage zu fördern, die den ersten Erwartungen eventuell widersprechen. In diesem Abschnitt soll allerdings überprüft werden, ob dieses überraschende Ergebnis belastbar oder vielleicht nur Folge verschiedener Operationalisierungen und (relativ) kleiner Fallzahlen ist. Der für die Metaanalyse mit Originaldaten zur Verfügung stehende aggregierte Datensatz bietet die Möglichkeit, tiefergehende Analysen durchzuführen und die Alterskonfiguration der Partnerschaft genauer zu differenzieren. Als Bezugsgröße dienen dabei Paare, bei denen beide Partner einen Altersunterschied von höchstens einem Jahr aufweisen. Im Vergleich dazu werden einerseits Paare betrachtet, bei denen die Frau zwischen 2 und 4 oder mehr als 4 Jahre älter ist, andererseits Paare, bei denen der Mann zwischen 2 und 4, zwischen 5 und 7, zwischen 8 und 10 oder mehr als 10 Jahre älter ist als die Frau, wobei wiederum die schon oben erwähnten Variablen kontrolliert wurden. In den ersten vier Modellen von Tabelle 3 werden für jede Studie gesonderte Analysen angestellt (der ALLBUS kann hier nicht berücksichtigt werden, weil hier keine Informationen zum Altersabstand vorliegen), wohingegen sich das letzte Modell auf den kumulierten Datensatz bezieht.

Die Ergebnisse zum Einfluss des Altersabstandes auf das Trennungsrisiko fallen uneinheitlich aus, je nachdem, ob Ehen aus dem SOEP (Modell 1), aus dem GGS (Modell 2), aus der Lebensverlaufsstudie (Modell 3) oder aus der Mannheimer Scheidungsstudie (Modell 4) analysiert werden. Im Vergleich zu Paaren, die (plus oder minus 1 Jahr) etwa gleich alt sind, geht ein großer Altersabstand von Frauen von einem jüngeren Partner (von mindestens 5 Jahren) nur in drei der vier Studien mit einem signifikant höheren Trennungsrisiko einher. In der Lebensverlaufsstudie ist dies nicht der Fall, was möglicherweise auf die vergleichsweise geringen Fall- und Ereigniszahlen zurückzuführen ist (siehe die zweit- und drittunterste Zeile in Tabelle 3). Ist der Mann wesentlich älter als seine Frau (um mehr als 10 Jahre), zeigt sich nur in zwei von vier Studien ein signifikant höheres Trennungsrisiko als bei etwa gleich alten Paaren. Für kleinere Altersabstände „zugunsten“ der Frau zeigt sich in keiner der Studien ein signifikanter Effekt, für kleinere Altersabstände „zugunsten“ des Mannes ist dies nur bei der Mannheimer Scheidungsstudie der Fall.

Ein anderes Bild ergibt sich bei einer Gesamtbetrachtung aller Ehen aus allen Studien (Modell 5). Die Ergebnisse fallen nun viel klarer aus: Altershomogame Paare und Paare mit einem 2 bis 4 Jahre

Tabelle 3 Effekte des Altersabstands und weitere Determinanten des Trennungsrisikos in den Einzelstudien und im kumulierten Datensatz (relative Risiken, Standardfehler in Klammern)

Parameter	SOEP	GG5	LV2	MS	alle
Ehedauer (b-Parameter)	0,910** (0,009)	0,910** (0,008)	0,920** (0,015)	0,940** (0,010)	0,916** (0,007)
<i>ln</i> Ehedauer (c-Parameter)	1,660** (0,176)	2,410** (0,205)	1,890** (0,309)	2,070** (0,193)	2,129** (0,161)
Heiratsalter (Durchschnitt von Mann und Frau) ¹	0,940** (0,007)	0,950** (0,008)	0,940* (0,023)	0,920** (0,011)	0,942** (0,004)
Folgeehe ²	1,890** (0,190)	2,820** (0,336)	2,710** (0,523)	2,980** (0,452)	2,442 (0,311)
Heiratsjahr ¹	1,030** (0,005)	1,030** (0,003)	0,980 (0,021)	1,060** (0,004)	1,033** (0,011)
Frau ist mehr als 4 Jahre älter ²	1,500** (0,278)	1,580** (0,285)	1,800 ⁺ (0,566)	1,900* (0,473)	1,642** (0,097)
Frau ist 2 bis 4 Jahre älter ²	1,120 (0,160)	1,140 (0,163)	1,250 (0,351)	1,230 (0,198)	1,178** (0,024)
Frau ist 1 Jahr älter bis 1 Jahr jünger (Referenzkategorie)	1	1	1	1	1
Mann ist 2 bis 4 Jahre älter ²	1,110 (0,100)	1,130 (0,089)	0,740 (0,149)	0,990 (0,088)	1,065 (0,046)
Mann ist 5 bis 7 Jahre älter ²	1,030 (0,116)	1,090 (0,108)	0,930 (0,188)	1,250 ⁺ (0,142)	1,113* (0,051)
Mann ist 8 bis 10 Jahre älter ²	1,260 (0,195)	1,020 (0,160)	1,140 (0,274)	1,840** (0,283)	1,312* (0,181)
Mann ist mehr als 10 Jahre älter ²	1,700** (0,282)	1,670** (0,274)	1,310 (0,393)	1,470 (0,372)	1,600** (0,067)
a-Parameter	-4,8**	-6,1**	-5,9**	-6,4**	-5,8**
Ereignisse	836	1014	231	2278	4359
Episoden	88401	135158	45732	78782	348073
Log-Likelihood	-4358,4	-5744,4	-1406,9	-3316,4	-14936,0

Signifikanzlimits: ** p < 0,01; * p < 0,05; + p < 0,10

Anmerkung: In Modell 5 wurden robuste Standardfehler berechnet, die ein mögliches Clustering auf Ebene der Einzelstudien berücksichtigen.

¹ Mittelwertzentriert. Die Mittelwerte betragen 25,6 Jahre (Heiratsalter) bzw. 1968,0 (Heiratsjahr).

² Dummy-Variable, die bei Vorliegen der genannten Ausprägung mit 1, ansonsten mit 0 kodiert ist.

Quelle: kumulierter Datensatz, eigene Berechnung

älteren Mann weisen im Vergleich zu allen anderen Konstellationen deutlich niedrigere Trennungsrisiken auf, während größere Altersunterschiede in beide Richtungen mit einer Erhöhung des Trennungsrisikos verbunden sind.

4. Fazit

Metaanalysen haben sich als Verfahren der Forschungssynthese in den letzten vier Jahrzehnten auf breiter Front etabliert und werden vielfach – besonders im Bereich der medizinischen und pharmakologischen Forschung, aber auch in der Psychologie

und teilweise der Pädagogik – als eine Art von Goldstandard betrachtet (vgl. Goldacre 2012). Andere Verfahren zur Synopse und Synthese der immer unüberschaubarer werdenden Forschungslandschaft wie beispielsweise das „vote counting“ oder rein narrative Essays weisen vergleichsweise große Probleme auf und stellen keine ernsthafte Alternative für eine methodisch sauber durchgeführte Metaanalyse dar. Vor diesem Hintergrund erstaunt es auf den ersten Blick, dass Metaanalysen in der Soziologie ein relativ selten angewandtes Verfahren sind und beispielsweise die verschiedenen Arbeiten von Wagner & Weiß (2003, 2006a, 2006b) in der Forschungslandschaft relativ folgenlos geblieben

sind. In zwei Hinsichten können Argumente geltend gemacht werden, die den bislang äußerst seltenen Einsatz der Metaanalyse in der Soziologie verständlich erscheinen lassen. Hierzu muss einerseits die Logik der bisher dominierenden Verfahren in der Metaanalyse noch einmal genauer betrachtet und andererseits auf die Logik der Datenanalyse innerhalb der Soziologie eingegangen werden. Beide Argumentationsstränge führen zu einer relativ strukturgleichen Antwort und einem Plädoyer für eine Metaanalyse mit Originaldaten (MAMO). Diese bleibt jedoch bei ihrer Anwendung in der Soziologie ebenfalls nicht ohne Probleme.

Betrachtet man zuerst die Diskussionen innerhalb der stark anwachsenden Literatur zur Metaanalyse, die vor allem in den verschiedenen Bereichen der „life sciences“ geführt wird, so werden in den letzten Jahren alternative Verfahren zu der lange Zeit üblichen „meta-analysis of the literature (MAL)“ (Stewart & Parmar 1993: 418) diskutiert. Als in vielerlei Hinsicht zu bevorzugendes Verfahren wird die „meta-analysis of individual patient data (MAP)“ (Stewart & Riley 1993: 418) betrachtet (vgl. Stewart & Tierney 2002; Simmonds et al. 2005; Riley et al. 2010; Broeze et al. 2010; Kivimäki & Kawachi 2013). Während eine herkömmliche MAL die publizierten Ergebnisse randomisierter, kontrollierter Studien zusammenfasst, werden in der MAP die entsprechenden Originaldaten noch einmal aufbereitet und gemeinsam statistisch analysiert. Dieses Vorgehen hat – wie in entsprechenden Simulationen (vgl. Stewart & Parmar 1993) gezeigt wurde – vielfältige Vorteile: So können über alle (Teil-)Datensätze hinweg konsistente Ein- und Ausschlusskriterien angewendet sowie fehlende Werte gleichartig behandelt werden. „The statistical analyses can be standardised across studies (...) and more appropriate or advanced methods can be applied where necessary“ (Riley et al. 2010: 3). Auch ethische Fragen können einheitlich angegangen und eventuell gelöst werden. Wie bereits Stewart & Parmar (1993: 419f.) zeigen, liefern MAP zudem die besseren statistischen Schätzer – ein Vorteil, der fast nicht zu überschätzen ist. Als Problem des MAP-Ansatzes werden die relativ hohen Kosten eines derartigen Vorgehens genannt (Riley et al. 2010: 4f.): „This approach is resource intensive, because substantial time and costs are required to contact study authors, obtain their individual participant data, input and ‚clean‘ the provided individual data, resolve any data issues through dialog with the data providers, and generate a consistent data format across studies.“

Wenn man diese Überlegungen und Entwicklungen nun auf die soziologische Forschung übertragen will, so sind vor allem zwei Beobachtungen bedeutsam, aus denen drei Schlussfolgerungen gezogen werden können. Die erste Beobachtung ist, dass die Ausgangslage aller Metaanalysen in der Regel randomisierte Experimente sein sollten. In einem experimentellen Forschungsdesign werden durch die Randomisierung störende Drittvariablen kontrolliert. Die unterschiedlichen Effekte verschiedener Experimente lassen sich deshalb mit Hilfe der in der Metaanalyse entwickelten Verfahren und Schätzmethoden synthetisieren. Soziologische Datenanalysen können hingegen in aller Regel nicht auf Experimente zurückgreifen, sondern beruhen meist auf einem Ex-post-facto-Design. In diesem Fall ist es notwendig – vor allem mit Hilfe so genannter Mediations- und Suppressionsanalysen – die nicht im Mittelpunkt stehenden alternativen Einflussfaktoren auf statistischem Wege zu kontrollieren. In der Folge dieser Analysestrategie berichten die entsprechenden Forschungsarbeiten meist eine Fülle einzelner Effektstärken, denen theoretisch jedoch häufig ein deutlich unterschiedliches Gewicht zukommt. So werden in einem ersten Brutto-Modell (vgl. Kopp & Lois 2012) oft bivariate Koeffizienten berichtet, die niemals den wahren Einfluss des betreffenden Faktors repräsentieren. Wenn man – wie in Metaanalysen üblich – alle empirischen (Regressions-) Gewichte gleichgewichtig behandelt, verstößt man gegen diese Grundlogik soziologischer Datenanalyse. In den wenigen in der Soziologie zur Anwendung gekommenen Metaanalysen werden diese Probleme durchaus gesehen, ebenso in der entsprechenden Literatur zur Methodik der Metaanalyse, aufgrund ihrer praktischen Unlösbarkeit jedoch meist schlicht ignoriert.

Aus dieser Beobachtung folgt, dass eine auf publizierten Regressionskoeffizienten beruhende Metaanalyse in der Soziologie (nahezu) keinen Sinn macht.¹⁶ Hier werden häufig selbst von den Autoren als unzureichend eingestufte Ausgangsmodelle mit theoretisch elaborierten Modellierungen verglichen, die die Realität besser widerspiegeln sollen. Diese Modelle gleichwertig zu behandeln, erscheint nicht zielführend.

Eine zweite Beobachtung für die soziologische empirische Forschung ist jedoch, dass immer mehr auf die gleichen Datenbestände wie beispielsweise das

¹⁶ Einen sicherlich unbestreitbaren Verdienst derartiger Metaanalysen stellt jedoch die systematische Suche nach entsprechenden Arbeiten dar, die meist schon an sich einen Wert darstellt.

Sozio-oekonomische Panel zurückgegriffen wird (vgl. Kopp et al. 2012). Auch in der als Ausgangspunkt der Diskussion dienenden Metaanalyse von Wagner & Weiß (2003) werden letztlich nur 7 Datensätze verwendet. Aus dieser Beobachtung folgt dann zweitens, dass das Hauptproblem einer auf die Originaldaten zurückgreifenden Metaanalyse für die Soziologie selten von Bedeutung ist. Die verwendeten Datenbestände sind in aller Regel gut zugänglich und dokumentiert. Stellt die Metaanalyse mit Originaldaten also möglicherweise zumindest in der Forschungssynthese einen neuen Königsweg dar? In den in Abschnitt 3 dargestellten Analysen zum Trennungsrisiko von Ehen wird deutlich, dass derartige Metaanalysen in der Soziologie durchaus möglich sind. Anhand von fünf Datensätzen, auf denen die meisten in Deutschland publizierten Untersuchungen zu sozialstrukturellen Determinanten der ehelichen Stabilität und der Ehescheidung beruhen, wurde beispielhaft untersucht, welche Möglichkeiten, aber auch welche Einschränkungen und Probleme eine Harmonisierung und Kumulation dieser Datenbestände und somit eine Metaanalyse mit Originaldaten als Verfahren der Forschungssynthese mit sich bringen. Hierbei ist zuerst festzuhalten, dass diese Synthese durch die Metaanalyse mit Originaldaten ein gangbarer Weg ist. Allerdings ist ein derartiges Vorgehen mit einem großen Aufwand bei der Harmonisierung der Datenbestände verbunden. Des Weiteren sind der Aufklärung von Heterogenität zwischen den Studien natürlich auch bei einer Analyse der Originaldaten Grenzen gesetzt, da sich auch mit Originaldaten kaum jemals alle relevanten Unterschiede zwischen Studien kontrollieren lassen. Als wesentlich bedeutsamer mag sich allerdings erweisen, dass Metaanalysen mit Originaldaten immer die Schwächen der am wenigstens elaborierten Originalstudie teilen. Wie am Beispiel der Scheidungsforschung gezeigt wurde, sind vor allem viele eher demographische Faktoren nahezu gleichartig in allen Datenbeständen erhoben. Komplexere theoretische Überlegungen bedürfen jedoch immer auch komplexerer Operationalisierungen. Im Rahmen der Scheidungsforschung finden sich beispielsweise vor allem in der Mannheimer Scheidungsstudie die entsprechenden Operationalisierungen für die Modellierung der verschiedenen theoretischen Überlegungen aus der Familienökonomie oder der Austauschtheorie. Da diese Konstrukte – etwa die Alternativen zu oder Barrieren in einer Beziehung – bei den anderen Datenbeständen nicht eingesetzt werden können, stellt die Metaanalyse mit Originaldaten hier keine Alternative dar und man bleibt auf eine detaillierte Analyse des betreffenden

Datenbestands angewiesen. Bei theoretisch einfacheren Überlegungen erscheint jedoch das hier vorgestellte Verfahren der Metaanalyse mit Originaldaten den wenigen klassischen Metaanalysen als Instrument der Forschungssynthese deutlich überlegen, und man kann mit geringer Übertreibung festhalten, dass – wenn Glass (1976) die Metaanalyse auf Grundlage publizierter Ergebnisse als dritten Forschungsweg bezeichnet – die Metaanalyse mit Originaldaten nicht nur für die Soziologie als vierter Weg gelten kann. Aufgrund der genannten Einschränkungen ist es jedoch weiterhin für viele wichtige Fragestellungen wohl immer noch der erfolgversprechendere Weg, eigenständig theoretisch angeleitete Forschungen durchzuführen.

Literatur

- Allison, P.D., 1995: *Survival Analysis Using the SAS System: A Practical Guide*. Cary, NC: SAS Institute Inc.
- Amato, P.R., B. Keith, 1991: Parental Divorce and the Well-Being of Children: A Meta-Analysis. *Psychological Bulletin* 110: 26–46.
- Baur, N., 2008: Was kann die Soziologie methodisch von der Geschichtswissenschaft lernen? *Historical Social Research* 33: 217–248.
- Berlin, J.A., J. Santanna, C.H. Schmid, L.A. Szczech & H.I. Fledman, 2002: Individual Patient- versus Group-Level Data Meta-Regressions for the Investigation of Treatment Effect Modifiers: Ecological Bias Rears its Ugly Head. *Statistics in Medicine* 21: 371–387.
- Bleck, M. & M. Wagner, 2006: Stadt-Umland-Wanderung in Nordrhein-Westfalen – eine Meta-Analyse. *Raumordnung und Raumforschung* 64: 104–115.
- Blossfeld, H.-P. & G. Rohwer, 2002: *Techniques of Event History Modeling. New Approaches to Causal Analysis*. Second Edition. Mahwah: Erlbaum.
- Borenstein, M., L.V. Hedges, J.P.T. Higgins & H.R. Rothstein, 2009: *Introduction to Meta-Analysis*. Chichester: Wiley.
- Broeze, K.A., B.C. Opmeer, F. van der Veen, P.M. Bossuyt, S. Bhattacharya & B.W.J. Mol, 2010: Individual Patient Data Meta-Analysis: A Promising Approach for Evidence Synthesis in Reproductive Medicine. *Human Reproductive Update* 16: 561–567.
- Brüderl, J., 2004: Meta-Analyse in der Soziologie: Bilanz der deutschen Scheidungsursachenforschung oder „statistischer Fruchtsalat“? *Zeitschrift für Soziologie* 33: 84–86.
- De Solla-Price, D.J., 1974: *Little Science, Big Science. Von der Studierstube zur Großforschung*. Frankfurt a.M.: Suhrkamp.
- Diekmann, A. & P. Mitter, 1983: The “Sickle Hypothesis”. A Time Dependent Poisson Model with Applications to Deviant Behavior and Occupational Mobility. *Journal of Mathematical Sociology* 9: 85–101.
- Dinkel, R.H. & I. Milenovic, 2002: Die Kohortenfertilität von Männern und Frauen in der Bundesrepublik

- Deutschland. Eine Messung mit Daten der empirischen Sozialforschung. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 44: 55–75.
- Dölling, D., H. Entorf, D. Herrmann, A. Häring, T. Rupp & A. Woll, 2006: Zur generalpräventiven Abschreckungswirkung des Strafrechts – Befunde einer Metaanalyse. *Soziale Probleme* 17: 193–209.
- Dölling, D., H. Entorf, D. Herrmann & T. Rupp, 2009: Is Deterrence Effective? Results of a Meta-Analysis. *European Journal of Criminal Policy Research* 15: 201–224.
- Dölling, D. & D. Herrmann, 2012: Zur generalpräventiven Abschreckungswirkung des Strafrechts bei jungen Menschen. S. 423–435 in: Deutsche Vereinigung für Jugendgerichte und Jugendgerichtshilfen e.V. (Hrsg.), *Achtung (für) Jugend! Praxis und Perspektiven des Jugendkriminalrechts. Dokumentation des 28. Deutschen Jugendgerichtstages vom 11. bis 14. September 2010*, Mönchengladbach.
- Eisend, M., 2004: Metaanalyse – Einführung und kritische Diskussion. Diskussionsbeiträge des Fachbereichs Wirtschaftswissenschaft der Freien Universität Berlin Nr. 2004/8. Manuskript. Berlin.
- Fulkerson, G.M. & G.H. Thompson, 2008: The Evolution of a Contested Concept: A Meta-Analysis of Social Capital Definitions and Trends (1988–2006). *Sociological Inquiry* 78: 536–557.
- Glass, G.V., 1976: Primary, Secondary, and Meta-Analysis of Research. *Educational Researcher* 5: 3–8.
- Goldacre, B., 2012: *Bad Pharma. How Drug Companies Mislead Doctors and Harm Patients*. London: Forth Estate.
- Hartmann, J., 1997: Komplexes Stichprobendesign und Ereignisanalyse: Zur Notwendigkeit einer Gewichtung bei diproportional geschichteter Stichprobe. *MZES Arbeitspapiere/Working Papers Arbeitsbereich I / Nr. 17: 1–46*. Mannheim: MZES.
- Herrmann, D., 2010: Die Abschreckungswirkung der Todesstrafe – ein Artefakt der Forschung? S. 791–808 in: D. Bölling, B. Götting, B.-D. Meier, T. Verrel (Hrsg.), *Verbrechen – Strafe – Resozialisierung. Festschrift für Heinz Schöch*. Berlin & New York: de Gruyter.
- Hill, P.B. & J. Kopp, 2006: *Familiensoziologie. Grundlagen und theoretische Perspektiven*. 4., überarbeitete Auflage. Wiesbaden: VS.
- Hunt, M., 1997: *How Science Takes Stock. The Story of Meta-Analysis*. New York: Sage Foundation.
- Jose, A., D. O’Leary & A. Moyer, 2010: Does Marital Cohabitation Predict Subsequent Marital Stability and Marital Quality? A Meta-Analysis. *Journal of Marriage and Family* 72: 105–116.
- Kalmijn, M. & G. Kraaykamp, 2007: Social Stratification and Attitudes: A Comparative Analysis of the Effects of Class and Education in Europe. *British Journal of Sociology* 58: 547–576.
- Kivimäki, M., I. Kawachi, 2013: Need for more Individual-level Meta-Analyses in Social Epidemiology: Examples of Job Strain and Coronary Heart Disease. *American Journal of Epidemiology* 177:
- Klein, T., 1995: Scheidungsbetroffenheit im Lebensverlauf von Kindern. S. 253–263 in: B. Nauck & H. Bertram (Hrsg.), *Kinder in Deutschland. Lebensverhältnisse von Kindern im Regionalvergleich*. Opladen: Leske + Budrich.
- Klein, T., 1999: Der Einfluß vorehelichen Zusammenlebens auf die spätere Ehestabilität. S. 143–158 in: T. Klein & J. Kopp (Hrsg.), *Scheidungsursachen aus soziologischer Sicht*. Würzburg: Ergon.
- Klein, T., 2003: Die Geburt von Kindern in paarbezogener Perspektive. *Zeitschrift für Soziologie* 32: 506–527.
- Klein, T., 2005: *Sozialstrukturanalyse. Eine Einführung*. Reinbek: Rowohlt.
- Klein, T. & J. Stauder, 1999: Der Einfluß ehelicher Arbeitsteilung auf die Ehestabilität. S. 159–177 in: T. Klein & J. Kopp (Hrsg.), *Scheidungsursachen aus soziologischer Sicht*. Würzburg: Ergon.
- Klein, T. & I. Rapp, 2010: Der Einfluss des Auszugs von Kindern aus dem Elternhaus auf die Beziehungsstabilität der Eltern. *Zeitschrift für Soziologie* 39: 140–150.
- Köbel, M., 2001: Das Wachstum der Wissenschaft in Deutschland 1650–2000. S. 113–128 in: Heinrich Parthey, Günter Spur (Hrsg.), *Wissenschaft und Innovation. Wissenschaftsforschung Jahrbuch 2001*. Berlin: Gesellschaft für Wissenschaftsforschung.
- Kopp, J. & D. Lois, 2012: *Sozialwissenschaftliche Datenanalyse. Eine Einführung*. Wiesbaden: VS.
- Kopp, J., J. Schneider & F. Timmler, 2012: Zur Entwicklung soziologischer Forschung. *Soziologie* 41: 293–310
- Lipsey, M.W. & D.B. Wilson, 2001: *Practical Meta-Analysis*. Thousand Oaks: Sage.
- Lois, D., 2008: Arbeitsteilung, Berufsorientierung und Partnerschaftsstabilität – Ehen und nichteheliche Lebensgemeinschaften im Vergleich. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 60: 53–77.
- MacKinnon, D.P., 2008: *Introduction to Statistical Mediation Analysis*. Milton Park: Routledge.
- Mann, C.M., 1990: Meta-Analysis in the Breech. *Science* 249: 476–480.
- Mann, C.M., 1994: Can Meta-Analysis Make Policy? *Science* 266: 960–962.
- Mosteller, F. & G.A. Colditz, 1996: *Understanding Research Synthesis (Meta-Analysis)*. *Annual Review of Public Health* 17: 1–23.
- Rapp, I., 2008: Wann werden Ehen getrennt? Der Einfluss der Ehedauer auf das Trennungsrisiko. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 60: 500–527.
- Rapp, I., 2012: In Gesundheit und Krankheit? Der Zusammenhang zwischen dem Gesundheitszustand und der Ehestabilität. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 64: 783–803.
- Rapp, I., 2013: *Ehestabilität in der zweiten Lebenshälfte. Eine Analyse von kumulierten sozialwissenschaftlichen Umfragedaten*. Wiesbaden: VS.
- Riley, R.D., P.C. Lambert & G. Abo-Zaid, 2010: *Meta-Analysis of Individual Participant Data: Rationale, Conduct, and Reporting*. *BMJ* 340: c221:
- Rosenthal, R. & M.R. DiMatteo, 2001: *Meta-Analysis: Recent Developments in Quantitative Methods for Literature Review*. *Annual Review of Psychology* 52: 59–82.
- Ruddat, M. & M. Schulz, 2010: Möglichkeiten und Grenzen einer integrativen Metaanalyse in der Soziologie. *Historical Social Research* 35: 289–322.

- Schnell, R., P.B. Hill & E. Esser, 2011: Methoden der empirischen Sozialforschung. 9. Auflage. München: Oldenbourg.
- Shor, E., D.J. Roelfs, M. Curreli, L. Clemow, M.M. Burg & J.E. Schwartz, 2012: Widowhood and Mortality: A Meta-Analysis and Meta-Regression. *Demography* 49: 575–606.
- Simmonds, M.C., J.P.T. Higgins, L.A. Stewart, J.F. Tierney, M.C. Clarke & S.G. Thompson, 2005: Meta-Analysis of Individual Patient Data from Randomized Trials: A Review of Methods Using in Practice. *Clinical Trials* 2: 209–217.
- Stauder, J., 2002: Eheleiche Arbeitsteilung und Ehestabilität. Eine Untersuchung mit den Daten der Mannheimer Scheidungsstudie 1996 unter Verwendung ereignisanalytischer Verfahren. Würzburg: Ergon.
- Stewart, L.A. & M.K.B. Parmar, 1993: Meta-Analysis of the Literature or of Individual Patient Data: Is there a Difference? *Lancet* 341: 418–422.
- Stewart, L.A. & J.F. Tierney, 2002: To IPD or not to IPD? Advantages and Disadvantages of Systematic Reviews Using Individual Patient Data. *Evaluation & the Health Professions* 25: 76–97.
- Sutton, A. & J.P.T. Higgins, 2008: Recent Developments in Meta-Analysis. *Statistics in Medicine* 27: 625–650.
- Sutton, A.J., D. Kendrick & C.A.C. Coupland, 2008: Meta-analysis of Individual- and Aggregate-Level Data. *Statistics in Medicine* 27: 651–669.
- Wagner, M. & B. Weiß, 2003: Bilanz der Scheidungsforschung. Versuch einer Meta-Analyse. *Zeitschrift für Soziologie* 33: 29–49.
- Wagner, M. & B. Weiß, 2006a: Meta-Analyse als Methode der Sozialforschung. S. 479–504 in: A. Diekmann (Hrsg.), *Methoden der Sozialforschung. Sonderheft 44 der Kölner Zeitschrift für Soziologie und Sozialpsychologie*. Wiesbaden: VS.
- Wagner, M. & B. Weiß, 2006b: On the Variation of Divorce Risk in Europe: Findings from a Meta-Analysis of European Longitudinal Studies. *European Sociological Review* 22: 483–500.

Autorenvorstellung

Thomas Klein, geb. 1955 in Heidelberg. Professor für Soziologie an der Universität Heidelberg.

Forschungsschwerpunkte: Sozialstrukturanalyse, Bevölkerungssoziologie, Familiensoziologie, Soziologie des Alters und Soziologie der Gesundheit.

Wichtigste Publikationen: „Durch Dick und Dünn“. Zum Einfluss von Partnerschaft und Partnermarkt auf das Körpergewicht. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 63, 2011; Sozialstrukturanalyse. Eine Einführung. Reinbek 2005; Love and Death in Germany: The Marital Biography and its Effect on Mortality, *Journal of Marriage and the Family* 66, 2004 (mit H. Brockmann); Die Geburt von Kindern in paarbezogener Perspektive, *Zeitschrift für Soziologie* 32, 2003; Der berufliche Wiedereinstieg von Müttern zwischen abnehmendem Betreuungsaufwand und zunehmender Dequalifizierung, *Zeitschrift für Soziologie* 24, 1995 (mit U. Braun); Zur Abhängigkeit zwischen konkurrierenden Mortalitätsrisiken, *Allgemeines Statistisches Archiv* 72, 1988; A Specification Separating Family Size and Individual Age Effects on Subjective Equivalence Scales – A Note, *The Review of Income and Wealth* 34, 1988; zuletzt in dieser Zeitschrift: Der Einfluss des Auszugs von Kindern aus dem Elternhaus auf die Beziehungsstabilität der Eltern (mit I. Rapp), *ZfS* 39, 2010: 140–150.

Johannes Kopp, geb. 1961. Professor für empirische Sozialforschung am Institut für Soziologie der Technischen Universität Chemnitz.

Forschungsschwerpunkte: Methoden der empirischen Sozialforschung und Datenanalyse, Familiensoziologie.

Wichtigste Publikationen: Familiensoziologie (mit P.B. Hill), Fünfte Auflage. Wiesbaden 2013; Sozialwissenschaftliche Datenanalyse (mit D. Lois). Wiesbaden 2012; Zur Entwicklung soziologischer Forschung (mit J. Schneider und F. Timmler), *Soziologie* 41 (2012): 293–310; Grundbegriffe der Soziologie (Hrsg. mit B. Schäfers). 10. Auflage. Wiesbaden 2010; Verlobt, verlobt, verheiratet (mit D. Lois, C. Kunz, O. Arránz-Becker), Wiesbaden 2010; Bildungssoziologie, Wiesbaden 2009; zuletzt in dieser Zeitschrift: Verhandlungstheoretische Modellierung des Übergangs zu verschiedenen Kinderzahlen (mit A. Kohlmann), *ZfS* 26, 1997: 258–274.

Rapp, Ingmar, geb. 1980 in Biberach an der Riss. Dr. rer. pol., Wissenschaftlicher Mitarbeiter am Max-Weber-Institut für Soziologie an der Universität Heidelberg.

Forschungsschwerpunkte: Familiensoziologie, Soziologie der Gesundheit.

Wichtigste Publikationen: Ehestabilität in der zweiten Lebenshälfte. Eine Analyse von kumulierten sozialwissenschaftlichen Umfragedaten. Wiesbaden 2013; In Gesundheit und Krankheit? Der Zusammenhang zwischen dem Gesundheitszustand und der Ehestabilität, *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 64, 2012; Wann werden Ehen getrennt? Der Einfluss der Ehedauer auf das Trennungsrisiko, *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 60, 2008; zuletzt in dieser Zeitschrift: Der Einfluss des Auszugs von Kindern aus dem Elternhaus auf die Beziehungsstabilität der Eltern (mit T. Klein), *ZfS* 39, 2010: 140–150.