

Meta-Analyse und "publication bias": eine empirisch-statistische Fallstudie

Fiebig, Joachim; Rüdebusch, Gerrit; Urban, Dieter

Veröffentlichungsversion / Published Version

Forschungsbericht / research report

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

SSG Sozialwissenschaften, USB Köln

Empfohlene Zitierung / Suggested Citation:

Fiebig, J., Rüdebusch, G., & Urban, D. (2012). *Meta-Analyse und "publication bias": eine empirisch-statistische Fallstudie*. (Schriftenreihe des Instituts für Sozialwissenschaften der Universität Stuttgart -SISS-, 1/2012). Stuttgart: Universität Stuttgart, Fak. 10 Wirtschafts- und Sozialwissenschaften, Institut für Sozialwissenschaften. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-407931>

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.

SISS:

**Schriftenreihe des Instituts für Sozialwissenschaften
der Universität Stuttgart**

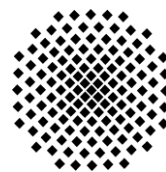
No. 1 / 2012

Meta-Analyse und "publication bias"

Eine empirisch-statistische Fallstudie

**Joachim Fiebig
Gerrit Rüdebusch
Dieter Urban**

**Universität Stuttgart
Institut für Sozialwissenschaften
Abteilung für Soziologie und
empirische Sozialforschung (SOWI-IV)
70174 Stuttgart**



SOWI

ISSN 0945-9197

**SISS:
Schriftenreihe
des Instituts für Sozialwissenschaften
der Universität Stuttgart: No. 1 / 2012**

Meta-Analyse und "publication bias".

Eine empirisch-statistische Fallstudie

Joachim Fiebig

Gerrit Rüdibusch

Dieter Urban

**Universität Stuttgart
Institut für Sozialwissenschaften
Abteilung für Soziologie und
empirische Sozialforschung (SOWI-IV)
70174 Stuttgart**

Joachim Fiebig, Gerrit Rüdebusch, Dieter Urban
Meta-Analyse und „publication bias“. Eine empirisch-statistische Fallstudie.

SISS–Schriftenreihe des Instituts für Sozialwissenschaften der Universität Stuttgart. No. 1/2012.
(<http://www.uni-stuttgart.de/soz/institut/forschung/2012.SISS.1.pdf>)

Z U S A M M E N F A S S U N G: Die Schrift berichtet über empirische Analysen zur Identifikation eines „publication bias“ in einer Meta-Analyse von 30 Studien zu den Effekten von pädosexueller Viktimisierung auf die Entstehung von pädosexueller Delinquenz. Zur Identifikation eines möglichen „publication bias“ werden verschiedene statistische Verfahren vorgestellt und erläutert (graphische und regressionsanalytische „funnel-plot“-Methodik sowie die „rank-correlation“-, „fail-safe numbers“- und „trim and fill“-Methodik). In den empirisch-statistischen Analysen kann kein Nachweis dafür gefunden werden, dass die Ergebnisse der Meta-Analyse durch einen „publication bias“ verfälscht wurden.

Joachim Fiebig, Gerrit Rüdebusch, Dieter Urban
Meta-Analysis and „Publication Bias“. An Empirical-Statistical Case Study.

SISS–Schriftenreihe des Instituts für Sozialwissenschaften der Universität Stuttgart. No. 1/2012.
(<http://www.uni-stuttgart.de/soz/institut/forschung/2012.SISS.1.pdf>)

A B S T R A C T: The paper reports about various empirical analyses to identify a publication bias in a meta-analysis of 30 studies investigating the effects of pedo-sexual victimization on pedo-sexual delinquency. The paper presents and explains several methods used to identify a publication bias in meta-analysis (methods of graphical and regression based funnel-plot-analysis, rank-correlation-, fail-safe numbers-, and trim and fill-analysis). No evidence can be found that the results of the meta-analysis were distorted by an existing publication bias.

1 Forschungskontext

Der vorliegende Report berichtet über empirisch-statistische Analysen, die im DFG-finanzierten Forschungsprojekt "Meta-Analyse zum Stand der empirischen Forschung über die psycho-sozialen Ursachen pädosexueller Delinquenz" (kurz: MAPS-Projekt) durchgeführt wurden. In diesem Projekt wurde vor allem die Gültigkeit der pädosexuellen Opfer-Täter-Transitionshypothese meta-analytisch untersucht. Nach dieser These steigt das Risiko, zum Täter pädosexuellen Missbrauchs zu werden, mit eigenen pädosexuellen Opfererfahrungen des Täters in dessen Kindheit oder Jugend deutlich an (vgl. dazu Urban/Fiebig 2010, 2011). Diese Transitionshypothese, die in den Sozial- und Sexualwissenschaften kontrovers diskutiert wird¹, sollte im MAPS-Projekt anhand von Meta-Analysen möglichst aller einschlägigen, weltweit in einem bestimmten Zeitraum erschienenen Primärstudien hinsichtlich ihrer empirischen und statistischen Gültigkeit untersucht werden.

Im MAPS-Projekt wurden durch automatisierte, suchstring-gesteuerte Recherchen in unterschiedlichsten Literatur-Datenbanken (u.a. in: SOWIPORT, PSYINDEX, SSCI, PsycINFO) und in zahlreichen, thematisch einschlägigen Fachzeitschriften (u.a. in: Child Abuse & Neglect, Journal of Family Violence etc.) für den Zeitraum von 1980 bis 2010 insgesamt ca. 64.000 Veröffentlichungen ermittelt, die thematisch im Gültigkeitsbereich der oben benannten Transitionsthese angesiedelt sind. Nach intensiver Auswertung der Abstracts dieser Veröffentlichungen und weiterer Publikationsinhalte reduzierte sich diese Zahl auf 122 Primärstudien, die thematisch für das MAPS-Projekt tatsächlich relevant waren.

Für die meta-analytischen Auswertungen, über die im Folgenden berichtet wird, konnten insgesamt 30 dieser 122 Forschungsveröffentlichungen benutzt werden. Diese Auswahl ergab sich vor allen Dingen daraus, dass nur in diesen 30 Studien ausreichend Informationen zur Verfügung gestellt werden, um Meta-Analysen durchführen zu können, deren Ergebnisse auf Verzerrungen (z.B. durch einen „publication bias“) getestet werden konnten. Allerdings ist eine Gesamtzahl von 30 Primärstudien im Vergleich zu anderen, inhaltlich und methodisch vergleichbaren Meta-Analysen im Bereich der pädosexuellen Transitionsforschung als überdurchschnittlich hoch einzuschätzen. Denn diese Meta-Studien verwendeten 7 bis maximal 19 Primärstudien.²

¹ Bestätigt wurde die Transitionsthese u.a. von Seghorn et al. 1987; Urban/Lindhorst 2003, 2004; Urban/Fiebig 2010; Dhaliwal et al. 1996. Widersprochen wurde der Transitionsthese u.a. von: Schneider 1999; Slater/McMillan et al. 2003; Hanson/Slater 1988; Freund et al. 1990.

² Vgl. Jespersen et al. 2009; Seto/Lalumiére 2010; Whitaker et al. 2008.

Wie bereits oben erwähnt, konzentrierte sich die MAPS-Studie auf Studien zum Zusammenhang zwischen pädosexuellen Viktimisierungserfahrungen und pädosexueller Delinquenz. Die diesen Zusammenhang postulierende Hypothese wird als pädosexuelle Opfer-Täter-Transitionshypothese (kurz: POTT-Hypothese) benannt.

Der Begriff „pädosexuell“ bezeichnet dabei sexuelle Interaktionen von Erwachsenen mit Kindern (nach StGB §176, sind das Personen bis 14 Jahre) im Unterschied zum Begriff „adultsexuell“, der sexuelle Interaktionen zwischen Erwachsenen und Personen über 13 Jahren belegt.

Für die statistischen Untersuchungen, bei welchen sowohl klassische Meta-Analyseverfahren als auch die Mehrebenenmodellierungen zur Anwendung kamen, wurden insgesamt 30 zwischen den Jahren 1980 und 2010 veröffentlichte Studien mit drei unterschiedlichen Kontrollgruppen herangezogen. Dabei bestand die Untersuchungsgruppe immer aus pädosexuellen Delinquenten, während die Kontrollgruppen adultsexuelle Straftäter (Kontrollgruppe 1), nichtsexuelle Gewaltstraftäter (Kontrollgruppe 2) oder Personen umfassten, die keine Straftaten begangen hatten (Kontrollgruppe 3).

Zentrales Ziel der Analysen war herauszufinden, ob sich die POTT-Hypothese im meta-analytischen Vergleich der Studienergebnisse bestätigen lässt, oder ob die Hypothese meta-analytisch zu verwerfen ist.

Um dies festzustellen, wurde zunächst über alle als themenrelevant ermittelten Primärstudien die mittlere Effektstärke des Einflusses von pädosexueller Viktimisierung auf pädosexuelle Delinquenz mittels Odds Ratio-Schätzung berechnet.

Die Ergebnisse der Analysen indizierten sehr eindeutig, dass die Hypothese unabhängig von den betrachteten Kontrollgruppen und unabhängig vom verwendeten Analyseverfahren nicht verworfen werden kann.

Die ermittelten Effektstärken schwankten zwischen Odds Ratios von 2,76 für Studien mit Kontrollgruppe 1 und 9,82 für Studien mit Vergleichsgruppe 3. Das bedeutet, dass eine pädosexuelle Viktimisierung unter pädosexuellen Straftätern deutlich häufiger zu beobachten ist als unter den Personen der jeweils betrachteten Vergleichsgruppe.³

Ein weiteres Ziel der Meta-Analysen (neben Prüfung der POTT-Hypothese) war es herauszufinden, inwiefern die Ergebnisse der einzelnen Studien aufgrund von unterschiedlichen

³ Der OR-Koeffizient lässt sich so interpretieren, dass Personen, die in ihrer Kindheit mindestens eine sexuelle Missbrauchserfahrung gemacht haben, im Unterschied zu Personen, die keine entsprechenden Erfahrungen gemacht haben, eine mindestens 2,7 bis maximal 9,8fach erhöhte Chance aufweisen, bei späterer Straffälligkeit ein pädosexuelles Delikt zu begehen.

Probanden- und Studienmerkmalen variieren.⁴ In den Meta-Analysen wurde mithin überprüft, ob bestimmte Studienmerkmale die Schwankungen der geschätzten Effektstärken zu einem signifikanten Ausmaß erklären können. Zur Erklärung wurden methodische und formale Merkmale der Primärstudien hinsichtlich ihrer Moderatorwirkung untersucht.

Als methodische Moderatorvariablen wurden die Operationalisierung der pädosexuellen Viktimisierung sowie der pädosexuellen Delinquenz (jeweils eng vs. weit), die Informationsquelle zur pädosexuellen Delinquenz (Selbst- vs. Fremdb bericht) und der Rekrutierungsort der Untersuchungsgruppe (Gefängnis vs. Therapie) benutzt. Als formale Moderatorvariablen wurden die Stichprobengröße, die Varianz der Effektstärke und die subjektiv eingeschätzte methodische Güte der Studien – allesamt (quasi)metrisch – getestet.

Bei allen durchgeführten bivariaten und multivariaten Analysen konnte zwar nur vereinzelt die Varianz der Effektstärke zwischen den Studien aufgeklärt werden, was bedeutet, dass keine eindeutige statistische Moderatorerklärung zu erreichen war. Viel wichtiger für den zentralen Befund war jedoch, dass die Signifikanz der mittleren Effektstärke für den in der POTT-Hypothese postulierten Zusammenhang von diesen Analysen nicht aufgehoben werden konnte.

Damit ließ sich unter Annahme einer unverzerrten Datenlage konstatieren, dass die POTT-Hypothese ungeachtet aller Moderator- und Kontrollvariablen und unabhängig von den betrachteten Primärstudien, nach „statistischem Ermessen“ auf Grundlage der durchgeführten Analysen nicht verworfen werden konnte.

Es stellte sich daher „nur“ noch die Frage, inwiefern sich die verwendete Datenlage als vertrauenswürdig, sprich als unverzerrt bezeichnen ließ. Eben diese Frage zu klären ist das Ziel der in den nachfolgenden Abschnitten erläuterten statistischen Verfahren und der dementsprechend durchgeführten Analysen.

Der vorliegende Forschungsbericht ist wie folgt gegliedert: Zunächst wird die Thematik des sogenannten „publication bias“ vorgestellt und auch erläutert, inwieweit sich dieser Bias in den Untersuchungen des MAPS-Projekt von den potenziellen Verzerrungen in anderen Meta-Analysen unterscheiden könnte (Abschnitt 2). Danach wird auf Diagnosemethoden eingegangen, die zur Identifikation des „publication bias“ eingesetzt werden können und die auch im MAPS-Projekt

⁴ Solche Merkmale seitens der Probanden können z.B. der Schweregrad der Missbrauchserfahrungen, die Häufigkeit dieser Erfahrungen oder das Alter sein, zu dem diese erlebt wurden. Weitere Variationen ergeben sich aufgrund methodischer Merkmale der einzelnen Studien, wie die Operationalisierung zentraler Variablen, z.B. der pädosexuellen Viktimisierung oder der pädosexuellen Delinquenz (die z.B. als „eng“, was ausschließlich Penetrationsdelikte/-erfahrungen meinen kann, oder als „weit“, womit jede Form pädosexueller Interaktion, also z.B. auch das Erleben oder Ausleben von Exhibitionismus, definiert werden kann) aber auch aufgrund formaler Merkmale wie die Stichprobengröße oder die statistische Güte der Studien.

angewendet wurden (Abschnitt 3). Daran anschließend werden die Ergebnisse, die mit diesen Diagnosemethoden im MAPS-Projekt erzielt wurden, vorgestellt (Abschnitt 4.1 bis 4.3). Eine Zusammenfassung der Ergebnisse unserer Analysen zum „publication bias“ im thematischen Kontext des MAPS-Projekts wird in Abschnitt 5 vorgetragen.

2 Zur Methodologie des „publication bias“

In der außerwissenschaftlichen Öffentlichkeit wird den Ergebnissen von Meta-Studien häufig eine höhere wissenschaftliche Glaubwürdigkeit zugesprochen als den Resultaten einzelner Primärstudien. Es wird davon ausgegangen, dass sauber durchgeführte Meta-Studien die Ergebnisse unterschiedlichster Primärstudien, die bisweilen mit ganz verschiedenen Erhebungs- und Auswertungstechniken sowie mit ganz unterschiedlichen Datenquellen operieren, zusammenführen und von ihren jeweiligen Besonderheiten bereinigen können, so dass in Meta-Studien der unverzerrte, allgemeingültige Ergebniskern vieler Primärstudien herausdestilliert und beschrieben werden könne. Eine solche meta-analytische Zusammenführung einzelner Studienergebnisse kann jedoch eine ganz eigene Problematik aufweisen: den sogenannten „publication bias“.

Kurz gefasst ist unter dem „publication bias“ eine Verzerrung der Ergebnisse von Meta-Analysen zu verstehen, die aufgrund einer systematischen Selektivität bei der Identifikation von Primärstudien entsteht und die durch soziale Normen des wissenschaftlichen Publizierens im Kontext der jeweiligen „scientific community“ hervorgebracht wird. Zu diesen sozialen Normen wissenschaftlichen Publizierens gehört z.B. die informelle soziale Regel, nach der solche Studien, die keine statistisch-signifikanten Ergebnisse erzielen, eher nicht publiziert werden sollten. Auch kann ein „publication bias“ in der meta-analytischen Forschung dadurch entstehen, dass bestimmte Primärstudien nicht in Literatur-Datenbanken aufgenommen werden und deshalb für Meta-Studien nicht zugänglich sind. Zu solchen „unsichtbaren“ Primärstudien gehören z.B. überhäufig nicht-englischsprachige Veröffentlichungen und Veröffentlichungen in eher unbekanntem Online-Fachzeitschriften. Auch sogenannte „graue Literatur“⁵ oder Forschungsergebnisse, die überhaupt nicht veröffentlicht wurden, lassen sich oftmals nur mit großem Aufwand, Vertrautheit mit dem Forschungsfeld und natürlich etwas Glück sowie detektivischem Gespür ausfindig machen.

Dass Studien, die signifikante Ergebnisse erzielen, viel häufiger publiziert werden als Studien mit Ergebnissen außerhalb der statistischen Signifikanz, wird in der meta-analytischen Forschung auch als „file drawer“ Problem bezeichnet. Damit ist Folgendes verbunden: Durch die statistische Definition von Signifikanz werden bei gleichbleibender Stichprobengröße die

⁵ Als „graue Literatur“ werden wissenschaftliche Texte bezeichnet, die nicht auf die klassische Weise publiziert wurden und somit evtl. nicht sehr weit verbreitet sind oder evtl. nicht einfach zu erhalten sind, wie z.B. Konferenzpapiere, Abschlussarbeiten oder unveröffentlichte Forschungsdaten (vgl. Hopewell et al. 2005: 56).

Ergebnisse eher dann signifikant, wenn ihnen eine größere Effektstärke zugrunde liegt. Somit ist davon auszugehen, dass Ergebnisse, insbesondere im Bereich kleinerer Stichproben, viel wahrscheinlicher dann publiziert werden – und damit auch einfacher gefunden werden können –, wenn in ihnen eine höhere Effektstärke aufgedeckt wurde. Wenn hingegen nur ein geringer oder kein Effekt gefunden wurde, besteht die Gefahr, dass dementsprechende Ergebnisse nicht der wissenschaftlichen Öffentlichkeit zur Verfügung gestellt werden.

Allerdings ist hinsichtlich des file drawer-Problems im hier untersuchten Forschungsfeld anzumerken, dass bzgl. des Zusammenhangs zwischen pädosexueller Viktimisierung und pädosexueller Delinquenz häufig gerade solche Studien, die keinen diesbezüglichen Zusammenhang finden, als ähnlich publikationswürdig angesehen werden wie Studien mit signifikanten Effekten. Denn gerade im Bereich der Forschung zur o.g. pädosexuellen Transitionsthese gibt es zahlreiche Autoren, die den genannten Zusammenhang für unwahrscheinlich bzw. schon die Annahme eines solchen Zusammenhangs für abwegig halten und dies auch über die Publikation nicht-signifikanter Forschungsergebnisse belegen wollen (vgl. die in Anmerkung 1 genannte Literatur).

Unabhängig davon, ob es also stimmt, dass im Forschungsfeld pädosexueller Delinquenz die „file drawer“-Problematik in spezieller Weise ausgeformt ist, muss auch dort jede Meta-Analyse hinsichtlich eines möglicherweise vorhandenen „publication bias“ untersucht werden. Denn auch dort besteht die Gefahr, dass ein „publication bias“ aufgrund nicht zu ermittelnder Forschungsergebnisse die Resultate dieser Analysen unbrauchbar macht.

Um das Ausmaß einer solchen Gefahr abschätzen zu können, wurden Methoden entwickelt, die versuchen, generelle Hinweise zur Existenz eines „publication bias“ und zu dessen erwartbaren Einflüssen auf die Ergebnisse von Meta-Analysen zu ermitteln. Mit diesen Methoden wird sowohl nach Indizien für verfälschende Ergebnisverschiebungen in Meta-Analysen als auch nach Hinweisen auf das ergebnisverzerrende Einflussgewicht von möglicherweise in Meta-Analysen fehlenden Studien gesucht (vgl. z.B. Borenstein et al. 2009: 278ff; Rothstein et al. 2005 oder Sutton et al. 2000). Im nächsten Abschnitt werden Logik und Technik einiger dieser Diagnosemethoden kurz vorgestellt.

3 Diagnosemöglichkeiten

Bei den Diagnosemethoden zum Nachweis eines publication bias handelt es sich sowohl um Methoden, die primär auf einer rein visuellen Überprüfung der gefundenen Ergebnisse in Abhängigkeit von den Stichprobengrößen der einzelnen Primärstudien basieren (Abschnitt 3.1), als auch um solche, die eher mit einer statistisch-logischen Methodik das Vorhandensein bzw. das

Ausmaß des publication bias bestimmen wollen.

Sterne et al. (2000) fanden in einer Simulationsstudie heraus, dass für eine Bias-Untersuchung mindestens zehn einzelne Studien in einer Meta-Analyse berücksichtigt werden müssen. Darüber hinaus empfehlen sie, in den Meta-Analysen sowohl grafische als auch statistische Analysen durchzuführen. Weiterhin fanden die Autoren heraus, dass zur Untersuchung des publication bias eine regressionsanalytische Vorgehensweise (s.u.) überwiegend besser abschneidet als die sogenannte Rank-Correlation-Methode (nach Begg und Mazumdar 1994, s.u.). Es wird zwar vermutet, dass die Rank-Correlation-Methode bei nicht-linearen Zusammenhängen bessere Ergebnisse liefert, aber solche Zusammenhänge werden eher als realitätsfern angesehen (vgl. Sterne et al. 2000: 1125ff).

In den folgenden Abschnitten werden die am häufigsten genutzten Methoden zur Analyse des publication bias erläutert, welche auch im MAPS-Projekt eingesetzt wurden und über deren Ergebnisse in Abschnitt 4 berichtet werden soll. Dazu gehören, wie bereits angesprochen, sowohl grafische Verfahren (Abschnitt 3.1) als auch regressionsanalytisch basierte Methodiken (Abschnitt 3.2). Zudem wird die auf Kendall's tau basierende „rank-correlation“-Methode nach Begg und Mazumdar kurz vorgestellt. Diese Methodik wird zu Validierungszwecken verwendet (Abschnitt 3.3). Auch wird erläutert, in welcher Weise die Anzahl von möglicherweise nicht identifizierten und daher in einer Meta-Analyse nicht berücksichtigten Primärstudien berechnet werden kann. Insbesondere interessiert dabei die Anzahl von nicht-signifikanten Studien, die notwendig wäre, um die im Rahmen des MAPS-Projektes entdeckten Befunde zum Zusammenhang zwischen pädosexueller Viktimisierung und pädosexueller Delinquenz zu kompromittieren (Abschnitt 3.4). Zuletzt wird noch eine Methode erörtert, mit der statistisch berechnet werden kann, für welche unbekannt gebliebenen Studien davon auszugehen ist, dass diese mit bestimmten Resultaten existieren bzw. (statistisch betrachtet) existieren müssten (Abschnitt 3.5). Natürlich gibt es darüber hinaus noch eine Vielzahl weiterer meta-analytischer Methoden, die hier nicht vorgestellt werden. Einen Überblick liefern u.a. Thornton/Lee (2000).

3.1 Grafische Analyse mittels „funnel-plot“

Der „funnel-plot“ (vgl. seine Darstellung in Abb. 4.1) ist eine Möglichkeit zur grafischen Analyse des publication bias von Meta-Analysen. Bei diesem Verfahren werden in einem Koordinatensystem die Effektstärken der für eine Meta-Studie herangezogenen Primäruntersuchungen (verortet auf der horizontalen X-Achse) gegen die Stichprobengröße der Primärstudien oder gegen eine davon abhängige Variante, die weiter unten beschrieben wird, (verortet auf der vertikalen Y-Achse) abgetragen. Davon ausgehend, dass die Präzision einer Studie

mit ihrer Stichprobengröße zunimmt, kann angenommen werden, dass in diesem Koordinatensystem bei hoher Stichprobengröße geringere Schwankungen der Effektstärken um die mittlere Effektstärke herum auftreten. Im Umkehrschluss müssten die Effektstärken von Studien mit geringer Fallzahl auch stärker um die „wahre“ bzw. mittlere Effektstärke schwanken bzw. würde, grafisch gesprochen, der Plot an diesen Stellen ausfransen. Dadurch entstünde eine Verteilung, die einem umgedrehten Trichter (engl.: funnel) entspricht: oben sehr schmal und nach unten sich immer stärker öffnend. Die Studien könnten dann innerhalb der Grenzen des erwarteten 95%-Konfidenzintervalls liegen (wie auch in der folgenden Abb. 4.1 dargestellt wird) (vgl. dazu Sterne et al. 2005: 75ff).

Bei Vorliegen eines publication bias ist davon auszugehen, dass diejenigen Studien nicht veröffentlicht bzw. nicht gefunden wurden, die eher eine geringe Präzision, also eine geringe Fallzahl, und eine geringe Effektstärke aufweisen. Denn eine Kombination von geringer Präzision und kleiner Fallzahl resultiert eher in nicht-signifikanten Ergebnissen. Somit würde bei einer grafischen Überprüfung des publication bias die Verteilung der Studien im Koordinatensystem über die verschobene y-Achse asymmetrisch wirken. Die Grafik würde weniger wie ein gerader Trichter aussehen (vgl. dazu Abb. 4.1), sondern eine Delle bzw. Verformung aufweisen.

Bei Verwendung des funnel-plots gibt es zwei weit verbreitete Darstellungsformen. Bei der einen werden die Odds-Ratio-Werte als Schätzwerte für die Effektstärken in logarithmierter Form auf der X-Achse abgetragen. Die andere Darstellungsform benutzt eine lineare Skalierung der Odds-Ratio-Werte auf der X-Achse. Auf der Y-Achse werden bei linearer Darstellung die Standardfehler in gedrehter Form abgetragen, um auf diese Weise die Präzision der Einzelstudien zu operationalisieren. Bei logarithmischer Darstellung wird dazu der Kehrwert des Standardfehlers benutzt (vgl. Abb. 4.1).

Welche Form zur visuellen Ermittlung eines publication bias besser geeignet ist, hängt von der Verteilung der Studien ab. Bei der logarithmischen Darstellung werden Studien mit geringer Präzision sehr viel weiter voneinander entfernt auf der X-Achse abgetragen, was vorhandene Asymmetrien deutlicher zum Vorschein bringt. Kommen Asymmetrien eher bei einem mittleren Standardfehler vor, so sind diese visuell eher über die lineare Darstellung auszumachen (vgl. Sterne/Egger 2001; Sterne et al. 2005).

3.2 Regressionsanalyse zum „funnel-plot“

Da die grafische Analyse von funnel-plots immer auch vom Auge des Betrachters abhängt, braucht es eine Methode, um den Zusammenhang zwischen Präzision und Effektstärke quantifizierbar und damit vergleichbar zu machen. Ein Wert, der dazu herangezogen werden kann, ist die „standard normal deviate“ (SND):

$$\text{SND} = \frac{\text{Effektstärke}}{\text{Standardfehler}}$$

wobei die Effektstärke für solcherart ausgewertete Studien durch das Odds Ratio gemessen wird. Dieses Verfahren kann sowohl bei ungewichteten als auch bei gewichteten Effektstärken (dann gewichtet mit der Inversen der Varianz) benutzt werden (vgl. Macaskill 2001: 644).

Der SND-Wert wird in diesem Verfahren mittels einer linearen Regressionsschätzung auf die Präzision der in den Primärstudien ermittelten Ergebnisse zurückgeführt, wobei diese, wie zuvor gesagt, gleichbedeutend mit der Inversen des Standardfehlers ist. Somit gilt für die Regressionsschätzung die Gleichung:

$$\widehat{\text{SND}} = a + b \cdot \text{Präzision}$$

Bei dieser Regressionsschätzung entspricht das Interzept „a“ dem Steigungskoeffizienten in einer gewichteten Regressionsschätzung, bei der die Effektstärke auf den Standardfehler zurückgeführt wird. Mit dieser Konstruktion haben Studien mit geringen Fallzahlen, die eher einen hohen Standardfehler aufweisen (jeweils der Nenner der SND und der Präzision), einen SND und eine Präzision, die beide nahe bei null liegen. Studien mit großer Fallzahl haben eine hohe Präzision und je nach Effektstärke eine hohe SND.

Wenn nun eine unverzerrte Studienausswahl vorliegt, in der Studien mit einer kleinen Fallzahl im Schnitt dieselbe Effektstärke aufweisen wie Studien mit einer höheren Fallzahl, dann läuft die Regressionsgerade durch den Ursprung des Koordinatensystems. Das ergibt sich daraus, dass kleine Studien durch einen hohen Standardfehler als Nenner sowohl bei der SND als auch bei der Präzision empirische Werte von nahe null aufweisen. Große Studien hingegen sollten bei ihrer hohen Präzision einen SND ausweisen, der in der Richtung und Stärke deutlich von der gemessenen Effektstärke abhängt.

Daher gilt, dass bei einer absolut unverzerrten Studienausswahl die Regressionsgerade der SND-Schätzung durch den Ursprung des Koordinatensystems läuft und das Interzept somit gleich null ist ($a = 0$). Je weiter hingegen der y-Achsen-Abschnitt vom Ursprung entfernt ist, desto größer ist die Asymmetrie der Studienausswahl (vgl. Egger et al. 1997: 629f).

3.3 „rank-correlation“-Methode

Statt oder zusätzlich zu der oben beschriebenen Regressionsanalyse nach Egger kann auch ein rank-correlation-Test nach Begg und Mazumdar (1994) zur Überprüfung der Beziehung zwischen Stichprobengröße und Effektstärke durchgeführt werden. Dieses Verfahren bietet somit eine andere, nicht-lineare Herangehensweise an die Bias-Problematik.

Die Methode bedient sich des Paar-Vergleich-Verfahrens (Kendall's-Tau). Ermittelt wird die Rang-Korrelation zwischen der standardisierten Effektstärke (t^*) und dem Standardfehler der Effektstärkenschatzung (v).⁶

Die einzelnen Studien werden jeweils nach t^* und v sortiert und in der Reihenfolge paarweise miteinander verglichen. Nach der Logik von Kendall's-Tau werden dabei die konkordanten (x) sowie die diskordanten Paare (y) ausgezählt. Konkordante Paare bilden jene Zweierkombinationen von Studien, die sowohl bei der Sortierung nach t^* als auch bei der Sortierung nach v in der gleichen Reihenfolge stehen. Diskordant ist ein Studienpaar dann, wenn sich die Reihenfolge bei der neuen Sortierung umkehrt. Diese Paarauszahlungen werden dann in die folgende standardisierte Teststatistik eingesetzt:

$$z = \frac{x - y}{\sqrt{k(k-1)(2k+5)/18}}$$

wobei k die Anzahl der Studien ist.

Der so ermittelte z -Wert kann anhand einer entsprechenden z -Verteilung auf zweiseitige Signifikanz getestet werden. Wenn ein Zusammenhang zwischen der Effektstärke und dem Standardfehler der Effektstärkenschatzung besteht, dann würde der Test positiv ausschlagen (vgl. Begg und Mazumdar 1994: 1089) und die Annahme bestätigen, dass die Asymmetrie im jeweiligen funnel plot (s.o.) durch einen publication bias verursacht wird.

Zu dieser Methode ist kritisch anzumerken, dass damit nur gute Ergebnisse erzielt werden können, wenn es eine deutliche Bandbreite der verschiedenen Studiengrößen gibt, da bei einer großen Zahl von Studien mit gleicher/ähnlicher Fallzahl bereits eine „natürliche“ Variation von Standardfehler und Effektstärke ausreichend sein kann, um entsprechende Ergebnisse zu erzielen. Auch wird von Borenstein (2005) angemahnt, dass diese Methode nur eine geringe Aussagekraft besitzt, solange kein großer publication bias vorhanden ist und somit die Abwesenheit eines signifikanten Testergebnisses nicht immer ein gültiger Beweis für die Abwesenheit eines publication bias sein kann (vgl. ebd., S.196). Daher verwendeten wir diese Methodik auch lediglich zu Validierungszwecken bei Einsatz der anderen hier genannten Methoden.

⁶ Zur Standardisierung der Effektstärken vgl. Begg/Mazumdar 1994: 1089.

3.4 „fail-safe numbers“-Methode

Eine ganz andere Logik zur Ermittlung eines möglichen publication effect verbirgt sich hinter der „fail-safe N“-Methode. Dabei wird die Anzahl von nicht-signifikanten Studien berechnet, welche zusätzlich in die betreffende Meta-Analyse aufgenommen werden müssten,⁷ damit die Gesamtsignifikanz eines gefundenen Effekts nicht mehr gegeben wäre. Gefragt wird also nach der Zahl von Studien, die aufgrund eines publication bias in eine betreffende Meta-Analyse nicht aufgenommen wurden, die aber notwendig wären bzw. existieren müssten, um die gefundenen Ergebnisse abzuändern. Wenn diese Zahl im Vergleich zur Anzahl der ausgewerteten Studien sehr hoch ist, kann davon ausgegangen werden, dass die gefundenen Ergebnisse auch Bestand haben.

Zum Verständnis dieser Methode sollte zunächst von einer sehr schlechten wissenschaftlichen Praxis ausgegangen werden. In einem solchen Fall würden nur solche Studien publiziert, die signifikante Ergebnisse zur Folge haben. Jedoch würden, selbst wenn in der Realität kein Effekt vorhanden wäre, bei fünf Prozent der Studien (bei entsprechend angenommener Irrtumswahrscheinlichkeit) signifikante Ergebnisse auf Grund des Fehlers erster Art ermittelt werden. Die restlichen 95 Prozent der Studien, selbst wenn sie die Realität besser wiedergäben, würden nach der publication-bias-These aufgrund der fehlenden Signifikanz nicht veröffentlicht werden.

Rosenthal (1979: 638) spricht in diesem Fall davon, dass diese Studien stattdessen in den Schubladen der Forscher verschwinden, weswegen er diese Problematik auch als „file drawer problem“ (s.o.) bezeichnet.

Dieser Annahme folgend würden bei Existenz eines extremen publication bias aus einer Meta-Analyse von Studien in einem Themengebiet mit nicht vorhandenem Variablenzusammenhang nur jene fünf Prozent aufgenommen werden, welche die Realität nicht richtig wiedergeben.

Leider kann auch bei Annahme einer guten wissenschaftlichen Praxis nicht ausgeschlossen werden, dass bestimmte Studien – aus welchen Gründen auch immer – nicht veröffentlicht werden. Es kann jedoch berechnet werden, wie viele unveröffentlichte Studien es ohne Zusammenhangsbefund geben müsste, damit der oben skizzierte Fall zuträfe und eine daraus resultierende Meta-Analyse vollständig dem Fehler erster Art unterliegen würde.

Es gibt mehrere Ansätze dazu, die Anzahl fehlender nicht-signifikanter Studien zu berechnen. Die ursprüngliche Methode wurde von Rosenthal (ebd.) vorgeschlagen. In dieser wird danach geschaut,

⁷ Das sind Studien, die aus irgendwelchen Gründen nicht Teil der Population bzw. Teil der Stichprobe der Meta-Analyse geworden sind und daher auch nicht in die Berechnung der Gesamteffekte eingingen.

wie viele Studien ohne Effekt (X) zu der Anzahl der tatsächlich eingebrachten Studien (k) hinzugefügt werden müssten, damit der kombinierte Z -Wert⁸ (Z_c) – also der SND-Wert, der über alle Studien gilt – erreicht wird. Auf diese Weise soll getestet werden wie viele Studien notwendig sind, um die Annahme in Zweifel zu ziehen, dass kein publication bias besteht.

Dazu wird folgende Formel verwendet:

$$Z_c = \frac{k\bar{Z}_k}{\sqrt{k+X}}$$

wobei \bar{Z}_k das arithmetische Mittel aller Z -Werte der eingebrachten Studien ist.

Bei Annahme von $\alpha = 5\%$ und demnach $Z_c = 1,645$ (einseitiger Test) ergibt sich folgende Formel (aufgelöst nach X), um die notwendige Anzahl von Studien zu berechnen (vgl. ebd., S. 639):

$$X = \left(\frac{k}{2,706} \right) \cdot [k(\bar{Z}_k)^2 - 2,706]$$

Alternativ hat Orwin (1983) diese Methode so abgewandelt, dass zusätzlich berücksichtigt werden kann, dass auch nur gering-signifikante Studien schwerer aufzufinden sind. Dafür ergänzt er die oben angegebene Formel in der Weise, dass dem X , also der Anzahl der gesuchten Studien, eine eigene Effektstärke mitgegeben werden kann, welche für diese nicht-gefundenen Studien angenommen wird. Diese Effektstärke sollte der geringsten Effektstärke entsprechen, bei der noch davon ausgegangen werden kann, dass entsprechende Ergebnisse nicht veröffentlicht werden.

Zudem ist es bei dieser Methode möglich, einen Schwellenwert festzulegen, auf den die Gesamteffektstärke minimal fallen darf, selbst wenn sie dann noch signifikant wäre. Die Methode berücksichtigt somit Faktoren, die aufgrund der genannten Eigenschaften näher an der tatsächlichen Forschungspraxis liegen. Welche Werte für die beiden Eigenschaften jedoch in die Berechnung einfließen, liegt alleine im Ermessen des jeweiligen Anwenders. Somit kann dieser Test zur Überprüfung von bestehenden Hypothesen oder im Vergleich mit anderen Meta-Analysen eingesetzt werden. Da beides im MAPS-Projekt nicht gegeben ist, wurde auf diese Art von Analysen verzichtet.

Beide Methoden, sowohl die von Rosenthal als auch die von Orwin, haben den Nachteil, dass sie keine Gewichtungen der Studien (z.B. in Bezug auf deren Fallzahl) erlauben. Gerade Meta-Analysen profitieren jedoch von eben solchen Gewichtungen, weswegen Rosenberg (2005) eine dritte Variante vorschlägt, die Gewichtungen bei fail-safe-N-Analysen erlaubt.⁹

⁸ Rosenthal (1979: 638) verwendet Z zur Bezeichnung des oben schon benannten SND. Doch Vorsicht: dieses Z darf nicht mit dem in der zuvor beschriebenen rank-correlation-Methode verwendeten kleinen z verwechselt werden. Rosenthal zeigt in einer Fußnote seines Artikels auch noch weitere Methoden auf, mit denen der SND zu berechnen ist.

⁹ Rosenberg hat zur einfachen Berechnung seines „fail-safe N“ eine Software zur freien Verfügung gestellt, welche im MAPS-Projekt eingesetzt wurde (s.u.). Diese Software wird in Rosenberg (2005) beschrieben.

3.5 „trim and fill“-Methodik

Die Methode des „trim and fill“ nach Duval und Tweedie (2000a, 2000b)¹⁰ basiert auf einem einfachen Symmetriegedanken bei Verteilungen von Werten (wie z.B. Effektstärken) um einen „wahren“ Wert, ähnlich dem Gedanken der Symmetrie bei einer Normalverteilung. Bei nicht vorhandener Symmetrie einer Werteverteilung wird entsprechend der „trim and fill“-Methodik mittels einer iterativen Technik die Anzahl fehlender Studien ermittelt, die die verlangte Symmetrie herstellen könnten. Als Ergebnis wird eine neue „tatsächliche“ Effektstärke berechnet, die auch bislang nicht-aufgefundene Studien berücksichtigt. Zugleich werden mit dieser Technik auch solche Studien zur Berechnung der „tatsächlichen“ Effektstärke berücksichtigt, die noch gar nicht durchgeführt wurden, die aber nach der Symmetrie-Logik in Zukunft durchgeführt werden müssten.

In der „trim and fill“-Methode wird zunächst iterativ bestimmt, wie viele der in der Meta-Analyse berücksichtigten Studien mit ihrer gemessenen Effektstärke so weit von einer als „wahr“ angenommenen Effektstärke (im Folgenden bezeichnet als „ θ “) abweichen, dass es zu ihnen auch Studien geben müsste, die genauso weit in die andere Richtung von θ abweichen. Wenn Studien ohne solche „Matching-Partner“ gefunden werden, wird ohne diese Studien ein neuer „wahrer“ Wert errechnet („trim“). Anschließend werden die nach dem Symmetriegedanken erwarteten Studien als vorhanden angenommen bzw. der Stichprobe der Meta-Analyse hinzugefügt („fill“), sodass eine neue durchschnittliche Effektstärke bestimmt werden kann, welche diese möglichen, jedoch nicht-veröffentlichten bzw. nicht aufgefundenen und damit auch nicht in der Meta-Analyse enthaltenen Studien berücksichtigt.

Das Verfahren läuft darauf hinaus, dass nach Berechnung eines θ auf der Basis aller Studien mit Matching-Partnern jene Studien dem Studienpool hinzugefügt werden, welche hypothetisch als „Matching-Partner“ existieren müssten. Es wird für jede Studie, die während der Trim-Phase wegfiel, eine solche hypothetische Studie angenommen, die denselben Standardfehler aufweist und eine Effektstärke von $ES_{\text{alt}} - \theta$ besitzt (vgl. z.B. Duval 2005).

In einer Untersuchung von Pham et al. (2001) wurden verschiedene Messmethoden bzgl. des publication bias getestet. Hierbei erwies sich die „trim-and-fill“-Methode als gutes Schätzverfahren zur Ermittlung von fehlenden Studien (ebd. S. 121).

Allerdings sind der funnel-plot und alle damit assoziierten Verfahren (also auch die „trim-and-fill“-Methode) nicht nur als Untersuchungsmethoden für einen möglichen publication bias oder andere, die Datengrundlage verzerrende Einflussgröße zu verstehen. Denn die Ergebnisse der jeweiligen Analysen können auch als Sensibilitätstest für die Robustheit der in den Meta-Analysen gefundenen Befunde verstanden werden (vgl. Duval 2005: 131).

¹⁰ In einer früheren Version auch Taylor/Tweedie (1998).

Das bedeutet, dass für den Fall, dass ein Fehler erster Art gemacht wird, also ein „publication bias“ ermittelt wird, obwohl ein solcher faktisch nicht vorliegt (vgl. u.a. Urban/Mayerl 2011: 138ff), das resultierende Ergebnis als Prüfung der Robustheit der Befunde verstanden werden kann. Am Beispiel der „trim-and-fill“-Methode ist dies so zu verstehen, dass, wenn die Ergebnisse nach Imputation noch signifikant sind, dann der tatsächliche nicht „korrigierte“ Effekt (vor Anwendung der „trim-and-fill“-Methode) robust ist. Und das gilt auch unabhängig davon, ob der implizierte „publication bias“ vorliegt oder dessen Vorliegen nur „fälschlicherweise“ angenommen wurde.

4 Analyseergebnisse

Im Folgenden wird über die Ergebnisse von Analysen zum Nachweis eines „publication index“ bei den im MAPS-Projekt (vgl. Abschnitt 1) durchgeführten Meta-Analysen berichtet. In diesen Analysen wurden die zuvor beschriebenen Verfahren eingesetzt.¹¹ Das Analysedesign sah dabei jeweils mehrere Schritte vor:

1. Schritt: In einem ersten Schritt wurden alle identifizierten Primärstudien (N=30) in die Untersuchung einbezogen. Dazu gehörten die Primärstudien, in denen überführte pädosexuelle Straftäter (=Untersuchungsgruppe) mit adultsexuellen Straftätern (= Kontrollgruppe 1), mit nicht-sexuellen Straftätern (=Kontrollgruppe 2) oder mit Personen aus der Allgemeinbevölkerung (=Kontrollgruppe 3) verglichen wurden.
2. Schritt: In einem zweiten Schritt wurden nur jeweils einzelne Untergruppen von Primärstudien oder Kombinationen von einzelnen Studien-Untergruppen in die Untersuchung einbezogen. Untergruppen entstehen dadurch, dass in diesen Studien die Untersuchungsgruppe „pädosexuelle Straftäter“ nur mit einer ganz bestimmten Kontrollgruppe (s.o.) verglichen wird. So gibt es:
 - (-) die Studien der Untergruppe 1 (N=16), in denen pädosexuelle Straftäter nur mit adultsexuellen Straftätern (aus der Kontrollgruppe 1) verglichen werden;
 - (-) die Studien der Untergruppe 2 (N=6), in denen pädosexuelle Straftäter nur mit nicht-sexuellen Straftätern (Kontrollgruppe 2) verglichen werden;
 - (-) die Studien der Untergruppe 3 (N=8), in denen pädosexuelle Straftäter nur mit nicht-straffällig gewordenen Personen aus der Normalbevölkerung (Kontrollgruppe 3) verglichen werden.

¹¹ Für die Berechnungen wurde die Software „Comprehensive Meta-Analysis (CMA) 2.0“ benutzt. Vgl. dazu Borenstein et al. 2009.

4.1 Analysen mit allen Primärstudien der Meta-Analyse

Für die Gruppe aller 30 Studien, die im MAPS-Projekt ausgewertet wurden, zeigt sich (ohne Unterscheidung nach Kontrollgruppen) der in Abbildung 4.1 präsentierte funnel-plot. Wie aus der Abbildung hervorgeht, ergibt sich eine leichte Asymmetrie um den mittleren Effektstärkewert von 1,23 auf der Log-Odds-Ratio-Skala¹², welcher durch die senkrechte Linie verdeutlicht wird und dessen 95%-Konfidenzintervall (0,95 bis 1,51) als weiße Raute dargestellt wird. Zu erkennen sind darüber hinaus einige weitere Studien (Abweichungen), die außerhalb der Grenzlinien des erwarteten 95%-Konfidenzintervalls liegen.

Um zu berechnen, ob diese Abweichungen die Effektstärke signifikant verzerren, und um die grafische Analyse näher zu qualifizieren, wird der in Abschnitt 3.2 erläuterte SND herangezogen. In der dabei geschätzten Regressionsgleichung liegt der y-Achenabschnitt bei $a = 0,21$. Diese Abweichung vom Ursprung ($a = 0$) ist mit einer Irrtumswahrscheinlichkeit von $\alpha = 0,05$ bei einem 2-seitigen Test nicht signifikant ($p = 0,37$). Somit kann von einer ausreichend symmetrischen Verteilung der in der Meta-Analyse berücksichtigten Studien ausgegangen werden.

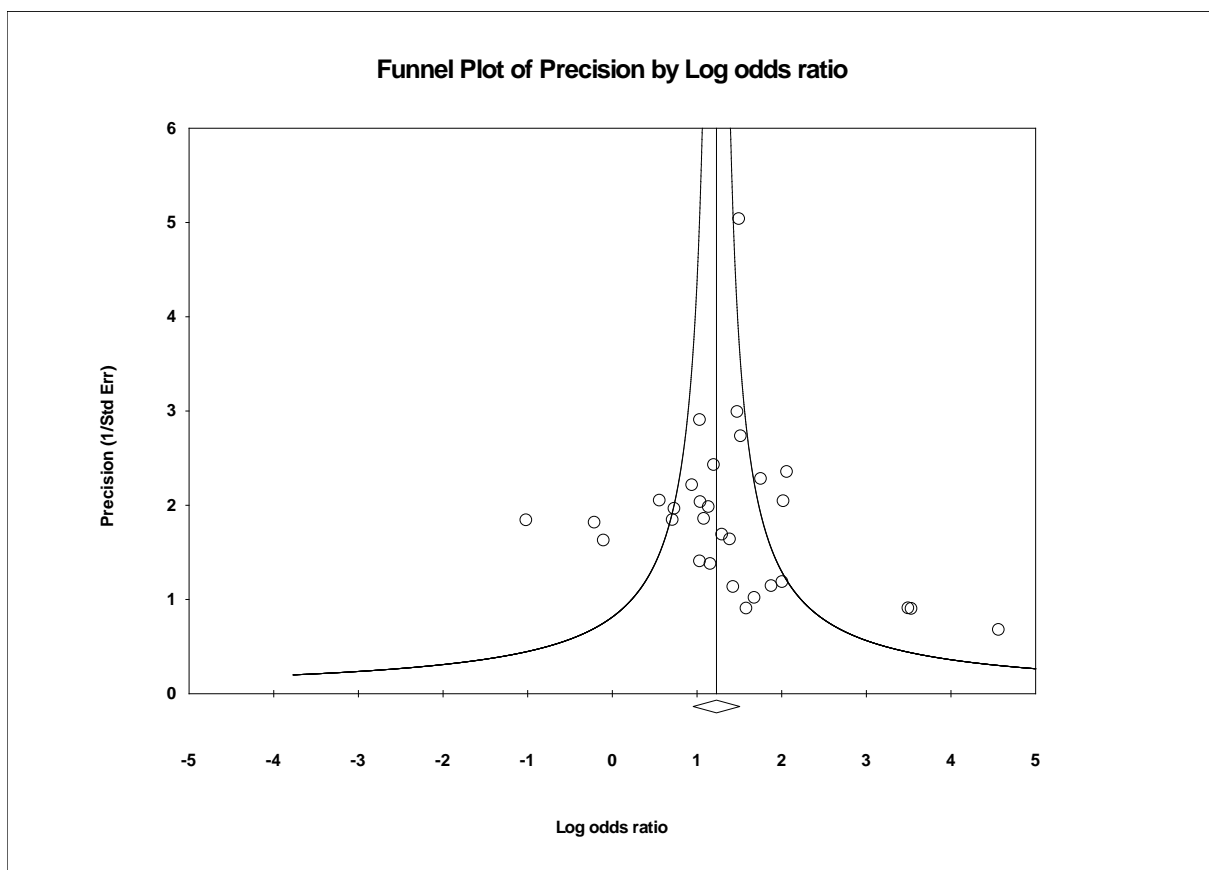


Abbildung 4.1: funnel-plot zu allen Studien der Meta-Analyse

Eine Analyse mit Hilfe des in Abschnitt 3.3 vorgestellten rank-correlation-Verfahrens von Begg

¹² Dieser Wert von 1,23 für das logarithmierte Odds Ratio entspricht dem in der Meta-Analyse über alle Studien ermittelten mittleren Odds Ratio von 3,41.

und Mazumdar erbringt einen z-Wert von 1,20, welcher bei einem 2-seitigen Test einem $p = 0,23$ entspricht. Damit ist auch der z-Wert deutlich nicht signifikant. Folglich kann nicht von einer Abhängigkeit zwischen der Teststärke und der Varianz ausgegangen werden, was wiederum ein Indiz ist, welches gegen einen möglichen publication bias spricht.

Auch die Berechnungen nach der fail-safe-N-Methode von Rosenthal (vgl. Abschnitt 3.4) bestätigen dieses Ergebnis. Demnach hätte es noch insgesamt 1295 Studien ohne Effekt geben müssen, um die ermittelte „wahre“ Effektstärke als „nicht-signifikant“ bezeichnen zu können. Abgesichert wird dieses Ergebnis auch mit der gewichteten Berechnung nach Rosenberg. Nach ihr müssten sogar 1514 Studien durchschnittlicher Gewichtung ohne gemessenen Effekt berücksichtigt werden, damit die in der kompletten Meta-Analyse gefundene Effektstärke ihre Signifikanz (mit $\alpha = 0,05$) verlieren könnte.

Somit wären für das gefundene Ergebnis selbst sehr viele, unberücksichtigt gebliebene Studien nicht gefährdend.

Mit der trim-and-fill-Methode (vgl. Abschnitt 3.5) werden insgesamt 8 Studien ermittelt, die unter Annahme eines publication bias existieren müssten, die aber im Datensatz der Meta-Analyse nicht zur Verfügung stehen. Diese werden in diesem Verfahren als Matching-Partner (s.o.) generiert und als solche in den weiteren Analysen berücksichtigt. Dargestellt werden sie in Abbildung 4.2 durch schwarze Punkte.

Die nach der trim-and-fill-Methode auf der Grundlage des erweiterten Studienpools berechnete „wahre“ Effektstärke, gemessen mittels der einfachen Odds-Ratio-Skala, würde dann von 3,41 (grafisch eingezeichnet ist der identische, lediglich logarithmierte Wert von 1,23) auf einen Wert von 2,63 (95%-KI: 1,96 bis 3,53) sinken. Dieser neue Wert wird in der Abbildung durch die Spitze der schwarzen Raute bzw. durch die senkrechte Linie beim logarithmischen Wert von 0,97 angezeigt. Anhand der Endpunkte der Raute ist zu erkennen, dass auch das neue 95%-Konfidenzintervall die null ausschließt.

Somit bleibt auch unter Annahme eines gewissen publication bias (dessen Auswirkungen durch die Berücksichtigung der ermittelten Matching-Partner simuliert werden) der signifikant positive Zusammenhang zwischen pädosexueller Viktimisierung und pädosexueller Delinquenz unberührt.

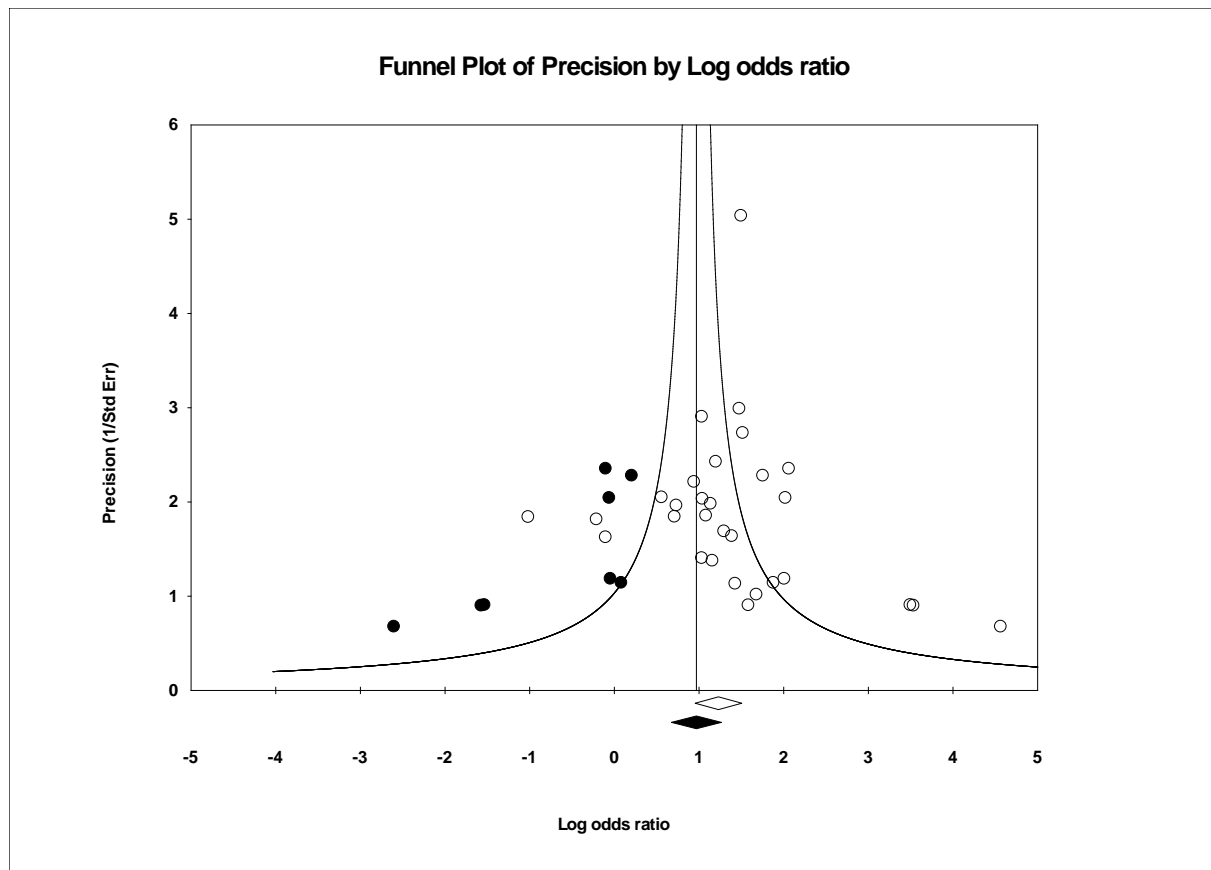


Abbildung 4.2: funnel-plot zu allen Studien der Meta-Analyse nach Einsatz der „trim and fill“-Methode

Im Folgenden werden diese Tests für verschiedene Studien-Untergruppen der Meta-Analyse mit spezifischen Kontrollgruppen (s.o.) separat durchgeführt. Dabei ist das Vorgehen identisch zu den oben beschriebenen Analysen.

4.2 Analysen mit den Primärstudien der Untergruppe 1

In diesem Abschnitt werden die Ergebnisse zu den Untersuchungen mit den Primärstudien der Untergruppe 1 (= Studien mit der Kontrollgruppe „adultsexuelle Straftäter“), bei welchen sich die Fallzahl auf 16 Studien reduziert, kurz vorgestellt. Genau wie bei den vorherigen Analysen wird auch hier mit der Auswertung der funnel-plots (Abbildungen 4.3 und 4.4) begonnen.

Bei Betrachtung des ersten funnel-plots (Abb. 4.3) fallen nur wenige Ausreißer auf. Lediglich im Bereich geringer bis mittlerer Präzision kann eine leichte Verschiebung zugunsten eines möglichen publication bias ausgemacht werden. Die geschätzte, mittlere Effektstärke wird wieder als Raute (bzw. senkrechte Linie) dargestellt und beträgt 1,01 mit einem Konfidenzintervall zwischen Werten von 0,66 und 1,36.

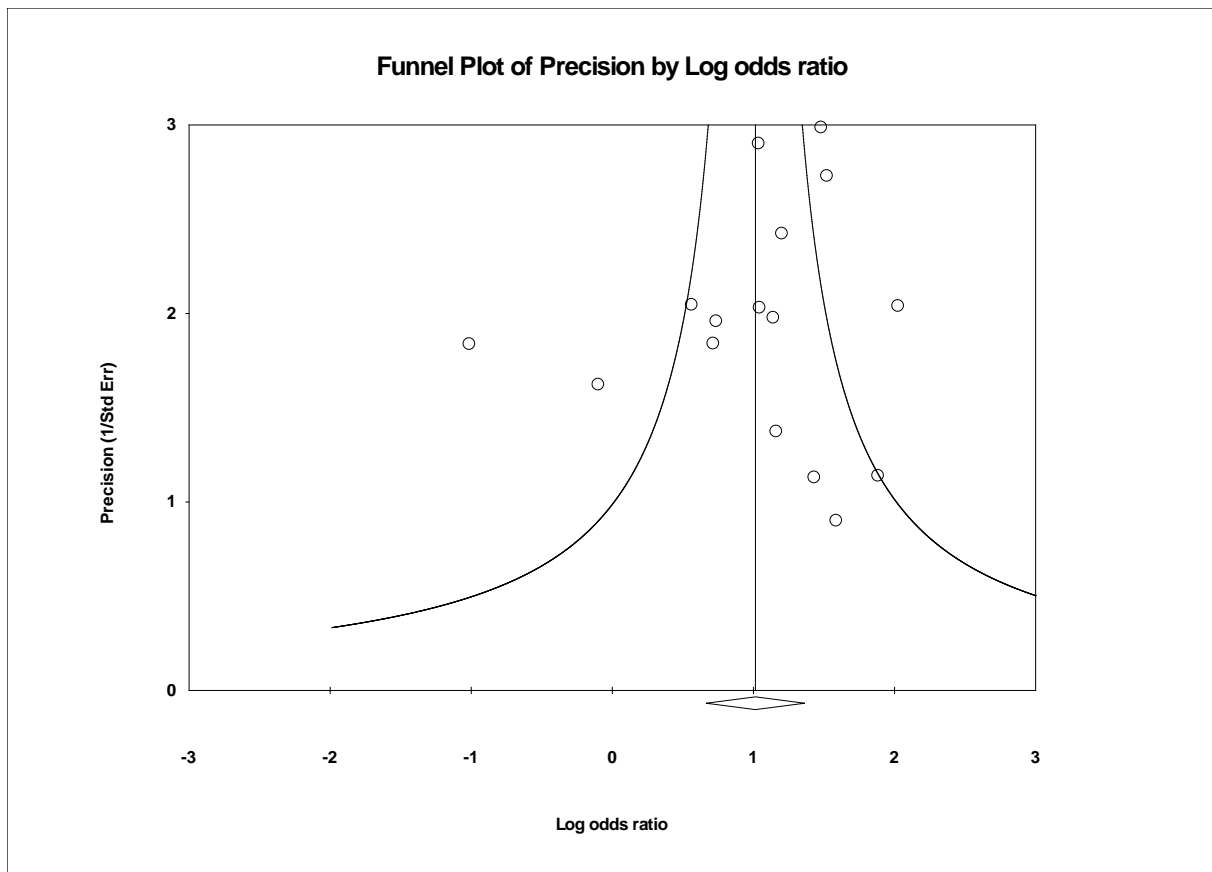


Abbildung 4.3: funnel-plot zu den Primärstudien der Untergruppe 1

Bei Anwendung der Regressionsschätzung liegt der y-Achsenabschnitt der Regressionsgeraden bei $a = -0,70$. Diese Abweichung vom Ursprung ist mit einer Irrtumswahrscheinlichkeit von $\alpha = 0,05$ bei einem 2-seitigen Test nicht signifikant ($p = 0,56$). Somit kann auch hier von einer ausreichend symmetrischen Verteilung ausgegangen werden (vgl. die Ausführungen zum SND in Abschnitt 3.2).

Das rank-correlation-Verfahren errechnet einen z-Wert von 0,54, welcher bei einem 2-seitigen Test eine Irrtumswahrscheinlichkeit von $p = 0,59$ aufweist. Damit ist auch dieser Wert deutlich als nicht-signifikant einzustufen. Folglich lässt sich auch hier keine Abhängigkeit zwischen Effektstärke und Fallzahl erkennen und kann somit auch in dieser Untergruppe kein Hinweis auf einen möglichen publication bias gefunden werden.

Nach der „fail-safe numbers“-Methode hätten in der Untergruppe 1 zusätzlich zu den hier analysierten Studien noch 247 nicht-signifikante Studien vorliegen müssen, um die in der Meta-Analyse gefundene „wahre“ Effektstärke auf ein nicht-signifikantes Niveau zu bringen.

Gemäß der gewichteten Berechnung nach Rosenberg müssten noch mindestens sieben Studien durchschnittlicher Gewichtung ohne gemessenen Effekt in die Analyse einbezogen werden,

damit die berechnete Effektstärke ihre Signifikanz ($\alpha = 0,05$) verliert.

Dies scheint eine sehr kleine Anzahl von Studien zu sein, jedoch muss bei einer solchen Einschätzung auch die geringe Fallzahl dieser Meta-Analyse (N=16) berücksichtigt werden. Zudem ist es, wie in Abschnitt 2 bereits erörtert, gerade im analysierten Themenbereich hochgradig unwahrscheinlich, dass Studien ohne signifikantes Ergebnis der file drawer-Problematik unterliegen. Daher ist die Zahl von 7 nicht publizierten und nicht aufgefundenen Studien eher als hoch bzw. bedeutsam einzuschätzen.

Nach der Methodik des trim-and-fill-Verfahrens werden drei Studien ausgemacht (siehe Abbildung 4.4), die zur Simulation der Symmetrie-Eigenschaft bei der Studienverteilung in der Analyse ergänzt werden sollten. Mit diesen würde aber der wahre Wert der Effektstärke nur geringfügig auf 2,59 der einfachen Odds-Ratio-Skala (95%-KI: 2,06 bis 3,27) reduziert werden (in der Grafik sind wiederum die logarithmierten Werte abgetragen).

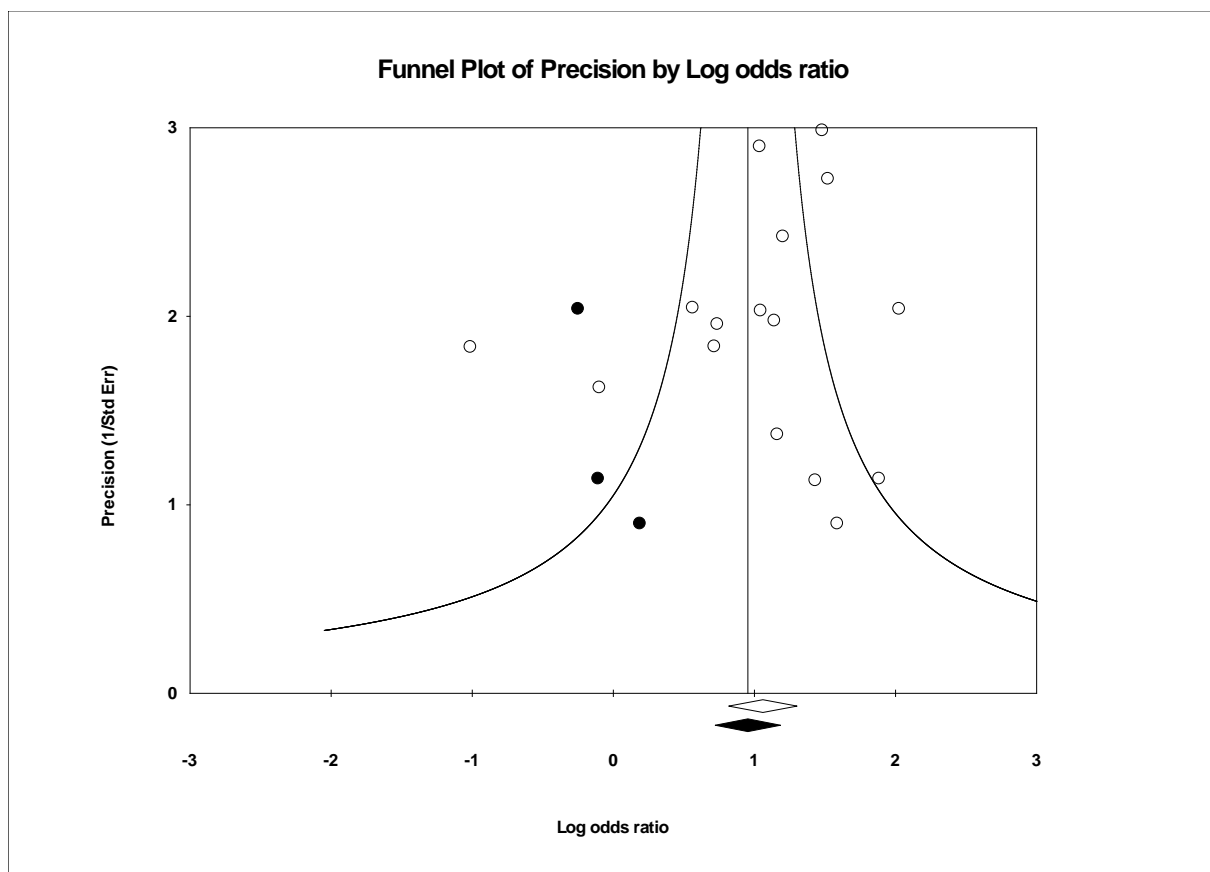


Abbildung 4.4: funnel-plot zu den Studien der Untergruppe 1 nach Einsatz der „trim and fill“-Methode

Abschließend werden nach demselben Vorgehen noch Ergebnisse von Analysen mit 8 Studien aus der Untergruppe 2 und 6 Studien auch der Vergleichsgruppe 3 vorgestellt.

4.3 Analysen mit den Primärstudien der Untergruppen 2 und 3

Die Studien, bei welchen die pädosexuellen Delinquenten mit nicht-sexuellen Gewaltstraftätern verglichen werden (N=8 Studien; Untergruppe 2), und die Studien, bei welchen nicht-straffällige Personen aus der Normalbevölkerung die Kontrollgruppe stellen (N=6 Studien; Untergruppe 3), können nur mit äußerster Vorsicht benutzt werden, um einen möglichen publication bias in den entsprechenden Meta-Analysen nachzuweisen. Denn wie Simulationsstudien gezeigt haben, sollten mindestens 10 Studien ausgewertet werden können, um dabei zu robusten Ergebnissen kommen zu können (vgl. Abschnitt 3). Dennoch sollen im Folgenden die Ergebnisse dieser Analysen zumindest der Vollständigkeit halber vorgestellt werden.

Mit den Studien der Untergruppe 2 ergibt sich bei einer regressionsanalytischen Auswertung des funnel-plots ein Interzept der Regressionsschätzung (s.o.) von $a = 0,01$. Bei Studien der Untergruppe 3 beträgt das Interzept 1,99. Wie zu erwarten, ist der Wert von 0,01 nicht signifikant. Anders ist dies hingegen bei einem Intercept von 1,99 bei den Studien in Untergruppe 3, bei welchen mit diesem Schätzergebnis die Annahme eines publication bias nicht verworfen werden kann. Dieses Ergebnis könnte jedoch ein methodisches Artefakt sein, das dadurch entstanden ist, dass die Rekrutierung von nicht straffällig gewordenen Vergleichspersonen aus der allgemeinen Bevölkerung je nach Studie sehr unterschiedlich durchgeführt wurde (auch hinsichtlich der methodischen Rigidität) und deshalb diese Studien ein Ausmaß an Heterogenität aufweisen, das es eigentlich nicht zulässt, sie in einer Kontrollgruppe zusammenzufassen.

In den beiden Untergruppen liegt der Wert für das fail-safe N nach der Berechnung von Rosenthal bei 147 (Untergruppe 2) und bei 54 (Untergruppe 3). Beides sind recht hohe Werte, die wiederum eher nicht für einen publication bias sprechen. Denn auch hier gilt, dass gerade im Rahmen des untersuchten Themenbereichs die Publikationschancen von Studien ohne signifikantes Ergebnis nicht deutlich verringert sein sollten und somit auch das file drawer-problem eher klein sein sollte (vgl. Abschnitt 2).

Nach der trim-and-fill-Methode würden in beiden Untergruppen jeweils drei Studien nicht berücksichtigt sein. Das führt bei den Studien in Untergruppe 2 zu einem Abfall des Odds Ratios von 3,01 auf 2,27 und bei den Studien in Untergruppe 3 von 9,82 auf 5,63. In beiden Fällen ist dies aber als nicht bedenklich einzustufen, da jeweils der positive Effekt ($OR > 1$) und damit ein positiver Zusammenhang zwischen pädosexueller Viktimisierung und pädosexueller Delinquenz auf signifikantem Niveau bestehen bleibt.

Damit lassen in den Studiengruppen 2 und 3 die Ergebnisse aller hier eingesetzten Verfahren (bis auf die Regression zum funnel-plot) den Schluss zu, dass nicht davon auszugehen ist, dass die Meta-Analysen mit diesen Primärstudien durch einen bedeutsamen publication bias verzerrt sein müssen.

5 Fazit

Wie eingangs erläutert können die Ergebnisse von Meta-Analysen von einem so genannten „publication bias“ betroffen sein. Der publication bias kann die Ergebnisse von Meta-Analysen verzerren, kann sie verfälschen oder kann sie zum reinen Artefakt werden lassen.

Der publication bias von Meta-Analysen ist die Folge einer systematisch verzerrten Auswahl derjenigen Primärstudien, die in eine meta-analytische Auswertung einbezogen werden. Für diese systematische Selektivität gibt es zwei Hauptursachen: zum einen können die grauen und/oder weißen Publikationskanäle der scientific community bestimmten wissenschaftlichen Ergebnisberichten unterschiedliche Publikationschancen einräumen. So werden z.B. häufig Berichte, in denen signifikante Ergebnisse mit beeindruckenden Effektstärken vorgestellt werden, eher veröffentlicht als Ergebnisse, die nicht inferenzstatistisch abgesichert werden können oder deren empirisches Ergebnis allein im Nachweis eines nicht-vorhandenen Effekts besteht. Zum anderen kann ein publication bias dadurch entstehen, dass zwar alle relevanten Forschungsergebnisse in einem wissenschaftlichen Themengebiet publiziert werden, dass diese aber nicht alle für eine Meta-Analyse identifiziert werden können, weil z.B. einige/viele überhaupt nicht in Literatur-Datenbanken aufgenommen wurden (wie z.B. Dissertationen oder „graue“ Literaturtitel) oder weil z.B. einige/viele nur in unbekanntenen Publikationsnischen (wie z.B. online-Journalen) erschienen sind.

Um die meta-analytischen Ergebnisse des MAPS-Projektes (=“Meta-Analyse zum Stand der empirischen Forschung über die psycho-sozialen Ursachen pädosexueller Delinquenz”) hinsichtlich ihrer möglichen Gefährdung durch einen real existierenden „publication bias“ zu überprüfen, wurden die Ergebnisse mit verschiedenen Verfahren zur statistischen Identifikation von „publication bias“-Effekten getestet. Dabei kamen Verfahren unterschiedlichster statistischer Provenienz und Methodik zur Anwendung (u.a.: „funnel-plot“- , „rank-correlation“- , „fail-safe numbers“- und „trim and fill“-Techniken).

Das besondere Augenmerk lag dabei einerseits a) auf einem möglichen publication bias in denjenigen Meta-Analysen, die mit der Gesamtauswahl aller als relevant identifizierten Primärstudien durchgeführt wurden, und andererseits b) auf einem möglichen publication bias in

den Meta-Analysen, die mit einer speziellen Auswahl von Primärstudien durchgeführt wurden. Bei dieser Auswahl (Untergruppe 1) handelt es sich um Primärstudien, bei denen die Untersuchungsgruppe der pädosexuellen Straftäter allein mit der Kontrollgruppe der adultsexuellen Straftäter verglichen wird.

In unseren Analysen wurden keine statistisch relevanten Hinweise auf die Existenz eines deutlich hervortretenden „publication bias“-Problems gefunden. So war z.B. nach der „fail-safe numbers“-Methode die Anzahl von Negativ-Studien, die notwendig gewesen wäre, um das in unseren Meta-Analysen (mit allen Studien) gefundene Ergebnis einer deutlich erhöhten Chance von pädosexueller Delinquenz nach vorhergehender pädosexueller Viktimisierung außer Kraft zu setzen bzw. als nicht-signifikanten Zusammenhang zu qualifizieren, mit Werten von $N=1.295$ bzw. $N=1.514$ so hoch, dass nicht davon auszugehen ist, dass unser Ergebnis auch durch die Hinzunahme einer realistisch großen Anzahl von zuvor nicht-identifizierten Studien mit widersprechendem Ergebnis hätte verändert werden können.

Auch der Test der im Projekt ermittelten Effektstärke für den Zusammenhang zwischen pädosexueller Viktimisierung und pädosexueller Delinquenz mittels „trim and fill“-Methode, die von einem im Hintergrund wirkenden, verzerrenden publication bias ausgeht, reduzierte den von uns berechneten Wert für die Effektstärke von 3,41 (Odds Ratio) auf lediglich einen Wert von 2,63 (Odds Ratio), so dass auch damit gezeigt werden konnte, dass die meta-analytisch ermittelten Schätzwerte als robust hinsichtlich von real nicht auszuschließenden „publication bias“-Effekten sind.

Ähnliches gilt auch für die Überprüfung der Primärstudien aus der Untergruppe 1 (s.o.). Bei dieser speziellen Gruppe von 16 Primärstudien, in denen pädosexuelle Straftäter mit adultsexuellen Straftätern verglichen werden, müsste zusätzlich noch einmal die unrealistisch hohe Anzahl von 247 Studien ohne Effekt gefunden werden, damit die von uns meta-analytisch ermittelten Ergebnisse unter ein kritisches Signifikanzniveau sinken würden.

Auch wenn bei Verwendung der gewichteten Methode von Rosenberg nicht mehr 247 sondern nur noch 7 zusätzliche Studien ohne Effektnachweis ausreichen würden, um in Untergruppe 1 (s.o.) das von uns ermittelte Ergebnis eines signifikanten Effekts von pädosexueller Viktimisierung auf pädosexuelle Delinquenz „zu kippen“, ist auch dies kein überzeugender Hinweis auf die Anfälligkeit unserer diesbezüglichen Meta-Analysen gegenüber „publication bias“-Effekten.

Denn, wie erläutert wurde, werden gerade im Themenfeld von pädosexueller Viktimisierung/pädosexueller Delinquenz nicht-signifikante Studien von vielen Forschern in besonderem Maße als veröffentlichungswürdig angesehen. Daher wäre gerade hier nicht mit einem starken file drawer-bias (siehe Abschnitt 2.1) zu rechnen. Dies auch gerade deshalb nicht, weil er

dann den extrem hohen Anteil von mutmaßlich 30% aller einschlägigen Studien betreffen würde.¹³

Auch erlaubt der mit Hilfe der „trim and fill“-Methode ermittelte, signifikante Odds-Ratio-Wert von 2,59 (als Indikator für die meta-analytisch geschätzte Effektstärke) den Schluss, dass auch in Untergruppe 1 unser Ergebnis nicht durch einen publication bias gravierend verzerrt wurde und dass unsere Befunde auch dort als substantiell bedeutsam für das untersuchte Forschungsfeld zu verstehen sind.

Zusammenfassend kann aus unseren Analysen abgeleitet werden, dass die Existenz eines publication bias bei den im MAPS-Projekt analysierten Gruppen von Primärstudien als eher unwahrscheinlich gelten kann und somit davon auszugehen ist, dass auch die Ergebnisse der im Projekt vorgenommenen Meta-Analysen eher nicht durch einen publication bias verzerrt sind.

Allerdings ist es, wie bereits in Abschnitt 3.5 erläutert, auch an dieser Stelle noch einmal wichtig zu betonen, dass die funnel-plot-Methodik (so wie die übrigen Verfahren, die darauf abzielen, den publication bias und andere verzerrende Einflussgröße zu identifizieren) auch als Sensibilitätstest für die Robustheit der in den Meta-Analysen erzielten Ergebnisse verstanden werden kann (vgl. Duval 2005: 131). Dementsprechend können die hier berichteten Ergebnisse auch so interpretiert werden, dass nach diesen Befunden davon auszugehen ist, dass die Gültigkeit unserer Meta-Analysen selbst bei unvollständiger oder systematisch beeinträchtigter Datenauswahl nicht eingeschränkt wird.

¹³ Die Zahl von 30% ergibt sich daraus, dass von uns nach sorgfältigem Screening von ca. 64.000 Studien zum Themenkomplex „Pädosexualität“ nur 16 Studien identifiziert werden konnten, die zur Untergruppe „pädosexuell Straffällige im Vergleich zu adultsexuell Straffälligen“ gehören. Gäbe es noch 7 Studien, die aufgrund der file drawer Problematik nicht von uns aufgefunden werden konnten, so würden insgesamt 23 Studien existieren, von denen wir 7 bzw. 30% nicht identifiziert hätten.

Literaturverzeichnis

- Begg, Colin B.; Mazumdar, Madhuchhanda (1994): Operating Characteristics of a Rank Correlation Test for Publication Bias. *Biometrics*, 50 (4), S. 1088–1101.
- Borenstein, Michael (2005): Software für Publication Bias In: Hannah R. Rothstein, Alexander J. Sutton und Michael Borenstein (Hg.): *Publication bias in meta-analysis. Prevention, assessment and adjustments*. Chichester, England; Hoboken, NJ: Wiley, S. 193–220.
- Borenstein, Michael; Hedges, Larry V.; Higgins, Julian P. T.; Rothstein, Hannah R. (2009): *Introduction to meta-analysis*. Chichester, U.K: John Wiley & Sons.
- Dhaliwal, Gurmeet K.; Gauzas, Larry; Antonowicz, Daniel H.; Ross, Robert R. (1996): Adult Male Survivors of Childhood Sexual Abuse: Prevalence, Sexual Abuse Characteristics, and Long-Term Effects. *Clinical Psychological Review*, 19 (7), S. 619–639.
- Duval, Sue J. (2005): The Trim and Fill Method. In: Hannah R. Rothstein, Alexander J. Sutton und Michael Borenstein (Hg.): *Publication bias in meta-analysis. Prevention, assessment and adjustments*. Chichester, England; Hoboken, NJ: Wiley, S. 127–144.
- Duval, Sue J.; Tweedie, Richard L. (2000a): A Nonparametric "Trim and Fill" Method of Accounting for Publication Bias in Meta-Analysis. *Journal of the American Statistical Association*, 95 (449), S. 89–98.
- Duval, Sue J.; Tweedie, Richard L. (2000b): Trim and Fill: A Simple funnel-plot-Based Method of Testing and Adjusting for Publication Bias in Meta-Analysis. *Biometrics*, 56 (2), S. 455–463.
- Egger, Matthias; Smith, George Davey; Schneider, Martin; Minder, Christoph (1997): Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal*, 315 (7109), S. 629–634.
- Freund, K.; Watson, R.; Dickey, R. (1990): Does Sexual Abuse in Childhood Cause Pedophilia: An Exploratory Study. *Archives of Sexual Behavior*, 19 (6), S. 557–568.
- Hanson, R.K.; Slater, S. (1988): Sexual victimization in the history of child sexual abusers: A Review. *Annals of Sex Research*, 1, S. 485–499
- Hopewell, Sally; Clarke, Mike; Mallett, Sue (2005): Gray Literature and Systematic Reviews. In: Hannah R. Rothstein, Alexander J. Sutton und Michael Borenstein (Hg.): *Publication bias in meta-analysis. Prevention, assessment and adjustments*. Chichester, England; Hoboken, NJ: Wiley, S. 49–72.
- Jespersen, A. F.; Lalumière, M. L.; Seto, M. C. (2009). Sexual abuse history among adult sex offenders and non-sex offenders: A meta-analysis. *Child Abuse & Neglect*, 33, 179–192.
- Macaskill, Petra; Walter, Stephen D.; Irwig, Les (2001): A comparison of methods to detect publication bias in meta-analysis. *Statistics in Medicine*, 20 (4), S. 641–654.
- Orwin, R. G. (1983): A Fail-Safe N for Effect Size in Meta-Analysis. *Journal of Educational and Behavioral Statistics*, 8 (2), S. 157–159.

Peters, Jaime L.; Sutton, Alexander J.; Jones, David R.; Abrams, Keith R.; Rushton, Lesley (2006): Comparison of Two Methods to Detect Publication Bias in Meta-analysis. *Journal of the American Medical Association*, 295 (6), S. 676–680.

Pham, Ba'; Platt, Robert; McAuley, Laura; Klassen, Terry P.; Moher, David (2001): Is There a "Best" Way to Detect and Minimize Publication Bias? An Empirical Evaluation. *Evaluation & the Health Professions*, 24 (2), S. 109–125.

Rosenberg, Michael S. (2005): The File drawer Problem Revisited: A General Weighted Method for Calculating Fail-Safe Numbers in Meta-Analysis. *Evolution*, 59 (2), S. 464–468.

Rosenthal, Robert (1979): The file drawer problem and tolerance for null results. *Psychological Bulletin*, 86 (3), S. 638–641.

Rothstein, Hannah R.; Sutton, Alexander J.; Borenstein, Michael (2005): Publication Bias in Meta-Analysis. In: Hannah R. Rothstein, Alexander J. Sutton und Michael Borenstein (Hg.): *Publication bias in meta-analysis. Prevention, assessment and adjustments*. Chichester, England; Hoboken, NJ: Wiley, S. 1–7.

Schneider, H.J. (1999): Kriminologie der Sexualdelikte – Teil 1. Neue Forschungsergebnisse zu Erscheinungsformen, Ursachen und Reaktionsmöglichkeiten. *Kriminalstatistik*, 53, S. 233–238.

Seghorn, Theoharis K.; Prentky, Robert A.; Boucher, Richard J. (1987): Childhood Sexual Abuse in the Lives of Sexually Aggressive Offenders. *The American Academy of Child and Adolescent Psychiatry*.

Seto, M. C.; Lalumière, M. L. (2010). What is so special about male adolescent sexual offending? A review and test of explanations through meta-analysis. *Psychological Bulletin*, 136, S. 526-575.

Slater, Daniel; McMillan, Dean; Richards, Mark; Talbot, Tiffany; Hodges, Jill; Bentovim, Arnon; Hastings, Richard; Stevenson, Jim; Skuse, David (2003): Development of Sexually Abusive Behaviour in Sexually Victimized Males: A Longitudinal Study. *The Lancet*, 361, S. 471-476.

Sterne, Jonathan A. C.; Becker, Betsy Jane; Egger, Matthias (2005): The Funnel Plot. In: Hannah R. Rothstein, Alexander J. Sutton und Michael Borenstein (Hg.): *Publication bias in meta-analysis. Prevention, assessment and adjustments*. Chichester, England; Hoboken, NJ: Wiley, S. 75–98.

Sterne, Jonathan A. C.; Egger, Matthias (2001): Funnel plots for detecting bias in meta-analysis. *Journal of Clinical Epidemiology*, 54 (10), S. 1046–1055.

Sterne, Jonathan A. C.; Egger, Matthias; Smith, George Davey (2001): Systematic reviews in health care: Investigating and dealing with publication and other biases in meta-analysis. *British Medical Journal*, 323 (7304), S. 101–105.

Sterne, Jonathan A. C.; Gavaghan, David; Egger, Matthias (2000): Publication and related bias in meta-analysis. *Journal of Clinical Epidemiology*, 53 (11), S. 1119–1129.

Sutton, Alexander J.; Duval, Sue J.; Tweedie, Richard L.; Abrams, Keith R.; Jones, David R. (2000): Empirical assessment of effect of publication bias on meta-analyses. *British Medical Journal*, 320 (7249), S. 1574–1577.

Taylor, Sue; Tweedie, Richard L. (1998): A Non-parametric "Trim and Fill" Method of Assessing Publication Bias in Meta-Analysis. Technical Report: Colorado State University, Department of Statistics.

Thornton, Alison; Lee, Peter (2000): Publication bias in meta-analysis its causes and consequences. *Journal of Clinical Epidemiology*, 53 (2), S. 207–216.

Urban, Dieter; Lindhorst, Heiko (2003): Vom Sexualopfer zum Sexualtäter? Unterscheiden sich pädosexuelle Straftäter von andern Sexualstraftätern durch ein erhöhtes Opfer-Täter-Risiko? - eine empirische Pilotstudie. *Soziale Probleme*, 14, S. 137-161.

Urban, Dieter; Lindhorst, Heiko (2004): Erhöht ein Opfer-Täter-Zyklus das Risiko, Sexualstraftaten als pädosexuelle Straftaten zu begehen? Ergebnisse einer ereignisanalytischen Pilotstudie. *SISS - Schriftenreihe des Instituts für Sozialwissenschaften der Universität Stuttgart*.

Urban, Dieter; Fiebig, Joachim (2010): Wird pädosexuelles Verhalten in jungen Jahren erlernt? Eine empirische Fall-Kontroll-Studie. *Soziale Probleme*, S. 226-248.

Urban, Dieter; Fiebig, Joachim (2011): Pädosexueller Missbrauch: Wenn Opfer zu Tätern werden. Eine empirische Studie. *Zeitschrift für Soziologie*, S. 42-61.

Whitaker, D.J.; Le, B.; Hanson, R.K.; Baker, C.K.; McMahon, P.M.; Ryan, G.; Klein, A.; Rice, D.D. (2008). Risk factors for the perpetration of child sexual abuse: A review and meta-analysis. *Child Abuse & Neglect*, 32, S. 529-548.

SISS: Schriftenreihe des Instituts für Sozialwissenschaften der Universität Stuttgart
<http://www.uni-stuttgart.de/soz/institut/forschung/siss.html>

bisher sind erschienen:

- No. 1/1994 "Vertrauen" - soziologisch betrachtet. Ein Beitrag zur Analyse binärer Interaktionssysteme.
Peter Antfang, Dieter Urban
- No. 2/1994 Report on the German Machine Tool Industry.
Frank C. Englmann, Christian Heyd, Daniel Köstler, Peter Paustian
with the assistance of Susanne Baur and Peter Bergmann
- No. 3/1994 Neue württembergische Rechtstatsachen zum Unternehmens- und Gesellschaftsrecht.
Udo Kornblum
- No. 4/1994 Rechtstatsachen zum Unternehmens- und Gesellschaftsrecht aus den neuen Bundesländern.
Udo Kornblum
- No. 1/1995 Die Bedeutung Neuronaler Netze in der Ökonomie.
Hermann Schnabl
- No. 2/1995 Regionale Strukturprobleme.
Sammelband der Beiträge zum Symposium vom 13. und 14. Oktober 1994.
Frank C. Englmann (Hrsg.)
- No. 3/1995 Latent Attitude Structures Directing the Perception of New Technologies.
An Application of SEM-Methodology to the Construction of Attitude
Measurement Models Related to Technologies of Prenatal Genetic
Engineering and Testing.
Dieter Urban
- No. 4/1995 Handbuch zur empirischen Erhebung von Einstellungen/Kognitionen zur Bio- und Gentechnologie (inklusive Diskette)
(zweite, überarbeitete und erweiterte Auflage)
Uwe Pfenning, Dieter Urban, Volker Weiss
- No. 5/1995 Social Indicators in a Nonmetropolitan County: Testing the Representativeness of a Regional Nonrandom Survey in Eastern Germany.
Dieter Urban, Joachim Singelmann
- No. 1/1996 Jugend und Politik im Transformationsprozeß. Eine Fallstudie zur Stabilität und Veränderung von politischen Einstellungen bei ostdeutschen Jugendlichen zwischen 1992 und 1995.
Dieter Urban, Joachim Singelmann, Helmut Schröder
- No. 2/1996 Einstellungsmessung oder Einstellungsgenerierung? Die Bedeutung der informationellen Basis bei Befragten für die empirische Rekonstruktion von Einstellungen zu gentechnischen Anwendungen.
Martin Slaby
- No. 1/1997 Gentechnik: „Fluch oder Segen“ versus „Fluch und Segen“.
Bilanzierende und differenzierende Bewertungen der Gentechnik in der öffentlichen Meinung.
Dieter Urban und Uwe Pfenning

(Fortsetzung ...)

SISS: Schriftenreihe des Instituts für Sozialwissenschaften der Universität Stuttgart
<http://www.uni-stuttgart.de/soz/institut/forschung/siss.html>

bisher sind erschienen (Fortsetzung):

- No.2/1997 Die soziale Vererbung von Ausländer“feindlichkeit“. Eine empirische Längsschnittanalyse der intra- und intergenerativen Transmission von sozialen Einstellungen.
Dieter Urban und Joachim Singelmann
- No. 3/1997 Politische Sozialisation im Transformationsprozeß: Die Entwicklung demokratiebezogener Einstellungen von ostdeutschen Jugendlichen und deren Eltern zwischen 1992 und 1996.
Barbara Schmidt, Dieter Urban, Joachim Singelmann
- No.1/1998 Bewertende Einstellungen zur Gentechnik: ihre Form, ihre Inhalte und ihre Dynamik. Kurzbericht zu Ergebnissen des Forschungsprojektes „Einstellungen zur Gentechnik“.
Dieter Urban, Uwe Pfenning, Joachim Allhoff
- No.2/1998 Technikeinstellungen: gibt es die überhaupt? Ergebnisse einer Längsschnittanalyse von Bewertungen der Gentechnik.
Dieter Urban
- No.3/1998 Zur Interaktion zwischen Befragten und Erhebungsinstrument. Eine Untersuchung zur Konstanz des Meinungsurteils von Befragten im Interviewverlauf.
Martin Slaby
- No.1/1999 Role Models and Trust in Socio-Political Institutions: A Case Study in Eastern Germany, 1992-96.
Joachim Singelmann, Toby A. Ten Ayck, Dieter Urban
- No.1/2000 Die Zufriedenheit von Stuttgarter Studierenden mit ihrer Lebens- und Wohnsituation. Erste deskriptive Ergebnisse einer sozialwissenschaftlichen Studie zu allgemeinen und bereichsspezifischen Zufriedenheiten der Studierenden des Campus Vaihingen und des Campus Hohenheim.
Projektgruppe Campus: Slaby, M.; Grund, R.; Mayerl, J.; Noak, T.; Payk, B.; Sellke, P.; Urban, D.; Zudrell, I.
- No.2/2000 Längsschnittanalysen mit latenten Wachstumskurvenmodellen in der politischen Sozialisationsforschung.
Dieter Urban
- No.1/2001 Unser „wir“ - ein systemtheoretisches Modell von Gruppenidentitäten.
Jan A. Fuhse
- No.2/2001 Differentielle Technikakzeptanz, oder: Nicht immer führt die Ablehnung einer Technik auch zur Ablehnung ihrer Anwendungen.
Eine nutzentheoretische und modell-statistische Analyse.
Martin Slaby, Dieter Urban
- No.3/2001 Religiosität und Profession. Longitudinale Analysen zur Entwicklung des religiösen Selbstbildes bei Erzieherinnen.
Heiko Lindhorst

(Fortsetzung ...)

bisher sind erschienen (Fortsetzung):

- No.4/2001 Ist Glück ein affektiver Sozialindikator subjektiven Wohlbefindens?
Dimensionen des subjektiven Wohlbefindens und die Differenz
zwischen Glück und Zufriedenheit.
Jochen Mayerl
- No.1/2002 Risikoakzeptanz als individuelle Entscheidung.
Zur Integration der Risikoanalyse in die nutzentheoretische
Entscheidungs- und Einstellungsforschung.
Martin Slaby, Dieter Urban
- No.2/2002 Vertrauen und Risikoakzeptanz. Zur Relevanz von Vertrauen
bei der Bewertung neuer Technologien.
Martin Slaby, Dieter Urban
- No.3/2002 Probleme bei der Messung individueller Veränderungsdaten.
13 empirisch und methodisch induzierte Effekte, die es schwierig machen,
Veränderungen von generalisierten Bewertungen zu ermitteln.
Dieter Urban
- No.1/2003 Systeme, Netzwerke, Identitäten. Die Konstitution sozialer Grenzziehungen
am Beispiel amerikanischer Straßengangs.
Jan A. Fuhse
- No.2/2003 Können Nonattitudes durch die Messung von Antwortreaktionszeiten ermittelt werden?
Eine empirische Analyse computergestützter Telefoninterviews.
Jochen Mayerl
- No.1/2004 Erhöht ein Opfer-Täter-Zyklus das Risiko, Sexualstraftaten als pädosexuelle Straftaten zu
begehen? Ergebnisse einer ereignisanalytischen Pilotstudie
Dieter Urban, Heiko Lindhorst
- No.1/2005 Persönliche Netzwerke in der Systemtheorie
Jan A. Fuhse
- No.2/2005 Analyzing cognitive processes in CATI-Surveys with response latencies:
An empirical evaluation of the consequences of using different
baseline speed measures.
Jochen Mayerl, Piet Sellke, Dieter Urban
- No.1/2006 Ist Bildung gleich Bildung? Der Einfluss von Schulbildung auf ausländerablehnende
Einstellungen in verschiedenen Alterskohorten.
Klaus Hadwiger
- No.2/2006 Zur soziologischen Erklärung individuellen Geldspendens.
Eine Anwendung und Erweiterung der Theory of Reasoned Action unter Verwendung von
Antwortlatenzzeiten in einem Mediator-Moderator-Design.
Jochen Mayerl
- No.1/2007 Antwortlatenzzeiten in TRA-Modellen. Zur statistischen Erklärung von (Geld)-
Spendenverhalten.
Dieter Urban, Jochen Mayerl
- No.1/2008 Berufseintritt und Berufssituation von Soziologieabsolventen der Universität Stuttgart.
Deskriptive Ergebnisse einer Absolventenbefragung aus dem Jahr 2007.
Jochen Mayerl, Dieter Urban

(Fortsetzung ...)

SISS: Schriftenreihe des Instituts für Sozialwissenschaften der Universität Stuttgart
<http://www.uni-stuttgart.de/soz/institut/forschung/siss.html>

bisher sind erschienen (Fortsetzung):

- No.1/2010 Der Bystander-Effekt in alltäglichen Hilfsituationen:
Ein nicht-reaktives Feldexperiment.
Katrin Alle, Jochen Mayerl
- No.2/2010 Das Working-Poor-Problem in Deutschland.
Empirische Analysen zu den Ursachen von Armut trotz Erwerbstätigkeit.
Leonie Hellmuth, Dieter Urban
- No.3/2010 Binär-logistische Regressionsanalyse.
Grundlagen und Anwendung für Sozialwissenschaftler.
Jochen Mayerl, Dieter Urban
- No.1/2011 Pädosexuelle Viktimisierung und pädosexuelle Straffälligkeit. Ein „harter“ empirischer
Hypothesentest.
Dieter Urban, Joachim Fiebig
- No.1/2012 Meta-Analyse und "publication bias". Eine empirisch-statistische Fallstudie.
Joachim Fiebig, Gerrit Rüdibusch, Dieter Urban