



universität  
wien

# MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Der Zusammenhang zwischen Geschwisterreihenfolge  
und Intelligenz: Eine multilevel Metaanalyse“

verfasst von / submitted by

Theresa Pichler, BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of  
Master of Science (MSc)

Wien, 2023 / Vienna 2023

Studienkennzahl lt. Studienblatt /  
degree programme code as it appears on  
the student record sheet:

UA 066 840

Studienrichtung lt. Studienblatt /  
degree programme as it appears on  
the student record sheet:

Masterstudium Psychologie UG2002

Betreut von / Supervisor:

Ass.-Prof. Mag. Dr. Jakob Pietschnig, Privatdoz.

### **Danksagung**

An dieser Stelle möchte ich mich bei allen Personen bedanken, die mich bei der Erstellung dieser Masterarbeit sowie in meinem gesamten Studium immer begleitet und unterstützt haben.

Vor allem möchte ich mich hier bei meinem Betreuer Ass.-Prof. Mag. Dr. Jakob Pietschnig bedanken, der von der Themenfindung bis hin zur fertigen Arbeit jederzeit ein offenes Ohr hatte und mir mit vielen Antworten zur Seite stand. Ohne die immer freundliche und kompetente Unterstützung hätte ich diese Arbeit nicht schreiben können.

Ein riesiges Dankeschön geht auch an Dany, welche das Lektorat übernommen hat und somit viele Teile der Arbeit lesbarer und verständlicher gemacht hat.

Ich möchte mich auch bei meiner Familie und meinen Freunden, vor allem bei Benedikt, Mathias und meiner Mama bedanken – ohne ihre Hilfe beim Durchlesen, Programmieren und der emotionalen Unterstützung wäre alles nur halb so gut geworden.

Danke!

## Inhaltsverzeichnis

Einleitung & theoretischer Hintergrund .....	5
Definitionen.....	5
Intelligenz .....	5
Geschwisterreihenfolge .....	6
Modelle.....	8
Confluence Modell .....	8
Resource dilution Modell .....	9
Admixture Hypothese und andere Einflussfaktoren.....	10
Bisherige Forschungsergebnisse .....	12
Fragestellungen und Hypothesen .....	14
Methode.....	16
Definitionen.....	16
Literaturrecherche .....	17
Ein- und Ausschlusskriterien.....	17
Suchstrategie.....	18
Kodierung .....	19
Analyse.....	21
Hauptmodelle.....	22
Moderatoranalysen .....	23
Publication Bias Methoden und Sensitivitätsanalysen .....	24
Sterne und Eggers Regression.....	25
Rank Correlation .....	25
p-curve.....	25
p-uniform.....	26
p-uniform* .....	26
Trim-and-Fill .....	26
Leave-One-Out.....	26
Ergebnisse .....	27
Deskriptive Beschreibung der inkludierten Studien.....	27

Hauptanalysen .....	31
Modell ohne Rauschen und Hypothese Eins .....	31
Umfassendes Modell und Hypothese Zwei .....	36
Familiengröße Modell und Hypothese Drei .....	40
Moderatoranalysen (Hypothesen Vier bis Acht) .....	45
Funnel Plot Symmetrie .....	47
Methoden mit p-Werten.....	53
Sensitivitätsanalysen.....	53
Diskussion .....	61
Hauptanalysen .....	61
Moderatoranalysen .....	63
Publication Bias.....	65
Limitationen und Implikationen für zukünftige Forschung .....	66
Conclusio.....	69
Literaturverzeichnis.....	70
Abbildungsverzeichnis .....	80
Tabellenverzeichnis.....	82
Anhang .....	83
Zusammenfassung .....	83
Abstract .....	84
Veränderungen zur Präregistrierung .....	85
Weitere graphische Darstellungen.....	86

### **Einleitung & theoretischer Hintergrund**

Im Jahr 2021 lebten in Österreich etwa 78 % der Kinder unter 17 Jahre mindestens mit einem weiteren Geschwisterkind im gleichen Haushalt. Nur 22 % waren Einzelkinder (Kaindl & Schipfer, 2022). Geschwister hat demnach fast jede\*r; sie sind großer Bestandteil des persönlichen Lebens.

Die Reihenfolge, in denen die Kinder in einer Familie geboren werden, also die Geschwisterreihenfolge, ist eine präsente Thematik. In der Vergangenheit wurde diese immer wieder mit unterschiedlichen Eigenschaften oder persönlichen Leistungen in Zusammenhang gebracht - so auch mit der kognitiven intellektuellen Leistung. Schon Galton beschrieb 1874 erstmals, dass die bedeutendsten Forscher\*innen häufiger Erstgeborene als Spätergeborene sind. Seitdem wird immer wieder versucht, den Geburtsrang und somit die Position in der Reihenfolge der Geschwister und dessen Einfluss auf Intelligenz empirisch zu beschreiben (Steelmann et al., 2002). Die Ergebnisse bisher waren jedoch immer wieder widersprüchlich. Viele Forscher\*innen sammelten Daten und kamen zu der Erkenntnis, dass die Beziehung zwischen den beiden Variablen extrem wichtig sei (z.B. Breland, 1974); andere wiederum konnten durch die Analyse ihrer Daten feststellen, dass der Geburtenrang nichts mit Intelligenz zu tun habe (z.B. Steelmann & Doby, 1983). Da die Thematik alle Personen mit Geschwistern betrifft, waren die Diskussionen bisher häufig emotional aufgeladen (Rodgers, 2001).

In dieser Arbeit soll erstmals eine umfassende Metaanalyse all der vorliegenden Studien durchgeführt werden. Dadurch können die empirischen Befunde eingeordnet und eine konkrete Zusammenfassung des Zusammenhangs von Intelligenz und Geburtenrang geliefert werden.

### **Definitionen**

Um die Thematik besser verstehen zu können, werden im Folgenden die zugrunde liegenden Konstrukte näher beschrieben und definiert.

#### ***Intelligenz***

Bis heute ist die Definition von Intelligenz nicht einheitlich. Weit verbreitet ist die CHC-Theorie nach Cattell, Horn und Carroll, welche die kognitive Fähigkeit einer Person durch mehrere flexible Ebenen oder Levels beschreiben. Das hierarchisch am höchsten stehende Level ist die allgemeine und generelle kognitive Fähigkeit, der g-Faktor. Viele unterschiedliche Testbatterien ergeben eine hohe Korrelation zwischen den gemessenen g-Faktoren. Die generelle Intelligenz beschreibt etwa die Hälfte an Varianz der kognitiven

Fähigkeiten. Als Maß für den g-Faktor wird häufig der Gesamt-IQ (Intelligenzquotient) gesehen (Deary, 2012; Schneider & McGrew, 2018).

Die unter der allgemeinen Intelligenz liegenden Level werden unterteilt in verschiedene Gruppen an hoch korrelierten kognitiven Funktionen. Die Anzahl, Namen und Inhalte dieser breiteren oder auch schmäleren Domänen variieren je nach Autor\*in oder IQ-Test. Einige häufig genannte Gruppen sind fluide und kristalline Intelligenz, Gedächtnis, Wahrnehmung, Visualisierung, verbale Fähigkeiten oder schlussfolgerndes Denken. Auf dem untersten Level befinden sich testspezifische Faktoren und Fähigkeiten. Diese können im Gegensatz zu den höheren Ebenen des Modells direkt beim Intelligenztest erfasst und gemessen werden (Deary, 2012; Schneider & McGrew, 2018).

Die Messung der Intelligenz kann so divers wie deren Definition erfolgen. Viele Forscher\*innen verwenden objektive standardisierte Intelligenztests, um die kognitiven Fähigkeiten der Testpersonen zu erfassen (Steelmann, 1985). Dabei spielen vor allem IQ-Skalen als Maßstab der generellen Intelligenz, wie z.B. in Belmont und Marolla (1973), Kanazawa (2012) oder Thurstone und Jenkins (1929) eine wichtige Rolle. In anderen Studien wird diese Gesamtintelligenz in verbale und quantitative oder mathematische Fähigkeiten zerteilt (beispielsweise in Damian & Roberts, 2015 oder Altus, 1965). Einen anderen Zugang zur Erfassung von Intelligenz im Sinne von schulischen Leistungen bietet der SAT (*Scholastic Aptitude Test*), welcher z.B. in Paulhus und Shaffer (1981) oder Zajonc und Bargh (1980a) angewandt wurde.

Intelligenz ist ein Prädiktor für Bildungsstand und Bildungserfolg (Deary, 2012). Um auf Intelligenz schließen zu können, wird so von manchen Forscher\*innen der akademische Erfolg, wie zum Beispiel die Anzahl der absolvierten Schuljahre, die Höhe des Bildungsgrades oder der Notendurchschnitt (*Grade Point Average*, GPA) herangezogen. In einigen Fällen schätzen Lehrpersonen die Fähigkeiten der Schüler\*innen. Diese Bewertung fällt jedoch sehr subjektiv aus. Nachteil der erwähnten Messmöglichkeiten ist die hohe Variation der Bildungsstandards über verschiedene Länder und Kulturen hinweg. In der Thematik der Intelligenzmessung wird zunehmend auf standardisierte Testverfahren zurückgegriffen (Steelmann, 1985).

### ***Geschwisterreihenfolge***

Familien kann man mithilfe unterschiedlichster Variablen beschreiben. Hierbei spielt die Geschwisterkonstellation bzw. Geschwisterkonfigurierung eine wesentliche Rolle. Diese kann mithilfe der Anzahl der Kinder als Familiengröße zusammengefasst werden. Die Geburtenfolge in einer Familie hingegen ist die ordinale Position in der Altersreihenfolge der

Geschwister. Jeder Mensch besitzt einen ganz spezifischen Geburtsrang, welcher zudem einfach zu erheben ist. Die meisten Personen wissen, ob sie Erst-, Zweit-, Dritt-, Viert- oder Spätergeborene\*r sind (Steelmann et al., 2002).

Geburtenreihenfolge ist im Alltag sehr gegenwärtig. Eltern stellen beispielsweise ihre Kinder mit „Das ist unser ältestes Kind“ vor. Die älteren und jüngeren Geschwister innerhalb einer Familie werden häufig miteinander verglichen, egal ob in schulischen Situationen („Deine ältere Schwester war besser in Mathematik als du.“) oder anderen Leistungen. Der Geburtenrang ist nicht nur oft präsent, sondern verändert sich für die meisten Menschen im Laufe ihres Lebens nicht (Rodgers, 2001).

Die Geburtenreihenfolge bietet für Eltern oder Mitmenschen eine stabile Erklärung für die Eigenschaften oder das Verhalten eines Kindes. Dass Eltern gerne diese Variable als Begründung heranziehen, liegt oft daran, dass sie keine direkte Kontrolle oder Verantwortung dafür haben, welchen Rang ihr Nachwuchs hat. Daher lässt sich schnell und einfach die Schuld darauf schieben („Sie ist so emotional, weil sie ein Einzelkind ist.“ oder „Er musste als Jüngster in der Familie immer zurückstecken, daher erreicht er seine Ziele nicht.“). Die zugeschriebenen Eigenschaften, welche vielleicht innerhalb der Familie beobachtet werden, werden auf andere Familien projiziert und somit generalisiert (Rodgers, 2001).

Der Rang der Geschwister und die generelle Anzahl der Kinder in einer Familie sind stark miteinander verbunden. Die Korrelation dieser beiden Variablen ist demnach hoch; nach Steelmann (1985) liege diese bei zirka  $r = .65$ . Aus diesem Grund sollte bei Untersuchung der einen Variable auch die andere berücksichtigt werden. Zur Verdeutlichung dieses Zusammenhangs könne als Beispiel der Vergleich zwischen Erstgeborenen und Viertgeborenen dienen. Die erstgeborene Person kann von jeder Familie mit Kindern stammen, wobei jemand mit Geburtsrang Vier aus einer Familie mit zumindest vier Geschwisterteilen sein muss. Bei einem Vergleich der Intelligenz unterschiedlicher Geburtenränge werden also kleine Familien großen gegenübergestellt, was zu Fehlinterpretationen führen kann (Steelmann, 1985).

Der Unterschied zwischen Familiengröße und Geschwisterreihenfolge liegt darin, dass die erste Variable Kinder verschiedener Familien und die zweite Geschwister innerhalb einer Familie unterscheidet (Steelmann, 1985).

Nicht nur Intelligenz soll mit der Reihenfolge der Geburt zusammenhängen, auch andere Faktoren wurden immer wieder auf deren Korrelation mit der Geburtenreihenfolge untersucht. So soll diese Einfluss auf die Wahrscheinlichkeit krimineller Aktivitäten und Verhaltensstörungen haben oder den akademischen und ökonomischen Status vorhersagen

(Steelman et al., 2002). Außerdem soll der Geburtsrang im Zusammenhang mit Alkoholkonsum (Smart, 1963), Allergien (Cullinan, 2006) oder Krebserkrankungen (Hemminki & Mutanen, 2001) stehen.

### **Modelle**

Der *Birth order* Effekt auf Intelligenz beschreibt die geringere durchschnittliche Intelligenz von später geborenen Personen im Vergleich zu früher geborenen Geschwistern (Kanazawa, 2012).

Um den Zusammenhang zwischen Geschwisterreihenfolge und kognitive Leistung zu beschreiben, bezieht sich die Literatur meist auf drei verschiedene Modelle. Dabei haben das *Confluence* Modell und das *Resource dilution* Modell als Grundaussage gemeinsam, dass die Intelligenz mit zunehmender Geschwisterreihenfolge sinke. Die erstgeborenen Kinder sind demnach intelligenter als die Zweitgeborenen, und diese wiederum intelligenter als Drittgeborene. Dies steht im Gegensatz zu der *Admixture* Hypothese, bei der Geburtenreihenfolge kein solch ausgeprägter Einfluss zugesprochen wird (Kanazawa, 2012).

#### ***Confluence Modell***

Das *Confluence* Modell wurde erstmalig 1975 von Zajonc und Markus als Reaktion auf eine niederländische Studie mit großer Stichprobe (Belmont & Marolla, 1973) vorgestellt. Dieses Modell nimmt eine negative Korrelation zwischen dem Geburtsrang und der Intelligenz von Personen, bedingt durch die bei der Geburt gegebene familiäre Stimulation an. Die Intelligenz des Kindes wird durch das intellektuelle Umfeld, in dem es sich befindet, geformt. Dieses setzt sich zusammen aus dem ungewichteten Durchschnitt des kognitiven Levels der einzelnen Familienmitglieder (Steelman, 1985; Steelman et al., 2002).

Geht man von einer in Europa typischen Familiensituation aus, so werden Erstgeborene in einer Familie geboren, die meist zu 100 % aus kognitiv voll entwickelten Erwachsenen besteht. In den meisten Fällen setzt sich die Kernfamilie eines erstgeborenen Kindes aus keinen anderen Personen außer den Eltern zusammen. Ist ein Kind jedoch das Zweitgeborene, so gibt es in der Familie nur noch 67 % vollentwickelte Erwachsene, da das ältere Geschwisterkind zumeist selbst noch ein Kind ist. Drittgeborene werden später in eine Familie mit nur 50 % Erwachsenen hineingeboren. Je mehr Erwachsene sich in einer Familie befinden, desto höher ist nach dieser Theorie die kognitive Stimulation aus dem familiären Umfeld, was die kognitive Entwicklung eines Kindes erheblich fördert. Daraus folgt ein weniger stimulierendes Umfeld bei höherer Geburtenreihenfolge (Kanazawa, 2012; Zajonc & Markus, 1975).



Nach Zajonc und Markus (1975) ist der Effekt der Geburtenreihenfolge ein Nebenprodukt des Zeitabstands, in welchem die Geschwister einer Familie geboren werden. Kinder, die eng nacheinander geboren werden, haben eine weniger hohe Chance, einem intellektuell profitablen Umfeld ausgesetzt zu sein. Der Geburtenabstand sei geringer, wenn die Geburtenrate einer Familie höher ist und dies führe zu einem geringeren intellektuellen Niveau (Steelman, 1985).

Vier Jahre nach Veröffentlichung der Theorie erweiterten Zajonc, Markus und Markus (1979) das *Confluence* Modell um den sogenannten *Tutoring* Effekt. Diese Adaption erfolgte deshalb, da die Studienlage schnell, sowohl unterstützende, aber auch gegensätzliche Ergebnisse für die ursprüngliche Theorie hervorbrachte. Die Autor\*innen meinen, die Erstgeborenen werden durch die Geburt einer Schwester oder eines Bruders in ihrer kognitiven Entwicklung vorerst gehemmt (*Sleeper* Effekt), da durch die Geschwister das Level an intellektueller Stimulation in der Familie herabgesetzt werde. Mit etwa 11 bis 13 Jahren würden sich die früher Geborenen wieder von diesem Defizit „erholen“ und ihre Geschwister intellektuell „überholen“, indem sie den Jüngeren unterschiedliche Dinge beibringen und so zu einer Art Tutor\*in werden. Das ist wiederum für den\*die Tutor\*in förderlicher als für den\*die Belehrt\*en. Diesem Effekt nach, hängt also die Korrelation von Geburtenrang und kognitiver Leistung auch mit dem Alter der getesteten Person zusammen. So finde sich dieser *Birth order* Effekt häufiger in Stichproben mit älteren Jugendlichen als mit Kindern (Kanazawa, 2012; Steelman, 1985; Steelman et al., 2002).

Die *Confluence* Theorie mache laut einigen Forscher\*innen theoretisch zwar Sinn, kann aber in der Praxis durch viele Datensätze nicht bestätigt werden – ganz im Gegensatz zu dem *Resource dilution* Modell, welches bestehende Daten gut erklärt (Steelman et al., 2002).

### ***Resource dilution Modell***

Ähnlich dem *Confluence* Modell, wird auch im *Resource dilution* Modell nach Blake (1981) davon ausgegangen, dass Spätergeborene eine geringere Intelligenz als ihre Geschwister aufweisen. Die Erklärung dafür sieht das Modell in der Verteilung der elterlichen Ressourcen innerhalb der Familie. Zu diesen zählen vor allem materielle Ressourcen, aber auch die Menge an Energie und ungeteilter Aufmerksamkeit, welche die Eltern für ihre Kinder aufbringen können. Weitere wichtige Ressourcen in dem Kontext sind beispielsweise die (interpersonale oder finanzielle) Förderung zur Verwirklichung schulischer Ziele, Investitionen in Bildungsmaterialien (Bücher, Schreibtisch als Lernort, Spiele) oder die Involvierung in den schulischen Alltag (Downey, 2001; Kanazawa, 2012; Steelman, 1985; Steelman et al., 2002).

Diese Ressourcen sind limitiert und werden unter dem Nachwuchs aufgeteilt. Ist jemand Erstgeborene\*r, so kommen dieser Person alle elterlichen Ressourcen zugute. Je später ein Kind also geboren wird, auf desto mehr Geschwister werden diese aufgeteilt, wodurch für jede\*n weniger zur Verfügung steht. Kinder mit höherem Geburtenrang leiden dem Modell nach an der Knappheit an zugeteilten Ressourcen, wodurch auch die kognitive Entwicklung eingeschränkt wird. Beispielsweise eine verringerte verbale Förderung durch Kommunikation mit Erwachsenen kann in einem Haushalt mit vielen Kindern für das jeweilige Kind kontraproduktiv wirken (Kanazawa, 2012; Steelman, 1985).

Die der *Resource dilution* Theorie zugrundeliegenden Effekte können auf die Anzahl der Geschwister zurückzuführen sein – je mehr Kinder in einer Familie, auf desto mehr Personen werden die Ressourcen aufgeteilt, wodurch jede\*r Einzelne weniger bekommt (Downey, 2001; Steelman, 1985; Steelman et al., 2002).

Die Verteilung der Mittel in der Familie hängt von der Art der Ressourcen ab. Der Anteil an erfahrender Zuwendung durch die Eltern sinkt linear mit der Anzahl an Geschwistern; ökonomische Mittel wie die Bezahlung der Ausbildung werden eher nach dem Muster  $1/\text{Anzahl der Kinder}$  aufgeteilt. Bei anderen Mitteln wie den Bildungsressourcen (z.B. Computer) sinkt die Verteilung bis zu einer Art unteren Schwelle. Die Familiengröße hat dabei zwar einen negativen Effekt auf den Anteil – jedoch nur bis zu einem bestimmten Punkt. Das *Resource dilution* Modell geht von familiären Mitteln aus, die förderlich für den Nachwuchs sind, dennoch können sich einige dieser auch negativ auf die Kinder auswirken (beispielsweise Missbrauch durch die Eltern). Dabei kann eine größere Familie förderlich sein, da der\*die Einzelne weniger abbekommt (Downey, 1995).

### ***Admixture Hypothese und andere Einflussfaktoren***

Nur einige Jahre nach Veröffentlichung des *Confluence* Modells publizierten Page und Grandon (1979) ihre *Admixture* Hypothese. Hierbei wird postuliert, der *Birth order* Effekt auf Intelligenz werde nur durch methodische Fehlinterpretationen sichtbar. Das Problem hierbei sei, dass zwischen-familiäre Daten verwendet werden um inner-familiäre Unterschiede zu beschreiben. Zudem liege die richtige Erklärung für den Zusammenhang außerhalb der Familie bei anderen Faktoren (Rodgers, 2001).

Eine mögliche Erklärung der Korrelation sieht die Hypothese in dem Fakt, dass Personen mit geringeren Intelligenzscore grundsätzlich eine größere Anzahl an Kindern bekommen und daher auch häufiger Nachwuchs mit höherem Geburtenrang haben. Da Intelligenz zu einem Teil vererbbar ist (Deary, 2012), haben diese Kinder nun auch eine eher geringere Intelligenz. Intelligentere Eltern, welche akademischem Erfolg eine höhere Stellung

einräumen, bekommen eher erst im späteren Alter Kinder und dadurch auch meist eine geringere Anzahl. Es zeigte sich in der Vergangenheit eine negative Korrelation zwischen elterlicher Intelligenz und Anzahl an Kindern (Kanazawa, 2012; Mercy & Steelman, 1982; Steelman, 1985).

Große Familien leben zumeist in weniger gut entwickelten Ländern, daher könnte ein weiterer Einflussfaktor auf den Zusammenhang zwischen kognitiver Leistung und Geschwisterrang die geographische Lage der Stichprobe einer Studie sein. Die meisten Studien werden in den USA oder im westlichen Europa durchgeführt. Unterschiedliche Kontinente und Länder unterscheiden sich erheblich in ihrem familiären Zusammenleben, im Zugang zu Bildung und dem Stand der wissenschaftlichen Möglichkeiten zur Testung von Intelligenz (Steeleman et al., 2002).

Mehr Nachwuchs zu haben, kann in Entwicklungsländern einen Vorteil darstellen, da dieser im Haushalt und der täglichen Arbeit als wichtige Arbeitskraft gilt. Nach der *Life history* Theorie leben sozioökonomisch schwächere Menschen nach der schnelleren „r-Strategie“. Dabei erhöht eine hohe Reproduktionsrate die Überlebenschancen der Familie in der unsicheren Umwelt mit niedriger Lebenserwartung (Sear, 2020). In westlicheren Regionen hingegen stellen viele Kinder häufig eine zusätzliche finanzielle Belastung dar (Steeleman et al., 2002).

Der Lebensort und der sozioökonomische Status gehen Hand in Hand. Der sozioökonomischen Status hat einen signifikanten Einfluss auf diverse Familienstrukturvariablen, darunter auch die Familiengröße. Familien mit eher niedrigeren Status haben häufig mehr Kinder als Familien höherer Klasse (O’Shields, 1980; Steelman, 1985).

Grundsätzlich spielt die Familiengröße eine wesentliche Rolle für die Höhe des Geburtsranges. Personen mit einer höheren Ordnungszahl müssen aus größeren Familien stammen, während kleinere Familien in den niedrigeren Geburtsrängen überrepräsentiert sind. So finden sich Personen mit dem Geburtsrang Vier nur in Familien mit mindestens vier Geschwistern, während Erstgeborene aus Familien mit einem, zwei oder sogar sechs Kindern stammen können. Erstgeborene gibt es in jeder Familie mit Kindern. Da vor allem in der westlichen Region der Welt die Familiengröße eher kleiner ist, ist die Zahl der Erst- oder Zweitgeborenen deutlich überrepräsentiert (Kanazawa, 2012).

Verwendet man in Datenanalysen nun zwischen-familiäre Daten, so werden Spätergeborene aus großen Familien mit Frühgeborenen aus meist eher kleinen Familien verglichen. In der *Admixture* Hypothese wird also davon ausgegangen, dass Kinder mit egal

welchem Geburtsrang von kleineren Familien im Durchschnitt intelligenter sind als Kinder mit egal welchem Geburtenrang aus größeren Familien. Unterschiede liegen nicht an der Geburtsreihenfolge, sondern an der Familiengröße. Wird für diese statistisch kontrolliert, verschwinde die negative Korrelation zwischen Geburtenreihenfolge und Intelligenz. Dies steht im Widerspruch zu den *Confluence* und *Resource dilution* Modellen, bei denen argumentiert wird, dass später Geborene innerhalb derselben Familie weniger intelligent sind als ihre älteren Geschwister (Kanazawa, 2012; Rodgers, 2001).

Generell sind die Korrelationen bei Familiengröße und Intelligenz häufig größer als jene bei Geburtenreihung mit derselben Variable. Bei zweiterer können diese Zusammenhänge in die negative und positive Richtung gehen (Steelmann et al., 2002).

Je nach Theorie gibt es noch einige andere Faktoren, die Einfluss auf den Zusammenhang zwischen Geschwisterreihenfolge und Intelligenz haben können. Dazu zählen beispielsweise Eigenschaften der Testpersonen wie die Ethnie (O'Shields, 1980; Page & Grandon, 1979), das Alter (Zajonc, Markus & Markus, 1979; siehe auch *Confluence* Modell und *Sleeper* Effekt) oder das Geschlecht (Paulhus & Shaffer, 1981; Steelmann & Mercy, 1983). Söhnen werden meist mehr Möglichkeiten zur schulischen Weiterbildung und somit mehr mit dieser im Zusammenhang stehenden familiären Ressourcen als Töchtern zugesprochen (Alexander & Eckland, 1974). Auch andere Faktoren in der Familie wie z.B. das Alter der Mutter bei Geburt des Kindes, fehlende Elternteile oder familiäre Zerrüttung werden als Einflussfaktoren der Intelligenz genannt (Zajonc, 1976).

### **Bisherige Forschungsergebnisse**

Die zahlreichen bisher vorliegenden Forschungsarbeiten zur Thematik Intelligenz und Geburtenreihenfolge sind vielfältig. Die Studie von Belmont und Marolla (1973) wird häufig als eine der einflussreichsten aufgrund ihrer Größe angesehen. Die Autor\*innen verwendeten eine Stichprobe bestehend aus 380 000 niederländische Männer im Wehrdienst. Die Ergebnisse zeigen, dass mit größer werdendem Geburtenrang das Intelligenzlevel sinke (Belmont & Marolla, 1973).

Auch Breland (1974) konnte mittels Daten von knapp 800 000 US-amerikanischen Teilnehmer\*innen bestätigen, dass Erstgeborene höhere Scores erzielen. In anderen Regionen der Welt lieferten groß angelegte Studien ähnliche Ergebnisse, wie beispielsweise Cheng et al. (2012) in Taiwan oder Black et al. (2011) in Norwegen.

Findet sich eine negative Korrelation zwischen Geburtenreihenfolge und Intelligenz, so ist sie meist sehr klein. In vielen Studien wurde dieser negative Zusammenhang sogar als nicht signifikant eingestuft. (z.B. Eysenck & Cookson, 1969; Steelmann & Doby, 1983).

Einige Untersuchungen, beispielsweise in Frankreich (Zajonc & Bargh, 1980a), konnten eine gegensätzliche, positive Korrelation zwischen den Hauptvariablen ausfindig machen. Spätergeborene haben demnach eine höhere Intelligenz. Dies zeigte sich auch bei Singh (2009) in Asien.

Diese unterschiedlichen und konträren Resultate spiegeln sich in Untersuchungen wider, bei denen je nach Stichprobe oder angewandtem Intelligenztest verschiedene Korrelationen gefunden wurden. In Israel etwa ergaben sich je nach örtlicher Herkunft positive oder negative Zusammenhänge (Davis et al., 1977). In anderen Studien zeigten sich neben Unterschieden je nach Testart auch Unterschiede hinsichtlich Intelligenz und Geschwisterreihenfolge zwischen den Geschlechtern (Altus, 1965) oder Altersgruppen (Witt, 1978). Der *Birth order* Effekt konnte beispielsweise bei männlichen Teilnehmern nicht beim mathematischen (MAT), jedoch beim verbalen Test (VAT) bestätigt werden. Die weiblichen Erstgeborenen zeigten bei beiden Tests höhere Leistungen als die spätergeborenen Frauen (Altus, 1965). Konsistent über zwei Altersgruppen (Schüler\*innen und Student\*innen) ergab sich zwar eine positive Korrelation mit fluider Intelligenz und eine negative Korrelation mit kristalliner. Die Altersgruppen unterschieden sich jedoch hinsichtlich der Signifikanz der Ergebnisse. Nur der negative Zusammenhang mit kristalliner Intelligenz in der Gruppe der Schüler\*innen und die positive Korrelation mit fluider Intelligenz der Student\*innengruppe zeigten sich ausreichend signifikant (Witt, 1978).

Ein weiterer häufig genannter Einflussfaktor auf den *Birth order* Effekt ist der sozioökonomische Status. Teilweise variiert die Korrelation mit diesem oder ist durch diesen bedingt (O'Shields, 1980; McCall & Johnson, 1972); andererseits zeigte Barclay (2015) anhand einer großen schwedischen Stichprobe keine Moderation durch den sozioökonomischen Status. Man erkennt also auch hier wieder die Variation und Ungereimtheiten in der Thematik der Geschwisterreihenfolge.

Wie zuvor beschrieben, kann die Familiengröße ein relevanter Faktor bei der interessierenden Korrelation sein. Dies konnte in verschiedenen Arbeiten auch empirisch bestätigt werden. So hatte die Aufteilung nach oder die Kontrolle für die Anzahl der Kinder in einer Familie vor allem Auswirkungen auf die Stärke der Korrelation (Abdel-Khalek & Lynn, 2008; Damrin, 1949; Zajonc & Bargh, 1980a), was zum Teil zu einer Nullkorrelation führte (Husain et al., 2019; Kanazawa, 2012).

### Fragestellungen und Hypothesen

Durch die Zusammenfassung einiger wichtiger Studienergebnisse im vorherigen Kapitel wird deutlich, wie unterschiedlich und uneinheitlich die Resultate von Studien der Thematik Geschwisterreihenfolge und Intelligenz sind. Differenzen hinsichtlich sozioökonomischen Status, geographischer Lage, Geschlecht, Alter oder Familiengröße sind erkennbar. Aus diesen Ungereimtheiten werden die Hypothesen und Fragestellungen dieser Metaanalyse abgeleitet.

Um auf das *Confluence* Modell und das *Resource dilution* Modell einzugehen, wird getestet, ob eine negative Korrelation zwischen den Variablen Intelligenz und Geburtenrang besteht. Daraus ergibt sich die erste Hypothese.

H1: Es besteht eine signifikant negative Korrelation zwischen Intelligenz und Geschwisterreihenfolge.

In vielen Untersuchungen werden nur mittlere Intelligenzwerte für den jeweiligen Geburtsrang berichtet. Die Umrechnung der Durchschnittswerte in Korrelationen geschieht, indem die nach Stichprobengröße zusammengefassten Mittelwerte und Standardabweichungen für alle Spätergeborenen mit jenen für Erstgeborene verglichen werden. Die dadurch resultierenden Mittelwertsvergleiche (Cohens *d*) ergeben durch Verwendung der biserialen Korrelation Zusammenhangsmaße. Durch den Vergleich von Erst- mit Spätergeborenen entsteht ein gewisses Rauschen, da nicht mehr jeder Geburtenrang innerhalb einer Familie berücksichtigt, sondern nur ein zusammengefasster Wert für alle jüngeren Geschwister inkludiert wird. Um auf mögliche Auswirkungen des inkludierten Rauschens auf die Ergebnisse der Metaanalyse zu testen, wird während der Datensammlung mittels Dummy-Kodierung angemerkt, ob die Korrelation durch solch eine Methode berechnet wird. Aus diesen Überlegungen ergibt sich Hypothese Nummer Zwei.

H2: Der Zusammenhang zwischen Intelligenz und Geschwisterreihenfolge ändert sich bei Einbezug jener Korrelationen, die mittels Mittelwertsvergleichen zwischen Erst- und Spätergeborenen berechnet werden (höheres Rauschen).

Die zentrale Aussage der *Admixture* Hypothese ist, dass die Korrelation zwischen kognitiver Fähigkeit und Geburtenrang verschwindet, wenn der Einfluss der Familiengröße herausgenommen wird. Dies wird durch Hypothese Drei überprüft. In der vierten Hypothese wird zusätzlich die Moderation durch die mittlere Familiengröße der einzelnen Studien untersucht.

H3: Der Zusammenhang zwischen Intelligenz und Geschwisterreihenfolge ist signifikant kleiner bei Studien mit Kontrolle für Familiengröße als bei jenen, die nicht für die Geschwisteranzahl kontrollieren.

H4: Der Zusammenhang zwischen Intelligenz und Geschwisterreihenfolge wird durch die mittlere Familiengröße der Studie moderiert.

Um den Einfluss anderer Variablen im Sinne der *Admixture* Hypothese zu überprüfen, werden in Hypothese Fünf bezugnehmend auf die Stichprobenszusammensetzung (Geschlecht, durchschnittliches Alter und Altersgruppe) Moderationsanalysen durchgeführt.

H5a: Der Zusammenhang zwischen Intelligenz und Geschwisterreihenfolge wird moderiert durch die Geschlechterverteilung (Prozentanteil männlicher Teilnehmer).

H5b: Der Zusammenhang zwischen Intelligenz und Geschwisterreihenfolge wird moderiert durch das Durchschnittsalter der Studienteilnehmer\*innen.

H5c: Der Zusammenhang zwischen Intelligenz und Geschwisterreihenfolge wird moderiert durch die Altersgruppe der Studienteilnehmer\*innen (rein minderjährige Gruppe unter 18 Jahren, reine Gruppe bestehend aus Erwachsene oder gemischte Stichprobe).

Auch eine mögliche Moderation durch die geographische Lage der Datenerhebung wird als Hinweis auf den sozioökonomischen Status der Testpersonen in der sechsten Hypothese berücksichtigt.

H6: Der Zusammenhang zwischen Intelligenz und der Geschwisterreihenfolge variiert je nach geografischer Lage der Studie (Moderation durch Kontinent).

Zusätzlich wird explorativ untersucht, ob das Bruttoinlandsprodukt je Einwohner\*in (GDP per capita) in dem Land und dem Jahr der Datenerhebung die Korrelation von Intelligenz und Geschwisterreihenfolge moderiert.

Für die Kontrolle auf den Einfluss des Jahres der Veröffentlichung der Studie und auf den Status der Publikation werden ebenso Moderationsanalysen als erste Hinweise auf einen möglichen *Publication Bias* durchgeführt.

H7: Der Zusammenhang zwischen Intelligenz und der Geschwisterreihenfolge wird durch das Jahr der Veröffentlichung moderiert.

H8: Der Zusammenhang zwischen Intelligenz und der Geschwisterreihenfolge der veröffentlichten Studien ist höher als jener der unveröffentlichten Studien (Moderation durch Publikationsstatus).

### Methode

Um die vorliegenden Hypothesen zu überprüfen, wurden mehrere multilevel Metaanalysen durchgeführt. Der Untersuchungsplan wurde vorab auf <https://osf.io/8rnkz/> präregistriert.

#### Definitionen

Die Geburtenreihenfolge als erste Hauptvariable wird in dieser Arbeit als Zahl angesehen, welche bei Erstgeborenen oder Einzelkindern die Ziffer 1 annimmt, bei Zweitgeborenen die Ziffer 2, bei Drittgeborenen 3 und so weiter. Nach oben hin begrenzt wird dieser Wert nur durch die Anzahl der Kinder in einer Familie bzw. der Familiengröße. Berichteten die gefundenen Studien nur die Anzahl der älteren Geschwister, so wurde die Geburtenreihenfolge durch diese Anzahl plus 1 dokumentiert.

Die zweite Hauptvariable, Intelligenz, wird in dieser Metaanalyse definiert durch die Verwendung eines standardisierten Intelligenztestverfahrens. Um möglichst eine breite Abdeckung an Untersuchungen der Thematik zu erreichen, war es hierbei nicht relevant, ob dies spezifische Fähigkeitstests oder umfangreiche Testbatterien waren.

Um den Zusammenhang zwischen den zwei Hauptvariablen Intelligenz und Geschwisterreihenfolge zu beschreiben, wurde die Pearson Korrelation  $r$  als Effektstärke verwendet. Diese Größe verdeutlicht anschaulich die lineare Veränderung einer Variable bei Änderung einer anderen. Beide Variablen sind hierbei kontinuierlich. Eine negative Korrelation liegt dann vor, wenn bei Erhöhung der einen Variable (beispielsweise Geburtsrang) eine andere sinkt (hier Intelligenzscore). Der Koeffizient kann dabei Werte von 0 bis  $\pm 1$  annehmen, wobei nach Cohen ein  $r$  von  $\pm .10$  einen kleinen Effekt, ein  $r$  von  $\pm .30$  einen mittleren und ein  $r$  von  $\pm .50$  einen großen Effekt darstellt (Cohen, 1969).

Explorativ wurde das Bruttoinlandsprodukt in die Untersuchung mitaufgenommen. Dabei wurden von *The World Bank* das für das jeweilige Land und Jahr der Datenerhebung die Werte für GDP (*Gross Domestic Product*) per capita und GDP per capita PPP (*Purchasing Power Parity*) dokumentiert. GDP per capita (*current US\$*) ist das Bruttoinlandsprodukt pro Einwohner\*in und somit ein Maß für das durchschnittliche Einkommen oder die Wirtschaftsleistung. Zusätzlich zum allgemeinen GDP wurde auch das GDP per capita PPP (*current international \$*) erhoben, um nähere Informationen zum Lebensstandard bzw. sozioökonomischen Status zu bekommen. Hierbei werden auch die Kaufkraft der Währungen und die Lebenserhaltungskosten in jedem Land berücksichtigt (*World Bank Open Data*, 2023).



## Literaturrecherche

### *Ein- und Ausschlusskriterien*

Für den Einbezug von Studien zum Thema Intelligenz und Geburtenrang in vorliegender Metaanalyse, mussten einige Kriterien erfüllt werden. Der in den Studien verwendete Intelligenztest, das Alter oder Geschlecht der Testpersonen oder die geographische Lage oder Zeit der Datenerhebung stellten keine Einschränkungen dar.

Es wurden nur englisch- und deutschsprachige Untersuchungen berücksichtigt.

Um eine Generalisierbarkeit der Ergebnisse auf die Gesamtbevölkerung zu ermöglichen, sollte die beforschte Population gesund und demnach ohne psychische oder körperliche Störungen sein.

Sehr häufig werden Familienstrukturvariablen im Zusammenhang mit Bildungserfolgen wie zum Beispiel Schulnoten, den höchsten erreichten Bildungsabschluss oder der Anzahl an Schuljahren untersucht (Steelmann et al., 2002). Diese Methoden stehen im Kontrast zu standardisierten Tests und messen womöglich nicht Intelligenz im engeren Sinne. Aus diesem Grund wurden sie nicht in vorliegender Arbeit berücksichtigt. Ebenso wurden nicht die durch eine andere Person geschätzte Intelligenz (beispielsweise durch Lehrer\*innen) oder *Theory of Mind* Messungen miteinbezogen.

Durch theoretische Überlegungen wurden die Exklusionskriterien zusätzlich durch Mehrlingsstudien, vor allem Zwillingsstudien erweitert. Nach der *Confluence* Theorie würde eine gleichzeitige Geburt mehrerer Geschwister das intellektuelle Level der Gesamtfamilie stärker senken als eine einzelne Geburt. Außerdem senkt ein weiteres Baby die Zeit für Kommunikation mit den Erwachsenen in der Familie (ähnlich wie eine höhere Familiengröße nach dem *Resource dilution* Modell; Zajonc & Markus, 1975). Zudem gibt es bei Mehrlingen durch die fast gleichzeitige Geburt keine eindeutige Geburtenreihenfolge. Infolgedessen wurden ganze Studien oder Gruppen mit Zwillingen oder Drillingen nicht in die Analyse inkludiert.

Weiters wurden keine Samples analysiert, welche den Geburtsrang X oder höher haben, wie beispielsweise Gruppen, für diese Mittelwerte für die Dritt- und Spätergeborenen berichtet werden. Dieses Zusammenfassen der Ranggruppen hätte zu Verfälschungen durch Ungenauigkeit geführt.

Ein weiteres Ausschlusskriterium in dieser Arbeit sind Einzelkinder. Nach Cicirelli (1978) zeigen diese, also erstgeborene Menschen ohne Geschwister eine geringere Intelligenz als bei Erstgeborenen erwartet. Auch andere Untersuchungen bestätigen, dass Einzelkinder nicht signifikant besser abschneiden als Personen mit einem Bruder oder einer Schwester

(Rodgers, 2001). Dies könnte darauf zurückzuführen sein, dass Personen ohne Geschwister nicht die Chance haben, als Tutor\*in jemand anderen in der Familie etwas beizubringen (Zajonc & Markus, 1975). Um Verzerrungen durch Einzelkinder zu vermeiden, wurden auch diese Stichproben nicht in die Analyse inkludiert.

Um eine mögliche Beeinflussung des IQs durch eine experimentelle Intervention auszuschließen, wurden in jenen Fällen nur Daten ohne vorangegangene Manipulation mitaufgenommen.

Ebenso wurde auf die Hinzunahme von Follow-up-Gruppen verzichtet, um eventuelle Lerneffekte bei der wiederholten Untersuchung derselben Personen zu verhindern. Dies würde zu Abhängigkeiten der Daten führen.

Bei Studien mit gleichen Untersuchungsteilnehmer\*innen und Tests wurden in vorliegende Arbeit jene inkludiert, welche die umfassendste Beschreibung des Datensatzes im Sinne der größten Stichprobengröße und der letzten Aktualisierung bereitstellen. So wurde anstatt anderer Erhebungen mit den Daten der *National Health Examination Survey* aufgrund der detaillierteren Beschreibung und eines größeren *N* nur die Studie von Steelman und Doby (1983) inkludiert. Aus ähnlichen Gründen wurde nur der Datensatz zur *National Longitudinal Survey of Youth* (NLSY79) von Pavan (2016) stellvertretend für diese Daten dokumentiert. Ebenso eine Aktualisierung der Daten der *Indonesian Family Life Survey* (IFLS) lieferten Botzet et al. (2021), wodurch nur diese Studie inkludiert wurde. Die Untersuchung von Kanazawa (2012) und keine anderen Datensätze der *National Child Development Study* (NCDS) wurde miteinbezogen, da hierbei umfassende Ergebnisse für alle Altersgruppen publiziert wurden. Bei anderen exakt gleichen Stichproben wurde nur die jüngste Studie in die Analyse miteingenommen.

### **Suchstrategie**

Um passende Studien zu finden, wurden verschiedene Methoden angewandt. Einerseits wurden auf Datenbanken mittels eines im Vorhinein festgelegten Suchbegriffs recherchiert. Andererseits wurden für jedes inkludierte Paper eine Rückwärtssuche im Literaturverzeichnis durchgeführt. Zusätzlich wurde anhand des Ankerpapers von Zajonc und Markus (1975), welches erstmals die *Confluence* Theorie beschreibt, eine Vorwärtssuche gemacht. Dabei wurden alle neueren Studien, welche dieses Paper zitiert haben, durchsucht. Dies wurde auf den Plattformen *Google Scholar*, *Scopus* und *Web of Science* ausgeführt.

Die englischen Begriffe für Geburtenrang „birth order“ und „order of birth“ wurden mit Begriffen für Intelligenz wie „intelligence“, „IQ“ und „cognitive“ kombiniert um den Suchstring „(intelligence OR IQ OR cognitive) AND ("birth order" OR "order of birth")“ zu

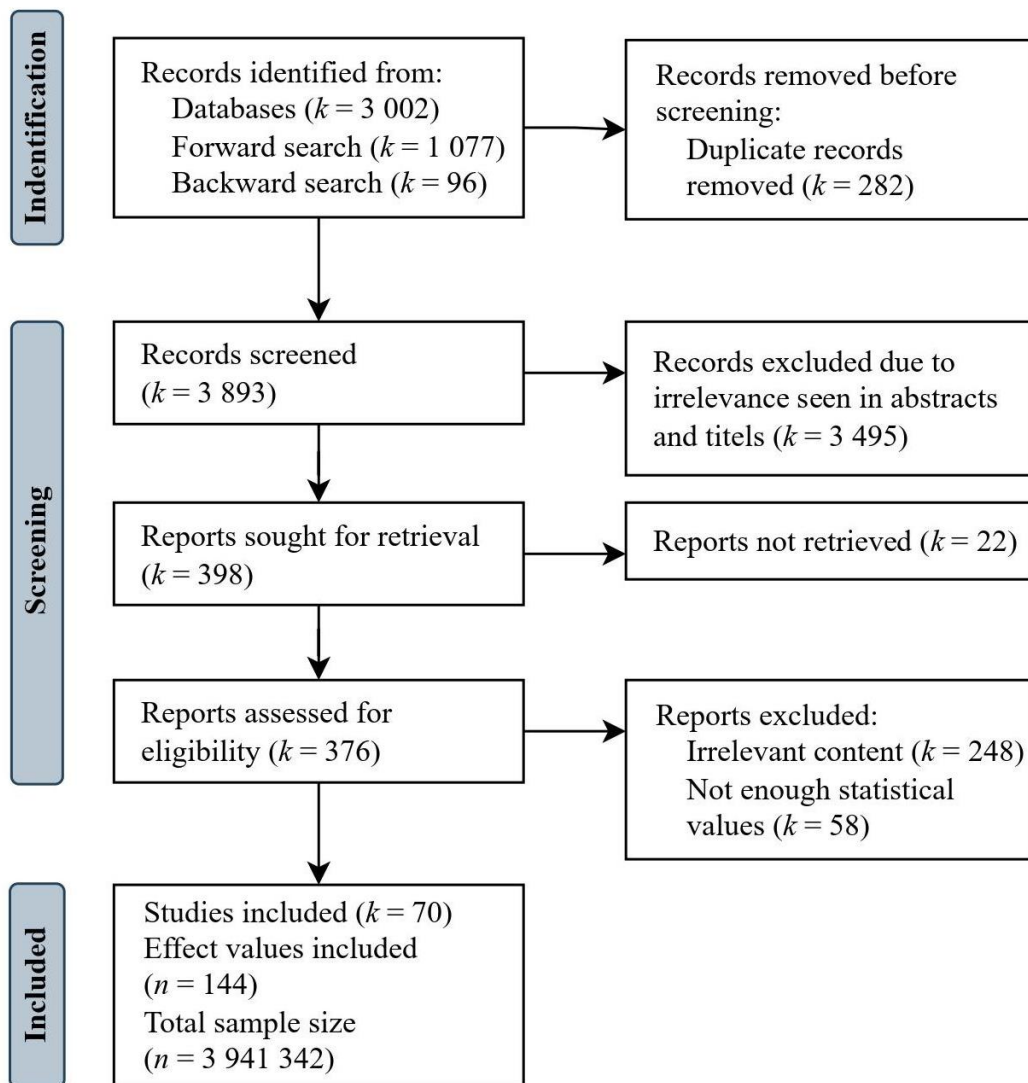
bilden. Dieser wurde auf den Datenbanken für publizierte oder auch graue Literatur *APA PsycInfo*, *Cochrane Library*, *ERIC*, *Open Access Theses and Dissertations*, *OpenGrey*, *ProQuest*, *PSYINDEX*, *PubMed*, *Scopus* und *Web of Science* angewandt. Die finale Prüfung auf weitere Studien fand im März 2023 statt.

In Abbildung 1 ist ersichtlich, wie viele Studien zu welchem Zeitpunkt und aus welchem Grund exkludiert oder inkludiert wurden. Die Autor\*innen jener Untersuchungen, die nicht abrufbar waren oder zu wenig statistische Werte berichteten, um Korrelationen zu extrahieren oder zu berechnen, wurden per E-Mail diesbezüglich kontaktiert. Zwei sowie vier Wochen nach Erstkontakt folgte eine Erinnerungsmail. Viele Autor\*innen berichteten, sie hätten die Datensätze nicht mehr, da die Studien schon lange zurück liegen. Aus diesem Grund oder wegen fehlender Rückmeldung bezüglich mangelnden Daten wurden 58 Studien exkludiert. 22 Untersuchungen waren auch nach Kontaktschreiben nicht abrufbar. Schlussendlich konnten 70 unterschiedliche Studien mit gesamt 144 Effektwerten in die Analyse miteinbezogen werden, was eine Gesamtstichprobengröße von etwa vier Millionen Proband\*innen ergab.

### **Kodierung**

Die Kodierung der inkludierten Studien erfolgte in Microsoft Excel. Dabei wurden folgende Variablen zwei Mal durch mich überprüft und dokumentiert:

- Studien-ID: Individuelle Kennzeichnung für jede Studie bestehend aus Erstautor und Publikationsjahr
- Sample-ID: Jede Stichprobe in einer Studie gekennzeichnet durch fortlaufende Zahlen
- Effekt-ID: Jede Effektstärke in einer Studie gekennzeichnet durch fortlaufende Zahlen
- Autor\*innen
- Publikationsjahr
- Publikationsstatus: Publiziert oder nicht publiziert
- Lage: Geographischer Ort bzw. Kontinent der Datenerhebung
- Ergebnis: Kurze Information über das Resultat der Studie hinsichtlich *Birth order* Effekt
- Datensatz: Name des Datensatzes
- Stichprobengröße  $N$
- Mittleres Alter angegeben in Jahren: Falls nur die Altersspanne angegeben war, wurde hier der Mittelpunkt dieser Spanne dokumentiert.
- Altersgruppe: Stichprobe bestehend rein aus Kindern, Erwachsenen oder aus beiden Gruppen
- Anteil männlicher Teilnehmer in Prozent: Maß für die Geschlechterverteilung

**Abbildung 1***Prisma Flow Diagramm zur Darstellung inkludierter und exkludierter Studien*

Last check for studies: 28.03.2023

- GDP per capita und GDP per capita PPP: Dabei wurde der jeweilige Wert für das Jahr und das Land der Datenerhebung entnommen. Falls kein genaues Erhebungsjahr angegeben wurde, wurde das GDP von dem Jahr vor der Publikation verwendet. Falls es für dieses Jahr keinen Wert gab (meist vor 1960 bzw. 1990) und der erste Eintrag nicht mehr als fünf Jahre vom Publikationsjahr entfernt liegt, so wurde dieser dokumentiert. Bei mehreren Jahren der Datenerhebung wurde das mittlere GDP über all diese Jahre gebildet.
- Intelligenztest: Testname

- Mittlerer Intelligenzscore inklusive Standardabweichung für jeden Geburtsrang: Die Standardabweichung *SD* wurde, falls nötig durch den Standardfehler *SE* berechnet:  

$$SD = SE * \sqrt{n}$$
- Mittlerer Geburtenrang und mittlere Familiengröße inklusive Verteilung (*n*): Falls der Mittelwert nicht angegeben wurde, wurde dieser mittels Gewichtung nach Gruppengröße berechnet.
- Pearson Korrelationskoeffizient *r* und falls angegeben Pearson *r* bei Kontrolle für Familiengröße inklusive *p*-Wert: Falls kein *r* angegeben war und keine Berechnung möglich ist, wurde ein *NR* (not reported) kodiert. Bei einem einfachen linearen Zusammenhang ist die standardisierte Steigung (beta) äquivalent zu *r*.
- Rauschen aufgrund der Berechnung der Korrelation durch den Vergleich von Erst- mit Spätergeborenen vorhanden: Ja oder Nein
- Ursprünglich berichteter Effektwert: Art und Wert
- Notizen zu Verständnis, Kodierung oder als Information

### Analyse

Die Berechnung fehlender Korrelationen und der Metaanalysen, sowie die Erstellung von Graphiken wurden mittels R Studio (RStudio Team, 2021) durchgeführt. In diesem Programm wurden die fehlenden Korrelationskoeffizienten mittels einer biserialen Korrelation berechnet, indem ein Mittelwertsvergleich (Cohens *d*) zwischen den Erstgeborenen und allen anderen Geburtsrängen (Spätergeborene) durchgeführt wurde. Um einen einzigen Mittelwert bzw. die Standardabweichung (*s*) für die Spätergeborenen zu erhalten, wurde das nach Gruppengröße *n* gewichtete Mittel berechnet. Dieses Vorgehen wurde mittels der Variable *Rauschen* gekennzeichnet, da hier nur durch zwei Punkte (Erst- und Spätergeborene) eine Korrelation berechnet wurde.

Die biserialen Korrelationen wurden verwendet, da diese die künstliche Dichotomisierung des Geburtsranges korrigiert. Sie nimmt eine kontinuierliche Verteilung hinter der dichotomisierten Variable an und kann in derselben Metaanalyse neben einer Produkt-Moment-Korrelation fehlerfrei verwendet werden (De Jonge et al., 2020; Jacobs & Viechtbauer, 2017).

Mittels Fishers *z*-Transformation wurden alle Korrelationswerte vor der Analyse in standardisierte *z*-Werte umgerechnet, um eine symmetrische Verteilung der Werte um den Populationsmittelwert zu erlangen. Dadurch können die Effektgrößen verschiedener Studien kombiniert und verglichen werden. Für den Bericht der Ergebnisse wurden die *z*-Werte wieder in *r* umgerechnet (Card, 2011).

### ***Hauptmodelle***

In dieser Arbeit wurde mit multilevel Metaanalysen gearbeitet. Diese können angewendet werden, wenn die Effektstärken hierarchisch organisiert bzw. ineinander verschachtelt und demnach abhängig voneinander sind (Cheung, 2014; Fernández-Castilla, 2020). Im Kontext dieser Arbeit wurden in einigen Studien mehrere Effektstärken berichtet. Diese sind innerhalb der Studien voneinander abhängig, da sie teilweise auf gleichen Stichproben oder verwendeten Intelligenztests beruhen. Um auf diese Abhängigkeiten einzugehen, wurde ein Modell mit mehreren Leveln für die Analyse verwendet. Auf dem ersten Level befinden sich nach dem Ansatz die Individuen und die Stichprobenvarianz für jede berichtete Effektstärke. Das zweite Level beschreibt die Effektstärken bzw. die Variation zwischen den einzelnen Korrelationen innerhalb der Studien. Das dritte und übergeordnete Level bezieht sich auf die einzelnen Studien und erklärt deren Varianz zwischen diesen. Demnach sind im vorhandenen Datensatz teilweise mehrere Effektstärken pro Studie kodiert.

Um eine gewisse Unabhängigkeit der Daten zu generieren, wurden aggregierte Datensätze gebildet, bei denen pro Stichprobe nur eine Korrelation berechnet wurde. Dies geschah, indem die *z*-transformierten Korrelationen mittels Gewichtung nach Stichprobengröße gemittelt wurden (Cheung, 2014). Diese Modelle wurden nur für jene Methoden verwendet, welche nicht für multilevel Daten anwendbar sind.

Für die Hypothesentestung wurden drei verschiedene Modelle verwendet. Diese bezogen sich auf die unterschiedlichen Korrelationskoeffizienten. Die Korrelation ohne Kontrolle für Familiengröße wurde in der ersten Hypothese und dem *Modell ohne Rauschen* sowie bei der zweiten Hypothese mit dem *umfassenden Modell* (Rauschen und kein Rauschen gemeinsam) untersucht. In der dritten Hypothese wurde die Korrelation mit Kontrolle für die Anzahl der Geschwisterkinder im sogenannten *Familiengröße-Modell* näher beleuchtet. Diese Kontrolle konnte durch partielle Korrelationen und Konstanthaltung der Familiengröße oder durch den genauen Bericht der mittleren Intelligenzwerte, aufgeteilt nach Gesamtanzahl der Kinder innerhalb einer Familie erfolgen.

Die Daten aller drei Modelle wurden mittels *Random Effects* Modellen analysiert, da sich die Primärstudien sowohl in der Methodik als auch in der untersuchten Population unterscheiden und somit systematische Unterschiede zwischen den Studien bestehen. *Random Effects* Modelle können gut auf den gesamten Forschungsbereich generalisiert werden (Card, 2011; Fernández-Castilla, 2020). In Studien, welche in diese Arbeit einbezogen wurden, verwendeten die Autor\*innen unterschiedliche Intelligenztests, bezogen verschiedene

Variablen mit ein, untersuchten diverse Alters-, Geschlechtergruppen oder Länder und prüften individuelle Thesen.

Da mit *Random Effects* Modellen gearbeitet wurde, wurde auch von einer hohen Heterogenität bzw. hohen Streuung zwischen den Studien ausgegangen. Eine heterogene Studienlage ist geprägt durch systematisch unterschiedliche und somit nicht vollständig vergleichbare Stichproben und suggeriert, dass die Variabilität der Effektgrößen größer ist als durch Schwankungen aufgrund Stichprobenunterschiede allein erklärt werden kann. Um auf Heterogenität und somit eine mögliche Generalisierbarkeit auf die Gesamtpopulation zu testen, wurde unter anderem Cochrans  $Q$  berichtet. Dieser Wert repräsentiert diese Heterogenität und mithilfe eines Signifikanztest kann getestet werden, wie die Effektstärken um deren Mittelwert verteilt sind. Liegt hierbei der Wert für  $p$  unter dem Signifikanzlevel von .05, so kann signifikante Heterogenität zwischen den Primärstudien angenommen werden (Card, 2011; Knippschild et al., 2015).  $\tau^2$  stellt eine ähnliche Prüfgröße für die Varianz zwischen den Studien dar. Dabei wird Cochrans  $Q$  auf einen Wertebereich  $\geq$  Null umgerechnet (Knippschild et al., 2015). Die prozentuale Verteilung der Variabilität über die Level der Metaanalyse wurde mittels  $I^2$  dargestellt. Ein Gesamtindex von 30 bis 60 % weist auf moderate Heterogenität hin; darunter ist diese zu vernachlässigen. Über 50 % ist dies ein Hinweis auf erhebliche und über 75 % auf beträchtliche Heterogenität (Higgins et al., 2022). Bei einer multilevel Metaanalyse wird diese Größe aufgeteilt in Variation zwischen (Level 3) und innerhalb (Level 2) der Studien.

Um für Hypothese zwei und drei die Vergleiche zwischen verschiedenen Modellen bzw. Metaanalysen und deren ergebnen Effektstärken zu erarbeiten durchzuführen, wurden Wald-Tests durchgeführt (Viechtbauer, 2023).

### ***Moderatoranalysen***

Um weitere mögliche Einflussfaktoren auf die Korrelation zwischen Intelligenz und Geschwisterreihenfolge zu identifizieren und auf Hypothesen des *Admixture* Modells einzugehen, wurden Moderatoranalysen durchgeführt. Dabei sollte die Erklärung der Heterogenität der Effektstärken durch Eigenschaften der Stichproben als Prädiktoren untersucht werden. Ziel war es hierbei, eine mögliche Variation der Korrelation der zwei Hauptvariablen über die verschiedenen Ausprägungen des Moderators zu erkennen (Card, 2011).

Je nach Art des Moderators wurden Subgruppenanalysen oder Metaregressionen ausgeführt. Erstere verwendet man bei kategorischen Moderatoren, wie in diesem Fall bei der Alterskategorie (Kinder, Erwachsene oder gemischte Stichprobe), Publikationsstatus

(publiziert oder nicht publiziert) und geographischer Lage (Kontinent) der Untersuchungen. Metaregressionen wiederum beziehen sich auf kontinuierliche Moderatoren wie hier die Geschlechterverteilung (Anteil männlicher Probanden), das mittlere Alter, das Publikationsjahr, die mittlerer Familiengröße und die beiden GDP-Werte. Signifikante Prädiktoren suggerieren eine lineare Variation entlang des Moderators. Es wurde überprüft, ob der Regressionskoeffizient der jeweiligen Moderatorvariable oder der Gruppenkategorie signifikant unterschiedlich zu Null ist und somit einen signifikanten Einfluss auf die Ergebnisse der Metaanalyse hat (Assink & Wibbelink, 2016; Card, 2011).

### ***Publication Bias Methoden und Sensitivitätsanalysen***

Der *Publication Bias* beschreibt eine Verzerrung der Ergebnisse einer Metaanalyse aufgrund fehlender Studien. Untersuchungen, die nicht signifikant sind und keinen Effekt finden oder ein unerwünschtes Ergebnis liefern, werden mit höherer Wahrscheinlichkeit nicht publiziert. Forscher\*innen sind bei diesen Resultaten häufig weniger motiviert und Zeitschriften akzeptieren solche Manuskripte seltener (Card, 2011). Aus diesem Grund könnten die publizierten Daten nicht repräsentativ für die gesamte Studienlage einer Thematik sein. Für Studien mit kleinen Stichproben ist es zudem aufgrund des Power-Problems schwierig, kleinere Effektstärken zu finden und zu erkennen. Daher sind diese häufig nicht signifikant. Durch einen *Publication Bias* können nicht nur Forscher\*innen, sondern auch Praktiker\*innen und andere interessierte Personen in die Irre geführt werden (Card, 2011; Dowdy et al., 2021).

Um diesen möglichen Fehler zu vermeiden, wurde in vorliegender Untersuchung auch graue, unveröffentlichte Literatur miteinbezogen. Leider ist diese schwer aufzuspüren und abzurufen. Für einen ersten Hinweis auf Unterschiede zwischen publizierten und nicht publizierten Daten, wurde eine Moderatoranalyse (siehe Hypothese Acht) durchgeführt (Dowdy et al., 2021). Zusätzlich wurde mittels verschiedener statistischer Methoden auf einen möglichen *Publication Bias* getestet. Diese werden aber nur anhand publizierter Studien und nicht mit Inklusion von grauer Literatur performt. Da multilevel Modelle bei Metaanalysen noch eine recht junge Möglichkeit für den Umgang mit abhängigen Daten sind, konnten die meisten der präregistrierten Methoden nicht klassisch angewendet werden. Um Abhilfe zu verschaffen, wurden diese Methoden mit den nach Stichproben aggregierten Modellen berechnet.

*Funnel Plots* sind ein Weg, einen möglichen *Publication Bias* graphisch darzustellen. Dabei werden die Datenpunkte als Punkt je nach Größe der Korrelation und Standardfehler eingeordnet. Umso kleiner die Stichprobe ist, desto größer ist der Standardfehler und die



Variabilität der Effektstärken. Aus diesem Grund sollten die einzelnen Daten bei einer umfangreichen Metaanalyse so verteilt sein, dass ein visueller Trichter (*Funnel*) um den gemeinsamen Effekt mit einer komprimierteren Ansammlung im oberen Bereich (größere Samples, kleinerer Standardfehler) entsteht. Wenn ein *Publication Bias* besteht, so kann eine visuelle Asymmetrie im *Funnel Plot* vorliegen (Card, 2011; Rodgers & Pustejovsky, 2021). Dabei zeigt sich häufig der sogenannte *small-study* Effekt, demnach werden bei kleineren Studien häufig nur größere Effekte als kleinere oder nicht-signifikante Ergebnissen publiziert. Dies führt zu einem geringeren Aufkommen der Studienpunkte im linken unteren Teil des *Funnel Plots*. Das Vorliegen eines *Publication Bias* ist jedoch nicht der einzige Grund für eine Asymmetrie, denn auch eine hohe Heterogenität oder gewisse Studiendesigns können das Aussehen der Graphik beeinflussen (Fernández-Castilla et al., 2021).

**Sterne und Eggers Regression.** Mit verschiedenen statistischen Methoden kann auf Asymmetrie des *Funnel Plots* getestet werden. Eine beliebte Methode für den Test auf Publikationsverzerrungen ist der Regressionstest nach Sterne und Egger (2000). Die ursprüngliche Testung besteht aus einer Regression der Effektstärken über den inversen Standardfehler. Gibt es hierbei ein signifikanter *Slope*, so ist dies ein Hinweis auf *Publication Bias*. Um bei einem multilevel Modell eine fehlerhafte Schätzung dieses Zusammenhangs aufgrund der Abhängigkeit der Daten zu vermeiden, wird der Standardfehler bzw. die quadrierte Varianz als Moderatorvariable mitaufgenommen (Egger MLMA). Ist diese Moderation signifikant, so sind die Studien nicht symmetrisch verteilt und es zeigt sich ein *Publication Bias*. Dieses Vorgehen hat leichte Vorteile hinsichtlich der Power im Vergleich zu anderen Variationen der Sterne und Egger Regression (Fernández-Castilla et al., 2021; Rodgers & Pustejovsky, 2021).

**Rank Correlation.** Des Weiteren wurde die Rangkorrelation nach Begg und Mazumdar (1994) auf die publizierten Daten angewandt. Dabei werden die Effektstärken und die Stichprobengröße korreliert. Ist dieser Zusammenhang signifikant, so weist dies auf eine Asymmetrie der Daten und einen möglichen *Publication Bias* hin. Da für diese Methode unabhängige Stichproben benötigt werden, wurden hier die aggregierten Modelle verwendet (Fernández-Castilla et al., 2021).

**p-curve.** Eine alternative Möglichkeit *Publication Bias* zu untersuchen, ist die *p-curve* (Simonsohn et al., 2014). Diese liefert Hinweise darauf, wie die signifikanten *p*-Werte der Studien in einer Metaanalyse verteilt sind und ob vielleicht post hoc Datenmanipulation stattgefunden hat (Dowdy et al., 2021). Ist die Kurve linksschief, also wenn es viele Studien gibt, die höhere *p*-Werte berichten, ist dies ein Hinweis auf *Publication Bias*. Ist der Test auf

Rechtsschiefe signifikant, so liegt keiner vor und ein wahrer Effekt existiert (van Aert et al., 2016).

**p-uniform.** P-uniform ist eine ähnliche Methode, die auch auf der Verteilung der signifikanten Ergebnisse beruht. Dabei kann auch ein Signifikanztest auf einen *Publication Bias* performt werden, welcher höhere Power als andere Verfahren hat (van Assen et al., 2015).

*P-curve* und *p-uniform* funktionieren auch bei einer geringeren Anzahl an signifikanten Studien oder bei unterschiedlich großen Stichproben. Ein Problem dieser Methoden ist jedoch, dass sie die Effektgröße bei hoher Heterogenität überschätzen. Aus diesem Grund werden diese Methoden als Schätzer für den nach *Publication Bias* korrigierten Mittelwert nicht für die Anwendung bei heterogenen Daten ( $I^2 \geq 50\%$ ) empfohlen. *P-uniform* zeigt bessere Leistung bei Heterogenität im Gegensatz zu anderen Techniken nur, wenn ein extremer *Publications Bias* vorliegt (van Aert et al., 2016; van Assen et al., 2015).

**p-uniform\*.** Als Erweiterung dieser Methoden wurde *p-uniform\** entwickelt. Bei der Berechnung eines für *Publication Bias* korrigierten Effektwertes werden signifikante, aber auch nicht signifikante Ergebnisse miteinbezogen und es ist ein besserer Umgang mit heterogenen Daten als bei *p-uniform* möglich (van Aert & van Assen, 2018).

Da die drei Methoden zu *p*-Werten unabhängige Datenpunkte voraussetzen, wurden in dieser Arbeit bei Anwendung dieser Methoden die nach Stichproben aggregierten Daten verwendet.

**Trim-and-Fill.** Die Methode Trim-and-Fill (Duval & Tweedie, 2000) korrigiert für einen möglichen *Publication Bias*, indem mögliche Datenpunkte, die Asymmetrie erzeugen könnten, identifiziert und durch Imputation weiterer gegenüberliegende Punkte in unterrepräsentierten Regionen Symmetrie im *Funnel Plot* hergestellt wird. Mit diesen neuen Daten kann dann ein bereinigter Gesamteffekt berechnet und mit dem ursprünglichen verglichen werden. Um zu testen, wie sich die einzelnen Stichproben und ihre Größe auswirken, wurden in dieser multilevel Metaanalyse die nach Sample aggregierten Daten verwendet (Card, 2011; Fernández-Castilla et al., 2021).

**Leave-One-Out.** Bei der *Leave-One-Out* Methode wird immer genau eine Studie oder in diesem Fall jeweils eine Stichprobe von der Berechnung ausgelassen. Die Durchführung einer Metaanalyse mit dieser Teilmenge zeigt das Ausmaß einzelner Gruppen und den individuellen Einfluss auf das Ergebnis der restlichen Studien. Die mittlere Effektstärke variiert demnach je nach gewählter Studie. Bei wenig Variation der errechneten Werte ist die

Metaanalyse robust gegen Ausreißer (*Random Effects / Sensitivity Analysis*, 2023; *Leave-one-out-Meta-Analyse*, 2023).

Erkennbar ist, dass die meisten vorgestellten Methoden mit den nach Stichproben aggregierten Daten angewandt wurden. Für hierarchisch geordnete Daten werden viele Methoden für den Test auf *Publication Bias* aufgrund der abhängigen Effektwerte und erhöhter Alphafehler nicht empfohlen. Daher sollten auch die Ergebnisse mit Vorsicht interpretiert werden. Zu beachten ist außerdem, dass bei einer multilevel Analyse die einzelnen Effektstärken und nicht die Stichproben oder ganze Studien von Bedeutung sind, was an der Aussagekraft der Methoden für aggregierte Daten zweifeln lässt (Fernández-Castilla et al., 2021). Allein die Egger MLMA kann einigermaßen gut mit der Datenabhängigkeit umgehen (Rodgers & Pustejovsky, 2021).

### Ergebnisse

Das verwendete Datenblatt mitsamt Notizen zu den einzelnen Studien und der finale R-Code können über <https://osf.io/8rnkz/> abgerufen werden.

#### **Deskriptive Beschreibung der inkludierten Studien**

Insgesamt konnten nach Ausschluss nicht relevanter Studien 70 Untersuchungen in die Analyse mitaufgenommen werden. Diese lieferten wiederum 144 verschiedene Korrelationen für den Zusammenhang zwischen Geschwisterreihenfolge und Intelligenz.

In Tabellen 1 und 2 wird die Verteilung der inkludierten Studien über unterschiedliche Variablen und deskriptive Faktoren zur Stichprobe ersichtlich.

Hervorzuheben ist außerdem die Stichprobengröße. Die gesamte Stichprobe aller Studien umfasste 3 941 342 Personen; wobei die Spanne von 30 bis 800 000 Proband\*innen reichte. Der Median lag bei 532 getesteten Teilnehmer\*innen. Der Großteil der Studien (fast 79 %) untersuchte ausschließlich Kinder. Dies wurde auch deutlich im Durchschnittsalter von etwa 11 Jahren. Die Stichproben waren zu ca. 54 % männlich.

Etwa 83 % aller Forschungsartikel wurden publiziert. Das Publikationsjahr lag zwischen 1926 und 2022 (siehe Abbildung 2) und der Median bei 1980. Bei zirka einem Drittel wurde die Korrelation durch Unterschiede zwischen den Erst- und Spätergeborenen berechnet und somit ein gewisses Rauschen in Kauf genommen. Mehr als die Hälfte aller Studien fand in Nordamerika statt, die wenigsten Untersuchungen mit knapp 1.5 % in Australien.

**Tabelle 1***Deskriptive Verteilung der inkludierten Studien*

	Anzahl Studien	% aller Studien	Anzahl unabhängiger Effektstärken
Rauschen durch Mittelwertsunterschiede			
Kein Rauschen	46	65.7	72
Rauschen	24	34.3	62
Publikationsstatus			
Nicht publiziert	12	17.1	17
Publiziert	58	82.9	117
Geographische Lage			
Afrika	4	5.7	6
Asien	9	12.9	20
Australien	1	1.4	1
Europa	16	22.9	24
Nordamerika	38	54.3	81
Südamerika	2	2.9	2
Altersgruppe			
Kinder	55	78.6	88
Erwachsene	8	11.4	13
Gemischte Gruppe	6	8.6	9
Keine Angabe	1	1.4	24

*Anmerkung.* Gesamtzahl Studien = 70. Gesamtzahl berichteter Effektwerte (unabhängiger Tests oder Stichproben) = 134. Gesamtzahl aller berichteten Effektstärken (inkl. für Familiengröße kontrollierte) = 144.

**Tabelle 2***Deskriptive Beschreibung der Stichproben*

	<i>N</i>	Mittelwert	Standard- abweichung	Median	Minimum	Maximum
Stichprobengröße	134	29 413	99 355.43	532	30	800 000
Mittleres Alter (Jahre)	75	11.23	5.38	11.50	0.25	25.00
Anteil männlicher Teilnehmer (%)	68	53.54	28.76	51.10	0	100
Publikationsjahr	134	1987.07	22.32	1980	1926	2022
GDP per capita	120	12 314.41	12 228.64	7 897.52	120.83	46 298.92
GDP PPP per capita	46	31 347.39	19 832.05	32 058.42	1 794.15	60 092.80
Mittlerer Geburtsrang	81	2.12	0.46	2.03	1.48	3.16
Mittlere Familiengröße	61	3.38	1.59	3.00	1.41	8.40
Korrelation <i>r</i>	94			-.10	-.90	.73
Korrelation <i>r</i> mit Kontrolle für Familiengröße	50			-.04	-.22	.47



Familiengröße lag zwischen  $-.22$  und  $.47$  und 50 % der Werte zeigten ein  $r$  zwischen  $-.08$  und  $.01$  (IQR =  $.09$ ).

### Hauptanalysen

Die numerischen Ergebnisse der multilevel Metaanalysen für die drei Hauptmodelle können Tabelle 3 entnommen werden.

**Tabelle 3**

*Ergebnisse der multilevel Metaanalysen*

	Modell ohne Rauschen	Umfassendes Modell	Familiengröße- Modell
Anzahl Studien	44	64	13
Anzahl Effektstärken	62	94	50
Gesamteffekt (Pearson $r$ )	$-.156$	$-.141$	$.004$
95 % KI für $r$	$[-.254, -.056]$	$[-.212, -.068]$	$[-.058, .065]$
$p$	$.002^*$	$<.001^*$	$.912$
$Q$ ( $df$ )	1 169 576.324 (61) $p <.001^*$	1 549 052.371 (93) $p <.001^*$	1083.392 (49) $p <.001^*$
$\tau^2$	0.053	0.047	0.011
$I^2$	Level 2: 59.16 % Level 3: 40.80 % Gesamt: 99.96 %	Level 2: 52.17 % Level 3: 47.80 % Gesamt: 99.97 %	Level 2: 8.90 % Level 3: 90.57 % Gesamt: 99.47 %
aggregiertes Modell nach unabhängigen Samples	$k = 48$ $r = -.145$ $[-.233, -.055]$	$k = 72$ $r = -.133$ $[-.198, -.067]$	$k = 35$ $r = -.031$ $[-.057, -.005]$

*Anmerkung.* KI = Konfidenzintervall.  $Q$  = Cochran's Q-Test für Heterogenität.  $\tau^2$  = Zwischen-Studien-Varianz.  $I^2$  = Prozentualer Anteil der Variabilität, welcher durch Heterogenität erklärt wird.

\*  $p < .05$

### *Modell ohne Rauschen und Hypothese Eins*

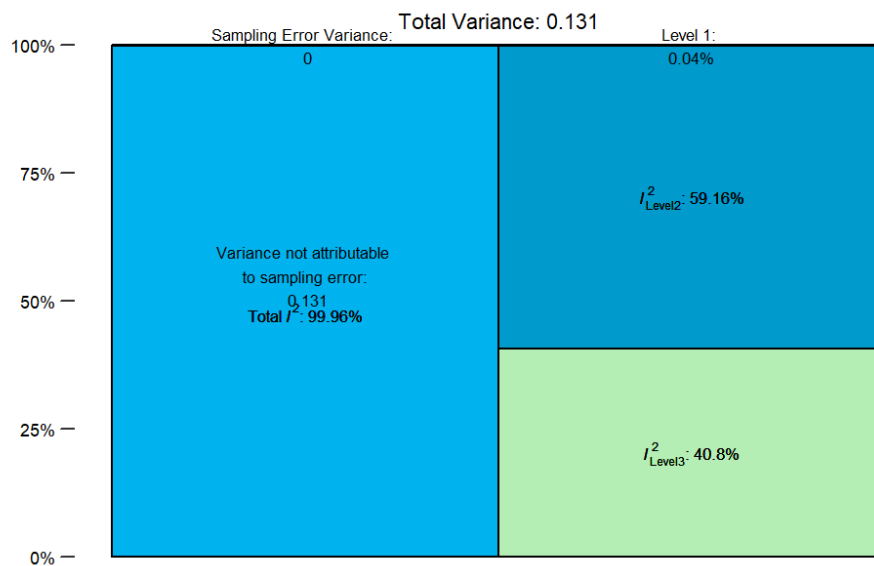
Der Zusammenhang zwischen Geschwisterreihenfolge und Intelligenz ohne Inklusion von Datensätzen mit Rauschen (44 Studien und 62 Korrelationen) ergab  $r = -.156$ . Cochrans  $Q$  und die Zwischen-Studien-Varianz  $\tau^2$  wiesen auf eine beträchtliche Heterogenität hin. Wie in Abbildung 4 ersichtlich, lag die Gesamtvarianz  $I^2$  bei über 99 %. Abbildung 5 zeigt die

einzelnen Effektstärken und den Gesamteffekt in einem *Forest Plot*. Im Anhang befinden sich weitere graphische Darstellungen, wie ein *Forest Plot* geordnet nach Publikationsjahr sowie ein sogenannter *Caterpillar Plot*, wobei die Effektstärken nach Größe sortiert wurden und somit die Streuung dieser gut ersichtlich ist.

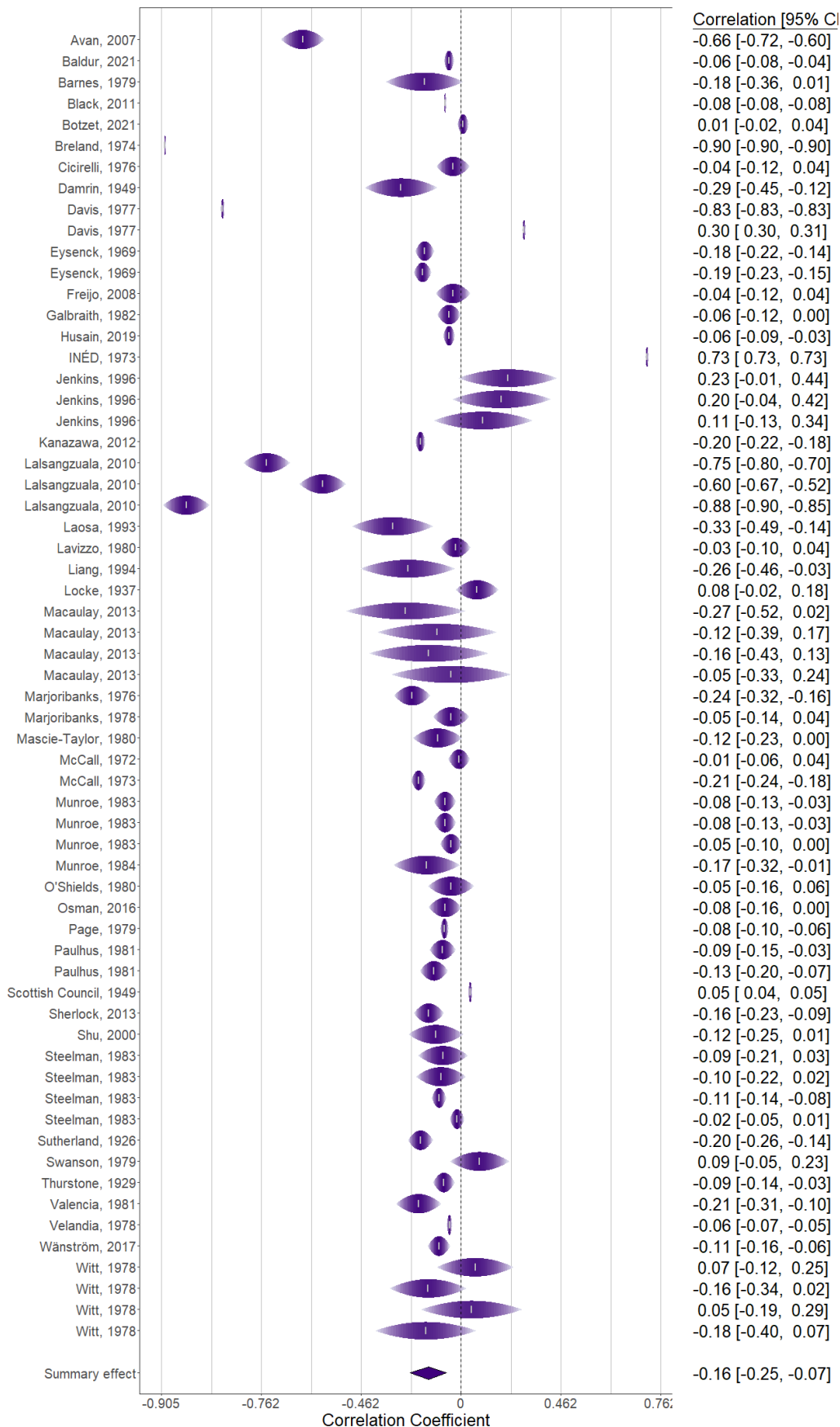
Des Weiteren können in Abbildung 6 die Effektstärken des für unabhängige Stichproben aggregierten Modells eingesehen werden. Der gemittelte Effekt für 48 Samples lag hier bei  $r = -.145$ .

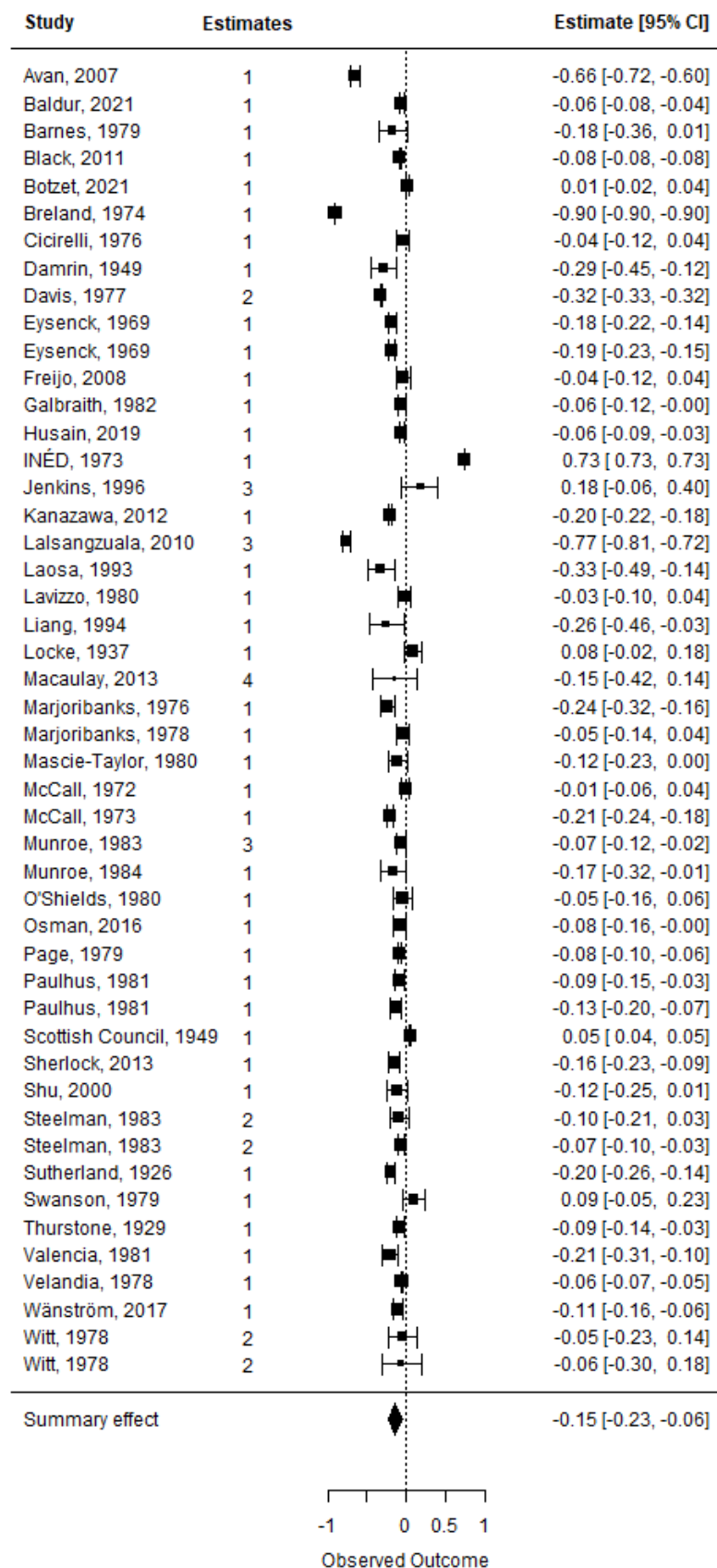
#### Abbildung 4

*Verteilung der Gesamtvarianz des Modells ohne Rauschen*





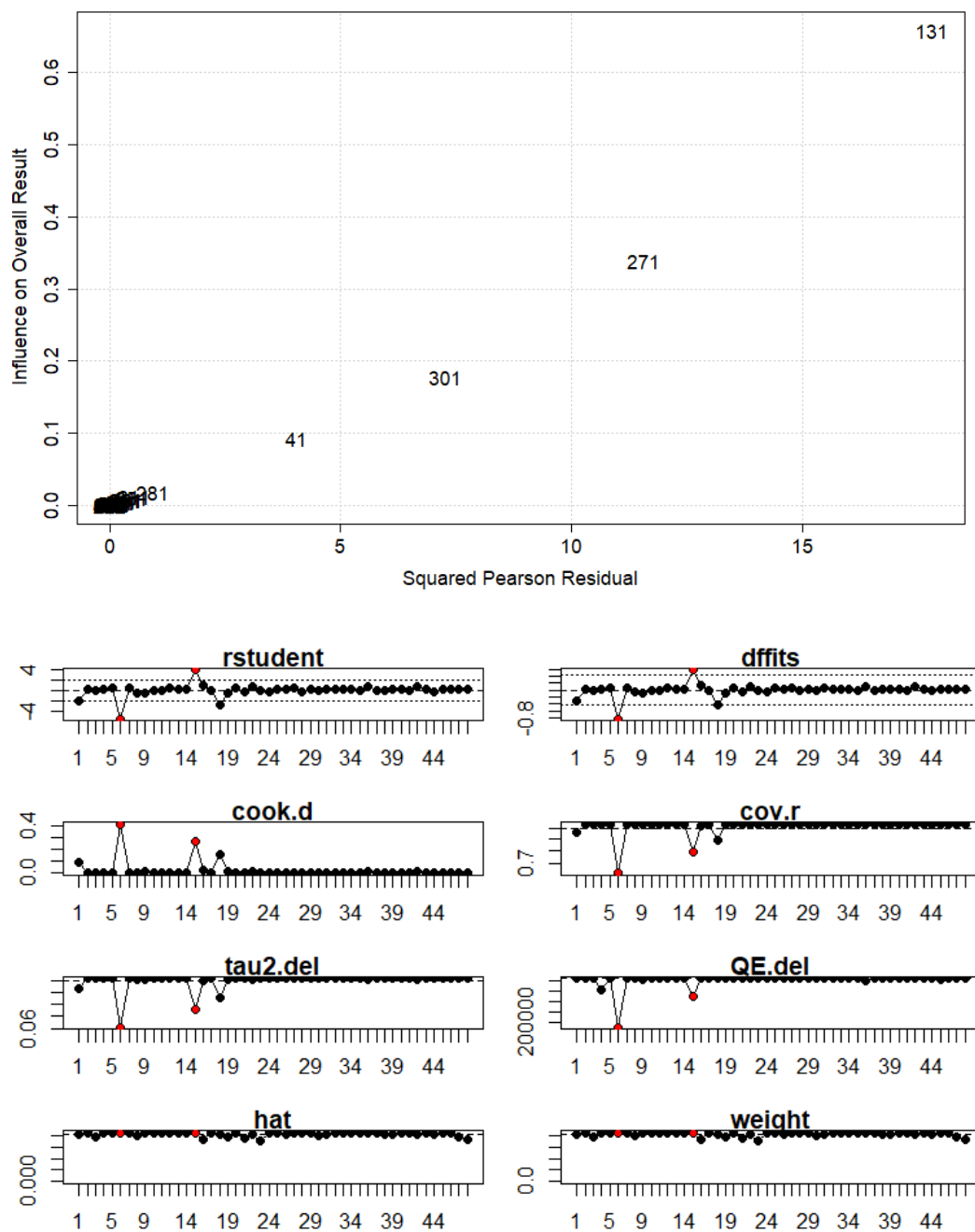
**Abbildung 5***Gemittelter Effekt des Modells ohne Rauschen (Rainforest Plot). Multilevel Modell*

**Abbildung 6***Gemittelter Effekt des Modells ohne Rauschen (Forest Plot). Aggregiertes Modell*

Außerdem konnten in dem aggregierten Modell zwei signifikante Ausreißer identifiziert werden (siehe Abbildung 7). Diese stammen aus den Studien von Breland (1974) und INÉD (1973, zitiert nach Zajonc & Bargh, 1980b). Bei Ausschluss dieser Samples reduzierte sich der gepoolte Effekt auf  $r = -.139$ .

### Abbildung 7

*Ausreißer des Modells ohne Rauschen (Baujat und Influence Plot)*



### *Umfassendes Modell und Hypothese Zwei*

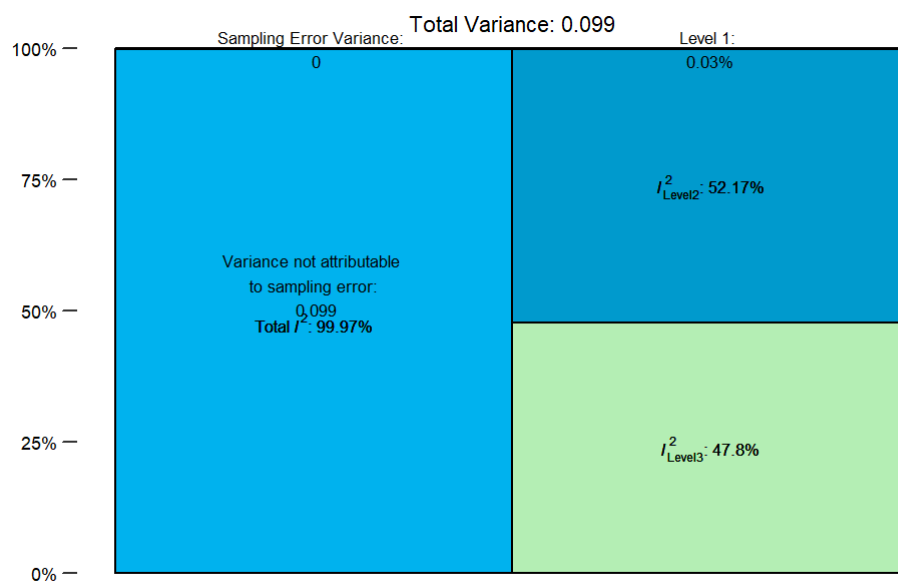
Das zweite Hauptmodell umfasste sowohl die Daten des zuvor vorgestellten Modells ohne Rauschen, als auch Korrelationen, die mithilfe von Mittelwertsunterschieden berechnet wurden. Erneut zeigte sich eine beträchtliche Heterogenität, wie bei der Verteilung der Gesamtvarianz  $I^2$  in Abbildung 8 ersichtlich wird.

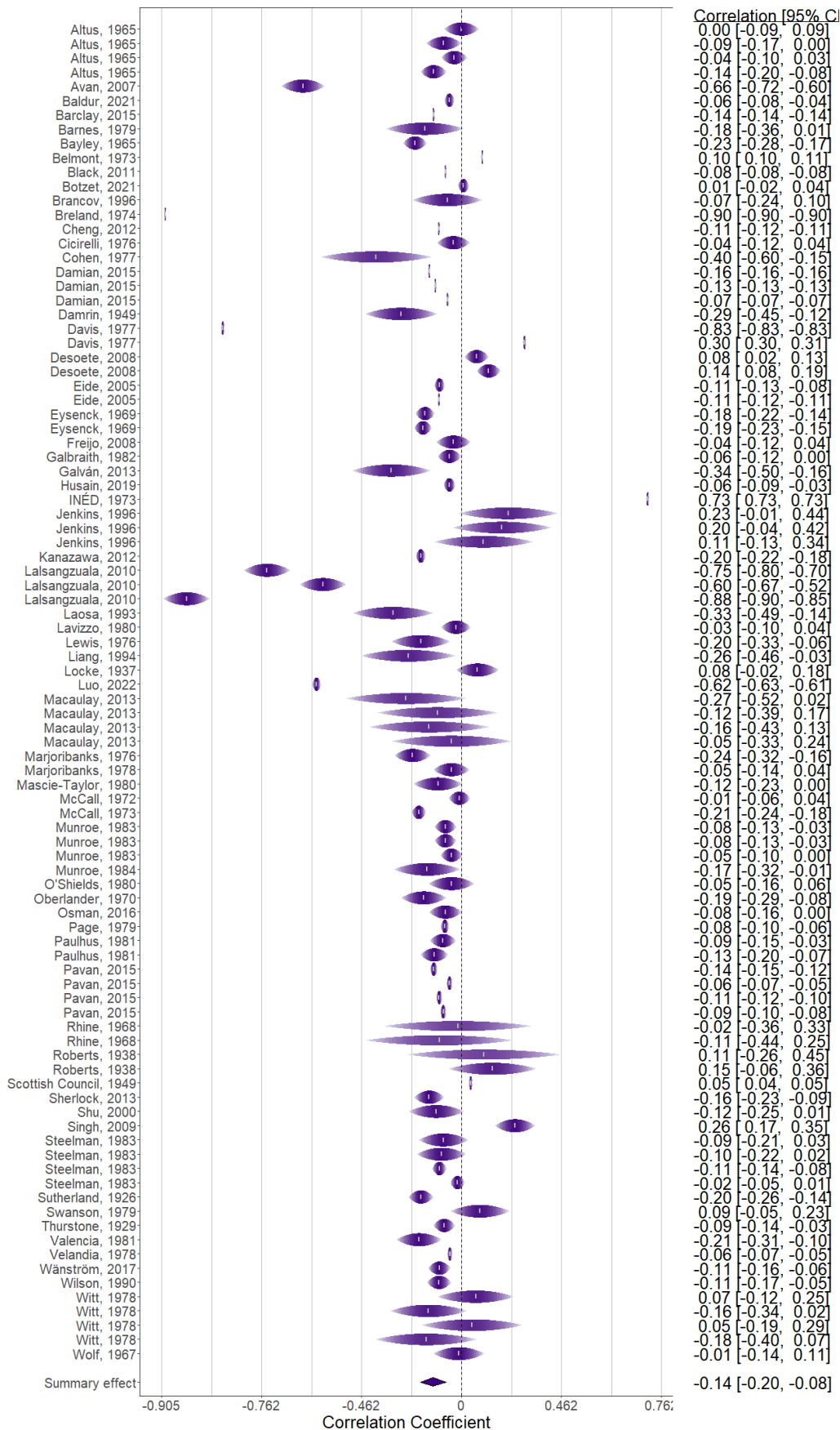
Dieses Modell (64 Studien, 94 Korrelationen) ergab einen Gesamteffekt von  $r = -.141$ , welcher in Abbildung 9 dargestellt wird. Im Anhang befinden sich weitere Graphiken zur Beschreibung der Verteilung der Effektstärken.

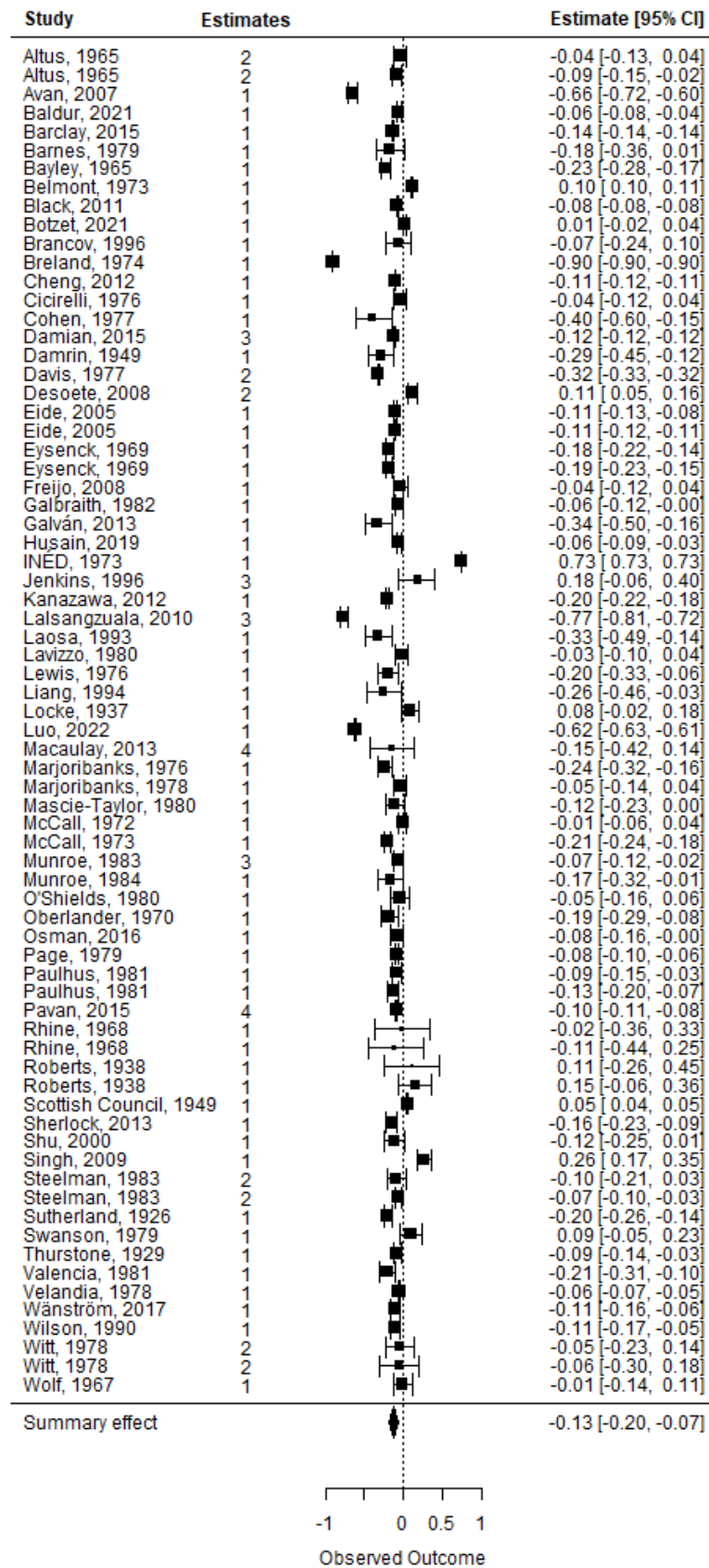
Bei Bildung eines aggregierten Modells mit 72 Stichproben ergab sich eine gemittelte Korrelation von  $r = -.133$  (siehe Abbildung 10).

### **Abbildung 8**

*Verteilung der Gesamtvarianz des umfassenden Modells*



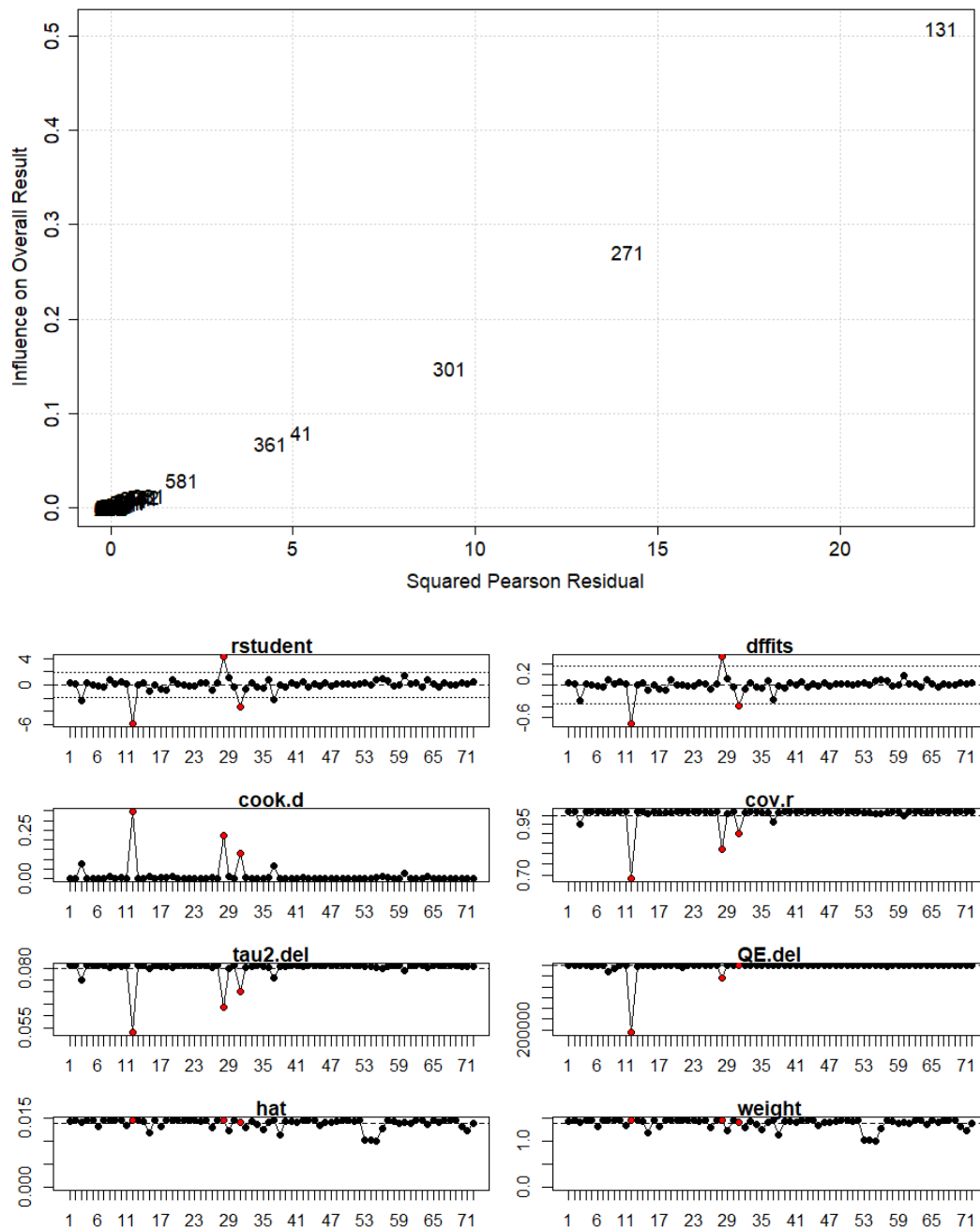
**Abbildung 9***Gemittelter Effekt des umfassenden Modells (Rainforest Plot). Multilevel Modell*

**Abbildung 10***Gemittelter Effekt des umfassenden Modells (Forest Plot). Aggregiertes Modell*

Wie in Abbildung 11 zu sehen, konnten abermals die selben Ausreißer wie im Modell ohne Rauschen und zusätzlich die Stichprobe von Lalsangzuala (2010) ausgeschlossen und ein bereinigter Effekt von  $r = -.116$  errechnet werden.

### Abbildung 11

*Ausreißer des umfassenden Modells (Baujat und Influence Plot)*



Für die Testung der zweiten Hypothese wurde das umfassende Modell mit dem Modell ohne Rauschen verglichen. Die Ergebnisse des Waldtest ( $\chi^2 = 0.246, p = .806$ ) ergeben keinen signifikanten Unterschied zwischen diesen beiden Modellen.

### ***Familiengröße Modell und Hypothese Drei***

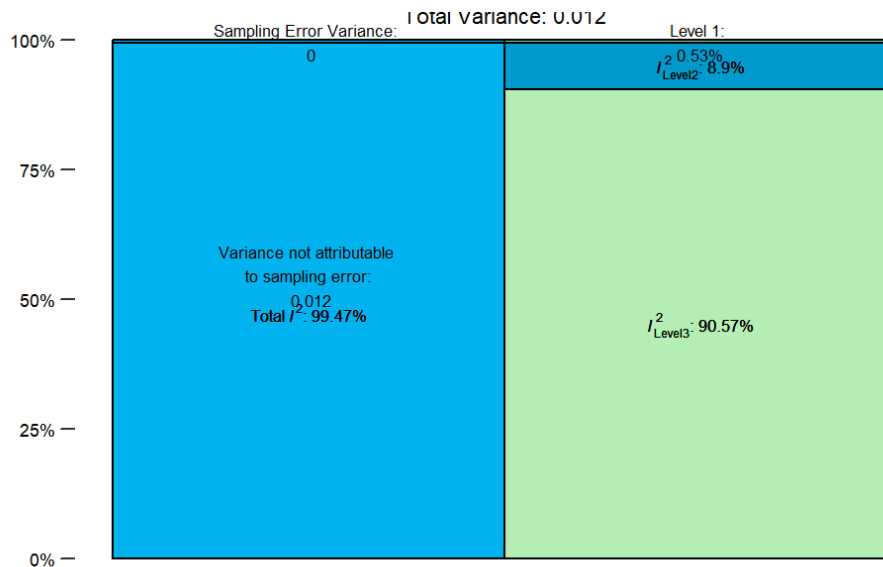
Für das dritte Hauptmodell wurden jene Effektstärken verwendet, die für die Anzahl der Kinder innerhalb einer Familie (Familiengröße) kontrollieren. Auch bei diesem Modell wurde beträchtliche Heterogenität erwiesen, wie in Abbildung 12 ersichtlich wird.

Mittels 13 Studien und 50 Effektstärken konnte ein Gesamteffekt von  $r = .004$  errechnet werden (siehe Abbildung 13).

Das aggregierte Modell hier mit 35 unabhängigen Samples zeigte einen Effekt von  $r = -.031$  (siehe Abbildung 14).

### **Abbildung 12**

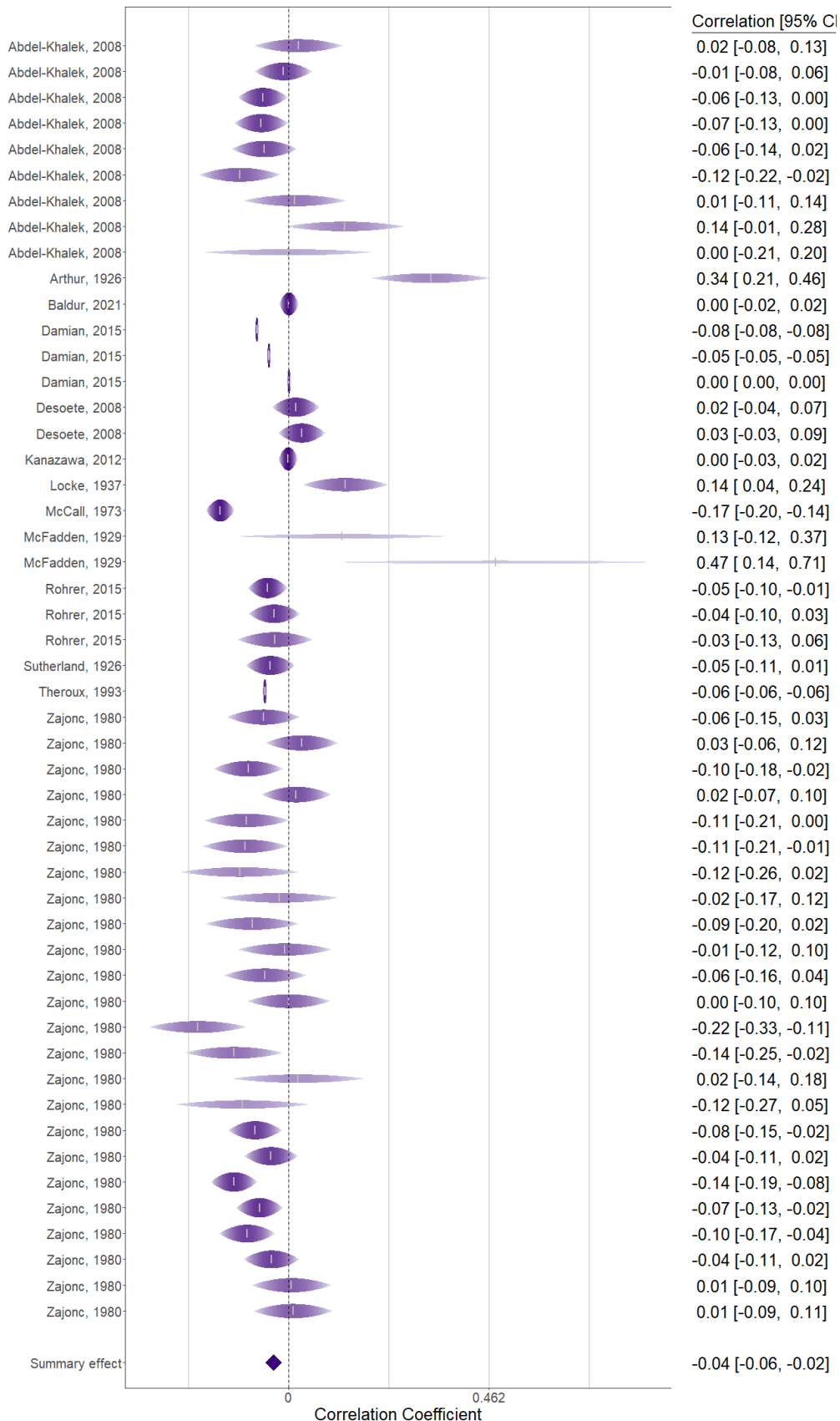
*Verteilung der Gesamtvarianz des Modells mit Kontrolle für Familiengröße*





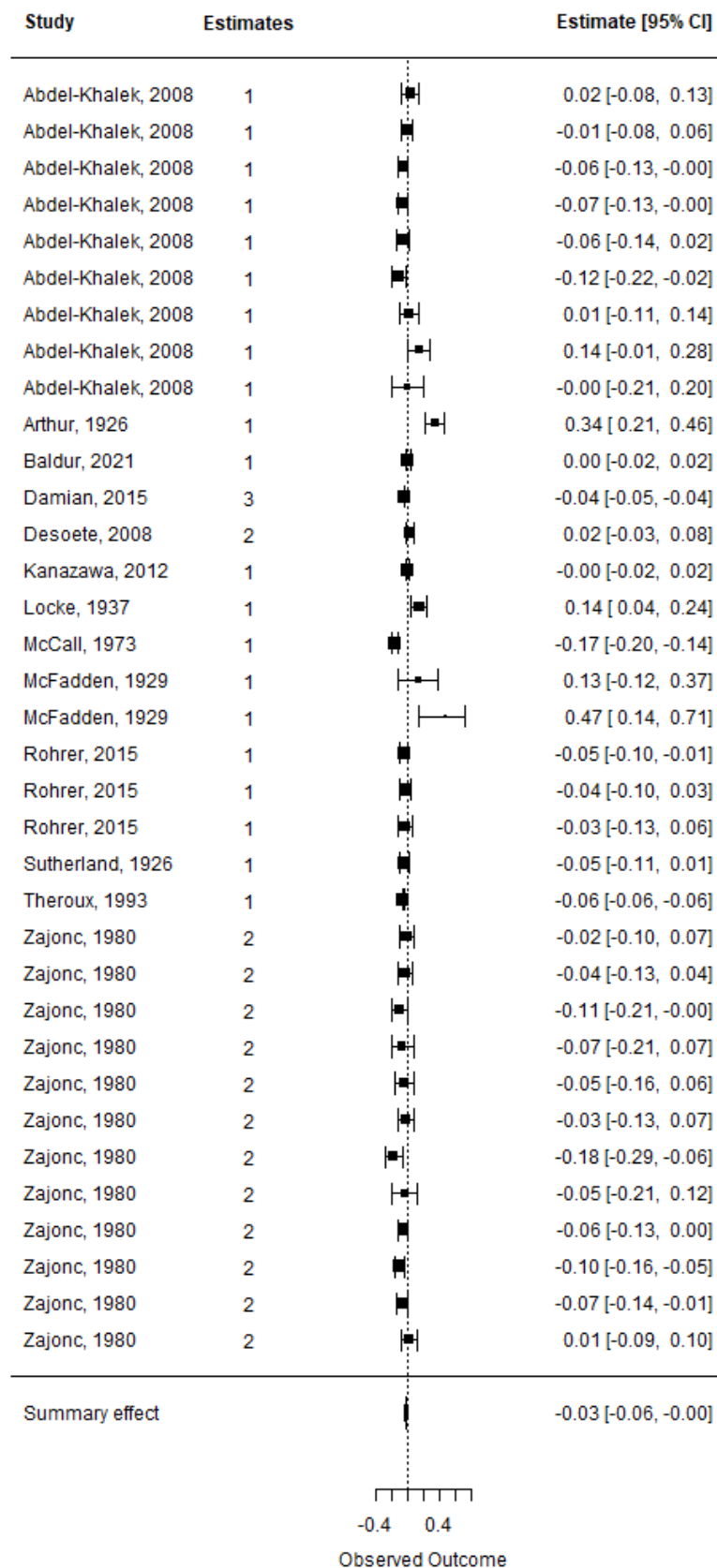
**Abbildung 13**

*Gemittelter Effekt des Modells mit Kontrolle für Familiengröße (Rainforest Plot). Multilevel Modell*



**Abbildung 14**

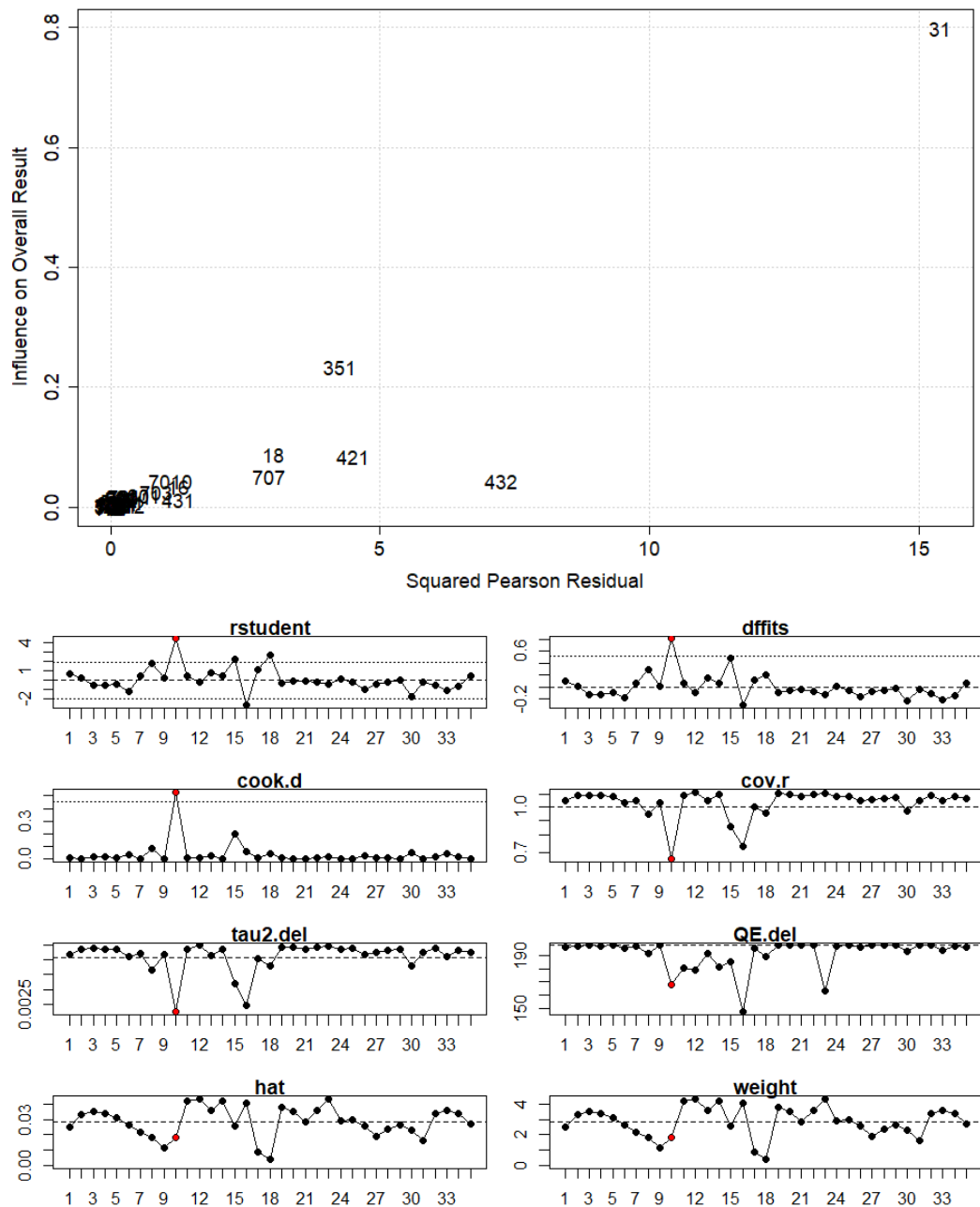
*Gemittelter Effekt des Modells mit Kontrolle für Familiengröße (Forest Plot). Aggregiertes Modell*



Ebenso wurde bei dem Modell mit Kontrolle für Familiengröße ein signifikanter Ausreißer identifiziert (Arthur, 1926; siehe Abbildung 15) und entfernt, was zu einem aggregierten Effekt von  $r = -.041$  führte.

### Abbildung 15

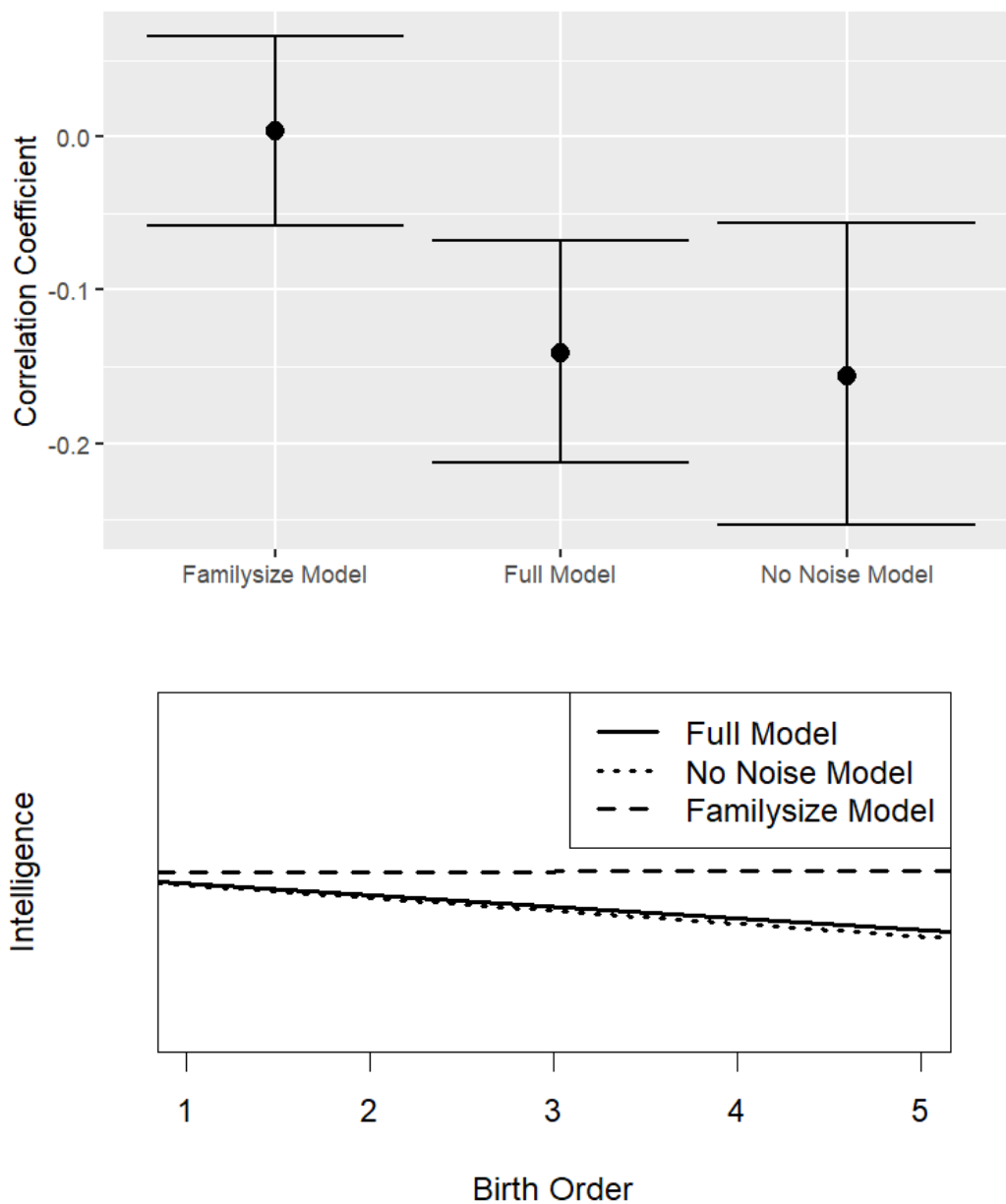
*Ausreißer des Modells mit Kontrolle für Familiengröße (Baujat und Influence Plot)*



Um Hypothese Nummer Drei zu überprüfen, wurden die zwei zuvor vorgestellten Modelle jeweils mit dem Familiengröße-Modell verglichen. Der Vergleich mit dem Modell ohne Rauschen ( $\chi^2 = -2.652, p = .008$ ), als auch jener mit dem umfassenden Modell ( $\chi^2 = -2.951, p = .003$ ) ergab signifikante Unterschiede. Die errechneten gemittelten Effektstärken sind demnach voneinander verschieden. Abbildung 16 zeigt die gepoolten Korrelationen der drei Modelle nochmals graphisch nebeneinander dargestellt.

**Abbildung 16**

*Gemittelte Korrelationen aller drei Modelle*



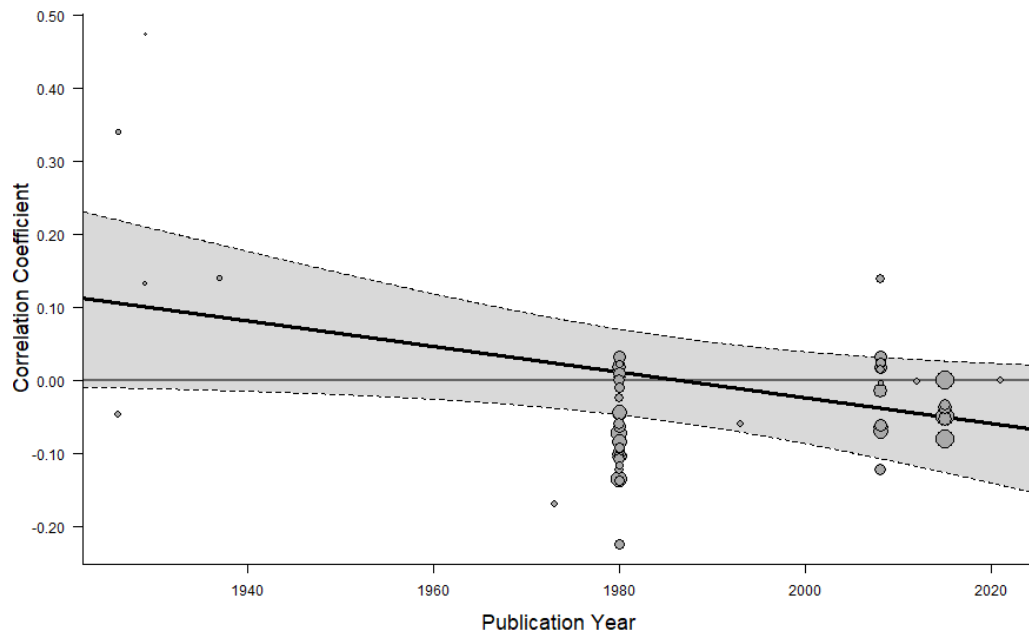
**Moderatoranalysen (Hypothesen Vier bis Acht)**

Alle Ergebnisse der Moderatoranalysen, also sowohl Subgruppenanalysen, als auch Metaregressionen können in Tabelle 4 eingesehen werden. Alle drei Hauptmodelle wurden auf Moderation durch die angegebenen Variablen getestet.

Einzig allein die Moderation durch das Jahr der Publikation beim Modell mit Kontrolle für Familiengröße zeigte sich signifikant ( $p = .042$ ) (siehe Abbildung 17). Alle anderen Moderationen wurden nicht bestätigt.

**Abbildung 17**

*Metaregression für Publikationsjahr für das Modell mit Kontrolle für Familiengröße*



**Tabelle 4***Ergebnisse der Moderatoranalysen*

	Modell ohne Rauschen			Umfassendes Modell			Familiengröße-Modell		
	<i>F(df)</i>	<i>p</i>	Koeffizient	<i>F(df)</i>	<i>p</i>	Koeffizient	<i>F(df)</i>	<i>p</i>	Koeffizient
Mittlere Familiengröße	0.014 (1,26)	.907	.002	0.116 (1,29)	.733	-.006	0.152 (1,33)	.697	-.004
Anteil männlicher Teilnehmer	0.130 (1,39)	.718	.000	0.584 (1,66)	.445	.001	1.099 (1,5)	.295	.016
Mittleres Alter	0.378 (1,42)	.539	-.009	0.174 (1,63)	.676	.004	0.002 (1,13)	.968	.000
Altersgruppe	0.822 (2,59)	.663	Kinder: -.182 Erwachsene: -.080 Gemischt: -.049	1.155 (2,91)	.561	Kinder: -.163 Erwachsene: -.062 Gemischt: -.065	0.047 (1,8)	.829	Kinder: -.094 Gemischt: -.060
Geographische Lage	8.446 (5,56)	.133	Afrika: -.092 Asien: -.486 Australien: -.050 Europa: -.043 Nordamerika: -.140 Südamerika: -.058	6.269 (5,88)	.281	Afrika: -.098 Asien: -.343 Australien: -.050 Europa: -.029 Nordamerika: -.150 Südamerika: -.197	0.412 (2,47)	.811	Asien: -.036 Europa: -.017 Nordamerika: .026
Publikationsjahr	0.553 (1,60)	.457	-.002	1.099 (1,92)	.295	-.002	4.130 (1,48)	.042*	-.002
Publikationsstatus	0.599 (1,60)	.439	unpubliziert: -.241 publiziert: -.138	0.045 (1,92)	.833	unpubliziert: -.160 publiziert: -.138	0.215 (1,48)	.643	unpubliziert: -.030 publiziert: .012
GDP per capita	0.336 (1,52)	.562	.000	0.001 (1,81)	.976	.000	1.287 (1,43)	.257	.000
GDP PPP per capita	1.428 (1,16)	.232	.000	0.322 (1,31)	.570	.000	0.218 (1,14)	.641	-.000

*Anmerkung.* \*  $p < .05$

### Tests auf Publication Bias und Sensitivitätsanalysen

Wichtig ist hier nochmals zu erwähnen, dass die Methoden, welche auf *Publication Bias* testen, meist unabhängige Stichproben benötigen und daher nicht die multilevel Daten verwendet werden konnten. Stattdessen wurden für jede publizierte Stichprobe aggregierte Korrelationen gebildet.

In Tabelle 5 können die Ergebnisse der diversen Tests für alle Modelle entnommen werden.

**Tabelle 5**

*Ergebnisse der Tests auf Publication Bias und Sensitivitätsanalysen*

	Modell ohne Rauschen	Umfassendes Modell	Familiengröße- Modell
Egger MLMA	$F(1,47) = 0.222$ $p = .640$	$F(1,76) = 0.259$ $p = .612$	$F(1,46) = 8.675$ $p = .005^p$
Rangkorrelation <sup>a</sup>	$\tau = -0.801$ $p < .001^p$	$\tau = -0.696$ $p < .001^p$	$\tau = 0.245$ $p = .046^p$
p-curve <sup>a</sup> (Test auf Rechts- Schiefe)	$pFull < .001$ $pHalf < .001$	$pFull < .001$ $pHalf < .001$	$pFull < .001$ $pHalf < .001$
p-uniform <sup>a</sup>	$ksig = 25$ $p < .001$	$ksig = 39$ $p < .001$	$ksig = 10$ $p = .347$
p-uniform * <sup>a</sup> (bereinigter Effekt)	$-.038^p$	$-.024^p$	$-.013^p$
Leave-One-Out <sup>a</sup> (Spannweite $r$ )	-0.166 bis -0.101	-0.155 bis -0.112	-0.041 bis 0.026
Trim-and-Fill <sup>a</sup>	15 Samples hinzugefügt <sup>p</sup> $r = -.257$	21 Samples hinzugefügt <sup>p</sup> $r = -.238$	0 Samples hinzugefügt $r = -.030$

Anmerkung. <sup>a</sup> auf unabhängige Samples aggregierte Daten. <sup>p</sup> Hinweis auf Publication Bias

### Funnel Plot Symmetrie

Anhand der Abbildungen 18 (a), 19 (a) und 20 (a) kann eine mögliche Asymmetrie bei der Verteilung der Datenpunkte im *Funnel Plot* eingesehen werden. Gut ersichtlich sind auch einige Ausreißer, sowie das höhere Aufkommen im oberen Bereich, wo geringere

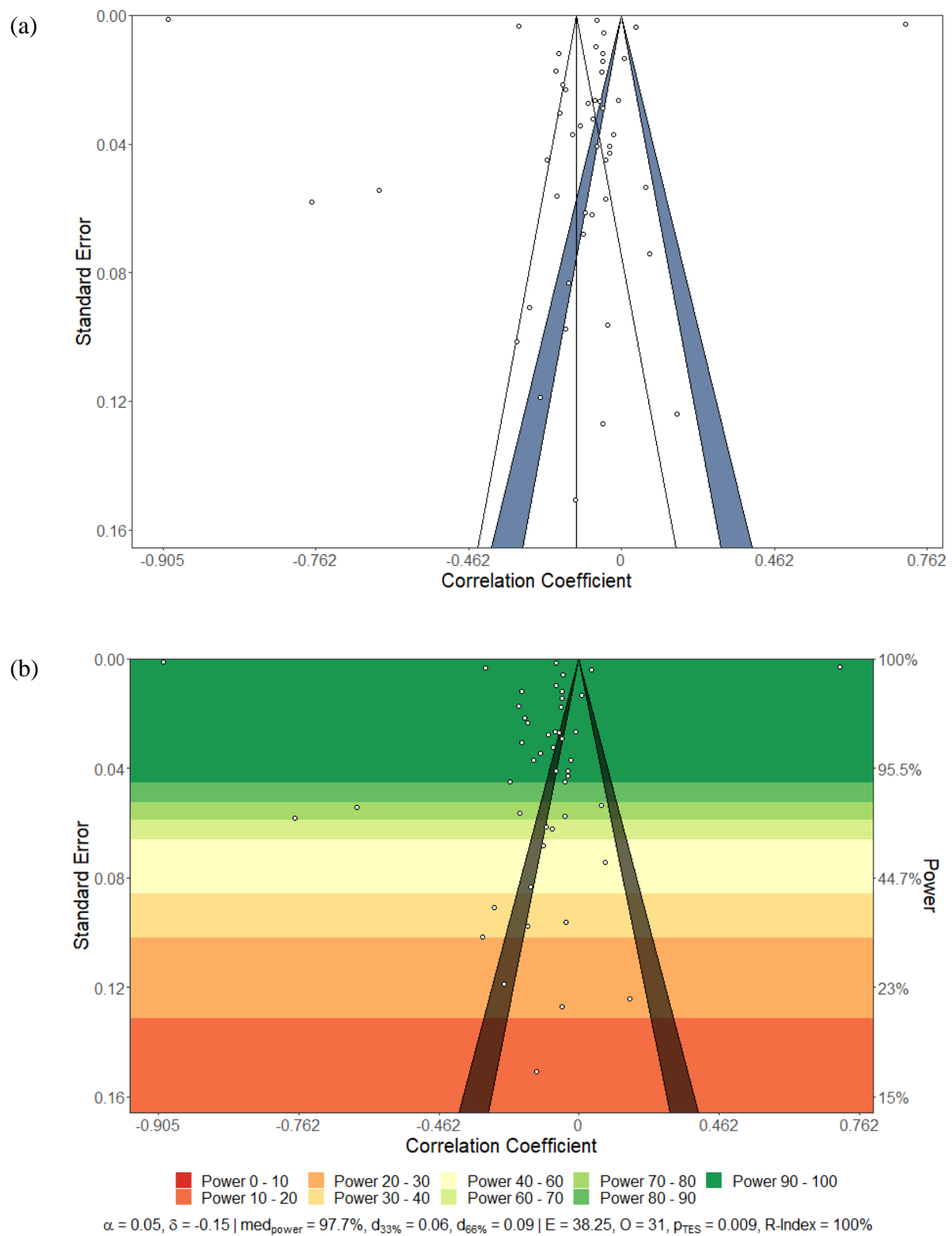
Standardfehler und größere Stichproben liegen. Zudem ist vor allem in der Graphik des Familiengröße-Modells ein Fehlen kleinerer Samples im unteren Teil und womöglich ein *small-study* Effekt zu erkennen. Der *Funnel Plot*, bei dem auch die Power der einzelnen Stichproben durch farbliche Abstimmungen dargestellt werden (*Sunset-Funnel Plots*, siehe Abbildungen 18 (b), 19 (b) und 20 (b) ), zeigte für die Untersuchungen des Modells ohne Rauschen und des umfassenden Modells gute Power über 90 % (Median 97.7 % und 95 %). Beim Modell mit Kontrolle für Familiengröße lagen jedoch die meisten Stichproben in einem niedrigen Bereich der Power (unter 20 %). Der Median der Powerverteilung lag hier nur bei 10.5 %.

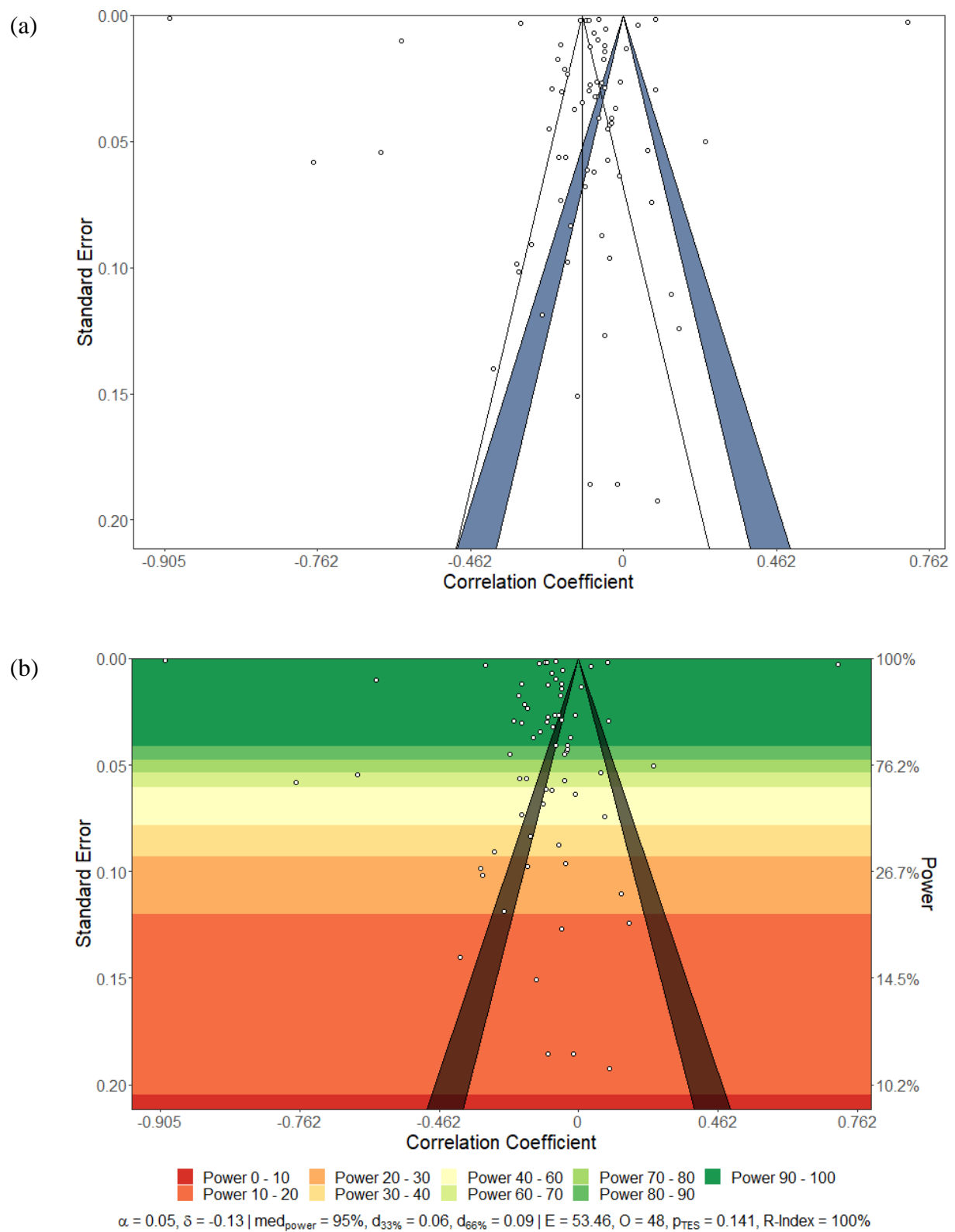
Um statistisch auf eine mögliche Asymmetrie zu testen, wurden folgende zwei Methoden angewandt.

Die Egger MLMA ergab als einzige Methode für multilevel Daten bei den Modellen ohne Kontrolle für die Familiengröße keinen Hinweis auf einen *Publication Bias*. Anders ist dies beim dritten Modell mit Kontrolle für die Anzahl der Kinder. Dabei ist die Moderation signifikant und daher ein Hinweis auf *Publication Bias* (siehe Abbildung 21).

Für die aggregierten Modelle wurde die Rangkorrelation angewandt. Diese zeigte bei allen drei Modellen signifikante Korrelationen zwischen Stichprobengröße und Effektstärke. Demnach sind die Datenpunkte im *Funnel Plot* nicht symmetrisch verteilt und ein *Publication Bias* ist wahrscheinlich.

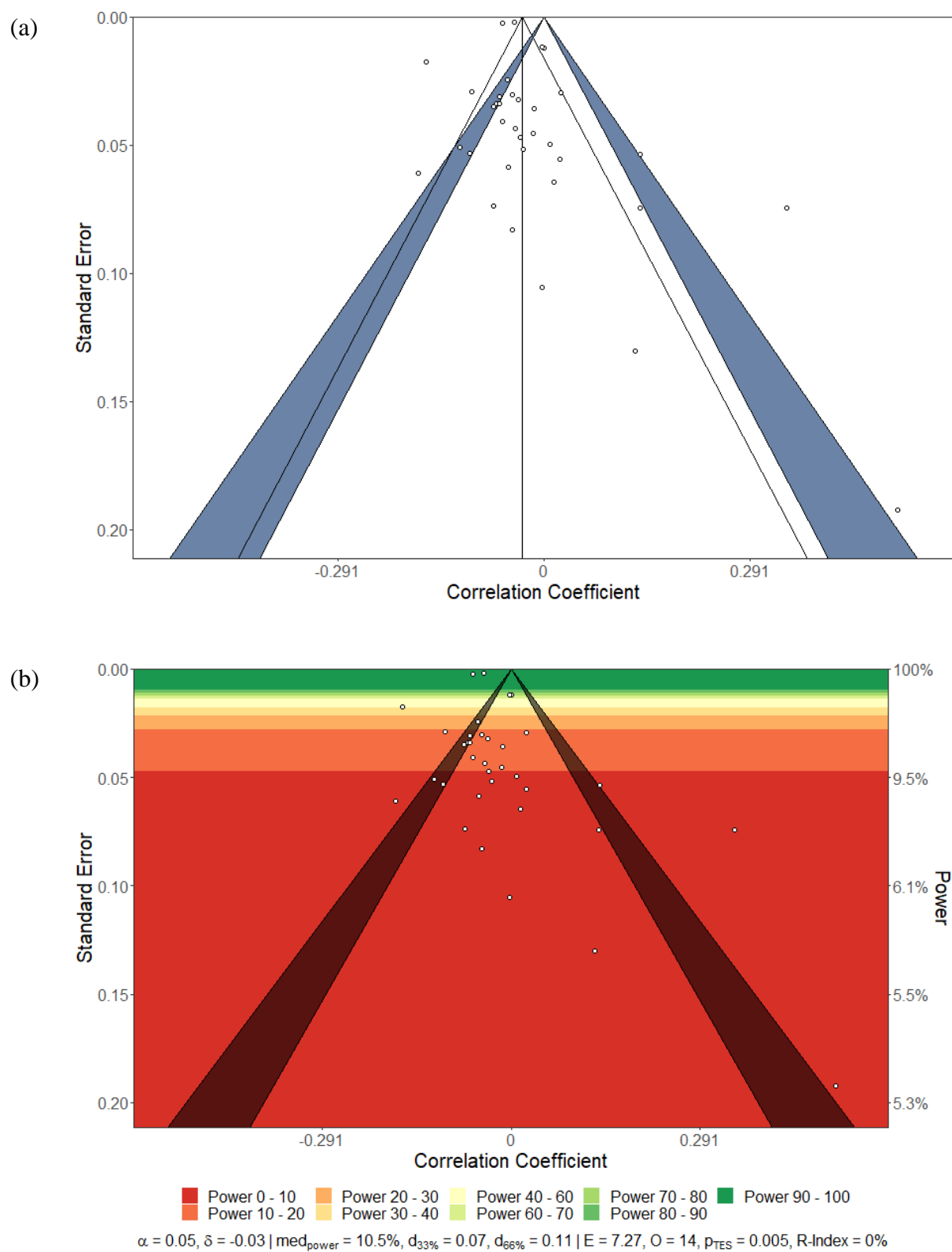


**Abbildung 18***Funnel Plot (a) und Sunset-Funnel Plot (b) des Modells ohne Rauschen*

**Abbildung 19***Funnel Plot (a) und Sunset-Funnel Plot (b) des umfassenden Modells*

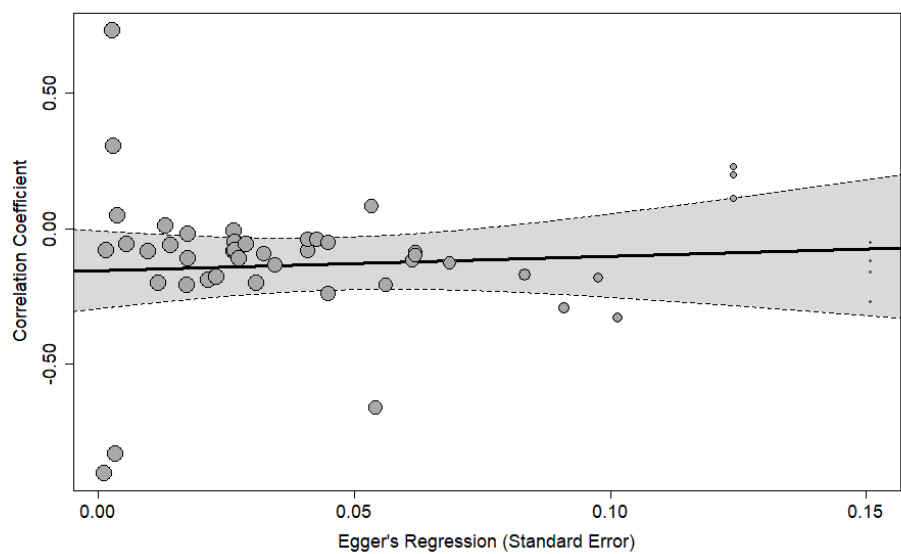
**Abbildung 20**

*Funnel Plot (a) und Sunset-Funnel Plot (b) des Modells mit Kontrolle für Familiengröße*

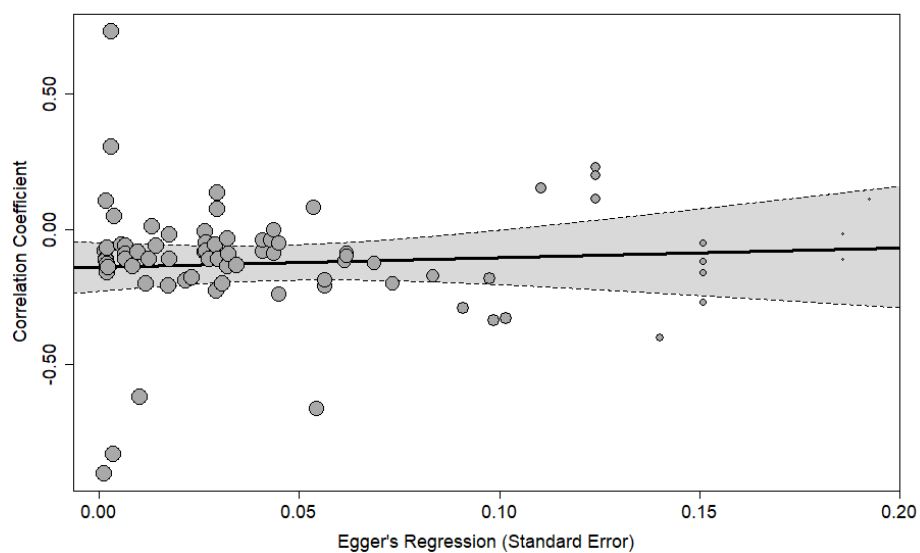


**Abbildung 21***Egger MLMA aller drei Hauptmodelle*

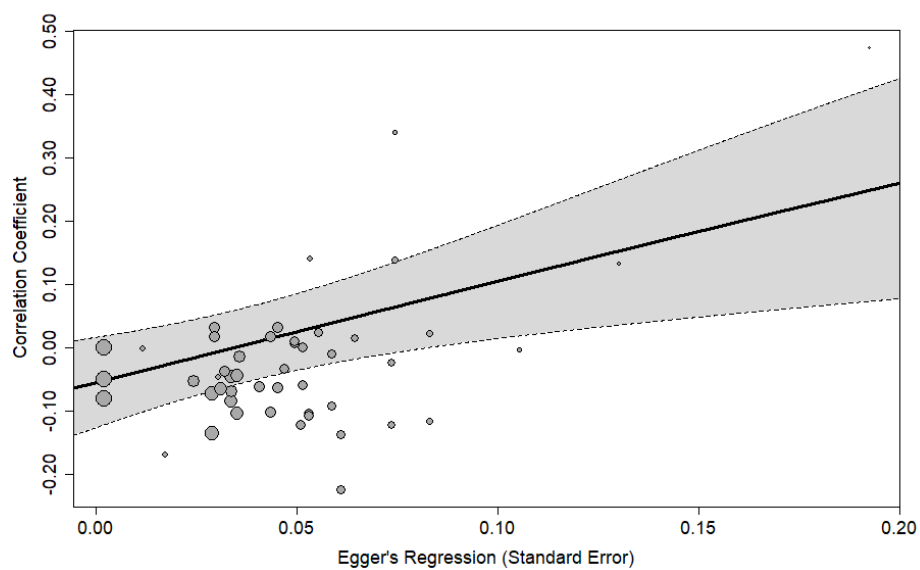
Modell ohne  
Rauschen



Umfassendes  
Modell



Modell mit  
Kontrolle für  
Familiengröße



### **Methoden mit *p*-Werten**

Die *p-curve* (siehe *p-curve* Abbildungen 22 (a), 23 (a) und 24 (a) ) zur Verteilung der signifikanten *p*-Werte der einzelnen unabhängigen Stichproben ergaben für alle drei Modelle eine rechtsschiefe Verteilung und somit keine Evidenz für *p*-Hacking.

Die Methode *p-uniform* ergab für das Modell ohne Rauschen und für das umfassende Modell jeweils einen signifikanten Wert beim Test auf *Publication Bias*. In Abbildungen 22 (b), 23 (b) und 24 (b) können abermals die Graphiken für diese Methode eingesehen werden. Der mithilfe der *p-uniform\** Methodik berechnete Effekt zeigte für alle drei Modelle einen geringeren als ohne Bereinigung. Bei Modell ohne Rauschen sank die gemittelte Effektstärke des aggregierten publizierten Modells von -.138 auf -.038. Beim umfassenden Modell war dies eine Senkung von -.136 auf -.024. Das Modell mit Kontrolle für die Anzahl der Kinder ergab eine Senkung des Effekts von -.030 auf -.013. Die Änderungen waren demnach sehr deutlich, was einen Hinweis auf einen *Publication Bias* bei allen drei Modellen lieferte.

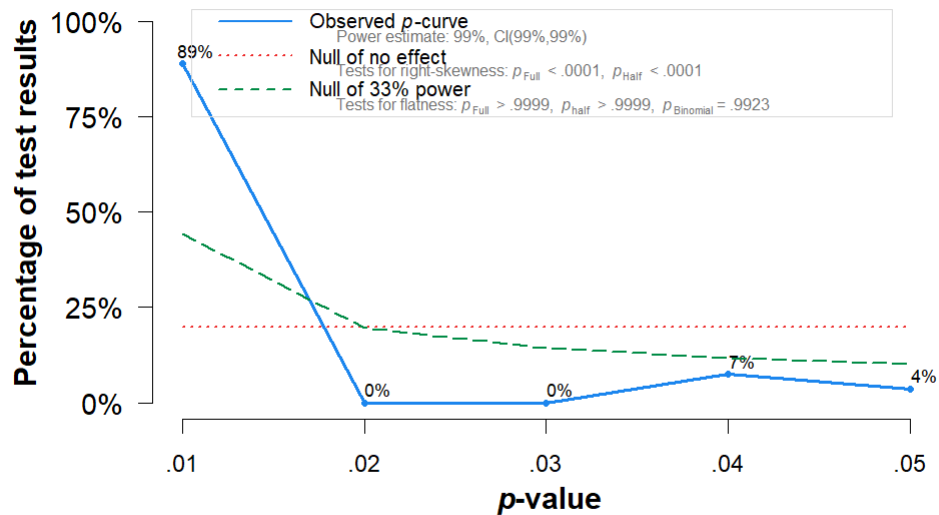
### **Sensitivitätsanalysen**

Auch die *Trim-and-Fill* Methode ergibt einen nach möglichen *Publication Bias* bereinigten Effekt. Vor allem bei den Modellen ohne Kontrolle für Familiengröße wurden einige Studien auf der linken Seite des *Funnel Plots* imputiert (siehe Abbildung 25). Diese hinzugefügten Samples ergaben beim Modell ohne Rauschen etwa 38 % mehr Stichproben als beim publizierten Originalmodell ( $k = 39$ ) und beim umfassenden ( $k = 60$ ) eine um ca. 35 % höhere Anzahl. Dies kann wieder als Hinweis auf einen vorliegenden *Publication Bias* im aggregierten Modell angesehen werden.

Um Hinweise über den Einfluss einzelner Stichproben zu erlangen, wurde die *Leave-One-Out* Methode durchgeführt. Die bei Herausnahme einzelner Samples errechneten Effekte ergaben keine auffälligen Schwankungen (siehe Abbildung 26) und blieben im selbigen Größenbereich der gemittelten Korrelation (kleine Effekte bzw. kein Effekt). Die Schwankungen zeigen nochmals gut die zuvor identifizierten Ausreißer.

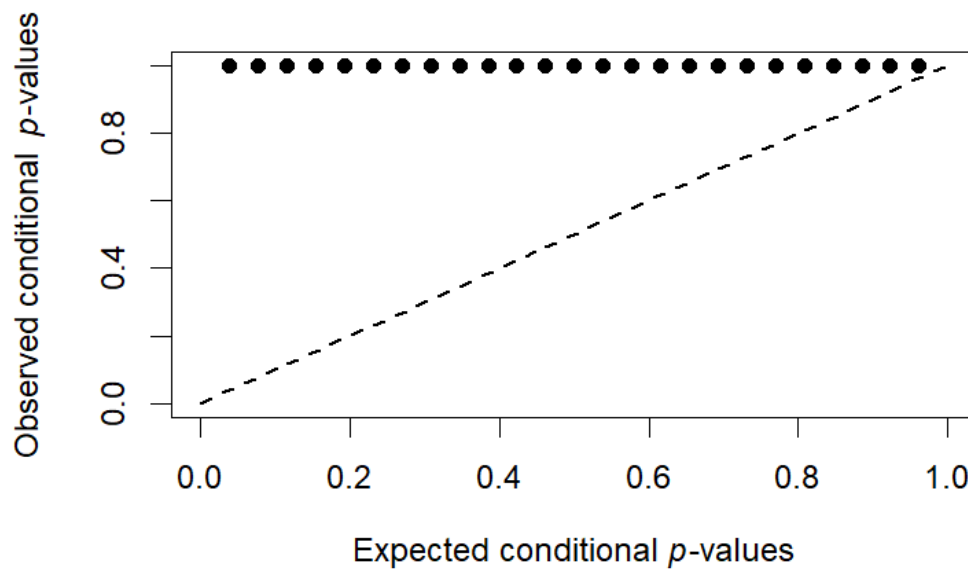
**Abbildung 22***p*-curve (a) und *p*-uniform (b) des Modells ohne Rauschen

(a)



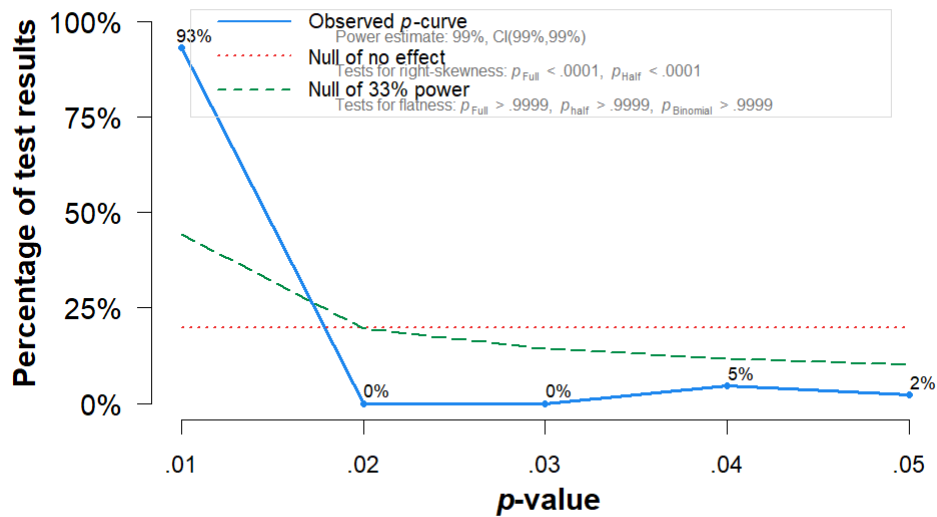
Note: The observed *p*-curve includes 27 statistically significant ( $p < .05$ ) results, of which 24 are  $p < .025$ . There were 12 additional results entered but excluded from *p*-curve because they were  $p > .05$ .

(b)



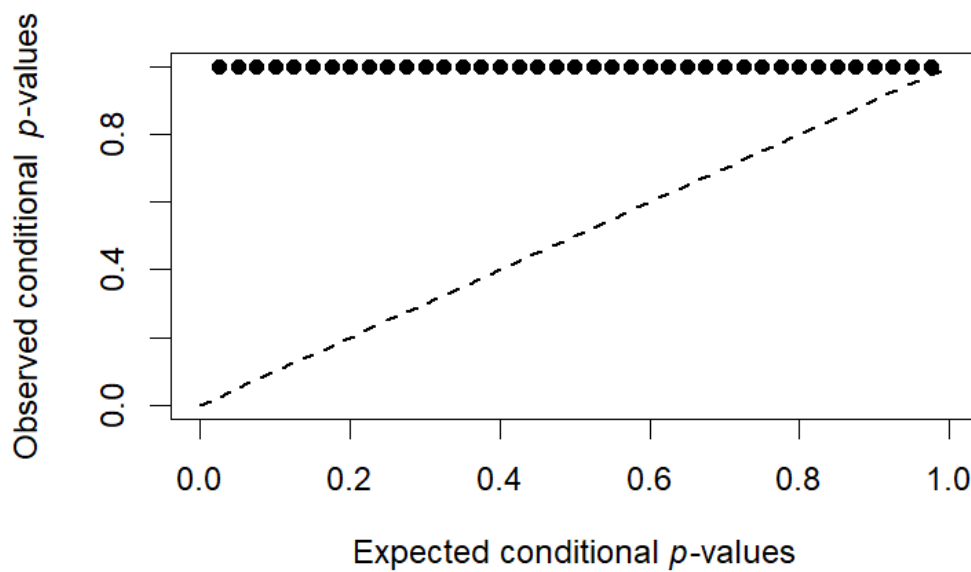
**Abbildung 23***p*-curve (a) und *p*-uniform (b) des umfassenden Modells

(a)



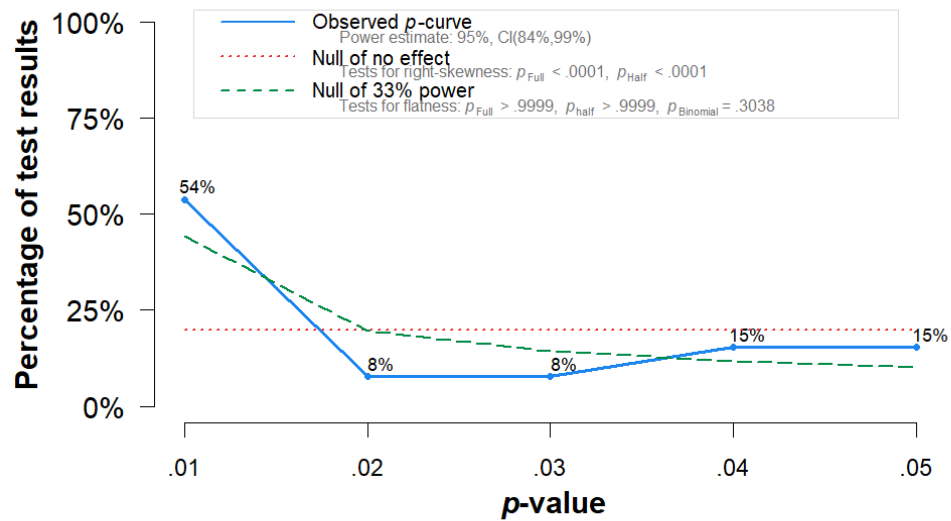
Note: The observed *p*-curve includes 43 statistically significant ( $p < .05$ ) results, of which 40 are  $p < .025$ . There were 17 additional results entered but excluded from *p*-curve because they were  $p > .05$ .

(b)



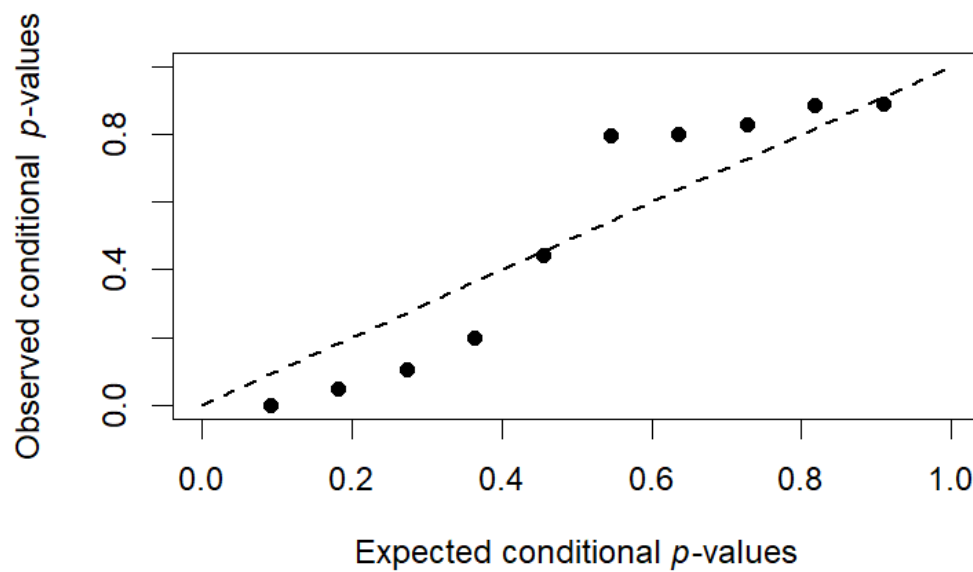
**Abbildung 24***p*-curve (a) und *p*-uniform (b) des Modells mit Kontrolle für Familiengröße

(a)



Note: The observed *p*-curve includes 13 statistically significant ( $p < .05$ ) results, of which 8 are  $p < .025$ . There were 20 additional results entered but excluded from *p*-curve because they were  $p > .05$ .

(b)

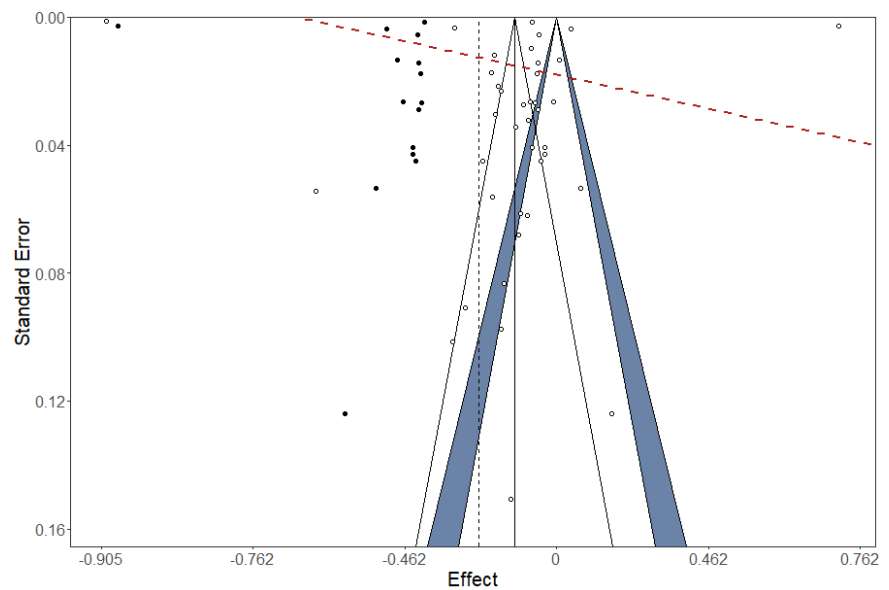




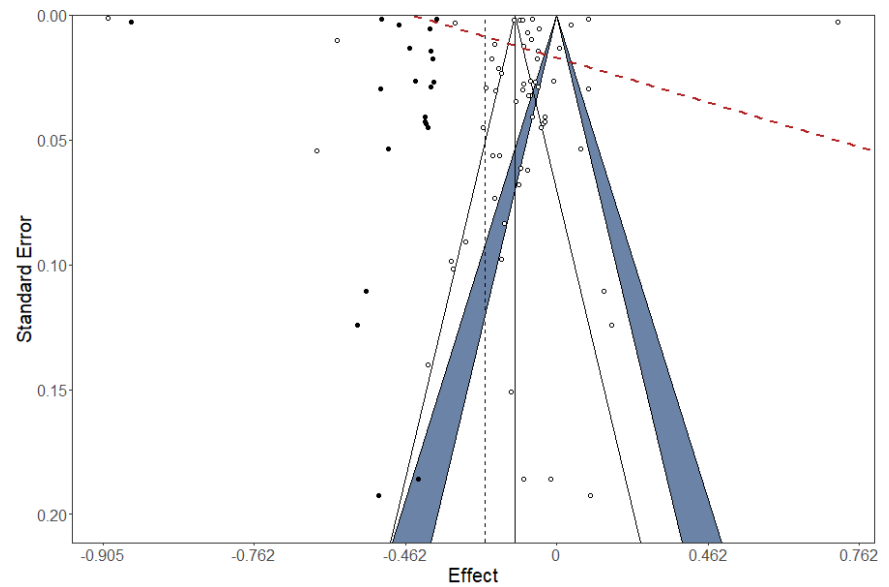
**Abbildung 25**

*Funnel Plot inklusive mit Trim-and-Fill hinzugefügte Studien und Eggers Regression*

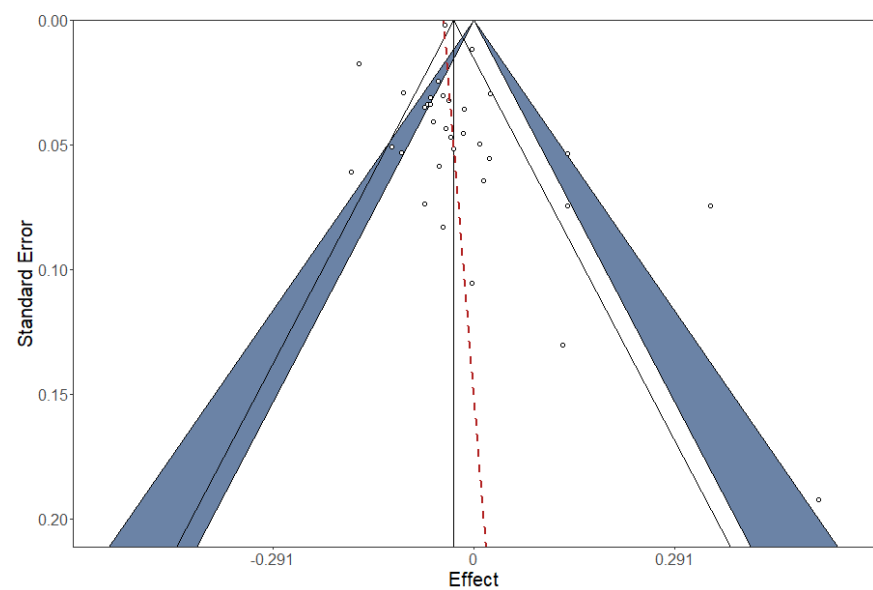
Modell ohne Rauschen



Umfassendes  
Modell



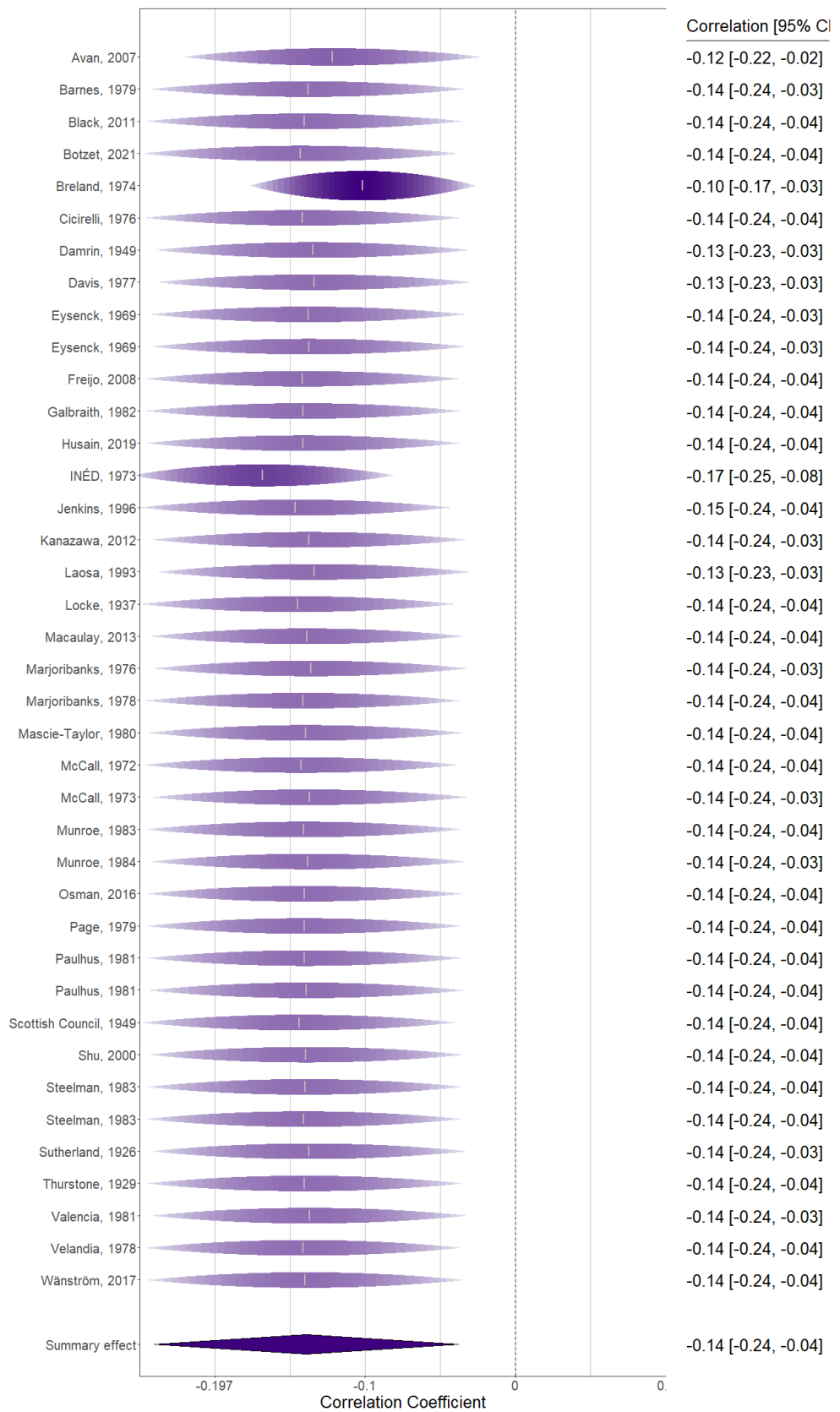
Modell mit Kontrolle  
für Familiengröße



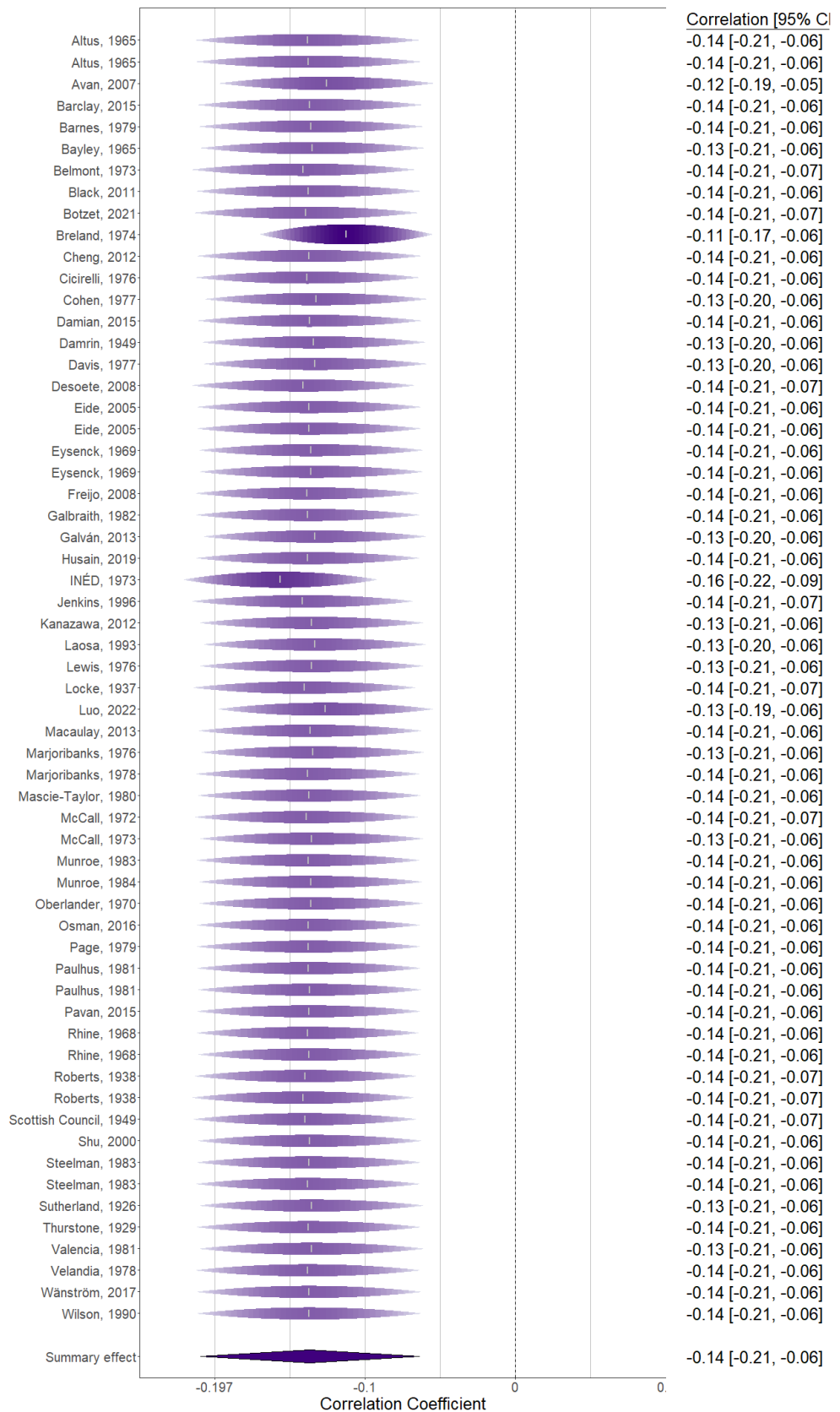
**Abbildung 26***Leave-One-Out Sensitivitätsanalysen aller drei Modelle*

Modell ohne

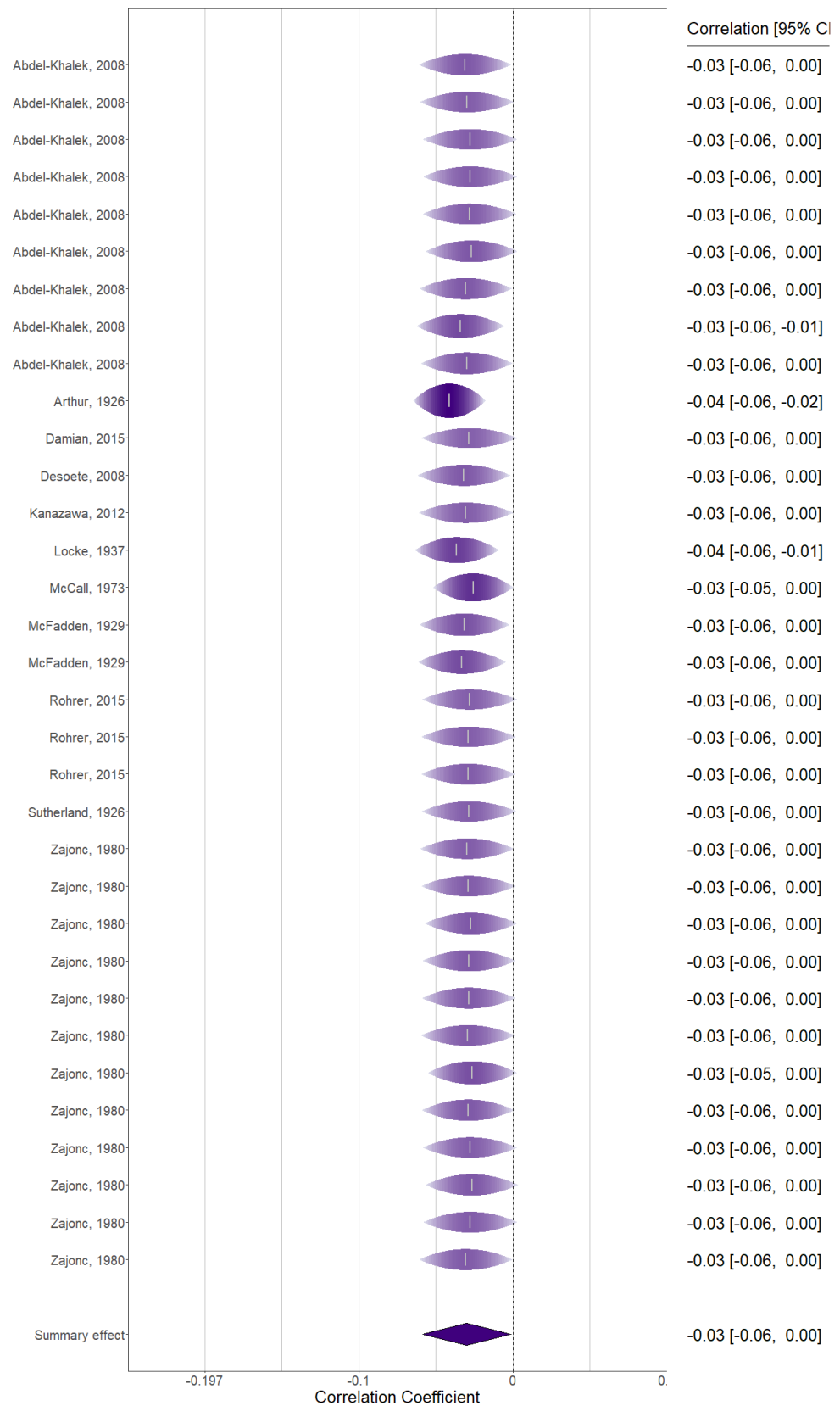
Rauschen



## Umfassendes Modell



Modell mit  
Kontrolle für  
Familiengröße



## Diskussion

Die vorliegende Arbeit stellt die erste bekannte Metaanalyse in der Thematik der Intelligenzforschung im Zusammenhang mit Geschwisterreihenfolge dar. Es wurden fast vier Millionen Proband\*innen in 70 Studien und 144 berichteten Effektstärken inkludiert. Diese Stichproben stammen aus verschiedenen Ländern, Altersgruppen und sozialen Schichten.

### Hauptanalysen

Die Studienlage stellte sich bisher als uneindeutig und widersprüchlich dar; in dieser Metaanalyse konnte erstmals ein gemittelter Effekt geschätzt und Antworten auf Fragen gefunden werden, deren Beantwortung in der Vergangenheit nicht einheitlich möglich war. Der Zusammenhang zwischen Geburtenrang und Intelligenz ergab einen kleinen negativen Effekt. Bei zusätzlicher Kontrolle für Familiengröße konnte dieser Effekt fast gänzlich herausgerechnet werden.

Die vorerst negative Beziehung konnte sowohl bei Inklusion bereinigter Korrelationen ( $r = -.156$ ), als auch bei Hinzunahme von Korrelationen gezeigt werden, die durch Mittelwertsunterschiede zwischen den Erst- und Spätergeborenen berechnet wurden ( $r = -.141$ ). Dadurch konnte die zuvor hypothesierte negative Korrelation angenommen werden. Der Zusammenhang zeigte zwar keine hohen Werte, jedoch sinke demnach der Intelligenzscore der inkludierten Stichproben mit steigendem Geburtenrang. Die älteren Geschwister erreichen höhere Scores als die jüngeren. Diese Ergebnisse stützen die Hypothese, dass Intelligenz mit steigender Geschwisterreihenfolge sinke, wie auch von der *Confluence* Theorie und der *Resource dilution* Theorie angenommen.

Beim Vergleich der bereinigten Korrelation mit jener mit Rauschen konnte kein signifikanter Unterschied gefunden werden. Die zweite Hypothese, welche eine Abweichung annahm, konnte deswegen nicht bestätigt werden. Den Ergebnissen zufolge habe es keinen Einfluss darauf, ob der Zusammenhang zwischen den Hauptvariablen durch Cohens  $d$  berechnet wurde oder durch eine lineare Korrelation.

Die dritte Hypothese bezieht sich auf den wohl wichtigsten Einflussfaktor für den vorliegenden Effekt – hierbei wurden nur Studien miteinbezogen, welche den Zusammenhang für Familiengröße kontrollierten. Diese Kontrolle wurde vor allem durch Verteilung der Effekte auf die verschiedenen Größen der Familien erreicht. Die multilevel Metaanalyse brachte ein Ergebnis von  $r = .004$ . Demzufolge gibt es keinen nennenswerten Zusammenhang zwischen Intelligenzscore und Geschwisterreihenfolge. Dieses Resultat wurde mit jenen verglichen, welche nicht für die Anzahl der Kinder innerhalb einer Familie kontrollierten. Es zeigten sich signifikante Unterschiede zum Modell mit Kontrolle dafür. Wird also die

Familiengröße in das Modell mitaufgenommen, so wird der Effekt und die negative Verbindung herausgerechnet und die Geschwisterreihenfolge steht nicht mehr in Bezug zur intellektuellen Leistung.

Aus dieser Aussage könnte abgeleitet werden, dass Familiengröße einen Einfluss auf die Intelligenz der Kinder haben kann. Die *Admixture* Hypothese stellte die Vermutung auf, dass kleinere Familien intelligentere Kinder haben. Dies kann jedoch in vorliegender Arbeit weder bestätigt noch verneint werden. Eine weitere Metaanalyse wäre notwendig, welche sich spezifisch mit dem Zusammenhang zwischen Familiengröße und Intelligenz beschäftigt. Mit dieser Arbeit konnte aufgezeigt werden, dass der Effekt von Geschwisterreihenfolge und Intelligenz bei Kontrolle für Familiengröße verschwindet. Dies könne als Ursprung haben, dass die beiden Familienstrukturvariablen, wie zuvor berichtet, stark miteinander korrelieren und kovariieren. Es ist schwierig, mit Sicherheit festzustellen, ob der Geburtsrang allein für die negative Korrelation zwischen Intelligenz und Geschwisterreihenfolge verantwortlich ist oder ob die Familiengröße eine beeinflussende Rolle spielt. Möglich ist ein gemeinsamer Beitrag von Geburtenreihenfolge und Familiengröße zur Erklärung dieses Effekts. In zukünftigen Studien sollten daher die eine Variable nicht ohne Berücksichtigung der jeweils anderen untersucht werden.

In bisheriger Forschung (beispielsweise in Belmont & Marolla, 1973 oder Black et al., 2011) wurde sehr häufig eine höhere negative Korrelation zwischen Intelligenz und Geschwisterreihenfolge als in dieser Metaanalyse beschrieben. Die nach Cohen (1969) als klein bewertete Effektstärke der Metaanalysen stehen im Widerspruch zu den zuvor publizierten Studien. Ein Grund für den bisher überschätzten Effekt könnte ein nicht erkannter *Publication Bias* sein. Durch die Popularität der *Confluence* Theorie sind viele Untersuchungen theoriegesteuert und gehen schon im Vorhinein von einem negativen Zusammenhang zwischen Intelligenz und Geschwisterreihenfolge aus. Wird dabei eine gegenteilige positive Korrelation oder eine Nulleffekt erhoben, so kann es sein, dass diese Studien nicht publiziert werden, da sie nicht mit den Erwartungen der *Confluence* Theorie übereinstimmen. Der Einfluss der Familiengröße auf diesen Effekt und eine vielleicht vernachlässigbare Korrelation zwischen den Hauptvariablen sind ebenso nicht mit dieser Theorie vereinbar. Unvoreingenommene umfangreiche Untersuchungen wären in Zukunft wichtig. Die vorliegende Metaanalyse ermöglicht eine bessere Kontrolle der Störvariablen, hat eine größere Stichprobe und nimmt somit eine repräsentative Perspektive ein.

Aus metaanalytischer Sicht hängen Intelligenz und Geschwisterreihenfolge sehr schwach zusammen, bei Kontrolle für die Anzahl der Kinder innerhalb einer Familie, der

Familiengröße, wird dieser Effekt vernachlässigbar klein. In einigen zuvor durchgeführten Untersuchungen konnte dies schon bestätigt werden. In der Untersuchung von Kanazawa (2012) sank die Korrelation von  $-.201$  auf  $-.002$ ; ebenso sanken in Damian (2015) bei einer Stichprobe mit über 257 000 Teilnehmer\*innen die ohnehin schon kleinen Korrelationen bei verbaler, mathematischer und räumlicher Fähigkeit auf  $-.08$ ,  $-.05$  bzw.  $-.00$ . Auch bei Baldur (2021) oder Husain et al. (2019) wurde die Korrelation vollständig durch Familiengröße mediert.

Auch wenn der Zusammenhang abgeschwächt wurde, zeigt die konsistente Tendenz in den inkludierten Studien, dass ein gewisser negativer Zusammenhang zwischen Geschwisterreihenfolge und Intelligenz besteht. Obwohl diese Korrelation in früheren Studien teilweise überschätzt wurde, ist sie dennoch thematisch relevant. Denkbar ist, dass der Geburtenrang und die Familiengröße in Kombination zur Erklärung des Gesamteffekts beitragen. Der spezifische Einfluss von Familiengröße sollte in zukünftigen Studien weiter untersucht werden.

### **Moderatoranalysen**

Um die Beziehung zwischen Geschwisterrang und Familiengröße weiter zu untersuchen, wurde eine Moderatoranalyse mit der mittleren Familiengröße für alle drei Modelle durchgeführt. Die Moderation und somit die vierte Hypothese konnten nicht bestätigt werden. Es können mehrere Faktoren dem zugrunde liegen, dass diese Moderation nicht signifikant war, aber die Hauptanalyse den Einfluss von Familiengröße deutlich zeigte. Beispielsweise wurde die mittlere Familiengröße sehr häufig nicht berichtet. Dies ergab sich häufiger bei den inkludierten Studien im Modell ohne Rauschen und dem umfassenden Modell als bei dem Modell mit Kontrolle für Familiengröße. Das liegt daran, dass sich wenig Studien mit der Anzahl der Kinder und mit Geschwisterreihenfolge gleichzeitig beschäftigen. Dadurch resultieren auch unterschiedliche Stichproben, die in diese beiden Analysen miteinbezogen wurden. Es ist möglich, dass diese eingeschränkten Teilmengen an Studien nicht repräsentativ für die Gesamtstichprobe waren und demnach die Ergebnisse verzerrten. Zudem beruhen die Metaregression und das Modell mit Kontrolle für Familiengröße auf unterschiedlichen statistischen Methoden und der Art und Weise, wie sie die Moderation durch Familiengröße erfassen. Während das Hauptmodell die Familiengröße direkt als Kontrollvariable einbezieht, analysiert die Metaregression die Beziehung zwischen der resultierenden Effektgröße und der mittleren Familiengröße in der Studie. Es ist möglich, dass die beiden Methoden unterschiedliche Aspekte der Familiengröße erfassen, was zu unterschiedlichen Ergebnissen führen kann.

Die eingeschlossenen Studien stammen aus verschiedenen Ländern oder Altersgruppen und somit unterschiedlichen Kontexten, was möglicherweise die Ergebnisse beeinflussen könnte. Für die Testung auf Einflussfaktoren nach der *Admixture* Hypothese, wie beispielsweise die Stichprobenkomposition oder das Publikationsjahr und der Publikationsstatus, wurden weitere Moderatoranalysen bei den drei multilevel Modellen durchgeführt. Die Hypothesen Fünf bis Acht konnten nicht bestätigt werden. Das Land der Datenerhebung oder das GDP als Hinweise auf den sozioökonomischen Status der Stichprobe konnten nicht als signifikante Moderatoren identifiziert werden.

In der Vergangenheit (beispielsweise in O'Shields, 1980) wurde der sozioökonomische Status als Einflussvariable auf die Korrelation zwischen Intelligenz und Geschwisterreihenfolge beschrieben. Die hier erwiesene nicht signifikante Moderation steht dazu im Gegensatz. Ein möglicher Grund für dies liegt in der uneinheitlichen Definition und Erhebung des sozioökonomischen Status. In Erststudien können Abstufungen dieser Variable genau definiert und gemessen werden, doch die Vergleichbarkeit der unterschiedlichen Ansätze in einer Metaanalyse ist nicht gegeben. Daher wurde hier vor allem das GDP als Marker für die Variable genutzt. Das vorliegende Ergebnis unterstützt jedoch eine Studie von Barclay (2015) und somit keine Variation des Effektes durch den sozioökonomischen Status.

Im Gegensatz zu den Aussagen des *Tutoring* Effekts der *Confluence* Theorie und dem damit verbundenen Einfluss des Alters der Person auf Intelligenzunterschiede, konnte in dieser Arbeit keine signifikante Moderation durch das Alter gefunden werden. Die Annahme der Autor\*innen der Theorie (Zajonc, Markus & Markus, 1979), dass der *Birth order* Effekt eher bei älteren Personen als bei Kindern vorliege, kann nicht unterstützt werden.

In einigen Studien (z.B. Alexander & Eckland, 1974) wird zudem männlichen Kindern grundsätzlich eine höhere Intelligenz zugeschrieben. In dieser Arbeit konnten keine Hinweise auf die Beeinflussung des Effekts durch das Geschlecht der Testpersonen gefunden werden.

Die einzige signifikante Moderation zeigte sich im Familiengröße-Modell durch das Jahr der Publikation. Mit steigender Jahreszahl sank der Wert des Korrelationskoeffizienten sehr leicht. Der *p*-Wert dieses Ergebnisses lag knapp unter der Signifikanzgrenze von .05. Wie auch in Abbildung 17 ersichtlich, ergaben die Studien vor 1940 auffallend hohe Korrelationswerte, was auf Hebeleffekte dieser schließen lässt. Dies konnte durch die Berechnung der *Hat* Werte jener Stichproben bestätigt werden. Die drei höchsten *Hat* Werte hatten Studien aus diesen Jahren. Ohne die Ergebnisse vor 1940 wurde die ohnehin sehr



kleine negative Steigung ( $b = -.002$ ) nicht nur relativiert, sondern sogar leicht positiv ( $b = .002, p = .005$ ). Durch die grenzwertige Signifikanz und die Hebeleffekte kann davon ausgegangen werden, dass sich die berichteten Korrelationen nach Jahr der Publikation nicht stark unterscheiden.

Bei den Variablen, welche als Moderatoren in die Analyse aufgenommen wurden, handelt es sich um gemittelte Werte für die gesamte Stichprobe oder Studie und um keine genauen Angaben für jede Testperson. Häufig wurden diese zudem nicht berichtet und es konnte nur mit einer kleineren Anzahl an Studien gearbeitet werden. Dies traf vor allem bei der Testung nach dem GDP PPP per capita oder der mittleren Familiengröße zu. Aus diesen Gründen ist es möglich, dass die Moderatorvariablen in Primärstudien trotzdem den Gesamteffekt signifikant beeinflussen können.

### **Publication Bias**

Die verschiedenen Methoden für das Testen auf einen *Publication Bias* in den einzelnen Modellen zeigten kein einheitliches Ergebnis. Für das multilevel Modell konnte alleinig die Egger MLMA angewandt werden. Dabei ergaben sich Anzeichen eines *Publication Bias* einzig beim Modell mit Kontrolle für die Familiengröße. Demnach sind die einzelnen Effektstärken nicht symmetrisch um den Gesamteffekt verteilt. Grund für dieses Ergebnis kann sein, dass die Studienzahl eher gering und der Nulleffekt bei kleiner Stichprobengröße schwer zu detektieren ist, was ein signifikantes Ergebnis und somit die Publikation dieser nicht signifikanten Evidenz unwahrscheinlicher macht.

Auch die aggregierten Modelle mit unabhängigen Samples wurden mit einigen Methoden auf *Publication Bias* getestet. Die Rangkorrelation ergab eine Asymmetrie bei allein drei Modellen und somit Hinweis auf diesen Fehler. Im Gegensatz dazu konnte bei Testung mit nur den signifikanten  $p$ -Werten mit der  $p$ -curve in keinem der drei Modellen eine linksschiefe Verteilung und dementsprechend kein Bias festgestellt werden. Bei der  $p$ -uniform Methode zeigte sich der integrierte Test auf *Publication Bias* nur bei dem Modell mit Kontrolle für Familiengröße nicht signifikant. In der graphischen Darstellung (siehe Abbildung 22 (b) und 23 (b) ) wird jedoch ein Problem des Algorithmus' beim umfassenden und beim Modell ohne Rauschen ersichtlich. Diese Veranschaulichung stimmt nicht mit den statistischen Ergebnissen überein. Aus diesem Grund wird hier die  $p$ -uniform Methode nicht weiter interpretiert. Hohe Änderungen zum ursprünglichen Effekt bei allen Modellen ergaben die bereinigten Effektschätzungen der  $p$ -uniform\* Funktion. Auch bei der *Trim-and-Fill* Methode mussten für den Ausgleich auf einen *Publication Bias* recht viele Stichproben imputiert werden.

Zusammengefasst ergab sich bei den aggregierten Daten ein uneinheitliches Bild hinsichtlich des Vorliegens eines *Publication Bias*. Die publizierten Modelle zeigten im Allgemeinen hohe Heterogenität, was zu Problemen bei einigen Methoden führen kann. Die *p-curve* oder *p-uniform* Methode können die adaptierten Effektstärken überschätzen. *P-uniform\** könne hingegen besser mit hoher Heterogenität umgehen (van Aert & van Assen, 2018). Es ist möglich, dass Asymmetrie der *Funnel Plots* nicht durch einen *Publication Bias* bedingt ist, sondern abermals durch die hohe Heterogenität (Fernández-Castilla et al., 2021).

Wichtig nochmals anzumerken ist, dass die multilevel und die aggregierten Modelle nicht identisch sind und dementsprechend nur mittels der Egger MLMA Rückschlüsse für die multilevel Metaanalyse zulässig sind. Ein *Publication Bias* könnte demnach vor allem beim Modell mit Kontrolle für Familiengröße vorliegen. Da die gemittelte Korrelation dieses Modells ohnehin vernachlässigbar klein ist, könnte diese durch einen *Publication Bias* unterschätzt werden. Um ein signifikantes Ergebnis beim Vorliegen eines sehr kleinen Effekts zu erlangen, müssten große Stichproben verwendet werden. Die Berechnung eines nach *Publication Bias* bereinigten Effekts ist beim multilevel Modell noch nicht möglich; jedoch könnte *p-uniform \** für das aggregierte Familiengröße-Modell einen Hinweis darauf liefern, in welche Richtung der Effekt unterschätzt werden könnte. Der bereinigte Effekt lag hier bei  $r = -.013$ , was auf einen abermals negativen Zusammenhang von Intelligenz und Geschwisterreihenfolge bei Kontrolle für Familiengröße schließen lässt. So kann auch bei Ausgleich eines *Publication Bias* und Einbezug der Familiengröße erneut geschlussfolgert werden, dass die Korrelation der beiden Hauptvariablen klein, jedoch negativ ist.

### **Limitationen und Implikationen für zukünftige Forschung**

Trotz des Versuchs der bestmöglichen Ausführung dieser multilevel Metaanalyse gibt es natürlich Aspekte, die in zukünftiger Forschung zu Intelligenz und Geschwisterreihenfolge verbessert oder verändert werden könnten. Auf erwägenswerte Einschränkungen wird im Folgenden eingegangen.

In den inkludierten Studien wird Intelligenz bzw. kognitive Fähigkeit verschiedenartig definiert. Teilweise wurde ein g-Faktor und in anderen Fällen Subtypen der Intelligenz berichtet. Dadurch unterscheidet sich auch die Art, wie getestet wurde. Durch die Inklusion von nur möglichst standardisierten Intelligenztests konnte dieses Problem in dieser Arbeit zu einem gewissen Maße eingeschränkt werden.

Eine weitere Erschwernis stellte zudem die unterschiedliche Definition und Erhebung von Geburtsrang dar. Die Fragen hierbei sind, ob erwachsene Geschwister, Mehrlingsgeburten oder totgeborene Geschwister dazugezählt werden sollten. Dies ist in

vielen Erststudien nicht ausreichend berichtet und kontrolliert worden. In dieser Untersuchung wurden deshalb, falls durch Dokumentation in den Studien möglich, Mehrlinge oder Einzelkinder ausgeschlossen. Die Erläuterungen davon, was Familie ist und was es bedeutet, Geschwister zu haben, sind abermals schwierig. Familien sind heutzutage sehr unterschiedlich aufgebaut – es gibt Patchworkfamilien oder Personen außerhalb der biologischen Familie, welche sich wie Eltern um die Kinder kümmern. Es ist nicht möglich die vielen unterschiedlichen Verständnisse davon, wer der Bruder oder die Schwester ist, miteinzubeziehen. In den meisten Studien wurden demnach nur Geschwister mit den gleichen leiblichen Eltern zur Berechnung des Geburtenrangs herangezogen. Durch eine einheitliche Verwendung der Definition von Geschwisterreihenfolge und Intelligenz in der Forschung wären jedenfalls übertragbare und einheitliche Ergebnisse möglich.

Des Weiteren könnten andere, zuvor bekannte oder auch unbekannte Störvariablen, welche hier nicht in den Moderatoranalysen kontrolliert wurden, die Ergebnisse verfälschen. Im Sinne der *Resource dilution* Theorie kann beispielsweise eine unterschiedliche Aufteilung der familiären Ressourcen auf die verschiedenen Geschwisterkinder Auswirkungen auf den Zusammenhang der Geburtenreihenfolge und Intelligenz haben. Vor allem immaterielle Ressourcen können nicht konkret gemessen werden, was die Untersuchung dieses Einflusses erschwert. Auch die Interaktionen innerhalb der Familie, sei es zwischen den Eltern und dem Kind oder zwischen den Geschwistern, sind als mögliche Moderatoren schwer zu berücksichtigen. Der Erwartungsdruck der Eltern, z.B. an Erstgeborene, oder das Lehren (siehe *Tutoring* Effekt des *Confluence* Modells) oder Lernen von anderen Geschwistern kann Leistung und Intelligenz beeinflussen. Die Intelligenz oder Bildung der Eltern und damit Auswirkungen auf die Bildungschancen oder der vererbten Intelligenz der Kinder im Allgemeinen konnte aufgrund der wenigen Forschungen zur vorliegenden Thematik nicht untersucht werden. Der sozioökonomische Status als wichtiger Einflussfaktor nach der *Admixture* Hypothese wird in Primärstudien unterschiedlich kategorisiert und konnte deshalb nur mithilfe des GDP und der geographischen Lage als Hinweise darauf miteinbezogen und so vergleichbar gemacht werden. Weitere denkbare Auswirkungen auf die Korrelation zwischen Intelligenz und Geschwisterreihenfolge sind die Struktur bzw. der Aufbau von Familien (Patchwork-Familien, Adoptivfamilien, Mehrgenerationenhaushalte etc.), das Geschlecht der Geschwister oder der zeitliche Abstand zwischen den Geburten der Geschwister. All diese Variablen sollten in Zukunft bei der Beschreibung der Korrelation berücksichtigt werden.

Ein weiteres Problem spezifisch für diese Thematik sind die teilweise sehr alten und nicht mehr verfügbaren Datensätze. Einige Autor\*innen hatten diese schon vor Jahrzehnten vernichtet oder sind leider selbst schon verstorben. Sehr frühe Studien konnten deshalb nicht immer in die Analyse miteinbezogen werden. Zudem bedienten sich einige Forscher\*innen derselben Daten von großen Erhebungen. Um Abhängigkeiten zu vermeiden, wurden nur die größten oder aktuellsten Stichproben miteinbezogen.

Auch gab es in der Vergangenheit wenige Untersuchungen, welche in diesem Kontext für Familiengröße kontrollierten. Um die Ergebnisse der Metaanalyse zu prüfen, sind weitere Studien nötig, welche zusätzlich diese Variable miteinbeziehen und mit den Ergebnissen ohne Kontrolle vergleichen. Der Effekt beim Familiengröße Modell ist außerdem sehr klein und ein *Publication Bias* möglich. Wichtig ist hierbei zudem, wie die Anzahl der Kinder konstant gehalten wird. Dies war bisher recht unterschiedlich und bedarf einer einheitlichen Herangehensweise.

Auswirkungen auf die Resultate der Metaanalyse kann außerdem die Wahl des Korrelationskoeffizienten zur Beschreibung des Zusammenhangs zwischen Geburtenreihenfolge und kognitiver Leistung haben. Die Verwendung von gepoolten Korrelationen als Maß dafür kann die Heterogenität in den Studien möglicherweise nicht vollständig berücksichtigen. Eventuell ist die Beziehung nicht linear. Vielerorts gibt es das Vorurteil, dass die sogenannten „Sandwich-Kindern“, also Personen, die sowohl ältere, als auch jüngere Geschwister haben, von Vernachlässigung durch die Eltern und deshalb niedrigerer Intelligenz betroffen sind (z.B. in Süddeutsche Zeitung, 2019). Dies könnte beispielweise durch eine U-Kurve beschrieben und untersucht werden, da nach diesem Vorurteil die Kinder mit mittleren Geburtenrängen niedrigere Intelligenzscores aufweisen müssten. In dieser Arbeit wurde die Pearson Korrelation verwendet, da diese durch ihre einfache Interpretierbarkeit gut zu verstehen ist. Zudem berichteten die meisten Untersuchungen die Ergebnisse mithilfe dieser.

Die Verwendung eines Modells mit mehreren Levels und somit ein besserer Umgang mit abhängigen Daten stellt im Sinne einer Metaanalyse eine recht neue Methode dar. Multilevel Metaanalysen sind dementsprechend noch jung und viele gängige Methoden für Metaanalysen noch nicht für diese angepasst. Hier konnten vor allem die Tests auf *Publication Bias* nur mit aggregierten Modellen berechnet werden, da diese Methoden unabhängige Datensätze voraussetzen. Außerdem sind viele der *Publication Bias* Methoden nicht für höhere Heterogenität konzipiert. In Zukunft sollte weiter an der Adaption für multilevel Analysen gearbeitet und vorliegende Ergebnisse mit diesen geprüft werden.

Für viele der genannten Limitationen konnte in dieser Untersuchung Abhilfe geschaffen werden. Aus diesem Grund stellt die hier verwendete Methodik eine Möglichkeit zur Anwendung der klassischen Methoden für Metaanalysen bei multilevel Berechnungen dar.

### **Conclusio**

Diese multilevel Metaanalyse enthält wichtige Ergebnisse und eine erstmalige Zusammenfassung der Thematik zu Geschwisterreihenfolge und Intelligenz. Es zeigte sich bei Zusammenfassung von 70 Studien zu dem Thema, dass es einen kleinen negativen Zusammenhang des Geburtenrangs und kognitiver Leistung gibt. Ältere Geschwister scheinen demnach intelligenter zu sein als jüngere. Wird jedoch für Familiengröße kontrolliert und dieser Einfluss herausgerechnet, geht die Korrelationen zwischen Geschwisterreihenfolge und Intelligenz gegen Null. Dies deutet darauf hin, dass der Zusammenhang einerseits klein, jedoch negativ ist und andererseits die hohe Korrelation der Familienstrukturvariablen maßgeblich für die Ergebnisse sind.

Vor allem für das Modell mit Kontrolle für Familiengröße besteht die Möglichkeit eines *Publication Bias*, wobei die gemittelte Korrelation in diesem Modell ohnehin vernachlässigbar klein ist und durch einen *Publication Bias* möglicherweise unterschätzt wird. Bei den auf unabhängige Stichproben aggregierten Daten gibt es keinen klaren Hinweis auf *Publication Bias*.

Es konnte außer der Moderation durch das Publikationsjahr bei dem Modell mit Kontrolle für Familiengröße keine signifikante Moderation gefunden werden. Diese grenzwertig signifikante Moderation kann jedoch auf Hebeleffekte durch Studien vor 1940 zurückgeführt werden und ist daher nicht weiter relevant. Auch das Alter, Geschlecht oder der sozioökonomische Status der Stichprobe und der Publikationsstatus haben keinen signifikanten Einfluss auf den Zusammenhang zwischen Geburtenrang und Intelligenz.

### Literaturverzeichnis

- \*Abdel-Khalek, A. M. & Lynn, R. (2008). Intelligence, family size and birth order: Some data from Kuwait. *Personality and Individual Differences*, 44(4), 1032–1038.  
<https://doi.org/10.1016/j.paid.2007.10.029>
- Alexander, K. L. & Eckland, B. K. (1974). Sex Differences in the Educational Attainment Process. *American Sociological Review*, 39(5), 668. <https://doi.org/10.2307/2094313>
- \*Altus, W. D. (1965). Birth order and scholastic aptitude. *Journal of Consulting Psychology*, 29(3), 202–205. <https://doi.org/10.1037/h0022158>
- Alwin, D. F. & Thornton, A. (1984). Family Origins and the Schooling Process: Early Versus Late Influence of Parental Characteristics. *American Sociological Review*, 49(6), 784.  
<https://doi.org/10.2307/2095530>
- \*Arthur, G. (1926). The relation of I. Q. to position in family. *Journal of Educational Psychology*, 17(8), 541–550. <https://doi.org/10.1037/h0072389>
- Assink, M. & Wibbelink, C. J. (2016). Fitting three-level meta-analytic models in R: A step-by-step tutorial. *The quantitative methods for psychology*, 12(3), 154–174.  
<https://doi.org/10.20982/tqmp.12.3.p154>
- \*Avan, B., Rahbar, M. & Raza, S. (2007). The role of family configuration in early childhood intellectual development in the context of an extended family system in Pakistan. *Journal of Postgraduate Medicine*, 53(1), 27. <https://doi.org/10.4103/0022-3859.30324>
- \*Baldur, P. (2021). *Effekte der Geburtenreihenfolge auf Intelligenz und Bildung in einer mexikanischen Stichprobe* [Unveröffentlichte Bachelorarbeit]. Georg-August-Universität Göttingen.
- \*Barclay, K. J. (2015). A within-family analysis of birth order and intelligence using population conscription data on Swedish men. *Intelligence*, 49, 134–143.  
<https://doi.org/10.1016/j.intell.2014.12.007>
- \*Barnes, J., Fisher, J. & Palmer, M. (1979). Family Characteristics and Intellectual Growth: An Examination By Race. *Educational and Psychological Measurement*, 39(3), 625–636. <https://doi.org/10.1177/001316447903900313>
- \*Bayley, N. (1965). Comparisons of Mental and Motor Test Scores for Ages 1-15 Months by Sex, Birth Order, Race, Geographical Location, and Education of Parents. *Child Development*, 36(2), 379. <https://doi.org/10.2307/1126464>
- Begg, C. B. & Mazumdar, M. (1994). Operating Characteristics of a Rank Correlation Test for Publication Bias. *Biometrics*, 50(4), 1088. <https://doi.org/10.2307/2533446>

- \*Belmont, L. & Marolla, F. A. (1973). Birth Order, Family Size, and Intelligence. *Science*, 182(4117), 1096–1101. <https://doi.org/10.1126/science.182.4117.1096>
- \*Black, S. E., Devereux, P. J. & Salvanes, K. G. (2011). Older and Wiser? Birth Order and IQ of Young Men. *CESifo Economic Studies*, 57(1), 103–120. <https://doi.org/10.1093/cesifo/ifq022>
- Blake, J. A. (1981). Family Size and the Quality of Children. *Demography*, 18(4), 421–442. <https://doi.org/10.2307/2060941>
- \*Botzet, L. J., Rohrer, J. M. & Arslan, R. C. (2021). Analysing effects of birth order on intelligence, educational attainment, big five and risk aversion in an Indonesian sample. *European Journal of Personality*, 35(2), 234–248. <https://doi.org/10.1002/per.2285>
- \*Brancov, I. (1996). *Birth order and family size effect on academic achievement among american indians* (Publikationsnummer DP18597) [Dissertation, University of South Dakota]. ProQuest Dissertations Publishing. <https://www.proquest.com/docview/856339846?pq-origsite=gscholar&fromopenview=true>
- \*Breland, H. M. (1974). Birth Order, Family Configuration, and Verbal Achievement. *Child Development*, 45(4), 1011–1019. <https://doi.org/10.2307/1128089>
- Card, N. A. (2011). *Applied Meta-Analysis for Social Science Research*. Guilford Publications.
- \*Cheng, C. C. J., Wang, W. L., Sung, Y. T., Wang, Y. C., Su, S. Y. & Li, C. Y. (2012). Effect modification by parental education on the associations of birth order and gender with learning achievement in adolescents. *Child: care, health and development*, 894–902. <https://doi.org/10.1111/j.1365-2214.2012.01427.x>
- Cheung, M. W. (2014). Modeling dependent effect sizes with three-level meta-analyses: A structural equation modeling approach. *Psychological Methods*, 19(2), 211–229. <https://doi.org/10.1037/a0032968>
- \*Cicirelli, V. G. (1976). Sibling structure and intellectual ability. *Developmental Psychology*, 12(4), 369–370. <https://doi.org/10.1037/0012-1649.12.4.369>
- Cicirelli, V. G. (1978). The Relationship of Sibling Structure to Intellectual Abilities and Achievement. *Review of Educational Research*, 48(3), 365–379. <https://doi.org/10.3102/00346543048003365>

- \*Cohen, S. E. & Beckwith, L. (1977). Caregiving Behaviors and Early Cognitive Development as Related to Ordinal Position in Preterm Infants. *Child Development*, 48(1), 152. <https://doi.org/10.2307/1128894>
- Cullinan, P. (2005). Childhood allergies, birth order and family size. *Thorax*, 61(1), 3–5. <https://doi.org/10.1136/thx.2004.039701>
- \*Damian, R. I. & Roberts, B. W. (2015). The associations of birth order with personality and intelligence in a representative sample of U.S. high school students. *Journal of Research in Personality*, 58, 96–105. <https://doi.org/10.1016/j.jrp.2015.05.005>
- \*Damrin, D. E. (1949). Family Size and Sibling Age, Sex, and Position as Related to Certain Aspects of Adjustment. *The Journal of Social Psychology*, 29(1), 93–102. <https://doi.org/10.1080/00224545.1949.9918944>
- \*Davis, D. J., Cahan, S. & Bashi, J. (1977). Birth Order and Intellectual Development: The Confluence Model in the Light of Cross-Cultural Evidence. *Science*, 196(4297), 1470–1472. <https://doi.org/10.1126/science.867047>
- Deary, I. J. (2012). Intelligence. *Annual Review of Psychology*, 63(1), 453–482. <https://doi.org/10.1146/annurev-psych-120710-100353>
- De Jonge, H., Jak, S. & Kan, K. (2020). Dealing With Artificially Dichotomized Variables in Meta-Analytic Structural Equation Modeling. *Zeitschrift für Psychologie*, 228(1), 25–35. <https://doi.org/10.1027/2151-2604/a000395>
- \*Desoete, A. (2008). Do birth order, family size and gender affect arithmetic achievement in elementary school? *Electronic Journal of Research in Education Psychology*, 6(14). <https://doi.org/10.25115/ejrep.v6i14.1274>
- Dowdy, A., Hantula, D. A., Travers, J. C. & Tincani, M. (2021). Meta-Analytic Methods to Detect Publication Bias in Behavior Science Research. *Perspectives on behavior science*, 45(1), 37–52. <https://doi.org/10.1007/s40614-021-00303-0>
- Downey, D. B. (1995). When Bigger Is Not Better: Family Size, Parental Resources, and Children's Educational Performance. *American Sociological Review*, 60(5), 746. <https://doi.org/10.2307/2096320>
- Duval, S. & Tweedie, R. L. (2000). Trim and Fill: A Simple Funnel-Plot-Based Method of Testing and Adjusting for Publication Bias in Meta-Analysis. *Biometrics*, 56(2), 455–463. <https://doi.org/10.1111/j.0006-341x.2000.00455.x>



- \*Eide, M. G., Øyen, N., Skjærven, R., Irgens, L. M., Bjerkedal, T. & Nilsen, S. T. (2005). Breech Delivery and Intelligence: A Population-Based Study of 8,738 Breech Infants. *Obstetrics & Gynecology*, 105(1), 4–11.  
<https://doi.org/10.1097/01.aog.0000149743.80837.d3>
- \*Eysenck, H. J. & Cookson, D. (1969). Personality in primary school children. *British Journal of Educational Psychology*, 39(2), 123–130. <https://doi.org/10.1111/j.2044-8279.1969.tb02055.x>
- Fernández-Castilla, B., Declercq, L., Jamshidi, L., Beretvas, S. N., Onghena, P. & Van Den Noortgate, W. (2021). Detecting Selection Bias in Meta-Analyses with Multiple Outcomes: A Simulation Study. *Journal of Experimental Education*, 89(1), 125–144.  
<https://doi.org/10.1080/00220973.2019.1582470>
- Fernández-Castilla, B., Jamshidi, L., Declercq, L., Beretvas, S. N., Onghena, P. & Van Den Noortgate, W. (2020). The application of meta-analytic (multi-level) models with multiple random effects: A systematic review. *Behavior Research Methods*, 52(5), 2031–2052. <https://doi.org/10.3758/s13428-020-01373-9>
- \*Freijo, E. B. A., Oliva, A., Olabarrieta, F., Martín, J. L., Manzano, A. & Richards, M. P. M. (2008). Quality of family context or sibling status? Influences on cognitive development. *Early Child Development and Care*, 178(2), 153–164.  
<https://doi.org/10.1080/03004430600685373>
- \*Galbraith, R. C. (1982). Sibling spacing and intellectual development: A closer look at the confluence models. *Developmental Psychology*, 18(2), 151–173.  
<https://doi.org/10.1037/0012-1649.18.2.151>
- \*Galván, M., Uauy, R., Corvalán, C., López-Rodríguez, G. & Kain, J. (2013). Determinants of Cognitive Development of Low SES Children in Chile: A Post-transitional Country with Rising Childhood Obesity Rates. *Maternal and Child Health Journal*, 17(7), 1243–1251. <https://doi.org/10.1007/s10995-012-1121-9>
- \*Husain, N. I. A. E., Meisenberg, G., Becker, D., Bakhiet, S. F., Essa, Y. A. S., Lynn, R. & Al Julayghim, F. M. H. (2019). Intelligence, family income and parental education in the Sudan. *Intelligence*, 77, 101402. <https://doi.org/10.1016/j.intell.2019.101402>
- Hemminki, K. & Mutanen, P. (2001). Birth order, family size, and the risk of cancer in young and middle-aged adults. *British Journal of Cancer*, 84(11), 1466–1471.  
<https://doi.org/10.1054/bjoc.2001.1811>

Higgins J.P.T., Thomas J., Chandler J., Cumpston M., Li T., Page M.J. & Welch V.A (2022).

*Cochrane Handbook for Systematic Reviews of Interventions version 6.3 (updated February 2022)*. Cochrane, 2022. [www.training.cochrane.org/handbook](http://www.training.cochrane.org/handbook)

Jacobs, P. & Viechtbauer, W. (2017). Estimation of the biserial correlation and its sampling variance for use in meta-analysis. *Research Synthesis Methods*, 8(2), 161–180.

<https://doi.org/10.1002/jrsm.1218>

\*Jenkins, J. M. & Astington, J. W. (1996). Cognitive factors and family structure associated with theory of mind development in young children. *Developmental Psychology*, 32(1), 70–78. <https://doi.org/10.1037/0012-1649.32.1.70>

Kaindl, M. & Schipfer, R. K. (2022). *Familien in Zahlen 2022 Statistische Informationen zu Familien in Österreich*. Projekt: Familienforschung in Österreich, 2022.

<https://doi.org/10.25365/phaidra.370>

\*Kanazawa, S. (2012). Intelligence, Birth Order, and Family Size. *Personality and Social Psychology Bulletin*, 38(9), 1157–1164. <https://doi.org/10.1177/0146167212445911>

Knippschild, S., Baulig, C. & Krummenauer, F. (2015). Heterogenität in Meta-Analysen – kein Vergleich von Äpfeln und Birnen erlaubt. *zzi*, 31(3). <https://www.online-zzi.de/archiv/ausgabe/artikel/zzi-3-2015/1634-heterogenitaet-in-meta-analysen-kein-vergleich-von-aepfeln-und-birnen-erlaubt/>

\*Lalsangzuala, C. (2010). *Impact of birth order on intellectual capacity, memory span, and concept formation among mizo children* [Dissertation, Mizoram University].

Shodhganga@INFLIBNET Centre. <http://hdl.handle.net/10603/234313>

\*Laosa, L. M. (1993). Family characteristics as predictors of individual differences in chicano children's emergent school readiness. *ETS Research Report Series*, 1993(2), i–62.

<https://doi.org/10.1002/j.2333-8504.1993.tb01545.x>

\*Lavizzo, N. I. (1980). *An Investigation of the Relationship Among Birth Order, Intelligence, Sex, Biorhythm and Academic Achievement* [Dissertation, Loyola University of Chicago]. Loyola University Chicago Dissertations.

[https://ecommons.luc.edu/luc\\_diss/2027](https://ecommons.luc.edu/luc_diss/2027)

*Leave-one-out-Meta-Analyse* (2023). SciExperts. Abgerufen am 23. Mai 2023, von

<https://sciexperts.com/leave-one-out-meta-analysis/?lang=de>

\*Lewis, M. & Gallas, H. (1976). COGNITIVE PERFORMANCE IN THE 12-WEEK-OLD INFANT: THE EFFECTS OF BIRTH ORDER, BIRTH SPACING, SEX, AND SOCIAL CLASS. *ETS Research Bulletin Series*, 1976(2), i–27.

<https://doi.org/10.1002/j.2333-8504.1976.tb01109.x>

- \*Liang, S. (1994). *Contributions of family size, birth order, socioeconomic status, and parent-child relationships to young children's intellectual development* [Masterarbeit, Oregon State University]. Library Oregon State.  
[https://ir.library.oregonstate.edu/concern/graduate\\_thesis\\_or\\_dissertations/4f16c537j](https://ir.library.oregonstate.edu/concern/graduate_thesis_or_dissertations/4f16c537j)
- \*Locke, N. M. & Goldstein, H. (1937). The Relation of Birth Order, Age of Mother, and Size of Family to Intelligence. *The Journal of Psychology*, 3(1), 89–96.  
<https://doi.org/10.1080/00223980.1937.9917481>
- \*Macaulay, C. E. & Ford, R. M. (2013). Family influences on the cognitive development of profoundly deaf children: Exploring the effects of socioeconomic status and siblings. *Journal of Deaf Studies and Deaf Education*, 18(4), 545–562.  
<https://doi.org/10.1093/deafed/ent019>
- \*Marjoribanks, K. (1976). Birth Order, Family Environment, and Mental Abilities: A Regression Surface Analysis. *Psychological Reports*, 39(3), 759–765.  
<https://doi.org/10.2466/pr0.1976.39.3.759>
- \*Marjoribanks, K. (1978). Birth Order, Age Spacing between Siblings, and Cognitive Performance. *Psychological Reports*, 42(1), 115–123.  
<https://doi.org/10.2466/pr0.1978.42.1.115>
- \*Mascie-Taylor, C. G. N. (1980). Family size, birth order and IQ components: a survey of a Cambridge suburb. *Journal of Biosocial Science*, 12(3), 309–312.  
<https://doi.org/10.1017/s0021932000012840>
- \*McCall, J. N. (1973). Birth-Order Differences in Special Ability: Fact or Artifact? *Psychological Reports*, 33(3), 947–952. <https://doi.org/10.2466/pr0.1973.33.3.947>
- \*McCall, J. N. & Johnson, O. G. (1972). The Independence of Intelligence from Family Size and Birth Order. *The Journal of Genetic Psychology*, 121(2), 207–213.  
<https://doi.org/10.1080/00221325.1973.10532661>
- \*McFadden, J. H. (1929). A further note on the differential IQ's of siblings. *Journal of Applied Psychology*, 13(1), 86–91. <https://doi.org/10.1037/h0071762>
- Mercy, J. A. & Steelman, L. C. (1982). Familial Influence on the Intellectual Attainment of Children. *American Sociological Review*, 47(4), 532. <https://doi.org/10.2307/2095197>
- \*Munroe, R. H. & Munroe, R. L. (1984). Birth Order and Intellectual Performance in Three East African Societies. *Journal of Social Psychology*, 123(2), 273–274.
- \*Munroe, R. L. & Munroe, R. H. (1983). Birth Order and Intellectual Performance in East Africa. *Journal of Cross-Cultural Psychology*, 14(1), 3–16.  
<https://doi.org/10.1177/0022002183014001001>

- \*Oberlander, M., Jenkin, N., Houlihan, K. & Jackson, J. (1970). Family size and birth order as determinants of scholastic aptitude and achievement in a sample of eighth graders. *Journal of Consulting and Clinical Psychology*, 34(1), 19–21.  
<https://doi.org/10.1037/h0028825>
- \*O'Shields, J. A. M. (1980). *Cognitive Effects of Birth Order and Parental Absence* [Dissertation, Loyola University of Chicago]. Layola University Chicago Dissertations. [https://ecommons.luc.edu/luc\\_diss/1923](https://ecommons.luc.edu/luc_diss/1923)
- \*Osman, H. A., Alahmadi, M., Lynn, R. & Lynn, R. (2016). Relation Between Intelligence and Family Size, Position, and Income in Adolescent Girls in Saudi Arabia. *Psychological Reports*, 119(3), 783–786. <https://doi.org/10.1177/0033294116665179>
- \*Page, E. B. & Grandon, G. M. (1979). Family Configuration and Mental Ability: Two Theories Contrasted with U.S. Data. *American Educational Research Journal*, 16(3), 257–272. <https://doi.org/10.3102/00028312016003257>
- \*Pavan, R. (2015). On the Production of Skills and the Birth-Order Effect. *Journal of Human Resources*, 51(3), 699–726. <https://doi.org/10.3368/jhr.51.3.0913-5920r>
- \*Paulhus, D. & Shaffer, D. R. (1981). Sex Differences in the Impact of Number of Older and Number of Younger Siblings on Scholastic Aptitude. *Social Psychology Quarterly*, 44(4), 363. <https://doi.org/10.2307/3033905>
- Random Effects / Sensitivity Analysis* (2023). PennState: Statistics Online Courses.  
Abgerufen am 23. Mai 2023, von <https://online.stat.psu.edu/stat509/lesson/16/16.8>
- \*Rhine, W. R. (1968). Birth Order Differences in Conformity and Level of Achievement Arousal. *Child Development*, 39(3), 987. <https://doi.org/10.2307/1127001>
- \*Roberts, C. S. (1938). Ordinal Position and its Relationship to Some Aspects of Personality. *The Pedagogical Seminary and Journal of Genetic Psychology*, 53(1), 173–213.  
<https://doi.org/10.1080/08856559.1938.10533805>
- Rodgers, J. L. (2001). What causes birth order–intelligence patterns? The admixture hypothesis, revived. *American Psychologist*, 56(6–7), 505–510.  
<https://doi.org/10.1037/0003-066x.56.6-7.505>
- Rodgers, M. A. & Pustejovsky, J. (2021). Evaluating meta-analytic methods to detect selective reporting in the presence of dependent effect sizes. *Psychological Methods*, 26(2), 141–160. <https://doi.org/10.1037/met0000300>
- \*Rohrer, J. M., Egloff, B. & Schmukle, S. C. (2015). Examining the effects of birth order on personality. *Proceedings of the National Academy of Sciences of the United States of America*, 112(46), 14224–14229. <https://doi.org/10.1073/pnas.1506451112>

- RStudio Team (2021). *RStudio: Integrated Development Environment for R*. RStudio, PBC, Boston, MA URL <http://www.rstudio.com/>
- Schneider, W. J., & McGrew, K. S. (2018). The Cattell–Horn–Carroll theory of cognitive abilities. In D. P. Flanagan & E. M. McDonough (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 73–163). The Guilford Press.
- Sear, R. (2020). Do human ‘life history strategies’ exist? *Evolution and Human Behavior*, 41(6), 513–526. <https://doi.org/10.1016/j.evolhumbehav.2020.09.004>
- \*Sherlock, L. (2013). *Early Life Influences on Cognition: Lanarkshire 1936* [Dissertation, University of Edinburgh]. Edinburgh Research Archive. <http://hdl.handle.net/1842/8674>
- \*Shu, B. C., Tien, A. Y., Lung, F. W. & Chang, Y. Y. (2000). Norms for the Wisconsin Card Sorting Test in 6- to 11-Year-Old Children in Taiwan. *The Clinical Neuropsychologist*, 14(3), 275–286. [https://doi.org/10.1076/1385-4046\(200008\)14:3;1-p;ft275](https://doi.org/10.1076/1385-4046(200008)14:3;1-p;ft275)
- Simonsohn, U., Nelson, L. D. & Simmons, J. P. (2014). p-Curve and Effect Size. *Perspectives on Psychological Science*, 9(6), 666–681. <https://doi.org/10.1177/1745691614553988>
- \*Singh, A. K. (2009). *A Comparative Study of Cognitive & Non-Cognitive Factors as a Sequel to Birth Order* [Unveröffentlichte Dissertation]. Veer Bahadur Singh Purvanchal University.
- Smart, R. G. (1963). Alcoholism, birth order, and family size. *The Journal of Abnormal and Social Psychology*, 66(1), 17–23. <https://doi.org/10.1037/h0041754>
- Steelman, L. C. (1985). A Tale of Two Variables: A Review of the Intellectual Consequences of Sibship Size and Birth Order. *Review of Educational Research*, 55(3), 353–386. <https://doi.org/10.3102/00346543055003353>
- \*Steelman, L. C. & Doby, J. T. (1983). Family Size and Birth order as Factors on the IQ Performance of Black and white Children. *Sociology of Education*, 56(2), 101. <https://doi.org/10.2307/2112658>
- Steelman, L. C., Powell, B. A., Werum, R. E. & Carter, S. L. (2002). Reconsidering the Effects of Sibling Configuration: Recent Advances and Challenges. *Annual Review of Sociology*, 28(1), 243–269. <https://doi.org/10.1146/annurev.soc.28.111301.093304>
- Sterne, J. A. C., Gavaghan, D. J. & Egger, M. (2000). Publication and related bias in meta-analysis. *Journal of Clinical Epidemiology*, 53(11), 1119–1129. [https://doi.org/10.1016/s0895-4356\(00\)00242-0](https://doi.org/10.1016/s0895-4356(00)00242-0)

Süddeutsche Zeitung. (2019, 5. Juli). Das schwere Los des Sandwichkindes? *Süddeutsche.de*.

Abgerufen am 22. Juli 2023, von <https://www.sueddeutsche.de/leben/familie-das-schwere-los-des-sandwichkindes-dpa.urn-newsml-dpa-com-20090101-190704-99-925618>

\*Sutherland, H. E. G. & Thomson, G. H. (1926). The correlation between intelligence and size of family. *British Journal of Psychology*, 17, 81–92.

\*Swanson, B. R. (1979). *The relationship of birth order, intelligence, academic achievement, and field dependence/field independence: A comparison among eighth grade science students* [Dissertation, University of Southern California]. ProQuest Dissertations Publishing.

<https://www.proquest.com/docview/302968169/BBFBCCDBEAB04369PQ/1?accountid=14682>

\*Theroux, N. L. (1993). *Birth order and its relationship to academic achievement and selected personal traits* [Dissertation, University of California Los Angeles]. ProQuest Dissertations Publishing. <https://www.proquest.com/docview/304057145?pq-origsite=gscholar&fromopenview=true>

\*Thurstone, L. L. & Jenkins, R. L. (1929). Birth order and intelligence. *Journal of Educational Psychology*, 20(9), 641–651. <https://doi.org/10.1037/h0071971>

\*Valencia, R. R., Henderson, R. W. & Rankin, R. J. (1981). Relationship of family constellation and schooling to intellectual performance of Mexican American children. *Journal of Educational Psychology*, 73(4), 524–532. <https://doi.org/10.1037/0022-0663.73.4.524>

van Aert, R. C. M. & van Assen, M. A. L. M. (2018). Correcting for Publication Bias in a Meta-Analysis with the P-uniform\* Method. *MetaArXiv Preprints*.

<https://doi.org/10.31222/osf.io/zqjr9>

van Aert, R. C. M., Wicherts, J. M. & Van Assen, M. A. (2016). Conducting Meta-Analyses Based on p Values. *Perspectives on Psychological Science*, 11(5), 713–729.

<https://doi.org/10.1177/1745691616650874>

van Assen, M. A., Van Aert, R. C. M. & Wicherts, J. M. (2015). Meta-analysis using effect size distributions of only statistically significant studies. *Psychological Methods*, 20(3), 293–309. <https://doi.org/10.1037/met0000025>

Viechtbauer, W. (2023). *Comparing Estimates of Independent Meta-Analyses or Subgroups*.

[http://www.metafor-project.org/doku.php/tips:comp\\_two\\_independent\\_estimates](http://www.metafor-project.org/doku.php/tips:comp_two_independent_estimates)



- \*Velandia, W., Grandon, G. M. & Page, E. B. (1978). Family Size, Birth Order, and Intelligence in a Large South American Sample. *American Educational Research Journal*, 15(3), 399–416. <https://doi.org/10.3102/00028312015003399>
- \*Wänström, L. & Wegmann, B. (2017). Effects of sibship size on intelligence, school performance and adult income: Some evidence from Swedish data. *Intelligence*, 62, 1–11. <https://doi.org/10.1016/j.intell.2017.01.004>
- \*Wilson, D., Mundy-Castle, A. & Panditji, L. (1990). Birth order and Intellectual Development among Zimbabwean Children. *The Journal of Social Psychology*, 130(3), 409–411. <https://doi.org/10.1080/00224545.1990.9924599>
- \*Witt, S. J. (1978). *Birth order effects and fluid/crystallized intelligence* [Dissertation, University of Florida]. University of Florida Library. <http://ufdcimages.uflib.ufl.edu/UF/00/09/74/78/00001/birthordereffect00wittrich.pdf>
- \*Wolf, K. M. (1967). *A comparison of sibling position and academic achievement in the elementary school* [Dissertation, The University of Oklahoma]. SHAREOK. <https://hdl.handle.net/11244/2207>
- World Bank Open Data. (2023). World Bank Open Data. Abgerufen am 2. April 2023, von <https://data.worldbank.org/>
- \*Zajonc, R. B. & Bargh, J. A. (1980a). Birth order, family size, and decline of SAT scores. *American Psychologist*, 35(7), 662–668. <https://doi.org/10.1037/0003-066x.35.7.662>
- \*Zajonc, R. & Bargh, J. (1980b). The confluence model: Parameter estimation for six divergent data sets on family factors and intelligence. *Intelligence*, 4(4), 349–361. [https://doi.org/10.1016/0160-2896\(80\)90028-9](https://doi.org/10.1016/0160-2896(80)90028-9)
- Zajonc, R. B. & Markus, G. B. (1975). Birth order and intellectual development. *Psychological Review*, 82(1), 74–88. <https://doi.org/10.1037/h0076229>
- Zajonc, R. B., Markus, H. R. & Markus, G. B. (1979). The birth order puzzle. *Journal of Personality and Social Psychology*, 37(8), 1325–1341. <https://doi.org/10.1037/0022-3514.37.8.1325>
- \* Inkludiert in die Metaanalyse

### Abbildungsverzeichnis

Abbildung 1: Prisma Flow Diagramm zur Darstellung inkludierter und exkludierter Studien.....	20
Abbildung 2: Verteilung der inkludierten Studien nach Publikationsjahr .....	30
Abbildung 3: Korrelation Familiengröße und Geburtenreihenfolge.....	30
Abbildung 4: Verteilung der Gesamtvarianz des Modells ohne Rauschen.....	32
Abbildung 5: Gemittelter Effekt des Modells ohne Rauschen (Rainforest Plot). Multilevel Modell .....	33
Abbildung 6: Gemittelter Effekt des Modells ohne Rauschen (Forest Plot). Aggregiertes Modell .....	34
Abbildung 7: Ausreißer des Modells ohne Rauschen (Baujat und Influence Plot).....	35
Abbildung 8: Verteilung der Gesamtvarianz des umfassenden Modells .....	36
Abbildung 9: Gemittelter Effekt des umfassenden Modells (Rainforest Plot). Multilevel Modell .....	37
Abbildung 10: Gemittelter Effekt des umfassenden Modells (Forest Plot). Aggregiertes Modell .....	38
Abbildung 11: Ausreißer des umfassenden Modells (Baujat und Influence Plot).....	39
Abbildung 12: Verteilung der Gesamtvarianz des Modells mit Kontrolle für Familiengröße	40
Abbildung 13: Gemittelter Effekt des Modells mit Kontrolle für Familiengröße (Rainforest Plot). Multilevel Modell.....	41
Abbildung 14: Gemittelter Effekt des Modells mit Kontrolle für Familiengröße (Forest Plot). Aggregiertes Modell.....	42
Abbildung 15: Ausreißer des Modells mit Kontrolle für Familiengröße (Baujat und Influence Plot) .....	43
Abbildung 16: Gemittelte Korrelationen aller drei Modelle.....	44
Abbildung 17: Metaregression für Publikationsjahr für das Modell mit Kontrolle für Familiengröße.....	45
Abbildung 18: Funnel Plot und Sunset-Funnel Plot des Modells ohne Rauschen .....	49
Abbildung 19: Funnel Plot und Sunset-Funnel Plot des umfassenden Modells .....	50
Abbildung 20: Funnel Plot und Sunset-Funnel Plot des Modells mit Kontrolle für Familiengröße.....	51
Abbildung 21: Egger MLMA aller drei Hauptmodelle.....	52
Abbildung 22: p-curve und p-uniform des Modells ohne Rauschen .....	54
Abbildung 23: p-curve und p-uniform des umfassenden Modells.....	55



Abbildung 24: p-curve und p-uniform des Modells mit Kontrolle für Familiengröße .....	56
Abbildung 25: Funnel Plot inklusive mit Trim-and-Fill hinzugefügte Studien und Eggers Regression .....	57
Abbildung 26: Leave-One-Out Sensitivitätsanalysen aller drei Modelle .....	58

**Tabellenverzeichnis**

Tabelle 1: Deskriptive Verteilung der inkludierten Studien .....	28
Tabelle 2: Deskriptive Beschreibung der Stichproben.....	29
Tabelle 3: Ergebnisse der multilevel Metaanalysen.....	31
Tabelle 4: Ergebnisse der Moderatoranalysen .....	46
Tabelle 5: Ergebnisse der Tests auf Publication Bias und Sensitivitätsanalysen.....	47

## Anhang

### Zusammenfassung

Der *Birth order* Effekt beschreibt eine höhere kognitive Leistung von früher geborenen Personen im Vergleich zu den später geborenen Geschwisterkindern. Die Studienlage dazu ergab bisher inkonsistente und auch konträre Befunde. Diese Arbeit ist die erste bekannte multilevel Metaanalyse im Bereich der Geschwisterreihenfolge und Intelligenz. Sie umfasst fast vier Millionen Teilnehmer\*innen aus 70 Studien und 144 gemeldeten Effektstärken aus verschiedenen Ländern, Altersgruppen und sozialen Hintergründen.

Die gemittelte Effektstärke mit Einbezug von Korrelationen, die mithilfe Mittelwertsunterschiede berechnet wurden, ergab einen kleinen negativen Effekt von  $r = -.141$  und somit Bestätigung des *Birth order* Effekts. Wird jedoch für die Anzahl der Kinder innerhalb einer Familie kontrolliert, so sank der gemittelte Effekt auf  $r = .004$ . Dies weist darauf hin, dass es keinen signifikanten Zusammenhang zwischen Geschwisterreihenfolge und Intelligenz bei Kontrolle der Familiengröße gibt. Durch vorherige Befunde kann jedoch festgestellt werden, dass der Zusammenhang in eine negative Richtung geht.

Außerdem konnten keine relevanten Moderatorvariablen, unter anderem das Alter, Geschlecht, Herkunft oder sozioökonomische Status der Stichprobe oder der Publikationsstatus identifiziert werden. Die verschiedenen Methoden für den Test auf *Publication Bias* wie die Egger MLMA, die Rangkorrelation, *p-curve*, *p-uniform* oder *p-uniform \** ergaben widersprüchliche Ergebnisse.

Hier konnte nicht geprüft werden, ob kleine Familien im Allgemeinen intelligentere Kinder haben. Dies sollte bei zukünftiger Forschung erarbeitet werden. Diese Arbeit zeigt jedoch, dass die hohe Korrelation von Geschwisterreihenfolge und Familiengröße Auswirkungen auf die Befunde im Sinne des *Birth Order* Effekts hat.

*Schlagwörter:* multilevel Metaanalyse, Geschwisterreihenfolge, Intelligenz, *Birth order* Effekt

**Abstract**

The birth order effect defines a higher cognitive performance of earlier-born people compared to their later-born siblings. Studies so far produced inconsistent and also contradictory findings. This paper is the first known multilevel meta-analysis in the area of sibling order and intelligence. It includes almost four million participants from 70 studies and 144 reported effect sizes from different countries, age groups and social backgrounds.

The summary effect including correlations calculated with mean differences showed a small negative effect of  $r = -.141$ , confirming the birth order effect. However, when controlling for the number of children within a family, the mean effect declined to  $r = .004$ . This indicates that there is no significant relationship between birth order and intelligence when controlling for family size and the negative birth order effect disappears. However, previous findings show that the correlation goes in a negative direction.

Furthermore, no relevant moderator variables could be identified, including the publication status and the age, gender, origin or socioeconomic status of the sample. The various methods for testing for publication bias, such as the Egger MLMA, the rank correlation, p-curve, p-uniform or p-uniform \* yielded contradictory results.

It could not be tested whether children from small families in general are more intelligent. This needs to be addressed in future research. This work shows that the high correlation between birth order and family size affect findings regarding the birth order effect.

*Keywords:* multilevel meta-analysis, birth order, intelligence, birth order effect

**Veränderungen zur Präregistrierung**

Im Vergleich zu der Präregistrierung wurde das Vorgehen im Laufe der Datensammlung und Datenanalyse entsprechend angepasst und verändert.

Die multilevel Metaanalyse wurde als Lösung für abhängige Daten innerhalb einer Studie aufgrund der Verwendung mehrerer Tests und Stichproben verwendet. Da jedoch viele Methoden noch nicht für multilevel Analysen verfügbar waren, wurden aggregierte Modelle gebildet. Dadurch wurde beispielsweise die Interpretation der *Publication Bias* Methoden auf die aggregierten Modelle eingeschränkt.

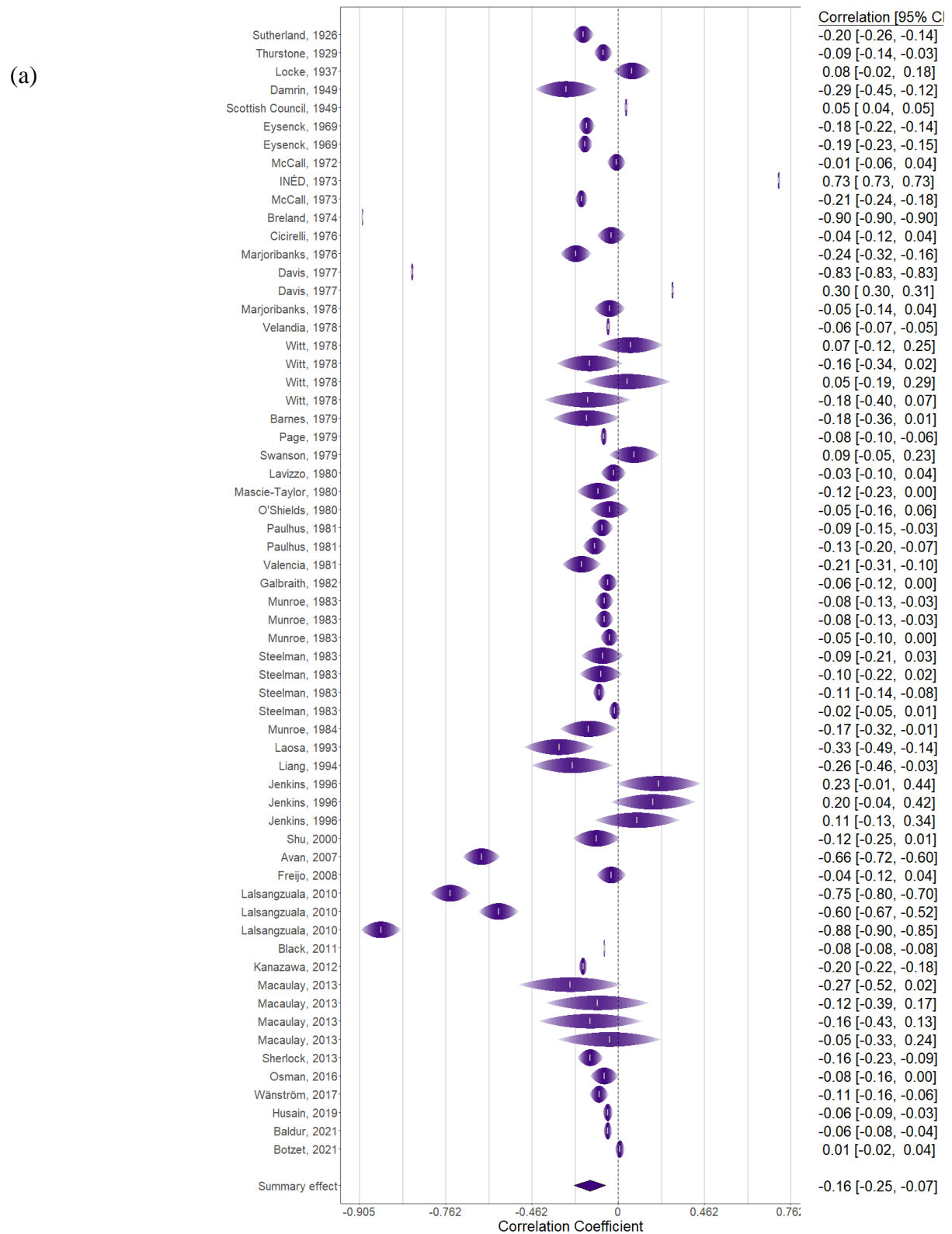
Des Weiteren wurden die Werte für GDP im jeweiligen Land und Jahr der Datenerhebung als weitere Moderatorvariable und Hinweis auf den sozioökonomischen Status nach der Präregistrierung hinzugefügt.

Die Recherchephase wurde aufgrund der groß angelegten Datenerhebung verlängert bis Ende März 2023.

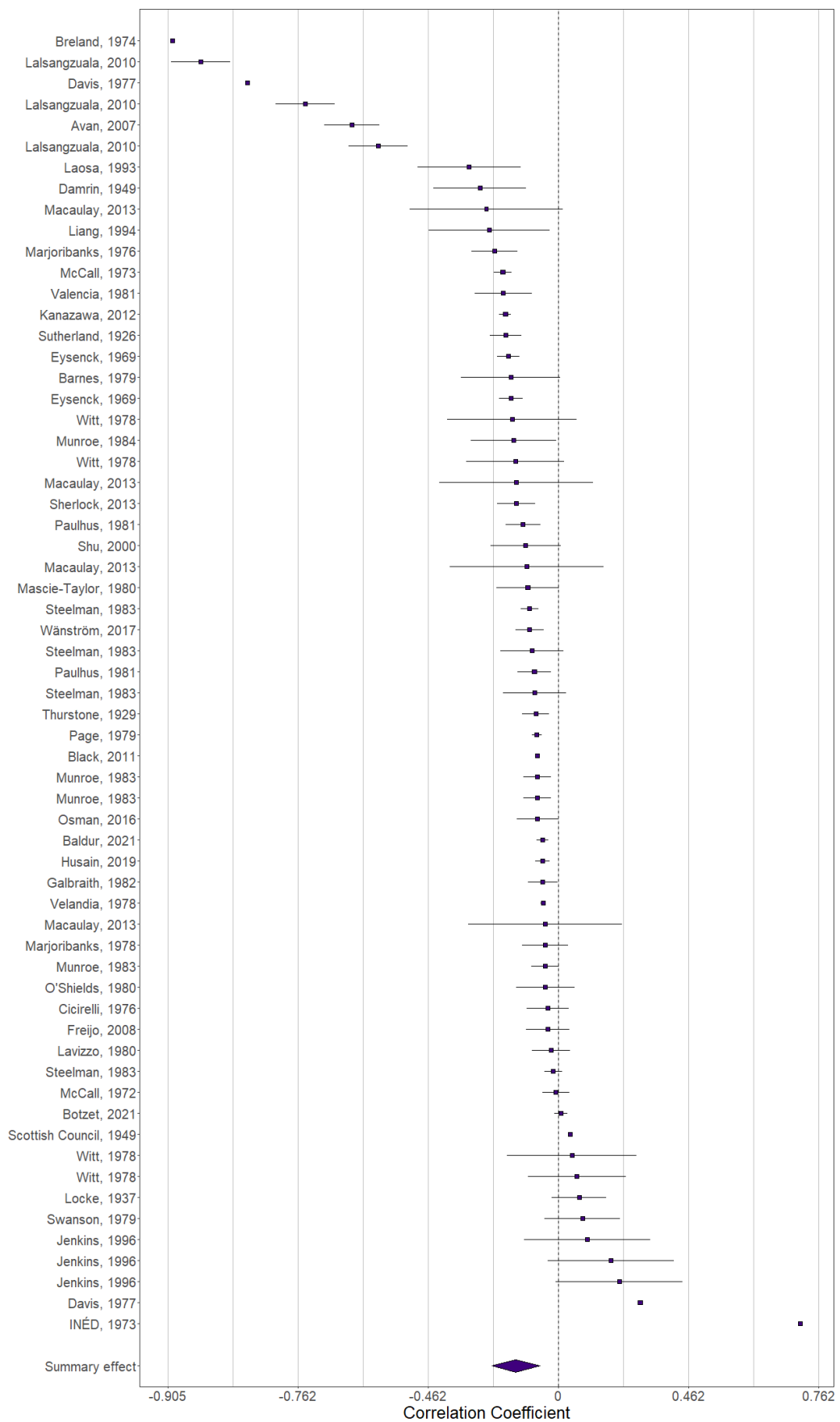
## Weitere graphische Darstellungen

### Modell ohne Rauschen

*Nach Publikationsjahr geordneter Rainforest Plot (a) und Caterpillar Plot geordnet nach Größe der Korrelation (b)*

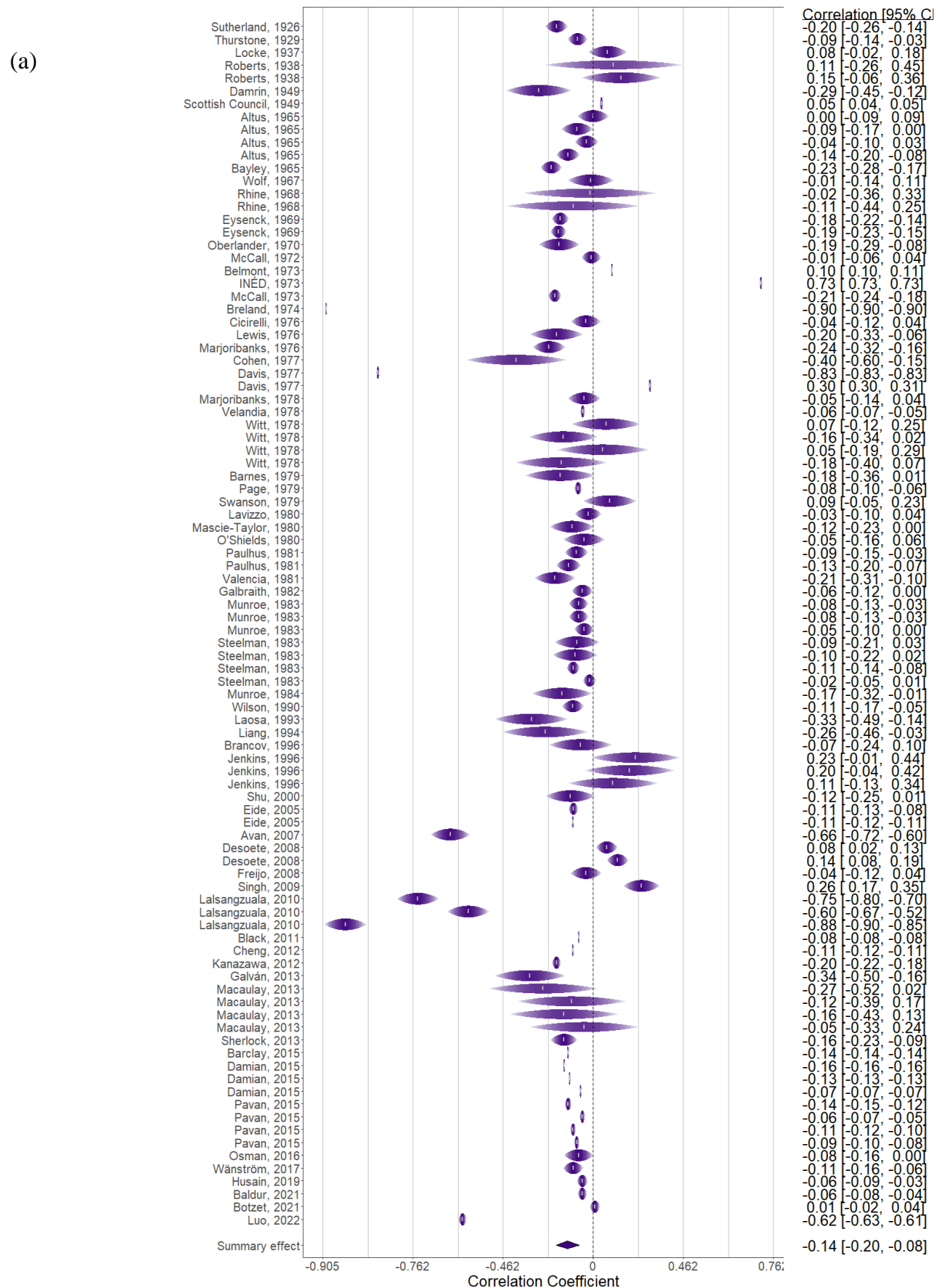


(b)



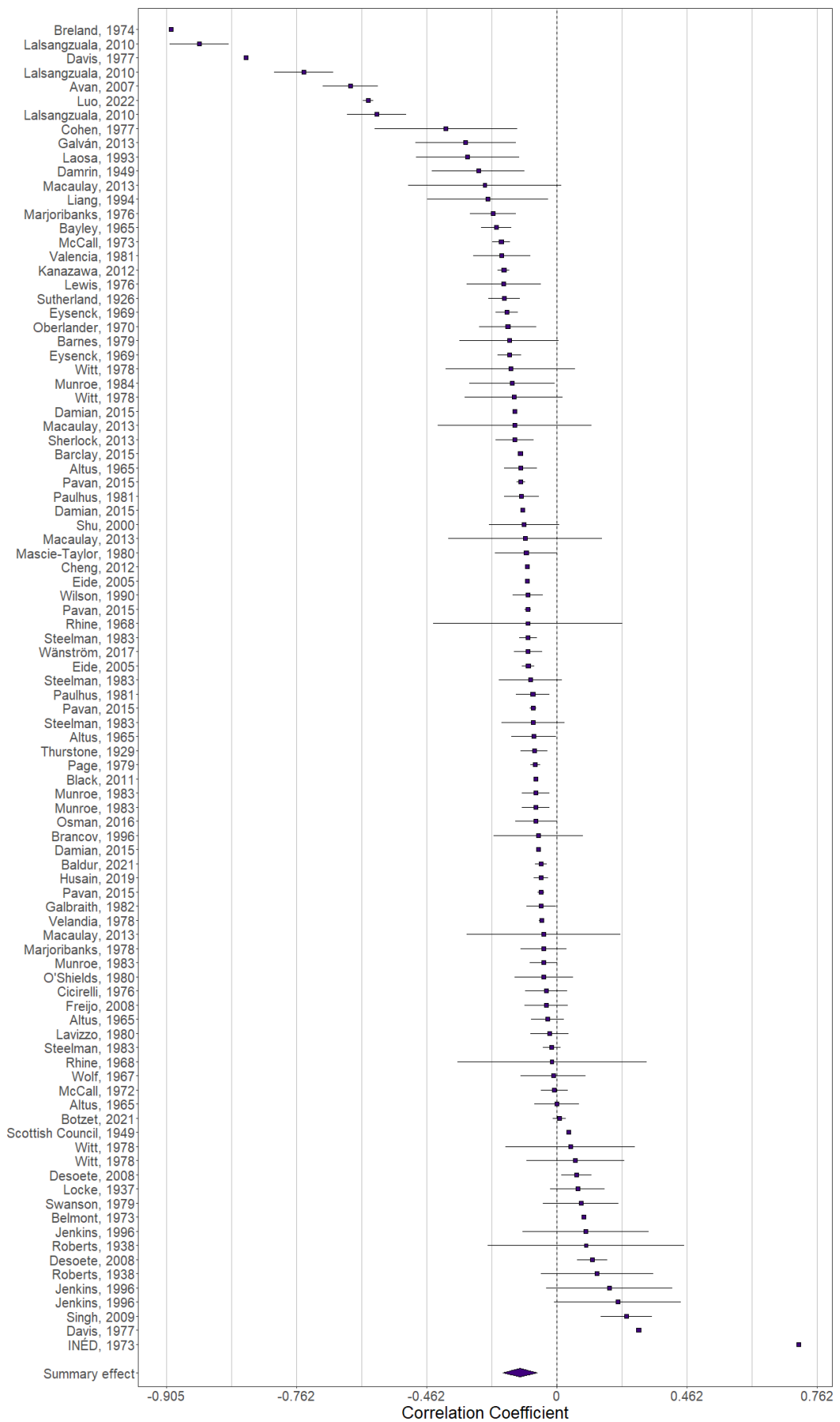
## Umfassendes Modell

*Nach Publikationsjahr geordneter Rainforest Plot (a) und Caterpillar Plot geordnet nach Größe der Korrelation (b)*



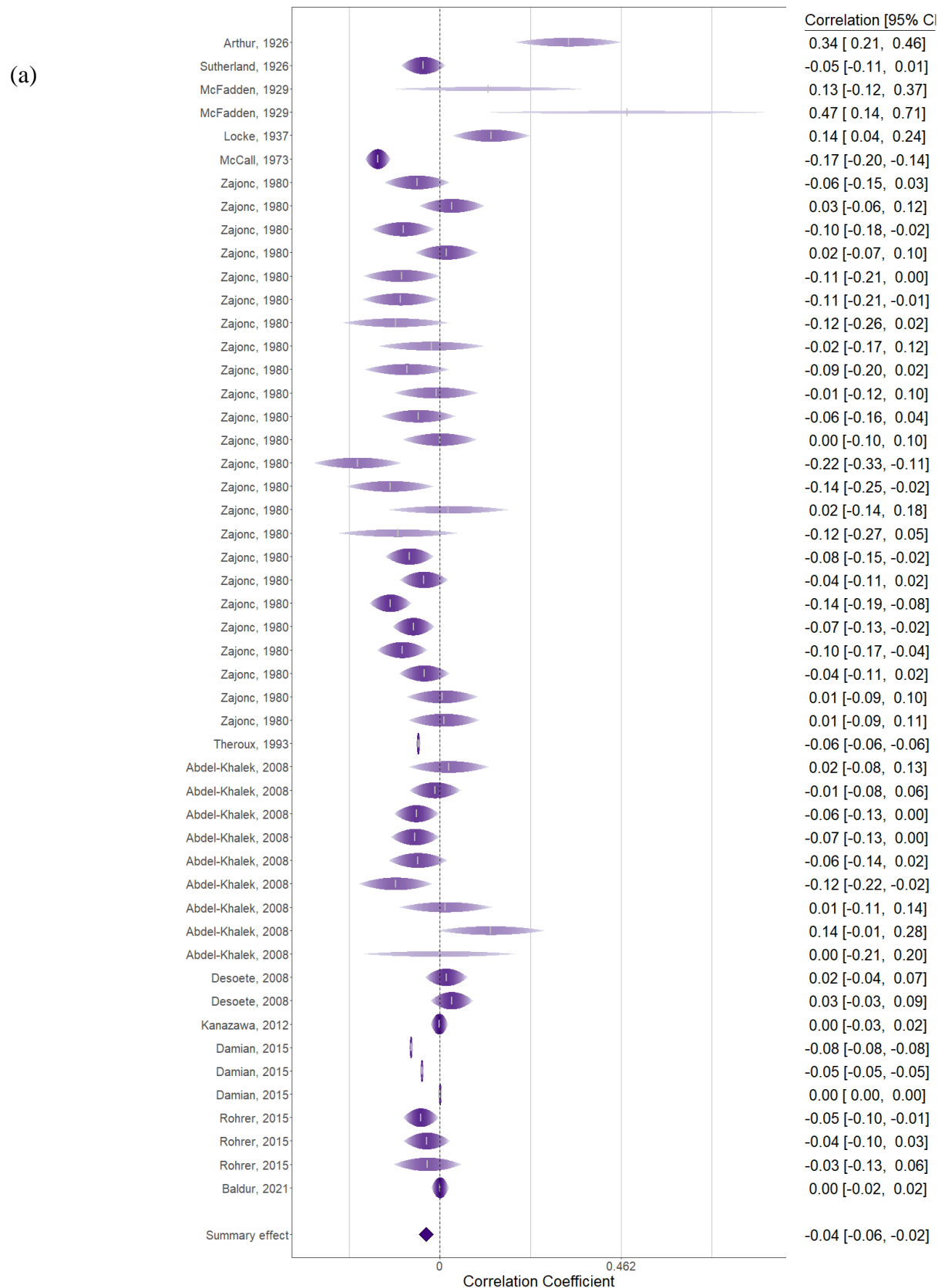


(b)



## Modell mit Kontrolle für Familiengröße

*Nach Publikationsjahr geordneter Rainforest Plot (a) und Caterpillar Plot geordnet nach Größe der Korrelation (b)*



(b)

