

ZWANZIGEINS VS. EINUNDZWANZIG:
PROFITIEREN RECHENSCHWACHE KINDER IN BEZUG
AUF DIE FEHLERZAHL VON EINER
STELLENWERTGERECHTEN ZAHLENSPRECHWEISE?

EINE EMPIRISCHE UNTERSUCHUNG

Vivien Hartwig

M.A. Integrative Lerntherapie

8. Semester

1. Prüfer: Prof. Dr. med. Dipl.-Päd. Michael G. von Aster

2. Prüfer: PD Dr. Peter Morfeld

1 INHALT

ABBILDUNGSVERZEICHNIS.....	4
-----------------------------------	----------

I THEORETISCHER TEIL

1 EINLEITUNG	6
1.1 Themenaufritt	6
1.2 Begriffe und Abkürzungsverzeichnis	9
2 LERN- UND RECHENSTÖRUNGEN	12
2.1 Lern- und Rechenstörungen – Begriff und Klassifikation.....	12
2.2 Rechenstörungen: Diagnostische Kriterien und Instrumente.....	14
2.2.1 Diagnostik von Rechenstörungen im deutschsprachigen Raum	16
2.3 Teufelskreis Lernstörung	18
2.3.1 Die Folgen von Lernstörungen	20
2.4 Prävalenz.....	21
2.4.1 Geschlechterverteilung.....	22
2.5 Komorbiditäten	23
2.5.1 Angst und ihre Auswirkungen.....	24
2.5.2 Entstehungskomponenten von Rechenangst.....	26
2.6 Ursachen von Lernstörungen	27
2.6.1 Unzureichende pädagogisch-didaktische Maßnahmen als Risikofaktor	27
2.6.2 Genetische Determinanten von Lernstörungen.....	29
2.7 Neuropsychologie normaler und abweichender Entwicklung der Zahlenverarbeitung....	30
2.7.1 Das <i>Triple Code Model</i>	31
2.7.2 Das Vier-Stufen-Modell der Entwicklung zahlenverarbeitender Hirnfunktionen.....	32
2.7.3 Abweichende Entwicklung zahlenverarbeitender Hirnfunktionen.....	34
2.7.4 Störung früher domänenübergreifenden Reifungsprozesse und Risiken auf Grund eingeschränkter kognitiver Funktionen.....	35
2.7.5 Arbeitsgedächtniskapazität und mathematische Leistung.....	36
3 ASPEKTE DER ZAHLENLINGUISTIK	38
3.1 Die Entwicklung und Verbreitung des dezimalen Stellenwertsystems	38
3.2 Verschiedene Sprachräume und ihr Umgang mit Zahlwortsystemen.....	39
3.3 Auswirkungen der invertierten Zahlensprechweise	41

4	ERHEBUNGSINSTRUMENTE UND STUDIENDESIGN	47
4.1	Eingesetzte Hardware & Berücksichtigung der unterschiedlichen Ansagedauer	47
4.2	Eingesetzte Software und App-Parameterfestlegungen	48
4.3	Grundstruktur und Studiendesign.....	50
4.3.1	Lerneffekte als Störgrößen	51
4.3.2	Der Zusammenhang von Eingabedauer und Fehlerzahl	53
4.4	Erläuterungen zur Durchführung	56
4.4.1	Responsvariablen	56
4.4.2	Auswahl von Kovariablen	57
4.5	Forschungsfrage und Hypothesen	59
4.6	Rekrutierung und Beschreibung der Untersuchungsgruppe.....	60
4.6.1	Rekrutierungsprozess für das Studienkollektiv 2.....	60
4.6.2	Beschreibung der Untersuchungsgruppe	61
4.7	Die Durchführung der Erhebung	64
4.8	Methodik der Auswertung	64
4.8.1	Hauptanalyse	65
4.8.2	Zusätzliche Analyse aller Daten	65
4.8.3	Bezeichnungen der Dateien und Graphen	65
5	DARSTELLUNG UND ANALYSE DER ERGEBNISSE ZU DEN FEHLERZAHLEN ..	66
5.1	Analysen der Mittleren Fehlerzahl (Responsvariable)	67
5.1.1	Deskriptive Analyse der Ergebnisse zu den Fehlerzahlen	67
5.1.2	Zufallskritische Analyse der Ergebnisse zur mittleren Fehlerzahl Y2 (Testdaten)	69
5.1.3	Zusammenfassung der Ergebnisse zur mittleren Fehlerzahl Y2.....	70
5.2	Analysen der fehlerfreien Durchgänge (Responsvariable)	71
5.2.1	Deskriptive Analyse der Ergebnisse zu den fehlerfreien Durchgängen	71
5.2.2	Zufallskritische Analyse und Zusammenfassung der Ergebnisse zu den fehlerfreien Durchgängen	72
5.3	Mittlere Fehlerzahl: Effekt des Alters.....	73
5.3.1	Deskriptive Analyse der Ergebnisse zum Effekt des Alters.....	73
5.3.2	Zufallskritische Analyse der Ergebnisse zur Kovariable Alter	76
5.3.3	Interpretation der Regressionsanalyse (Kovariable Alter).....	79
5.3.4	Ergebnisse der Auswertung zur Kovariable Klassenstufe	80
5.4	Weitere Befunde unter Berücksichtigung der Kovariablen.....	82
5.4.1	Ergebnisse der Auswertung zur Kovariable Fördereinrichtung	83
5.4.2	Ergebnisse der Auswertung zur Kovariable Geschlecht	85
5.4.3	Ergebnisse der Auswertung zur Kovariable Erstsprache.....	87
5.4.4	Ergebnisse der Auswertung zur Kovariable LRS.....	89
5.4.5	Ergebnisse der Auswertung zur Kovariable Sequenz.....	91
5.4.6	Ergebnisse der Auswertung zur Kovariable Schulwechsel.....	92

5.4.7	Interpretation der Befunde nach Auswertung der Kovariablen.....	93
5.5	Basisbefund bei Auswertung aller Perioden.....	94
5.5.1	Auswertung aller Perioden: Interpretation	95
6	REFLEXIONEN DER KINDER.....	96
7	ZUSAMMENFASSUNG UND DISKUSSION.....	104
8	LITERATURVERZEICHNIS	110
9	PROGRAMM & APP	120
10	ANHÄNGE	121
Anhang 1	Gesprächsleitfaden Einweisung.....	121
Anhang 2	<i>Einverständnis zur Teilnahme an der Studie.....</i>	123
Anhang 3	Umsteigeliste und randomisierte Sequenzen (Vorlage).....	124

Abbildungsverzeichnis

ABBILDUNG 1: 4-STUFEN-MODEL NUMERISCHER KOGNITION NACH SHALEV & VON ASTER.....	33
ABBILDUNG 2: Y1_PERIOD_ALL.TIF: DAUER / AUSWERTUNG ALLER PERIODEN	52
ABBILDUNG 3: Y1_PERIOD_TEST.TIF: DAUER / AUSWERTUNG TESTPERIODEN	53
ABBILDUNG 4: Y1_Y2_ALL.TIF: ZUSAMMENHANG Y1 UND Y2 / AUSWERTUNG ALLER PERIODEN.....	54
ABBILDUNG 5: Y1_Y2_AK_ALL.TIF: ZUSAMMENHANG Y1 UND Y2 / AUSWERTUNG DER STUDIENGRUPPE AK.....	55
ABBILDUNG 6: Y1_Y2_VH_ALL.TIF: ZUSAMMENHANG Y1 UND Y2 / AUSWERTUNG DER STUDIENGRUPPE VH.....	56
ABBILDUNG 7: BAR_Y2_TEST.TIF: AUSWERTUNG DER MITTLEREN FEHLERZAHL SPRECHWEISE: TV. VS. ZE.....	67
ABBILDUNG 8: BOX_DIFF_Y2_TEST.TIFF: VERTEILUNG DER FEHLERZAHL-DIFFERENZEN	68
ABBILDUNG 9: BAR_Y2NEGP_TEST.TIF: ANTEILE DER FEHLERFREIEN TESTUNGEN	71
ABBILDUNG 10: BAR_Y2_ALT_KAT_TEST.TIF: MITTLERE FEHLERZAHL / AUSWERTUNG DER KOVARIABLE ALTER.....	73
ABBILDUNG 11: BOX_DIFF_Y2_KAT_TEST.TIF: FEHLERDIFFERENZ / AUSWERTUNG DER KOVARIABLE ALTER	74
ABBILDUNG 12: BAR_Y2NEGP_ALT_KAT_TEST.TIF: FEHLERFREIE DURCHGÄNGE / AUSWERTUNG DER KOVARIABLE ALTER	75
ABBILDUNG 13: BAR_Y2_KLASSEN_TEST: MITTLERE FEHLERZAHL / AUSWERTUNG KOVARIABLE KLASSENSTUFEN	80
ABBILDUNG 14: BAR_Y2NEG_KLASSEN_TEST: FEHLERFREIE TESTS / AUSWERTUNG KOVARIABLE KLASSENSTUFEN	81
ABBILDUNG 15: BAR_Y2_FOERDER-EINRICHTUNG_TEST: MITTLERE FEHLERZAHL / AUSWERTUNG DER KOVARIABLE FÖRDEREINRICHTUNG	83
ABBILDUNG 16: BAR_Y2NEGP_FOERDER-EINRICHTUNG_TEST: FEHLERFREIE TESTS / AUSWERTUNG DER KOVARIABLE FÖRDEREINRICHTUNG	84
ABBILDUNG 17: BAR_Y2_GESCHL_TEST: MITTLERE FEHLERZAHL / AUSWERTUNG DER KOVARIABLE GESCHLECHT.....	85
ABBILDUNG 18: BAR_Y2NEG_GESCHL_TEST: FEHLERFREIE DURCHGÄNGE / AUSWERTUNG DER KOVARIABLE GESCHLECHT.....	86
ABBILDUNG 19: BAR_Y2_ERST_21_TEST: MITTLERE FEHLERZAHL / AUSWERTUNG DER KOVARIABLE ERSTSPRACHE ...	87
ABBILDUNG 20: BAR_Y2NEG_ERST_21_TEST: FEHLERFREIE DURCHGÄNGE / AUSWERTUNG DER KOVARIABLE ERSTSPRACHE	88
ABBILDUNG 21: BAR_Y2_LRS_TEST: MITTLERE FEHLERZAHL / AUSWERTUNG DER KOVARIABLE LRS	90
ABBILDUNG 22: BAR_Y2NEGP_LRS_TEST: FEHLERFREIE TESTS / AUSWERTUNG DER KOVARIABLE LRS	91
ABBILDUNG 23: BAR_Y2_SEQU_TEST: FEHLERFREIE TESTS / AUSWERTUNG DER KOVARIABLE SEQUENZ.....	92
ABBILDUNG 23: BAR_Y2_SCHULWECHSEL_TEST: MITTLERE FEHLERZAHL / AUSWERTUNG DER KOVARIABLE SCHULWECHSEL.....	93
ABBILDUNG 25: BAR_Y2_ALL.TIF: MITTLERE FEHLERZAHL / AUSWERTUNG ALLER PERIODEN	95
ABBILDUNG 26: BAR_Y2NEGP_ALL.TIF: FEHLERFREIE TESTS / AUSWERTUNG ALLER PERIODEN	96

I THEORETISCHER TEIL

1 EINLEITUNG

1.1 THEMENAUFRISS

Arithmetische Fertigkeiten gehören zu den wesentlichen Kulturtechniken der Menschheit, Sie fördern die individuelle Entwicklung in besonderer Weise und ebnen als Werkzeug des Alltags den Weg für persönliche und berufliche Möglichkeiten (Eichler 2006, S. 1). Mit dem Eintritt in die Grundschule werden die bis dahin überwiegend spielerischen Alltagserfahrungen im Bereich Mathematik aufgegriffen und die mathematischen Kompetenzen systematisch weiterentwickelt (Benz et al. 2017, S. 1). Zum Arithmetiklernen der Grundschule gehören nicht nur das Erlernen von Rechenoperationen, geometrischem Grundwissen und das Erschließen und Bearbeiten von Sachaufgaben, sondern gleichermaßen die Entwicklung von „Selbstvertrauen in die eigenen mathematischen Kompetenzen“ sowie die Bewahrung von „Interesse und Neugier gegenüber mathemathikhaltigen Phänomenen“ (Schmidt 2009, S. 77). Deutschsprechende Kinder stehen allerdings bei der Aneignung des Zahlbegriffs und dem damit verbundenen Aufbau von Zahlvorstellungen vor einer besonderen Herausforderung: Die Sprechweise der Zahlen, die sie bis zum Vorschulalter gelernt haben, stimmt nicht mit der indo-arabischen Notation, die sie nun mit Schuleintritt lernen, überein. Die gebräuchliche Verdrehung der Sprechweise besteht darin, dass die Zehnerstelle zwar vor der Einerstelle geschrieben, aber nach der Einerstelle gesprochen wird: Für 21 sagen wir „einundzwanzig“ statt „zwanzigeins.“

Das internationale Klassifikationssystem von Krankheiten ICD-10 der Weltgesundheitsorganisation, Kapitel V (dt. Dilling et al. 2013), differenziert Rechenleistungen in Zahlensinn, Merkfähigkeit für arithmetisches Faktenwissen, exaktes und flüssiges Rechnen sowie rechnerisches Problemlösen (vgl. von Aster et al. 2021, S. 2). Demnach muss das Ziel sowohl des mathematischen Unterrichts in der Primarstufe als auch der therapeutischen Intervention in Mathematik der sichere Umgang mit Zahlen als Grundlage für das Mengen- und Zahlenverständnis und den Erwerb arithmetischer Kompetenzen sein. Voraussetzung hierfür ist die sichere Beherrschung der symbolischen Repräsentationen von Zahlen und die Entwicklung von Zahlraumvorstellungen, deren Entwicklung ein komplexer neuropsychologischer Prozess ist. Bei Kindern mit Rechenstörungen verläuft der

Entwicklungsprozess zahlenverarbeitender Hirnfunktionen atypisch. Betroffene Kinder und Jugendliche entwickeln sekundär vielfältige psychische Symptome und Erkrankungen und stehen daher in ihrem Alltag oft unter einem hohen Leidensdruck. Das chronische Misserfolgserleben kann z. B. zur Entwicklung von Ängsten und depressiven Symptomen führen, die „langfristig oft zu einem größeren Problem werden als das eigentliche schulische Handicap“ (von Aster et al. 2021, S. 14). Rechenschwache Kinder leiden zudem häufig unter Störungen der Aufmerksamkeits-, der Emotions- und der Verhaltensregulation sowie unter einer eingeschränkten Arbeitsgedächtniskapazität.

Die das Arbeitsgedächtnis zusätzlich belastende inverse Zahlensprechweise im Deutschen bedeutet für viele Kinder – nicht nur für jene mit persistierenden Rechenschwierigkeiten – eine zusätzliche Hürde, die aus Sicht der Initiative Zwanzigeins e. V. beseitigt werden könnte: Parallel zur inversen Sprechweise der deutschen Zahlen setzt sich Zwanzigeins e. V. für eine stellenwertgerechte Sprechweise ein, die nicht nur im schulischen und lerntherapeutischen Kontext sinnvoll erscheint, sondern eine logische Konsequenz der Forschung der vergangenen Jahre und nicht zuletzt der Zwanzigeins-App-Studien wäre, wie das Positionspapier des Vereins (Zwanzigeins e. V. 2024a) ausführt. Schmid (2023) zeigte in ihrer auf der Vereinswebsite frei zugänglichen Masterarbeit in einer ersten empirischen Untersuchung mit Hilfe der Zwanzigeins-App (Kap. 4.2), dass Kinder aus 2. Klassen statistisch gesichert von einer stellenwertgerechten Zahlensprechweise profitieren. Sowohl die Eingabedauer als auch die Anzahl der Fehler verringerten sich im Vergleich zu den Ergebnissen bei Verwendung der inversen Sprechweise signifikant, d. h. die Transkodierfähigkeit verbesserte sich bei Verwendung der stellenwertgerechten Sprechweise. Diese steht im engen Zusammenhang zur mathematischen Leistung (Kap. 3.3).

Ziel der vorliegenden Arbeit ist es, die komplexen Schwierigkeiten, mit denen Kinder mit Lern- und Rechenstörungen sowie ihre Eltern und Lehrkräfte konfrontiert sind, darzustellen und sie mit den spezifischen Problemen des Erlernens der inversen Zahlensprechweise und deren Transkodierung in die indo-arabische Notation in Beziehung zu setzen. Es soll der Frage nachgegangen werden, ob sich durch eine stellenwertgerechte Zahlensprechweise ein Vorteil für rechenschwache Kinder

beim Transkodieren im Vergleich zur derzeit im deutschsprachigen Raum gebräuchlichen inversen Sprechweise ergibt.

Die Hypothesenbildung erfolgte auf Basis der Forschungstätigkeit von Mitgliedern und Unterstützern des Vereins Zwanzigeins e. V. einerseits und der Verdichtung der Zwanzigeins-Hypothese durch neuropsychologische Erkenntnisse andererseits. Diese werden mit Blick auf unsere Untersuchungsgruppe rechenschwacher Kinder im ersten, theoretischen Teil der Arbeit thematisiert, wobei zunächst die Ursachen von Lern- und Rechenstörungen, ihre Diagnostik, komorbide Störungen, Auftretenshäufigkeit und die Folgen für Betroffene und die Gesellschaft erläutert werden. Anschließend werden die indo-arabische Notation von Zahlen und das dezimale Stellenwertsystem aus historischer Sicht sowie zahlenlinguistische Problemstellungen näher betrachtet, um der Frage nachzugehen, warum die inverse Sprechweise im deutschen Sprachraum gebräuchlich ist und welche Lernhürden dies für rechenschwache Kinder impliziert.

Im zweiten empirischen Teil wurde mit Hilfe der Zwanzigeins-App untersucht, ob rechenschwache Kindern von einer stellenwertgerechten Zahlensprechweise profitieren. Die Annahme konnte durch unsere gemeinsam mit Anna Kuhl durchgeführte Untersuchung mit statistisch signifikanten und hochsignifikanten Ergebnissen bestätigt werden. Die statistische Auswertung, die durch PD Dr. Morfeld erfolgte, korrespondiert mit den Reflexionen vieler Kinder, deren Feedbacks dokumentiert wurden und in die abschließende Diskussion einfließen, wobei Positionen, die gegen die Einführung einer stellenwertgerechte Zahlensprechweise argumentieren, aufgegriffen werden.

Im Interesse einer erleichterten Lesbarkeit werden Abkürzungen verwendet, die in der nachfolgenden Aufstellung zusammen mit grundlegenden Begriffen kurz erläutert werden.

1.2 BEGRIFFE UND ABKÜRZUNGSVERZEICHNIS

AWMF: Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften. Sie ist Herausgeberin der S3-Leitlinie für die Diagnostik und Förderung von Rechenstörungen unter der Federführung der >> DGKJP.

Dezimales Stellenwertsystem: Detaillierte Darstellung im Kapitel 3.

DGKJP: Deutsche Gesellschaft für Kinder- und Jugendpsychiatrie, Psychosomatik und Psychotherapie e. V.; federführende Fachgesellschaft der S3-Leitlinie Rechenstörung.

DSM-V (abweichende Schreibweise DSM-5) ist die Abkürzung für die 5. Auflage des von der Amerikanischen Psychiatrischen Gesellschaft herausgegebenen Klassifikationssystems *Diagnostic and Statistical Manual of Mental Disorders*. Deutsch: Falkai et al. 2015 (s. Literatur).

Intelligenz-Diskrepanz-Kriterium: Das so genannte (Intelligenz-) Diskrepanzkriterium kommt in der Diagnostik von Lernstörungen zum Tragen. Es beschreibt das Phänomen, dass schulische Fertigkeiten (Lesen, Rechtschreiben, Mathematik) sich nicht normal entwickeln, sondern deutlich unter dem Niveau liegen, das aufgrund des Alters, der allgemeinen Intelligenz und der Beschulung zu erwarten wäre. Die Anwendung des Diskrepanzkriteriums führt zu einer Differenzierung zwischen lernschwachen Kindern mit höherer vs. niedrigerer Intelligenz. Die Rechtfertigung für diese Differenzierung ist die zugrunde liegende theoretische Annahme, dass den Lernstörungen, die eine Diskrepanz zur Intelligenz aufweisen, ein basales primäres Defizit zugrunde liege, das unabhängig von der Intelligenz sei. Diese Annahme ist Gegenstand einer anhaltenden Kontroverse, wobei das Diskrepanzkriterium zunehmend infrage gestellt wird (Mähler 2017). Siehe auch Kap. 2.2 der vorliegenden Arbeit.

ICD: *International Statistical Classification of Diseases and Related Health Problems*, herausgegeben von der Weltgesundheitsorganisation. Die Ziffer 10 (ICD-10) bezeichnet die 10. Revision der Klassifikation, die Ziffer 11 (ICD-11) die 11. Revision, die seit dem 01.01.2022 in Kraft ist. Die Einführung auf internationaler Ebene ist auf Grund von Übersetzungs- und lizenzrechtlichen Prozessen ein

komplexer und langwieriger Prozess. In Deutschland erfolgt die Diagnostik psychischer Erkrankungen daher i. d. R. noch nach der ICD-10. Deutsch: Dilling et al. 2013 (s. Literaturverzeichnis).

Rechenschwäche (RS), Rechenstörung, Dyskalkulie: Die Begriffe werden in der Literatur nicht einheitlich verwendet. Hier werden die Begriffe Rechenstörung und Dyskalkulie synonym gebraucht; Rechenschwäche bezeichnet hingegen eine schwächere Ausprägung (Kap. 2.1). Wir möchten darauf hinweisen, dass das abgeleitete Adjektiv „rechenschwach“ auch für solche Kinder mit einer stärkeren Ausprägung der Lernstörung verwendet wird. Alle Kinder, die von schulischer oder fachärztlicher Seite festgestellte Schwierigkeiten haben, dem Regelunterricht im Fach Mathematik zu folgen, werden in der vorliegenden Arbeit als *rechenschwache Kinder* bezeichnet. Die englischsprachige Literatur verwendet neben *dyscalculia* auch *Developmental Dyscalculia*, kurz *DD*.

Repräsentationen von Zahlen: Im neurowissenschaftlichen Sinn handelt es sich bei einer Repräsentation um ein Set von internalen Einheiten oder Ereignissen, die für „etwas anderes“ stehen (Greco 1995). Die verschiedenen Zahlenrepräsentationen beziehen sich auf die Wortform (verbale Repräsentation), die indo-arabische Notation (symbolische Repräsentation) und die numerische Repräsentation, die die Abbildung von Zahlen und / oder Anzahligkeit in der inneren Realität umschreibt (von Aster et al. 2013c, S. 39; Lorenz 2017).

Transkodieren: Hier: das Übertragen von Zahlen aus der verbalen (Wort)form in die indo-arabische Notation. Beispiel: „fünf“ \gg 5.

Tv: Traditionell verdreht, kurz Tv, steht hier für die aktuell im deutschsprachigen Raum gebräuchliche verdrehte (inverse), nicht stellenwertgerechte Sprechweise mehrstelliger Zahlen.

Zahlbegriff / Aspekte des Zahlbegriffs: Es gibt zwei unterschiedliche Wege der mathematischen Präzisierung dafür, was unter den Zahlen 1, 2, 3 ... verstanden werden kann. Eine mögliche Präzisierung ist die mengentheoretische Beschreibung (Kardinalzahlaspekt): *Menge* und *Gleichmächtigkeit* sind hier grundlegende Begriffe. Eine andere mögliche Beschreibung bezieht sich auf den Ordinalzahlaspekt und stellt die Reihenfolge der Zahlen und die Nachfolger-

Beziehung in den Fokus. Weitere Aspekte des Zahlbegriffs umfassen Maßzahlaspekt, Operatoraspekt, Rechenzahlaspekt sowie Codierungsaspekt. Andere Publikationen nennen zusätzlich den geometrischen, narrativen, relationalen sowie den Systemzahlaspekt (vgl. Padberg & Benz 2021, S. 8). Dem Verständnis der Zahl als Kardinalzahl, Ordinalzahl und Relationalzahl (Beziehung zwischen zwei Zahlen) kommt in der Therapie von Rechenschwäche z. B. mit dem computerbasierten Lernprogramm *Calcularis* eine besondere Bedeutung zu (Käser & von Aster 2013, S. 266).

Zahlwortbildung im Deutschen: Anders als in vielen anderen Sprachen wie beispielsweise dem Englischen, Russischen, Türkischen oder Chinesischen, in denen die Zahl 123 in der in unserem Sprachraum üblichen Leserichtung von links nach rechts *ehnhundert-zwanzig-drei* ausgesprochen wird, ist die deutsche Zahlensprechweise invertiert: Man sagt *ehnhundert-drei-und-zwanzig*. Die Bildung der Zahlwörter widerspricht somit der Lese- und Schreibrichtung und der Logik des Stellenwertsystems (vgl. u. a. Morfeld & Schuppener 2024, S. 23, Position Morfeld).

Ze: Zehner-Einer, kurz Ze, steht hier für die stellenwertgerechte Sprechweise der deutschen Zahlen, so wie das Positionspapier des Vereins Zwanzigeins e. V. (2024a) sie vorschlägt: z. B. *zwanzigeins* für 21, *sechzigfünf* für 65.

Zwanzigeins-Hypothese: Die Zwanzigeins-Hypothese besagt, dass die inverse Zahlensprechweise im Deutschen zu Leistungsdefiziten bei Grundschulkindern in Mathematik führt (Zwanzigeins e. V. 2024a).

2 LERN- UND RECHENSTÖRUNGEN

2.1 LERN- UND RECHENSTÖRUNGEN – BEGRIFF UND KLASSIFIKATION

Die meisten Kinder lernen im Verlauf ihrer Grundschulzeit problemlos das Lesen, Schreiben und Rechnen. Einigen Kindern bleibt der Zugang zu diesen grundlegenden Kulturtechniken jedoch verwehrt: Entweder sie haben auf Grund allgemeiner kognitiver Entwicklungseinschränkungen ein deutlich erhöhtes Risiko, an den Lernanforderungen einer Regelschule zu scheitern oder sie entwickeln bei guten allgemeinen Lernvoraussetzungen eine signifikante Diskrepanz zur erwartbaren spezifischen Leistung. Diese Kinder sind nicht in der Lage, ihrem kognitiven Leistungsniveau entsprechende Fähigkeiten im Lesen, Schreiben und / oder Rechnen zu entwickeln.

Die in früheren internationalen Klassifikationssystemen als „Schulleistungsstörung“ oder „Teilleistungsstörung“ bezeichneten Defizite beim Erwerb schulischer Fertigkeiten werden in den aktuellen Versionen ICD-10 bzw. ICD-11 und DSM-V als Lernstörungen bezeichnet. Beide Diagnosemanuale kategorisieren Rechenstörungen als domänenspezifischen Subtyp von Lernstörungen unter dem Sammelbegriff „neurodevelopmental disorders“ (von Aster et al. 2021, S. 2), d. h. es handelt sich sowohl bei Rechenstörungen als auch bei Lese-Rechtschreibstörungen um neuropsychologische Entwicklungsstörungen. Die ICD-11 bezeichnet Lernstörungen (inklusive Rechenstörungen) in der weiteren Ausdifferenzierung als „developmental learning disorder“, während das DSM-V den Begriff „specific learning disorder“ verwendet. Beide Diagnosemanuale verwenden für die Rechenstörung den Zusatz „with impairment in mathematics“ (vgl. *ibid.*), wobei das DSM-V dann weiter in Schweregrade ausdifferenziert. Der Schweregrad der Minderleistung manifestiert sich dabei in dem Leistungsunterschied, der mittels eines standardisierten Rechentests zwischen dem individuellen Testwert des betroffenen Kindes und der entsprechenden Alters- und / oder Klassennorm feststellbar ist. Strengere Cut-off-Werte von 10-15% sind indikativ für eine *Lernstörung*, während bei großzügigeren Cut-off-Werten von bis zu 25% das Vorliegen einer *Lernschwäche* angenommen wird (*ibid.*).

Es gibt jedoch bislang keine einheitliche und allgemein anerkannte Definition dieser begrifflichen Umschreibungen, so dass sowohl in der Literatur als auch in der

Diagnostik neben Lernschwäche und Lernstörung (bzw. Rechenschwäche / Rechenstörung und Lese-Rechtschreib-Schwäche / Lese-Rechtschreib-Störung) weiterhin Teilleistungsstörung, Dyskalkulie bzw. Dyslexie oder gemäß ICD-10 „umschriebene Entwicklungsstörungen schulischer Fertigkeiten“ synonym verwendet werden. Auch die Begriffe Lern- bzw. Rechenschwierigkeiten sind nach wie vor gebräuchlich, auch wenn Schipper (2001, zitiert nach Kaufmann & Wessolowski 2021, S. 9) zu Recht kritisiert, dass nur von Rechenschwierigkeiten zu sprechen „zu wenig deutlich [mache], dass es hier um außerordentliche Schwierigkeiten beim Erlernen des Rechnens geht, die ein Kind nicht mehr ohne besondere Unterstützung überwinden kann, und nicht um solche, die zu jedem Lernprozess gehören.“ Auch Kaufmann et al. (2013, S. 1) betonen: „Any theory of DD [developmental dyscalculia, Anm. d. Verf.] must account for differences between DD and individual differences in arithmetic in the general population.“

Der unterschiedliche Gebrauch der Begriffe ist also unter anderem auf die verschiedenen Erscheinungsformen der Dyskalkulie und den Schweregrad der Symptome zurückzuführen. Die individuellen Ausprägungen ergeben sich unter anderem aus der Tatsache, dass arithmetische Fähigkeiten aus verschiedenen Komponenten bestehen: „Arithmetic ability consists of many components (e.g., memorizing facts, executing procedures, understanding, and using arithmetical principles (Desoete et al. 2004; Dowker 2005, 2008)), each subject to individual differences that continue into adulthood (Dowker 2005; Kaufmann et al. 2011a)“ (ibid.). Demzufolge ist die Rechenstörung kein homogenes Syndrom, sondern durch heterogene Erscheinungsformen geprägt, die sich nicht durch ein Kerndefizit erklären lassen. Hierbei spielen auch Umwelteinflüsse eine Rolle: „ (...) heterogeneity of DD and other mathematics difficulties is also fostered by environmental factors, ranging from cultural factors (e.g., nature and extent of schooling, characteristics of the counting system) to the effects of pre-/postnatal illness or socio-emotional adversity (e.g., math anxiety)“ (ibid.). Die verschiedenen Ausprägungen der Rechenstörung lassen sich also auch auf kulturelle Faktoren, z. B. Art und Umfang der Schulbildung, die Merkmale der Zahlensprechweise, aber auch auf Auswirkungen prä- und postnataler Krankheiten oder sozio-emotionale Schwierigkeiten (z. B. Angst vor Mathematik) zurückführen.

2.2 RECHENSTÖRUNGEN: DIAGNOSTISCHE KRITERIEN UND INSTRUMENTE

Grundlage für die Diagnose einer Rechenstörung in Deutschland ist auf Grund der komplexen Prozesse bezüglich der Einführung der ICD-11 derzeit noch die ICD-10: „Die ICD-11 ist seit ihrem Inkrafttreten am 01.01.2022 grundsätzlich einsetzbar, jedoch ist die Entwurfsfassung der ICD-11 in Deutsch aus lizenzrechtlichen Gründen noch nicht nutzbar“ (Bundesinstitut für Arzneimittel und Medizinprodukte 2024). Weiterhin ist laut BfArM seit dem 1. Januar 2024 die ICD-10-GM (German Modification) in der Version 2024 anzuwenden (ibid.).

Da die ICD-10 sich bezüglich der Rechenstörung ausschließlich auf Defizite in den Grundrechenarten stützt, wurde sie auf der Basis aktueller Erkenntnisse hinterfragt (Haberstroh & Schulte-Körne 2019). Die Deutsche Gesellschaft für Kinder- und Jugendpsychiatrie, Psychosomatik und Psychotherapie (DGKJP) entwickelte daher in Zusammenarbeit mit 19 weiteren Fachgesellschaften und Verbänden die weltweit erste evidenz- und konsensbasierte S3-Leitlinie zur Diagnostik und Behandlung der Rechenstörung. Ziel war es, klare, empirisch fundierte Handlungsanweisungen für eine vereinheitlichte Diagnostik der Rechenstörung bereitzustellen und über die Wirksamkeit aktueller Präventions- sowie Fördermethoden aufzuklären. Die Leitlinie, deren Gültigkeit Ende Februar 2023 abgelaufen ist und die sich derzeit in Überarbeitung befindet, hat zu einer beachtlichen Verbreitung des evidenzbasierten Wissens zur Diagnostik und Behandlung bei der Rechenstörung beigetragen.

Die Anwendung des Intelligenz-Diskrepanz-Kriteriums wird hier ausdrücklich nicht empfohlen (AWMF: S3-Leitlinie, S. 21), was im Widerspruch zu den Kriterien der ICD-10 steht, die Lernstörungen (inklusive Rechenstörungen) „über das Vorliegen (erwartungswidriger) Minderleistungen im Bereich des Lesens, des schriftlichen Ausdrucks (einschließlich der Rechtschreibung) und des Rechnens definiert“ (zitiert nach Hasselhorn 2022, S. 2). Die Minderleistung muss demnach deutlich unter dem auf Grund des Alters, der bisherigen Beschulung und der allgemeinen Intelligenz zu erwartenden Niveau liegen. Unter folgenden Bedingungen ist somit laut ICD-10 die Diagnose „Lernstörung“ zu vergeben:

- eine lernbereichsspezifische Minderleistung (Diskrepanz zu dem auf Grund des Alters und des Beschulungsstands zu erwartenden Leistungsstand)
- eine allgemeine Intelligenz mit einem Gesamt-IQ über 70

- eine bedeutsame Diskrepanz zwischen der allgemeinen Intelligenz und den schulbezogenen Teilleistungen

Lernstörungen werden demzufolge über zwei Erwartungswidrigkeiten – auch als Diskrepanzen bezeichnet – definiert: (i) die Minderleistung im jeweiligen Lernbereich und (ii) die im Vergleich zur Leistung sehr viel bessere Intelligenz des betroffenen Kindes (vgl. *ibid.*). Diese Grundannahme ist jedoch umstritten.

So wie die S3-Leitlinie verzichtet auch das amerikanische Klassifikationssystem DSM-V auf das zweite strikte Intelligenz-Diskrepanz-Kriterium (vgl. von Aster et al. 2021, S. 2). Bemühungen, diesen Verzicht auch in der ICD-11 vorzunehmen, blieben erfolglos (Hasselhorn 2022, S. 2), so dass hier weiterhin kein einheitliches System bezüglich der diagnostischen Kriterien besteht.

Beide Klassifikationssysteme – sowohl das DSM-V als auch ICD-10 und ICD-11 – fordern, dass die schulischen Entwicklungsstörungen nicht durch unangemessene Beschulung erklärbar sein dürfen. Diese Forderung steht im Widerspruch zu der Tatsache, dass die Strukturen und Funktionen des Gehirns weitgehend erfahrungsabhängig und neuroplastisch reifen. Fehlende oder maladaptive Erfahrungen können also durchaus das Aus- oder Zurückbleiben solcher Reifungsprozesse bedingen. (Siehe auch das Vier-Stufen-Modell der Entwicklung zahlenverarbeitender Hirnfunktionen, Kap. 2.7.2.)

Es stellt sich zudem die Frage, wie eine qualitative Bewertung der Beschulung bzw. des Unterrichts in der Praxis umsetzbar sein sollte, da „unangemessene“ vs. „angemessene“ Beschulung weder definiert noch überprüfbar ist, was sowohl für den Einzelfall als auch für übergreifende Merkmale von pädagogischen und didaktischen Konzepten gilt (vgl. von Aster et al. 2021, S. 2). Pädagogische Konzepte werden länderübergreifend kontrovers diskutiert: So beinhalten japanische und chinesische Mathematikstunden im Vergleich zu amerikanischen Schulen mehr Erklärungen und fördern somit das rechnerische Verständnis, worauf einige Autoren die besseren Mathematikleistungen zurückführen (Perry 2000, zitiert nach Delazer et al. 2013, S. 236). Unangemessene Beschulung kann sich u.a. aber auch in häufigen Schul- und Lehrerwechseln ausdrücken oder in unzureichenden pädagogisch-didaktischen Kenntnissen der Lehrkraft (Kap. 2.6.1).

2.2.1 Diagnostik von Rechenstörungen im deutschsprachigen Raum

Wie oben erläutert, bemühte sich die S3-Leitlinie u.a. um eine klare Handlungsanweisung zu einer vereinheitlichten Diagnostik von Rechenstörungen. Laut Leitlinie sind Rechenstörungen durch Minderleistungen in den Bereichen Basiskompetenzen, Grundrechenarten und Textaufgaben gekennzeichnet, wobei sich die verminderte Leistung auf die Dauer und die Fehlerzahl bezieht, d. h. Menschen mit Rechenstörung machen bei mathematischen Aufgaben mehr Fehler und benötigen länger zum Lösen einer Rechenaufgabe als Menschen ohne Rechenstörung. Betroffen ist überdies das Arbeitsgedächtnis, und zwar vor allem der visuell-räumliche Notizblock: Menschen mit Rechenstörung weisen demzufolge Schwierigkeiten auf, räumliche und visuelle Informationen kurzzeitig zu speichern. Auch die Zentrale Exekutive ist betroffen (AWMF: S3-Leitlinie, S. 20 ff; Kap. 2.7.5). In der Praxis kommt trotz dieser Bemühungen um ein einheitliches System auch die ICD-10 zum Einsatz, die lediglich ein Defizit in den Grundrechenarten als Grundlage für die Diagnose definiert (ibid.) und, wie oben dargestellt, zusätzlich das Intelligenz-Diskrepanz-Kriterium fordert. Daher liegt bis heute weder im internationalen Kontext noch im deutschsprachigen Raum ein einheitlicher Begriff oder ein einheitliches Konzept zur Diagnostik der Rechenstörung vor: „(...) criteria for defining and diagnosing dyscalculia remain ambiguous“ (Mazzocco & Meyers 2003).

Die unterschiedlichen Diagnostikverfahren für die Rechenstörung im deutschsprachigen Raum sollen nachfolgend kurz erläutert werden.

Für eine Diagnostik kommen in der Praxis verschiedene standardisierte Rechentests zum Einsatz, die sich überwiegend an der S3-Leitlinie orientieren. Die Anforderungen, die eine solche Testbatterie mathematischer Funktionen demnach erfüllen soll, beinhalten einfache Zähl- und Transkodierungsoperationen, Grundrechenarten, perzeptive und kontextuelle Mengenoperationen, Textaufgaben und Aufgaben zum mathematisch-logischen Verständnis. Diese werden in Subtests evaluiert und die Ergebnisse im Vergleich zu einer Kontrollgruppe als normal, grenzwertig oder pathologisch bewertet (vgl. von Aster et al. 2013b, S. 85). Es sollen überdies die Leistungen im Bereich des visuell-räumlichen Arbeitsgedächtnisses und im Bereich der Exekutiven Funktionen erfasst werden (AWMF: S3-Leitlinie Rechenstörung, S. 15 f.). Weiterhin soll die Diagnostik einer Rechenstörung gemäß S3-Leitlinie ganzheitlich erfolgen und neben psychometrischen Verfahren ebenso eine klinische Untersuchung sowie qualitative Kriterien (Anamnese und Exploration)

beinhalten. Die klinische Untersuchung beinhaltet neben den körperlichen / neurologischen, sensorischen und intellektuellen Funktionen den biographischen Entwicklungsverlauf, die Familien- und die Schulsituation, die Auswirkungen der Leistungsdefizite auf die psychische und soziale Entwicklung, die schulische Integration sowie die gesellschaftliche Teilhabe (AWMF: S3-Leitlinie Rechenstörung, S. 20 ff.).

Bei der Durchführung der psychometrischen Tests ist der Zeitfaktor umstritten: Einerseits ist erwiesen, dass „besonders Kinder mit Teilleistungsstörungen unter Stress einen überproportionalen Abfall ihrer Leistungen haben“ (von Aster et al. 2013, S. 85), d. h. weniger gute Ergebnisse erzielen, als wenn sie ohne Zeitdruck arbeiten. Andererseits kann das Bearbeitungstempo auf Grund schulischer und beruflicher Gegebenheiten im späteren Leben nicht außer Acht gelassen werden, d. h. es ist keine Option, den Zeitfaktor unberücksichtigt zu lassen.

Auch die Normierung nach Altersklassen vs. Klassenstufen wird in den jeweiligen Tests in unterschiedlicher Weise umgesetzt. Da das Alter der Kinder einer Klassenstufe je nach Geburtsdatum, Einschulungsalter im jeweiligen Bundesland, Rückstellung (d. h. Wiederholung des letzten Kindergartenjahres) und etwaiger Wiederholung einer Klassenstufe sehr stark variieren kann, erfolgte die Normierung der Neuropsychologischen Testbatterie für Zahlenverarbeitung und Rechnen bei Kindern (kurz: ZAREKI-R; von Aster et al. 2013d) nach Altersklassen, so dass keine Verzerrung auf Grund des Alters stattfinden kann. Allerdings ist dies aus testökonomischer Sicht kritikwürdig, da z. B. die Durchführung als Gruppentest im Klassenverband erschwert wird. Die meisten anderen Tests wie z. B. auch der Heidelberger Rechentest (kurz: HRT 1–4; Haffner et al. 2005) wurden nach Klassenstufen gegliedert normiert. Hier wird wiederum der sehr straffe Zeitrahmen kritisiert, da dieser – wie oben erläutert – die Ergebnisse besonders bei Kindern mit Lernstörungen verzerren kann.

Kaufmann & Wessolowski (2021) halten standardisierte Tests zwar grundsätzlich für angezeigt, um „einen Vergleich der Leistungen des Einzelnen mit der Gruppe Gleichaltriger der Normierungsstichprobe möglich zu machen“; sie geben jedoch gleichzeitig zu bedenken, dass es nicht möglich sei, aus den Vergleichswerten dieser Tests konkrete Fördermaßnahmen abzuleiten, da diese „produktorientiert“ seien. Die Denkwege der Kinder würden nicht beachtet, „obwohl richtige Ergebnisse nicht immer für Verständnis und sinnvolle Lösungswege stehen und falschen

Lösungen häufig ein richtiger Kern zugrunde liegt“ (S. 17). Sie plädieren daher für den Einsatz des von ihnen konzipierten informellen Verfahrens, um die Vorgehensweisen der Kinder besser nachvollziehen zu können. Grundlegend für die Einsicht in das Denken der Kinder sei das „laute Denken“ des Kindes im Rahmen des „diagnostischen Gesprächs“ und die daraus resultierende qualitative Fehleranalyse, da Fehler „nicht zufällig oder durch flüchtiges Verrechnen (entstehen), sondern Ergebnisse subjektiver Strategien sind“ (ibid., S. 18 ff).

Obwohl Kriterien und diagnostische Instrumente unterschiedlich sind und zweckgebunden eingesetzt werden, besteht doch weitgehend Einigkeit darin, dass „die Diagnose einer Rechenstörung dann gerechtfertigt ist, wenn signifikante und anhaltende Schwierigkeiten beim Erlernen der rechnerischen Fähigkeiten vorliegen, die deutlich unter dem Niveau liegen, das in Bezug auf die Alters- bzw. Klassennorm zu erwarten wäre“ (von Aster et al. 2021, S. 2). Kaufmann & Wessolowski (2021, S. 9) spezifizieren die rechnerischen Fähigkeiten und verstehen unter Rechenstörungen bzw. Rechenschwäche, „dass Kinder auf Grund (noch) fehlender Voraussetzungen kein Verständnis für Zahlen, Rechenoperationen und Rechenstrategien aufbauen konnten.“

2.3 TEUFELSKREIS LERNSTÖRUNG

Unabhängig davon, ob eine Rechen- bzw. Lernstörung von fachlicher Seite diagnostiziert wurde oder nicht, ergeben sich in vielen Fällen weitreichende belastende Folgen für die Betroffenen und ihr Umfeld.

Betz & Breuninger beschreiben bereits 1982 negative Kreisläufe aus Versagen und Misserfolgserwartung, Stress, Blockierung und Angst: „Das Kind nimmt selbst die Minderleistung vorweg und bekommt Angst. Diese führt zur Blockierung der kognitiven Prozesse infolge von Stressreaktionen, so dass die Leistung gerade dann abfällt, wenn sie gefordert wird“ (S. 40). Abwehr- und Vermeidungsmechanismen, erneute Lernausfälle, Konzentrationsstörungen, die Entwicklung von Schuldgefühlen und einem negativen Selbstkonzept sind weitere Faktoren, die zu einer Stabilisierung der negativen Lernstruktur mit Misserfolgserwartung führen. Durch Unverständnis und Zuschreibungen des sozialen, schulischen und familiären Umfelds können diese Symptomatiken

systematisch verstärkt werden: Laut Betz & Breuninger werden Konzentrationsprobleme und Verhaltensauffälligkeiten demnach nicht als Folge, sondern als Ursache für die Lernprobleme identifiziert, pädagogische Hilfen deshalb zu spät und an falscher Stelle eingesetzt. Die Enttäuschung und Entmutigung von Eltern und Lehrern durch ausbleibende Erfolge trotz angebotener Hilfe führe dazu, dass Vorurteile gegen diese Kinder aufgebaut werden, die die Entwicklung der negativen Lernstruktur wiederum auf verheerende Weise begünstigen: „Die Ursachen bilden ein vernetztes System, in welchem sie sich gegenseitig aufrechterhalten und verstärken (Teufelskreise)“ (ibid., S. 47). Die Folgen, die sich aus beständigen Misserfolgserlebnissen und psychosozialen Belastungen ergeben, führen häufig zu weiteren Störungen, wie Angststörungen, Somatisierungsstörungen wie z. B. häufige Kopf- und Bauchschmerzen oder Störungen aus dem externalisierenden Spektrum (Kap. 2.5), was wiederum zu einer möglichen Stigmatisierung beitragen kann.

Lernstörungen sind demnach meist nicht auf eine Ursache zurückzuführen, sondern es wirken mehrere Faktoren zusammen (s. auch Kap. 2.6). Ein mangelndes Selbstwertgefühl (emotionale Ebene) und ein negatives akademisches Selbstkonzept tragen häufig zur Entstehung negativer Kreisläufe bei. Dies ist insofern fatal, als dem schulischen Selbstkonzept eine entscheidende Rolle bei der Bewältigung von Lernaufgaben zugesprochen wird, wie Schuchardt et al. (2015, S. 522) betonen. Zwischen spezifischen Lernschwächen und zugehörigen niedrigen Selbstbewertungen besteht laut ihrer Studie eine hohe Koppelung (ibid., S. 523), wobei sie einen Unterschied feststellten zwischen Kindern mit isolierten Lernschwächen – diese wiesen allein in dem betroffenen Lernbereich niedrige Selbstbewertungen auf – und solchen Kindern mit kombinierten Lernschwächen, deren Selbstkonzeptschätzung in allen schulischen Bereichen schwach ausfiel. Es handelt sich dabei um ein zeitlich überdauerndes Phänomen: Bei allen Kindern hatten sich zum Ende der 4. Klasse die Selbsteinschätzungen auf Grund jahrelanger negativer Erfahrungen „stabil gefestigt“ (ibid.).

Auch Schulte-Körne (2016, S. 186) betont die Relevanz negativer Selbstbewertungen von Kindern mit Lernstörungen. Negative Schulerfahrungen und Misserfolgserlebnisse führen zu Versagensängsten und einem reduzierten Selbstwert. Außerdem findet sich eine erhöhte Komorbidität mit depressiven Störungen, Angststörungen, hyperkinetischen Störungen und sozial-emotionalen

Störungsbildern. Auf Grund einer „hohe(n) Persistenz der Kernsymptomatik“ und der psychischen Belastungen kommt es häufiger zu Klassenwiederholungen, Schulwechseln und Schulabbrüchen und Betroffene erreichen häufig einen deutlich schlechteren Schulabschluss als aufgrund der kognitiven Leistungsfähigkeit zu erwarten wäre.

2.3.1 Die Folgen von Lernstörungen

Bei einer Rechen- oder Lese-Rechtschreib-Störung handelt es sich um einen chronischen Zustand, der unbehandelt bis ins Erwachsenenalter persistiert und demnach mit hohen Risiken für die Bildungs-, Gesundheits- und Persönlichkeitsentwicklung einhergeht (vgl. von Aster et al. 2021, S. 2). Ohne spezifische Förderung führt Rechenschwäche häufig zu Schulversagen und -absentismus (Haberstroh & Schulte-Körne 2019).

Die Folgen unbehandelter Lernstörungen haben aber nicht nur für Betroffene und ihre Familien weitreichende Konsequenzen, sondern sind auch gesellschaftlich relevant: Es gibt Hinweise darauf, dass von Lernstörungen betroffene Kinder später häufiger Stigmatisierung (Geiger & Brewster, zitiert nach Hasselhorn 2022, S. 2) und sozialen Ausschluss (Daley & Rappolt-Schlichtmann 2018, zitiert *ibid.*) erfahren. Lernbeeinträchtigungen etwa beim Erwerb der Schriftsprache erhöhen auch das Risiko für antisoziales Verhalten im Erwachsenenalter (Simonoff et al. 2004, zitiert *ibid.*).

Haberstroh & Schulte-Körne kritisieren, dass länderübergreifend keine schulische Gleichbehandlung der Rechenstörung mit der LRS stattfindet. Fördermaßnahmen (auch Nachteilsausgleich) sind beispielsweise nur auf die Grundschule beschränkt oder werden grundsätzlich nicht angeboten. Ebenso werden die Kosten einer außerschulisch notwendigen Lerntherapie von den Krankenkassen nicht übernommen, was für die betroffenen Familien eine deutliche finanzielle Belastung über Jahre hinweg darstellt mit der Folge, dass die betroffenen Kinder und Jugendlichen oft nicht adäquat gefördert und behandelt werden.

Bezüglich der ökonomischen Folgen persistierender Rechenschwäche verweisen Haberstroh & Schulte-Körne (2019, S. 107) auf eine große Kohortenstudie in England, die zeigte, dass eine geringe Rechenleistung mit großen psychosozialen und ökonomischen Risiken – sowohl aus individueller als auch aus

volkswirtschaftlicher Sicht – verbunden ist: 70–90% der Betroffenen verließen die Schule vorzeitig im Alter von 16 Jahren; mit 30 Jahren war der Anteil an Vollzeitbeschäftigten sehr gering. Die Wahrscheinlichkeit, arbeitslos zu sein und depressive Symptome zu entwickeln, war doppelt so hoch wie bei Personen ohne Rechenstörung. Für Großbritannien werden die durch massive Beeinträchtigung der Rechenfertigkeit entstehenden Kosten auf 2.4 Milliarden Pfund jährlich geschätzt.

Anlass zur Sorge geben laut Hasselhorn auch die hartnäckige Persistenz sowie die jüngeren epidemiologischen Untersuchungen zufolge deutlich höheren Auftretenswahrscheinlichkeiten von Lernstörungen als noch vor wenigen Jahrzehnten (Hasselhorn 2022, S. 12).

2.4 PRÄVALENZ

Die Auftretenswahrscheinlichkeit und Bedeutung von Rechenstörungen wird laut Haberstroh & Schulte-Körne (2019, S. 1) nach wie vor unterschätzt: 2-7% aller Grundschulkinder, Jugendlichen und Erwachsenen leiden unter Dyskalkulie, was in Deutschland bei einer Grundschulpopulation von ca. 3 Millionen Kindern (Statistisches Bundesamt) geschätzten 84 000–195 750 rechenschwachen Grundschulern entspricht. Auch andere Quellen geben die Prävalenzraten länderübergreifend mit etwa 3-6% der Grundschulpopulation an, womit sie in etwa gleich hoch wie diejenige von Lese-Rechtschreib-Störungen und ADHS sind (Shalev et al. 2000, S. 59; Delazer et al. 2013, S. 231). Es scheint hinsichtlich der Auftretenswahrscheinlichkeit kaum länderspezifische Unterschiede zu geben (Shalev et al., *ibid.*), wobei sich der Ländervergleich an dieser Stelle auf Studien aus den USA, England, Deutschland, der Schweiz und Israel bezieht. Dass in dem DSM-IV (DSM 4) die Dyskalkulie noch als „rare learning disability“ mit einer Prävalenzrate von lediglich 1% bezeichnet wurde, ist angesichts aktueller Forschungsergebnisse hinfällig und wurde in dem DSM-V entsprechend korrigiert.

Die Rechenschwäche weist eine hohe Komorbidität mit Lese- und / oder Rechtschreib-Schwierigkeiten auf (Kap. 2.5). Die Wahrscheinlichkeit, dass zusätzlich zur Rechenschwäche eine Lese-Rechtschreibschwäche vorliegt, geben von Aster et al. mit ca. 33-40% an (2021, S. 3). Andere Quellen gehen davon aus,

dass 48% der Kinder mit einer Rechenstörung auch die Kriterien für eine Störung im Lesen und / oder Rechtschreiben erfüllen (Fischbach et al. 2013, zitiert nach Hasselhorn 2022, S. 4). Jedes 7. bis 8. Grundschulkind erfüllte in der repräsentativen Studie von Fischbach et al. (2013) die WHO-Kriterien für eine Lernstörung.

Hasselhorn geht davon aus, dass etwa jedes dritte Kind in Deutschland im Laufe der Grundschule unter besonderen Lernschwierigkeiten beim Erwerb von Lesen, Rechtschreiben und Rechnen leidet (2022, S. 1). Dies bedeutet, dass vermutlich mehr Grundschulkindern beim Erwerb von Schriftsprache und Arithmetik beeinträchtigt sind, als auf Grund fehlender Diagnose statistisch erfasst werden. Zum einen werden nicht alle betroffenen Kinder an entsprechender Stelle vorstellig, was verschiedene Gründe haben kann, wie z. B. nicht ausreichende fachärztliche Versorgung bzw. monatelange Wartezeiten auf einen Termin. Unsicherheit bei den Eltern, ob die Schwierigkeiten bereits behandlungsbedürftig sind, oder fehlendes Wissen über die Versorgungsangebote sowie sprachliche und zum Teil kulturelle Hürden bei Familien mit Migrationshintergrund können mitursächlich sein.

Zum anderen erfüllen nicht alle Kinder die diagnostischen Kriterien – auch, da zum Teil noch das Diskrepanzkriterium zur Anwendung kommt. Dass die diagnostischen Kriterien nach wie vor nicht einheitlich umgesetzt werden und für Betroffene wenig transparent sind, verkompliziert den Prozess zusätzlich auf individueller und gesellschaftlicher Ebene: Die Kinder bekommen zum Teil ihnen zustehende Hilfen nicht, weil sie keine Diagnose haben, und die statistische Erfassung dieser Kinder ist nicht gegeben, womit die gesellschaftliche Bedeutung unterschätzt und die Planung und Durchführung von Förderangeboten erschwert wird.

2.4.1 Geschlechterverteilung

Inkonsistente Befunde zeigen sich auch hinsichtlich der Geschlechterverteilung. Shalev et al. gehen davon aus, dass länderübergreifend mehrheitlich Jungen von allgemeinen Lernschwierigkeiten betroffen sind, während es bei Rechenstörungen ein ausgeglichenes Verhältnis zwischen den Geschlechtern gibt (2000, S. 59). Eine in Deutschland durchgeführte Studie mit 546 Drittklässlern kam jedoch 1992 zu dem Ergebnis, dass bei einer Prävalenz von insgesamt 4,4% „a slight overrepresentation of girls“ besteht (Klauer 1992, zitiert *ibid.*). Zwei weitere Studien aus den 90er

Jahren kamen wiederum zu Ergebnissen mit ausgeglichener Geschlechterverteilung bei einer Prävalenz von 6,6% (zitiert *ibid.*). Auch Haberstroh & Schulte-Körne (2019, S. 107) kommen mit Bezug auf eine länderübergreifende Untersuchung von Lewis & Fisher (2015) zu dem Schluss, dass das Geschlechterverhältnis bei der Rechenstörung ungefähr ausgeglichen ist, wobei tendenziell mehr Mädchen betroffen sind.

Dem widersprechen aktuelle Studien mit repräsentativen Stichproben im deutschsprachigen Raum, die zeigen, dass Mädchen zwei bis drei Mal häufiger von einer Rechenstörung betroffen sind als Jungen (Fischbach et al. 2013; Moll et al. 2014; beide zitiert nach Hasselhorn 2022). Die uneinheitlichen Ergebnisse zu dieser Frage führt er darauf zurück, dass die Stichproben teilweise nicht repräsentativ sind, d. h. nur Kinder, die bereits in Einrichtungen zur Diagnose und Therapie vorstellig wurden, berücksichtigt wurden (*ibid.*).

Auch Krinzinger & Günther (2013) gehen davon aus, dass Mädchen häufiger betroffen sind und führen verschiedene mögliche Gründe an: Es konnte zum Beispiel gezeigt werden, dass Jungen häufiger räumliche Strategien wie räumliches Visualisieren für das Lösen von mathematischen Problemen (v. a. Textaufgaben) verwenden, Mädchen hingegen häufiger verbale bzw. algorithmische Strategien, die weniger effektiv seien. Als weitere Faktoren führen die Autoren die bei Mädchen ausgeprägtere Rechenangst sowie Stereotypisierungen an: Erwartungshaltungen von Eltern und Lehrern, dass Jungen in Mathematik talentierter seien als Mädchen, könnten verschiedenen Studien zufolge den Charakter einer selbsterfüllenden Prophezeiung annehmen, so dass sich die Rechenleistungen von Mädchen weiter verschlechtern. Laut einer aktuellen Studie (Wagner et al. 2025), die auf einer Umfrage mit 124 Kinder der dritten und vierten Schulstufen basiert, schätzen Mädchen im Vergleich mit Jungen ihre eigenen Rechenleistungen schlechter ein. Obwohl das Grundinteresse an Mathematik laut Selbsteinschätzung ähnlich ist, scheinen geschlechtsspezifische Leistungsvorstellungen in den Köpfen von Kindern immer noch vorhanden zu sein.

2.5 KOMORBIDITÄTEN

Wie oben beschrieben, treten Lese-Rechtschreib-Schwierigkeiten und Rechenstörungen nicht nur in Form von isolierten Teilleistungsstörungen auf; in

vielen Fällen werden beide Störungen diagnostiziert. Zudem sind in Verbindung zu Lernstörungen häufig auch neurologisch-neuropsychiatrische Krankheitsbilder beobachtbar wie z. B. Sprachentwicklungsstörungen, grob- und feinmotorische Auffälligkeiten sowie Epilepsien, Stoffwechselstörungen wie Phenylketonurie und genetische Störungen wie fragiles X-Syndrom (Delazer et al. 2013, S. 231; Bachmann 2008, zitiert nach Bender et al. 2017, S. 66).

Weitere häufige Begleitstörungen von Lern- bzw. Rechenstörungen umfassen ADHS mit einer Merkmalstrias von eingeschränkter Aufmerksamkeit, Überaktivität und impulsivem Handeln sowie externalisierende Symptome wie Störungen des Sozialverhaltens, z. B. aggressives Verhalten.

Andere Kinder entwickeln depressive Symptome mit Niedergeschlagenheit, Selbstentwertung und sozialem Rückzug, also Symptome aus dem internalisierenden Störungsspektrum. Des Weiteren können sich Angststörungen entwickeln, die entweder auf bestimmte Bereiche beschränkt bleiben (Mathe- oder Prüfungsangst) oder sich zu einer generellen Schulunlust oder -angst ausdehnen können (von Aster et al. 2021, S. 3). Daniel Pennac beschreibt in seinem autobiografischen, mit dem Prix Renaudot ausgezeichneten Roman *Chagrin d'école* (dt. Übersetzung: Schulkummer) die Situation eines Schülers, der bei dem Versuch, seine Hausaufgaben zu erledigen, zwei Stunden in „mathematischer Stumpsinnigkeit (...), in einer geistigen Lähmung“ verbringt. Die Angst bezeichnet er als großes Thema seiner Schuljahre, als eine „gigantische Barriere.“

2.5.1 Angst und ihre Auswirkungen

Laut Vloet & Romanos sind Angststörungen „mit einer durchschnittlichen Prävalenz von ca. 10% die häufigste psychische Störung im Kindes- und Jugendalter (Merikangas et al. 2010) und stellen im Sinne einer sequenziellen Komorbidität einen bedeutenden Risikofaktor für die Entwicklung weiterer psychischer Störungen im Erwachsenenalter dar“ (2021, S. 430). Eine Metaanalyse (Nelson & Harwood 2011) zur Häufigkeit von Ängsten bei Kindern und Jugendlichen mit Lernstörungen (Lese-Rechtschreib-Störung, Dyskalkulie) zeigte ein signifikant häufigeres Auftreten von Ängsten ($d = 0,61$; $p < 0,001$), unabhängig von Geschlecht und Klassenstufe (Schulte-Körne 2016, S. 186). Den engen Zusammenhang

zwischen allgemeinen Ängsten, Lernmotivation und Mathematikleistungen belegen verschiedene Studien (Ramirez et al. 2018; Xie et al. 2019).

Angst kann als „habitual tendency to perceive stressful situations as threatening“ beschrieben werden (Luttenberger et al. 2018). Sie drückt sich im Sinne von Langs *triple response framework of emotion* auf drei Ebenen aus, die auch auf eine spezifische Mathematik-Angst zutreffen:

„According to the triple response framework of emotion (Lang, 1968), anxiety is expressed in three main behavioral systems: physiological, verbal (cognitive), and overt (behavioral). The emotional experience of math anxiety is consistent with this approach; math-anxious individuals exhibit physiological arousal (Faust, 1992), have faulty beliefs about their problem-solving abilities (F. C. Richardson & Woolfolk, 1980), enroll in fewer math courses (Hembree, 1990), and avoid majoring in mathematics and the physical sciences (Chipman, Krantz & Silver, 1992)“ (Hopko et al. 1998, S. 344).

Mathe- oder Rechenangst ist eine psychopathologische Störung, für die eine entsprechende Behandlung indiziert ist. Tatsächlich spielt die Rechenangst eine wichtige Rolle bei der Entstehung und Aufrechterhaltung der Dyskalkulie (von Aster et al. 2021, S. 7).

So zeigten Untersuchungen, dass sich Angst vor Mathematik und / oder Zahlen und Zahlbeziehungen negativ auf die Rechenfertigkeiten auswirkt. Individuen mit hohem (Mathematik-)Angstniveau erzielen schlechtere Rechenleistungen als Individuen mit niedrigem Angstniveau (Ashcraft & Kirk 2001, zitiert nach Delazer et al. 2013, S. 237). Es zeigte sich weiterhin, dass sich ein hohes Angstniveau speziell negativ auf das Arbeitsgedächtnis auswirkt (für kontroverse Resultate bei Grundschulkindern siehe Krinzinger et al. 2009, zitiert *ibid.*), welches eine besondere Rolle bei dem Erwerb mathematischer Fähigkeiten spielt: Unter anderem ist es notwendig für den Faktenabruf, für die Anwendung von prozeduralem Wissen (also das Lösen komplexer, mehrstelliger Rechnungen) sowie für das Lösen von Textaufgaben (*ibid.*; s. auch Kap. 2.7.4). Ist also das ohnehin bei vielen rechenschwachen Kindern defizitäre Arbeitsgedächtnis zusätzlich durch Angst belastet, hat dies zwangsläufig negative Auswirkungen auf die mathematischen Leistungen.

Saga und KollegInnen (2022) fanden in einer kürzlich erschienenen Studie, dass die Rechenleistungen von Kindern mit Dyskalkulie von multiplen Faktoren beeinflusst werden, die nicht nur das schulische und familiäre Umfeld betreffen,

sondern auch die subjektiv empfundene Lebensqualität der Kinder (vgl. Kucian 2023, S. 229).

2.5.2 Entstehungskomponenten von Rechenangst

Die verschiedenen Entstehungskomponenten für Rechenangst fassten Ashcraft et al. (2007; zitiert nach von Aster et al. 2021, S. 8) in einem Strukturmodell zusammen, das auf verschiedenen Studienergebnissen basiert. Den Ausgangspunkt bilden die persönliche Lerngeschichte sowie eine transgenerationale Transmission für generelle Ängstlichkeit. Daraus resultieren im ungünstigen Fall unzureichende rechnerische Fertigkeiten, eine geringe Motivation sowie Defizite im Arbeitsgedächtnis, woraus wiederum Leistungsdefizite und Vermeidungsverhalten resultieren. Kognitive Verzerrungen in Bezug auf mit Zahlen und Rechnen assoziierte Inhalte können eine weitere Folge sein: Es kommt zu einer erhöhten Selbstaufmerksamkeit, d. h. die Aufmerksamkeit wird auf negative Ereignisse in Bezug auf das Rechnenlernen gelenkt, negatives Feedback zu einer schlechten Rechenleistung wird internalisiert. Dies führt generell zu einer negativen Einstellung gegenüber dem Rechnen und schließlich zu einer erhöhten Vulnerabilität für Angst (vgl. *ibid.*). Somit erfüllt die Rechenangst alle Kriterien einer Angststörung. Das Modell verdeutlicht, dass sich Rechenangst in einer komplexen Wechselwirkung verschiedener Faktoren ungünstig auf die Lern- und Leistungsentwicklung auswirkt.

Bei der Entstehung von Rechenangst spielt offenbar neben nicht altersgerechten Fähigkeiten und einem Kreislauf aus negativen Erfahrungen die soziale Transmission eine tragende Rolle: Kinder, deren Eltern eine negative Einstellung gegenüber oder wenig Erfahrung mit dem Schulfach Mathematik haben, scheinen besonders gefährdet zu sein, Rechenstörungen und Rechenangst zu entwickeln (von Aster et al. 2021, S.7).

Eltern, die selbst unter Mathe-Angst leiden, gefährden laut einer Studie ihre Kinder, wenn sie bei den Mathe-Hausaufgaben regelmäßig unterstützen:

We found that when parents are more math anxious, their children learn significantly less math over the school year and have more math anxiety by the school year's end—but only if math-anxious parents report providing frequent help with math homework. Notably, when parents reported helping with math homework less often, children's math achievement and attitudes were not related to parents' math anxiety.

provide evidence of a mechanism for intergenerational transmission of low math achievement and high math anxiety (Maloney et al., 2015).

Chinesische Forscher fanden, dass insbesondere die Vater-Kind-Beziehung Auswirkungen auf die Mathe-Angst von Kindern haben könnte:

The important result was that the father–child relationship rather than the mother–child relationship longitudinally predicted the math anxiety of children, even after controlling for a set of covariates, including the gender of children, age of children, learning anxiety of children, social anxiety of children, learning motivation of children, math performance of children and math anxiety of parents (Ma et al. 2021, S. 6).

Neben der Rolle der Eltern ist eine etwaige negative Einstellung oder Angst seitens des Lehrpersonals problematisch: Die wachsende Literatur zu diesem Thema zeigt, „dass die „mathematischen Einstellungen der Lehrer_innen und Mathematikangst sich auf die mathematischen Einstellungen, Ängste und mathematischen Leistungen der Schüler_innen auswirken“ (Kucian 2022, S. 239). Eine Befragung unter 173 Grundschullehrpersonen kommt beispielsweise zu dem Ergebnis, dass „räumliche Kompetenzen, räumliche Ängste und Fähigkeiten räumlicher Aspekte der Mathematik der Lehrer_innen im engen Zusammenhang zur Angst, Mathematik zu unterrichten, der Einstellung gegenüber Mathematik sowie der Wirksamkeit, Mathematik zu unterrichten beziehungsweise zu lernen, stehen“ (Burte et al. 2020, zitiert ibid.).

Eigene Mathe-Angst, fehlendes Wissen oder mangelnde Kompetenzen seitens der Eltern und / oder des Lehrpersonals können demnach Risikofaktoren für die Entstehung von Rechenangst und -störungen sein, wobei dies keine Schuldzuweisung, sondern Anregung zu (bildungs-)politisch initiierten Veränderungen im Bildungssystem sowie für Themenschwerpunkte in Aus- und Weiterbildungen sein soll.

2.6 URSACHEN VON LERNSTÖRUNGEN

2.6.1 Unzureichende pädagogisch-didaktische Maßnahmen als Risikofaktor

Es besteht also ein Konsens darüber, dass unzureichende pädagogisch-didaktische Maßnahmen für Kinder mit Lernschwierigkeiten dazu führen können,

dass Schwierigkeiten persistieren und sich z. B. zu einer Rechenschwäche entwickeln (u.a. Ginsburg et al. 1998; Reusser 2000; beide zitiert nach Delazer et al. 2013, S. 231 und S. 236). Neben den oben beispielhaft beschriebenen Einstellungen und Kompetenzen von Lehrpersonal, die mitursächlich für Angst und somit Versagen sein können, kann der Unterricht selbst „unbedacht und ungewollt bei einzelnen Kindern Lernprobleme in Mathematik mitverursachen“ (Kaufmann & Wessolowski, S. 11). Kleine Lücken können auf Grund des hierarchischen Aufbaus des Stoffes zu großen Verständnisproblemen führen oder aber der Unterrichtsstil bzw. die Unterrichtsmethodik passen nicht zu den Lernbesonderheiten des Kindes. Schwierigkeiten eines Kindes beim Erwerb arithmetischer Fähigkeiten können daher auch als Passungsproblem zwischen Unterricht und den Lernbesonderheiten eines Kindes zu sehen sein (ibid.).

Lorenz führt aus, dass „ungünstige Szenarien“ im schulischen Kontext „dem Kind in seiner arithmetischen Entwicklung nicht nur nicht helfen, sondern es im Gegenteil behindern“, was letztendlich dazu führe, dass Kinder „nicht durch, sondern trotz des Mathematikunterrichts lernen“ (2013, S. 182). Konkret kritisiert werden an anderer Stelle zum Beispiel veraltete Unterrichtsmethoden, in denen „die in Deutschland übliche, sehr lehrerzentrierte und kleinschrittige Behandlung von Textaufgaben (...) die Schüler entmündige,“ wie Stern formuliert (2013, S. 150). Die Autorin kritisiert eine verbreitete Grundhaltung: „Es wird davon ausgegangen, dass das Wissen des Lehrers direkt auf den Schüler übertragen wird.“

Hasselhorn verweist auf den Zusammenhang zwischen der Häufigkeit, mit der die Diagnose Rechenstörung gestellt wird, und Umfang und Qualität des Unterrichtsangebots (2022, S. 4). Er bezieht sich dabei auf eine Studie von Fuchs und KollegInnen (2005) mit 564 Kindern der ersten Klasse, bei denen zu Schulbeginn ein erhöhtes Risiko für Schwierigkeiten beim Erwerb des Rechnens festgestellt worden war. In der Teilstichprobe, die drei Mal wöchentlich gezielten Förderunterricht in Mathematik bekam, fiel die Prävalenzrate von Kindern mit niedrigen Mathematikleistungen am Ende der ersten Klasse mit 4.3% deutlich niedriger aus als bei denjenigen Kindern, die keine spezifische Förderung erhielten: Hier lag die Prävalenzrate bei 5.3%. Ein qualitativ hochwertiges und auf die Bedürfnisse der Kinder abgestimmtes Angebot reduziert also das Risiko einer diagnostizierten Rechenstörung.

Auch Bender et al. (2017, S. 66) benennen erfahrungsabhängige und lerngeschichtliche Einflüsse wie das Lernangebot und -klima der Schule, kognitive Anregungen, aber auch die emotionale Unterstützung und die Leistungserwartungen der Eltern als Wirkfaktoren für Lern- und Rechenschwächen.

2.6.2 Genetische Determinanten von Lernstörungen

In der Verhaltensgenetik werden Lernstörungen als erheblich genetisch prädestiniert eingestuft. Auf den Zusammenhang zwischen persistierenden Rechenstörungen und positiven Familienanamnesen wird in der Literatur verschiedentlich hingewiesen (von Aster 2013a, S. 25). Lernstörungen verteilen sich demnach nicht zufällig in einer Population (Hasselhorn 2022, S. 4). Das Risiko, Schwierigkeiten beim Erwerb von Rechnen, Lesen oder Schreiben auszubilden, ist bei Kindern, in deren Familien bereits eine diagnostizierte Lernschwäche aufgetreten ist, deutlich erhöht. Beispielsweise entwickelt ein Kind, das mindestens ein Geschwisterkind mit einer diagnostizierten Lese-Rechtschreibstörung hat, mit einer Wahrscheinlichkeit von ca. 43-60% auch eine LRS (Scerri & Schulte-Körne 2010, zitiert *ibid.*). Die Heritabilität von Dyslexie wird von einem internationalen Forschungsteam in einer aktuellen Studie mit 40-60% angegeben (Gialluisi et al. 2021, S. 3004). Zu ähnlichen Ergebnissen kommen Shalev et al. (2001, zitiert nach Hasselhorn 2022, S. 4) für den Bereich Dyskalkulie. Es bleibt jedoch bis heute unklar ob und in welchem Umfang genetische oder nicht-genetische, epigenetische Mechanismen der Vererbung eine Rolle spielen. Epigenetische Veränderungen bilden sich, anders als genetische, erfahrungsabhängig und sind grundsätzlich reversibel. Sie werden auch transgenerational wirksam.

Die Betonung angeborener genetischer Effekte fördert „deterministische Reflexe des Unabänderlichen und untergräbt damit das Bemühen um therapeutische Veränderung“, wie von Aster (2023, S. 2) herausstellt. Den Anteil genetischer Effekte an der Entstehung von Teilleistungsstörungen schätzt er im Vergleich zu den erfahrungsabhängigen, epigenetischen Wirkungen als „sehr gering“ ein. Das neuroplastische Veränderungspotential im lernenden menschlichen Gehirn sei enorm und schulisches Scheitern daher kein einfach hinzunehmendes Schicksal. Neuronale Plastizität steht für die Erkenntnis, dass die neuronalen Netzwerke nicht starr und invariabel sind, sondern im Prozess des Lernens entstehen und sich fortlaufend verändern (Kucian & von Aster 2013; von Aster et al. 2021).

2.7 NEUROPSYCHOLOGIE DER NORMALEN UND ABWEICHENDEN ENTWICKLUNG DER ZAHLENVERARBEITUNG

Rechnen erfordert das Zusammenspiel zahlreicher verschiedener Hirnfunktionen. Störungen bzw. Entwicklungsverzögerungen spezifischer kognitiver Komponenten der Zahlenverarbeitung gelten demnach neben biogenetischen Voraussetzungen und Umweltfaktoren als Risikofaktoren für die Entstehung von Dyskalkulie. Im Folgenden soll daher zunächst der reguläre Erwerb arithmetischer Kompetenzen aus neuropsychologischer Perspektive erläutert werden, um anschließend zu untersuchen, in welchen Bereichen Störfaktoren auftreten können.

Die aktuellen Annahmen über die Organisation von Rechenprozessen basieren einerseits auf Läsionsstudien, die Rückschlüsse auf verschiedene Funktionen und Verarbeitungsmechanismen des Gehirns zulassen; zum anderen konnte die moderne Neurowissenschaft mit unterschiedlichen Methoden der Bildgebung nachweisen, dass verschiedene Gehirnareale an den komplexen Prozessen der Zahlenverarbeitung und des Rechnens beteiligt sind. So eröffnet zum Beispiel die funktionelle Magnet-Resonanz-Tomographie (fMRI) die Möglichkeit, aktivierte Gehirnareale zu lokalisieren und bildlich darzustellen. Mit dieser Visualisierung von Gehirnaktivitätsmustern kann die neuronale Basis des Zahlenverständnisses erforscht und zu einem besseren Verständnis assoziierter Prozesse beigetragen werden (van Eimeren & Ansari 2009, S. 26).

1985 entstanden erstmals Bilder vom rechnenden Gehirn (Roland & Friberg 1985), die beim Kopfrechnen erhöhte Aktivitäten im präfrontalen Kortex sowie in einem eingeschränkteren inferior-parietalen Areal nachwiesen, und zwar sowohl in der linken als auch in der rechten Hemisphäre. Auch wenn die Methode von Roland und Friberg inzwischen veraltet ist – sie arbeiteten mit radioaktivem Xenon –, so können doch die Ergebnisse von damals auch unter Verwendung moderner bildgebender Verfahren im Kern bestätigt werden. Beim Lösen mathematischer Aufgaben greift das Gehirn offenbar auf ein bildlich-räumliches und ein sprachliches Zahlensystem zurück. Soll beispielsweise die größere von zwei Zahlen genannt werden, so werden die Zahlen offenbar auf einem inneren Zahlenstrahl¹ lokalisiert und die

¹ Zum so genannten SNARC-Effekt (Dehaene et al. 1993) und der darauf basierenden verbreiteten Annahme eines „mental number line“ (*mental number line*) siehe u. a. Lorenz 2017, S.130 ff., Grond et al. 2013, S. 47 ff. und van Eimeren & Ansari 2009, S. 26. Für eine kontroverse Meinung s. Gaidoschik 2015.

Entscheidung auf der Basis ihrer relativen räumlichen Distanz getroffen. Auf den Hirnbildern leuchten in diesem Fall Regionen im linken und rechten Parietallappen auf, das Gebiet für räumliches Denken. Bei einfachen Rechnungen hingegen wie $5 + 7$ oder 2×3 greift das Gehirn auf auswendig gelerntes Faktenwissen zurück; in diesem Fall sind die linkshemisphärischen präfrontalen Gebiete der Sprachverarbeitung aktiv. Auch eine dritte Region, der posterior-superiore Parietallappen, ist bei vielen Aufgaben aktiv, die numerische Manipulationen beinhalten. Er scheint vor allem in Lösungsprozesse von visuell-räumlichen Aufgaben involviert (Kucian & von Aster 2013, S. 65 ff.).

Bis heute sind nicht alle Prozesse der zahlenverarbeitenden Hirnfunktionen erforscht. So scheint es noch immer im Detail unklar, in welcher Weise bestimmte Gehirnareale konkret beteiligt sind, wie u. a. Pinheiro-Chagas und KollegInnen herausstellen: „(...) the precise role of pITG (posterior inferior temporal gyrus) in mathematical cognition remains largely elusive“ (2018, S. 1758).

2.7.1 Das Triple Code Model

Dehaene entwickelte 1992 das *Triple Code Model* als Zahlenverarbeitungsmodell, das neben den auditiv-sprachlichen und visuell-indo-arabischen Repräsentationen von Zahlen (d. h. Zahlwörter und deren indo-arabische Notation) ein drittes analoges Modul postuliert, das für eine mentale Zahlenraumvorstellung steht:

„The triple-code model of number processing predicts that, depending on the task, three distinct systems of representation may be recruited: a quantity system (a nonverbal semantic representation of the size and distance relations between numbers, which may be category specific), a verbal system (where numerals are represented lexically, phonologically, and syntactically, much like any other type of word), and a visual system (in which numbers can be encoded as strings of Arabic numerals)“ (Dehaene 2003, angelehnt an Dehaene 1992; Dehaene & Cohen 1995).

Das analoge Modul erfasst also die Semantik, d. h. die (größenmäßige) Bedeutung einer Zahl. Es fungiert somit als Basis für Größenvergleiche von Zahlen und Mengen, da hier die Zahl auf einem räumlich konfiguriertem inneren Zahlenstrahl repräsentiert wird. Fähigkeiten wie das Einschätzen einer Menge, das Überschlagen von Rechnungsergebnissen und das Vergleichen von Zahlen werden diesem Modul der analogen Mengenrepräsentation zugeschrieben (Kucian & von Aster 2013, S. 65).

2.7.2 Das Vier-Stufen-Modell der Entwicklung zahlenverarbeitender Hirnfunktionen

Ausgehend von Dehaenes *Triple Code Model* entwickelten von Aster & Shalev (2007) das vierstufige Entwicklungsmodell zahlenverarbeitender Hirnfunktionen. Demnach handelt es sich aus entwicklungspsychologischer und neurowissenschaftlicher Sicht bei der Entwicklung spezifisch zahlenverarbeitender Hirnfunktionen um einen stufenweisen neuroplastischen Reifungsprozess, der im Verlauf von Kindheit und Jugend zu einem komplexen, spezialisierten neuronalen Netzwerk führt. Dieser Prozess ist mit der komplementären Reifung domänenübergreifender Denkinstrumente wie Aufmerksamkeit, Arbeitsgedächtnis, Sensomotorik, Sprache und räumlicher Vorstellung verbunden. Die auf Mengen und Zahlen bezogenen Denkfunktionen erfordern also neben verbal-phonologischen und visuell-räumlichen Verarbeitungsprozessen auch Aufmerksamkeits- und Arbeitsgedächtnisleistungen.

Das Vier-Stufen-Modell numerischer Kognition

Verschiedene Studien und Experimente haben gezeigt, dass die Erfassung und Verarbeitung von diskreten Mengen bei Tieren, Babys, Kindern und Erwachsenen vergleichbar sind (van Eimeren & Ansari 2009, S. 27). Die Fähigkeit, Mengen (Quantitäten) zu erfassen, hat auf Grund der Relevanz für die Futtersuche, das soziale Verhalten und für die Reproduktion einer Art eine lange evolutionäre Geschichte (Grond et al. 2013, S. 41). Bereits wenige Monate alte Säuglinge können bei angemessener Stimulation Mengen unterscheiden, was mittels Habituationsstudien nachgewiesen werden konnte. Kleinkinder verfügen über konzeptuelles Wissen zu arithmetischen Prinzipien, ohne dass eine Anzahlerfassung hierfür notwendig ist (Weißhaupt & Peucker 2009, S. 56). Diese sogenannten protoquantitativen Schemata umfassen das Schema des Mengenvergleichs (wobei die Anzahl nicht durch Zählen ermittelt wird), das Schema der Mengenveränderung (Vermehren durch Hinzufügen und Vermindern durch Wegnehmen) sowie das protoquantitative Teil(e)-Ganzes-Schema (Resnick 1992, zitiert *ibid*). Ein grundlegendes Verständnis von Quantitäten ist also angeboren (van Eimeren & Ansari 2009, S. 27).

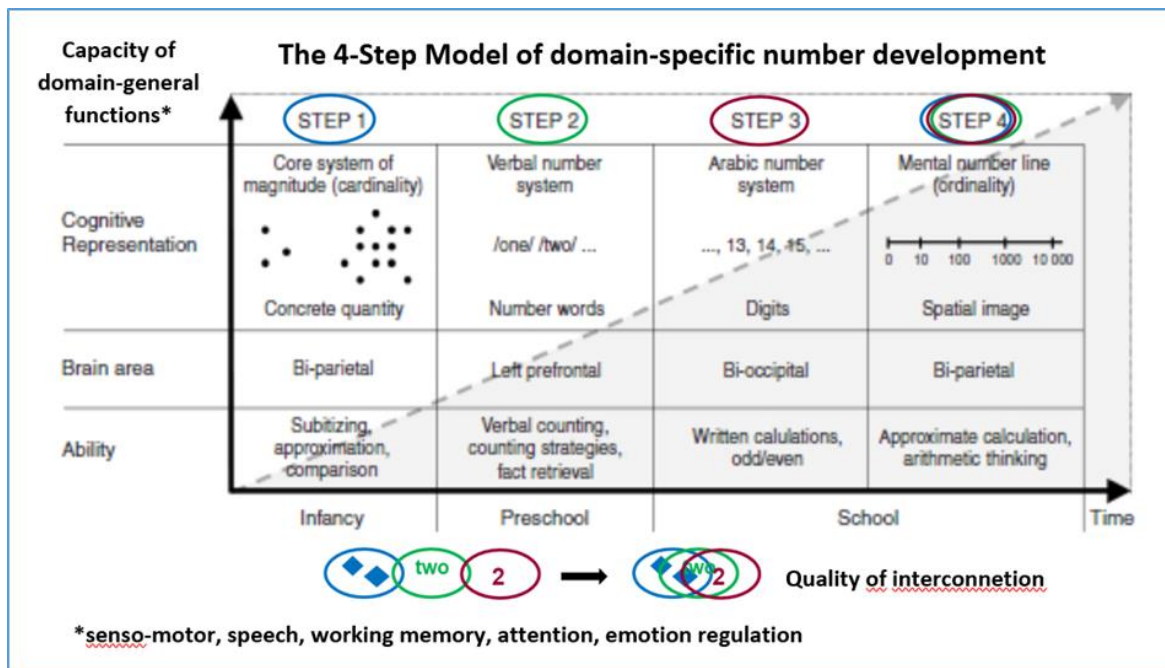


Abb. 1: Das Vier-Stufen-Modell numerischer Kognition (nach von Aster & Shalev 2007)

(Stufe 1) Das Vier-Stufen-Modell (Abb. 1) geht entsprechend von einer angeborenen Grundkompetenz zur Erfassung und Unterscheidung konkreter Mengengröße aus. Dabei können Mengen mit bis zu 4 Items exakt (*subitizing*) und größere Mengen ungefähr eingeschätzt werden.

(Stufe 2) Im frühen Vorschulalter setzt die Fähigkeit zur sprachlichen Symbolisierung von Merkmalen anzahliger Größe ein. Mit Beginn der Sprachentwicklung lernen Kinder die Zahlwortsequenz, Zählprinzipien (z. B. Eins-zu-Eins-Zuordnung, stabile Reihenfolge), das Zu- und Wegzählen zum Verändern von Mengen und Begriffe wie „mehr“ und „weniger“ kennen sowie einfache arithmetische Operationen (Additionen und Subtraktionen) durch Zählstrategien auszuführen. Die vorschulischen Entwicklungsschritte der Zahlenverarbeitung erfolgen ohne systematische Unterrichtung im Kontakt mit dem sozialen und familiären Umfeld und sind eng an den anschaulichen sensomotorischen Gebrauch der Finger gebunden.

(Stufe 3) Mit der einsetzenden systematischen Beschulung wird eine rein visuelle (nicht linguistische) „Zahlensprache“, das indo-arabische Notationssystem, erlernt, welches eine eigene, stellenwertbezogene Syntax besitzt und sich somit im Deutschen stark vom linguistischen Zahlwortsystem unterscheidet. Kinder müssen nun das Transkodieren erlernen, was in Sprachen mit nicht stellenwertgerechter

Zahlensprechweise fehleranfällig ist (Kap. 3). Ein erhöhtes Risiko besteht hier zudem „für Kinder, die zwei- oder mehrsprachig aufwachsen und bei denen es bei der Übertragung aus bzw. in verschiedene Zahlenlinguistiken zu Verwirrung und Unsicherheit kommt“ (von Aster 2017, S. 110).

(Stufe 4) Auf der Grundlage der Stufen 2 und 3 entwickelt sich im Verlauf der ersten Grundschuljahre eine Zahlenraumvorstellung und ein so genannter innerer Zahlenstrahl als Voraussetzung für das mentale Operieren mit Zahlen. Diese geistige Zahlenrepräsentation ist in der Regel durch indo-arabische Landmarken gegliedert.

Die Entwicklung der kognitiven Zahlenrepräsentationen verläuft demnach hierarchisch: Um die Kompetenzen der Stufe 4 zu erreichen, sind die Beherrschung der Zahlwörter und der entsprechenden indo-arabischen Notation sowie das sichere und automatisierte Übersetzen (Transkodieren) zwischen diesen Systemen unabdingbar. Von den sprachgebundenen Risiken für die Entwicklung dieses sicheren Transkodierens handelt die vorliegende Arbeit.

2.7.3 Abweichende Entwicklung zahlenverarbeitender Hirnfunktionen

Die Entwicklung des Zahlenverständnisses kann auch als „Aufbau immer komplexerer Verbindungen zwischen den verschiedenen Bedeutungen der Zahl“ verstanden werden (Weißhaupt & Peucker 2009, S. 52). Bei Kindern mit Dyskalkulie verläuft die Entwicklung zahlenverarbeitender Hirnfunktionen abweichend. Sie unterscheiden sich von Kindern ohne Dyskalkulie in ihrer Hirnentwicklung in verschiedenen Bereichen: Festgestellt werden konnten Defizite im neuronalen Volumen und in der funktionellen Aktivierung im intraparietalen Sulcus sowie in der Entwicklung der neuronalen Verbindungsbahnen zwischen den vorderen (frontalen) und hinteren (parietalen) Hirnabschnitten (Kaufmann et al. 2014; Kucian et al. 2014, zitiert nach von Aster 2017, S. 111). Auch laut Hasselhorn „dominiert derzeit die Auffassung, dass sehr unterschiedliche kognitive Funktionsdefizite eine Rechenstörung auslösen können“ (2022, S. 7).

Empirisch gut abgesichert ist die Annahme, dass ein Defizit basisnumerischer Fertigkeiten, also im basalen Verständnis von Anzahlen und Zahlen,

mitverantwortlich für die Entstehung von Rechenstörungen ist. Diese machen sich schon im Kindergartenalter durch erhebliche Schwierigkeiten bei der Unterscheidung von Mengen, beim raschen Erkennen kleiner Mengen (*subitizing*) sowie im Zahlenverständnis bemerkbar (ibid.).

2.7.4 Störung früher domänenübergreifenden Reifungsprozesse und Risiken auf Grund eingeschränkter kognitiver Funktionen

Als mögliche Ursachen für Schwierigkeiten beim Erwerb arithmetischer Fähigkeiten kommen auch Störungen früher domänenübergreifender Entwicklungs- und Reifungsprozesse, z. B. durch organische pre-, peri- oder postnatale Risiken, in Frage. Weiterhin können ein sehr niedriges Geburtsgewicht oder stressinduzierte epigenetische Fehlregulationen ursächlich sein, was mit Lerndefiziten in anderen kognitiven Domänen und in verhaltensregulierenden Funktionen einhergehen kann (von Aster 2013a, S. 25). Manche Kinder sind von frühestem Kindesalter an irritabel und affektlabil. Sie zeigen neben einer schwachen Aufmerksamkeit auch eine geringe Fähigkeit, Bedürfnisse und Belohnungen aufzuschieben, was in der Folge zu Störungen in der Beziehungsfähigkeit und des Sozialverhaltens und im weiteren Verlauf auch zu depressiven Symptomen führen kann. Auffällig werden diese Kinder in der Schule neben Lernschwierigkeiten in Mathe und Deutsch zunächst wegen ihrer Verhaltensprobleme, nicht selten erhalten sie eine ADHS-Diagnose. Als ätiopathogenetischen Mechanismus für solche Störungsmuster können laut von Aster (2013a, S. 26) heute auch frühkindliche Stresserfahrungen vermutet werden, die sich über epigenetische Wirkfaktoren negativ auf die Entwicklung zentraler domänenübergreifender Funktionen (z. B. Arbeitsgedächtnis, Aufmerksamkeit, Affektregulation) auswirken.

Kinder mit Lernschwierigkeiten haben also häufig nicht nur domänenspezifische zahlenbezogene Defizite, sondern auch Schwierigkeiten im Bereich domänenübergreifender kognitiver Fähigkeiten. Die Letzteren werden für den neuronalen Aufbau der Ersteren benötigt. Die Entwicklung der Kapazität und Regulationsfähigkeit von Aufmerksamkeit, Arbeitsgedächtnis, Emotion, Sensomotorik, Sprache und visuell-räumlicher Verarbeitung kann beeinträchtigt oder verzögert sein (von Aster et al. 2021, S. 11). Die S3-Leitlinie (Kap. 2.2) berichtet zudem von verminderten Inhibitionsleistungen: Demzufolge haben die Kinder Mühe, ablenkende Reize zu unterdrücken, um eine Aufgabe zügig

bearbeiten zu können. Im schulischen Umfeld fallen sie daher nicht nur durch ihre fachlichen Defizite, sondern auch durch geringe Motivation, Anstrengungsbereitschaft und ein herabgesetztes Instruktionsverständnis auf, woraus sich im Zusammenspiel mit dem chronischen Ausbleiben von Erfolgen, herabgesetztem Selbstwert, Stress, Ängsten und Blockaden ein negativer Kreislauf ergeben kann (Kap. 2.3). Die aus der Vermeidung resultierenden Übungsdefizite schwächen schließlich laut von Aster (2013a, S. 35) jene neuroplastischen Hirnreifungsprozesse (Modularisierungen), die für eine erfolgreiche weitere domänenspezifische Lernentwicklung notwendig sind.

2.7.5 Arbeitsgedächtniskapazität und mathematische Leistung

Das Arbeitsgedächtnis ist an sämtlichen höheren kognitiven Leistungen wesentlich beteiligt (Alloway et al. 2009; Baddeley 2012). Die Kapazität des Arbeitsgedächtnisses variiert jedoch stark zwischen Individuen: Wie oben angeführt, sind u. a. Rechenstörungen häufig von einem defizitären Arbeitsgedächtnis begleitet, wobei die S3-Leitlinie von verminderten Inhibitionsleistungen, die der Zentralen Exekutive zuzuordnen sind, berichtet² (s. auch Kap. 2.7.4). Eine reduzierte verbale Merkspanne wiederum kann zu Minderleistungen beim Rechnen führen: Wenn es Defizite in der phonologischen Schleife gibt, können sprachliche Einheiten, z. B. mehrere Zahlwörter oder Zahlwortreihen, nicht im verbalen Kurzzeitspeicher zur Bearbeitung gehalten werden. Bei der Phonologischen Schleife zeigt sich allerdings nur eine geringe Effektstärke, während beim visuell-räumlichen Notizblock ein signifikant hoher Effekt festzustellen ist (S3-Leitlinie Rechenstörung, S. 15).

Eine defizitäre visuell-räumliche Verarbeitung bedeutet, dass z. B. das Merken einer räumlichen Sequenz nicht gelingt (von Aster et al. 2021, S. 12). Konkret erschwert

²Für die Darstellung des Zusammenhangs zwischen defizitärem Arbeitsgedächtnis und mathematischen Leistungen legen wir das *working memory model* von Baddeley (2012, S. 6; erstmals veröffentlicht von Baddeley & Hitch 1974) zu Grunde, welches von einer Phonologischen Schleife, einem visuell-räumlichen Notizblock sowie der Zentralen Exekutive (auch Zentrales Kontrollsystem genannt) ausgeht.

stellen sich Aufgaben wie beispielsweise das Platzieren von Zahlen auf einem Zahlenstrahl dar, da hier die Verknüpfung von numerischem Wissen und räumlichen Leistungen erforderlich ist. Die inverse Zahlensprechweise erschwert diese Aufgabe in besonderem Maße. Helmreich et al. (2011) ermittelten, dass die Ergebnisse deutschsprachiger Kinder bei der Angabe von Zahlen auf einem leeren Zahlenstrahl signifikant schlechter ausfielen als diejenigen von italienischsprachigen Kindern, die die Zahlen stellenwertgerecht aussprechen. Auch beim Größenvergleich von zweistelligen Zahlen wurde eine erhöhte Fehleranfälligkeit bei Sprachen mit verdrehter Zahlensprechweise nachgewiesen (Pixner et al. 2011). Es konnte überdies gezeigt werden, dass die Fehler häufig auf Zahlendreher zurückzuführen sind. Die inverse Zahlensprechweise behindert somit das Erfassen von Zahlengrößen, das grundlegend für den Erwerb weiterer arithmetischer Kompetenzen ist. Auch Zuber et al. (2009) konnten in ihrer Studie nachweisen, dass invers strukturierte Zahlwörter den Erwerb numerischer Kompetenz erschweren können. Entsprechend dem Modell von Baddeley zeigten sich größere Nachteile für Kinder mit geringerer Arbeitsspeicherkapazität. Diese Ergebnisse bestätigen van der Ven et al. (2017), die in einer breit angelegten niederländischen Studie zur inversen Zahlensprechweise mit 25 620 teilnehmenden Kindern eine hohe Korrelation zwischen Arbeitsspeicherkapazität (visuell-räumlicher Speicher und Zentrales Kontrollsystem) und mathematischer Leistungsfähigkeit nachwiesen. Diese Beziehung wird durch die individuelle Transkodierfähigkeit moduliert, d. h., dass bei gleicher Arbeitsspeicherkapazität eine höhere Fehlerrate in den Transkodieraufgaben mit einer niedrigeren mathematischen Leistungsfähigkeit einherging. Jede syntaktische Regel zur korrekten Transkodierung belastet die Arbeitsspeicherkapazität zusätzlich und kann zur Überlastung und somit fehlerbehaftetem Transkodieren führen (s. auch Kap. 3.3).

3 ASPEKTE DER ZAHLENLINGUISTIK

3.1 DIE ENTWICKLUNG UND VERBREITUNG DES DEZIMALEN STELLENWERTSYSTEMS

Länderübergreifend hat sich in der Neuzeit die Darstellung von Zahlen im Dezimalsystem durchgesetzt. Es zeichnet sich durch ein durchgängiges Konstruktionsprinzip aus und hat sich – wie anzunehmen ist – im Vergleich zu anderen Systemen als praktikabler vor allem im Umgang mit größeren Anzahlen erwiesen, und zwar auch in linksläufigen Sprachen wie dem Hebräischen und Arabischen (Morfeld & Gerritzen 2024, S. 5). Anders als das römische Zahlssystem oder die seit Ende des 4. Jahrtausends v. u. Z. gebräuchliche ägyptische Hieroglyphen-Zahlschrift ermöglicht es eine sehr ökonomische visuelle Symbolisierung auch sehr großer Zahlen sowie ein vorteilhaftes und bequemes Operieren mit Größen (Schwank 2013, S. 95).

Seine vollständige Entwicklung erfuhr das Stellenwertsystem zur Basis 10 (= Dezimalsystem) spätestens vor 1500 Jahren in Indien (Ifrah 1991; zitiert nach Morfeld & Gerritzen 2024, S. 7). Um 825 u. Z. erarbeitete al-Ḥwārizmī in Bagdad, wie man die Grundrechenarten mit den indischen Ziffernzahlen ausführen kann. Diese Verfahren wurden in Anlehnung an seinen Namen später Algorithmen genannt (Morfeld & Gerritzen 2024, S. 1 und 7).

In Europa und dem deutschen Sprachraum, wo das römische Zahlssystem genutzt wurde, stand man seitens der Autoritäten den indo-arabischen Zahlen zunächst ablehnend gegenüber. Vor allem die römisch-katholische Kirche versuchte über Jahrzehnte die Nutzung zu unterbinden, konnte aber aus Gründen der alltagstauglichen Praktikabilität vor allem im Handels- und Finanzsektor die letztendliche Verbreitung nicht verhindern. Die indo-arabische Zahldarstellung sichert ihre semantische Eindeutigkeit durch ihre sprachübergreifende, universal gültige Schreibregel ab: rechtsläufig, also von links nach rechts, und mit abfallenden Potenzen, also z. B. Hunderter-Zehner-Einer-Zehntel-Hundertstel (ibid., S. 5). Damit kann das dezimale Stellenwertsystem als Verwirklichung einer globalen Sprache gelten (Barrow 2005, S. 148, zitiert ibid.) und verdient laut Antrag des Zwanzigeins e. V. die Anerkennung als immaterielles Kulturerbe der UNESCO (2017).

3.2 VERSCHIEDENE SPRACHRÄUME UND IHR UMGANG MIT ZAHLWORTSYSTEMEN

Die Zahlensprechweise im Deutschen bildet dieses weltweit als logisch und alltagstauglich wahrgenommene dekadische Stellenwertsystem jedoch nicht strukturiert ab. Dem Phänomen der globalen Verbreitung des indo-arabischen Zahlensystems und der länderübergreifend einheitlichen Notation stehen paradoxerweise die uneinheitlichen Sprechregeln gegenüber. Diese haben sich in den verschiedenen Sprachräumen unterschiedlich entwickelt: In einigen Sprachen, wie z. B. dem Chinesischen, Japanischen, Italienischen und Spanischen korrespondiert die Sprechweise mit der universal gültigen indo-arabischen Schreibregel; in anderen Sprachen entwickelten sich insofern irreguläre Sprechweisen, als sie dem durchgängigen Konstruktionsprinzip des Dezimalsystems nicht gerecht werden – so (neben dem Deutschen) unter anderem im Niederländischen, Dänischen, Arabischen und im Englischen von 11 bis 19: „Between ten and twenty, English is a mess“ (Bellos 2010, S. 64). Die verbalen Zahlensysteme korrespondieren in den genannten Sprachen nur teilweise mit der Notation des indo-arabischen Zahlensystems.

Die Verwendung dieser Sprechweisen wurde bereits in der Vergangenheit vor allem von Pädagogen verschiedentlich kritisiert: Da die französische Umgangssprache noch ungeeignete Zahlwörter zur Verfügung stellt als die deutsche, wurde beispielsweise in den „Instructions Officielles“ bereits 1923 der Vorschlag unterbreitet, im Arithmetikunterricht eine Fachsprache mit logisch aufgebauten Zahlnamen einzuführen, „um durch die gegebenen Anomalien der umgangssprachlichen Zahlnamen nicht unnötig behindert zu werden“ (nach Letierrier 1956, zitiert nach Schwank 2013, S. 131).

Dass sich diese Idee nicht durchgesetzt hat, verwundert insofern, als Kinder in Ländern mit stellenwertgerechter Sprechweise nachweislich in der Grundschule bessere arithmetische Fertigkeiten entwickeln, wie u. a. die TIMSS-Studie 2015 zeigt (Wendt et al. 2016). Andere Studien kommen zu ähnlichen Ergebnissen: Chinesische Kinder sind im Vergleich zu amerikanischen beim Zahlen- und Rechnenlernen im Vorteil, was unter anderem auf die für diese Zwecke unterschiedlich gut geeigneten Zahlwörter zurückzuführen ist (Miller et al. 1995, Azar 1999; beide zitiert nach Schwank 2013, S. 131). Ng & Rao kommen bei ihrer Auswertung verschiedener Studien, die die mathematischen Leistungen von

ostasiatischen (d. h. chinesischen, koreanischen und japanischen) vs. Englisch sprechenden Kindern vergleichen, zum gleichen Schluss: „Chinese number words are described as being clear, transparent, straightforward, and conceptually well designed, as well as a logical representation of the base-ten system. In contrast, the English number system is nonsystematic, has several exceptions, is relatively poorly designed for counting, and is perceived to be harder to learn than Chinese“ (Ng & Rao 2010, S. 182).

Eine in den USA und China durchgeführte Längsschnittstudie legt den Schluss nahe, dass chinesische Kinder bereits im Alter von 3-4 Jahren einen klaren Vorteil beim Zählen und bei der Bildung von Zahlwörtern haben: „By the age of 4, the Chinese children were progressing rapidly in generalizing number names up to 100 after being able to recite numbers up to approximately 40. Fewer American children were able to master the rules for generating larger numbers in counting to 100“ (Miller et al. 2005, zitiert *ibid.*, S. 185). Auch die Länge der Zahlwörter scheint eine Rolle zu spielen: Diese sind im Chinesischen einsilbig und belasten so das Arbeitsgedächtnis weniger als zwei- und mehrsilbige Zahlwörter. Dies ist nicht nur beim Erlernen der Zahlwörter, sondern vor allem bei dem Erwerb von anspruchsvolleren Zähl- und Rechenstrategien ein Vorteil: „The lighter memory load associated with Chinese number names helps Chinese-speaking children to proceed at a faster rate from concrete object counting strategies to more advanced abstract strategies, compared to non-Chinese speakers“ (*ibid.*, S. 185 f.).³

Die Annahme, dass die stellenwertgerechte Zahlensprechweise das Transkodieren erleichtert, konnten Pixner et al. (2011) in einer Studie mit Kindern in Tschechien nachweisen, wo nach einer Reform neben der inversen auch eine stellenwertgerechte Sprechweise zugelassen ist. Das Forschungsteam fand, dass sich die Inversionsfehler bei Schulkindern häufen, wenn die Zahlen in dem Zahlwortsystem mit Inversion präsentiert werden. Auch in anderen Ländern haben

³ Die Autoren verweisen allerdings einschränkend darauf, dass der klare Vorteil des chinesischen Zahlwortsystems vor allem in den frühen Jahren des Erwerbs arithmetischer Fähigkeiten zum Tragen kommt. Die in der späteren Schullaufbahn vergleichsweise schwächeren Mathematikleistungen nicht-chinesisch sprechender Kinder führen sie neben den linguistisch begründeten Nachteilen auch auf andere Faktoren wie den Lehrplan, Unterrichtsmethoden und den Einsatz von Strategien zurück.

sich Reformbemühungen zugunsten einer stellenwertgerechten Sprechweise durchsetzen können: In Norwegen ist man der inversen Zahlensprechweise bereits 1951 offensiv durch einen parlamentarischen, einstimmig angenommenen Beschluss begegnet, der eine geänderte, stellenwertgerechte Sprechweise im Rahmen des Schulunterrichts vorsah. Diese Sprechweise hat sich heute weitgehend auch in der norwegischen Alltagssprache durchgesetzt (Voigt 2008, S. 106 ff.).

3.3 AUSWIRKUNGEN DER INVERTIERTEN ZAHLENSPRECHWEISE

Voraussetzung für das Rechnen (d. h. das Operieren mit Zahlen) ist die Entwicklung einer tragfähigen Zahlraumvorstellung, die wiederum auf dem automatisierten Übersetzen (Transkodieren) zwischen der verbalen und der indo-arabischen Zahlenrepräsentation basiert, wie das Vier-Stufen-Modell der Entwicklung zahlenverarbeitender Hirnfunktionen zeigt (Kap. 2.7.2). Verschiedene Studien belegen den Zusammenhang zwischen Transkodierfähigkeit und mathematischer Leistung: Im Vergleich zu normal entwickelten Kindern machen rechenschwache Kinder mehr Transkodierfehler (Moura et al. 2013) und sind langsamer (van Loosbroek, Dirx, Hulstijn & Janssen 2009), während Kinder mit guten Transkodierfähigkeiten bessere Mathematiknoten erzielten (Imbo, van den Bulcke & Fias 2014; alle zitiert nach van der Ven 2017, S. 488, 500). Einen statistisch signifikanten Zusammenhang zwischen der Transkodierfähigkeit im ersten Schuljahr und der mathematischen Leistungsfähigkeit und den Schulnoten in Mathematik im dritten Schuljahr stellten Möller et al. (2011) fest.

Beim Transkodieren können zwei Arten von Regeln unterschieden werden: (i) lexikalische Regeln, d. h. die Entsprechung zwischen Zahlwörtern und den Zahlensymbolen (z. B. die deutsche Zahlwortbezeichnung „acht-und-vierzig“ und die entsprechende Übertragung in die indo-arabische Notation 48) und (ii) syntaktische Vorgaben, die die Bildung von mehrstelligen Zahlen regeln (Power & Dal Martello 1990). Das Erlernen dieser Regeln ist ein langwieriger Prozess. Transkodierfehler in den Klassen 1 und 2 treten häufig auf (Zuber et al. 2009), wobei lexikalische Fehler die geringste Fehlerquote aufweisen (Power & Dal Martello 1990). 25% der Fehler, die deutschen Zweitklässlern beim Schreiben von zweistelligen Zahlen unterlaufen, sind sogenannte Zahlendreher – französische

Kinder kennen diese Art von Fehlern praktisch nicht (Koumoura et al. 2004; von Aster et al. 1997, zitiert nach von Aster 2013, S. 29). Auch Nürk et al. (2005) stellen die Häufigkeit von Inversionsfehlern bei deutschen Erstklässlern im Vergleich zu japanischen Kindern heraus: „German-speaking 1st graders not only committed significantly more transcoding errors in general as compared to Japanese-speaking children but more inversion-related errors in particular“ (zitiert nach Klein et al. 2013, S. 2).

Viele Kinder behelfen sich beim Transkodieren mit einer Inversion der Schreibrichtung. Sie notieren in Umkehrung der üblichen Schreibrichtung zunächst die Einer und anschließend die Zehner links vom Einer, was der Sprechweise entspricht. Obschon dies kurzfristig hilfreich sein kann, können beim Schreiben mehrstelliger Zahlen – da man eine oder mehrere Lücken lassen muss – sowie bei der Nutzung numerischer Tastaturen wie z. B. dem Taschenrechner oder dem Computer Probleme auftreten (Padberg & Benz 2021, S. 83). Daher wird vielfach empfohlen, auf eine stellenwertgerechte Schreibrichtung Einfluss zu nehmen, was „wenn es rechtzeitig geschieht, leicht zu bekämpfen“ sei (ibid.). Daraus ergibt sich allerdings beim Mitdenken oder lauten Mitsprechen z. B. der Zahl 21 das Paradoxon: „Ich spreche *einund* und schreibe 2, ich spreche *zwanzig* und schreibe 1.“ Für Kinder mit nichtdeutscher Erstsprache und / oder Rechenstörung kann dies vor allem bei einer reduzierten Arbeitsgedächtnisleistung, wie dies bei Letztgenannten häufig der Fall ist (Kap. 2.7.5), oder auch bei Links-Rechts-Problematiken das Transkodieren und die Erschließung des dezimalen Stellenwertsystems zusätzlich erschweren. Kontraproduktiv ist darüber hinaus, dass in der Fachdidaktik häufig „in eine Lerneinheit gepackt (wird), was drei durch stoffdidaktische Analyse isolierbare, sachlogisch aufeinander aufbauende Lerninhalte ausmacht:

- Zehner als Bündelung von zehn Einern und zweistellige Zahlen als Zusammensetzungen aus Zehnern und Einern verstehen (Bündelungsprinzip);
- verstehen, dass die Anzahlen von Zehnern und Einern mit denselben Zeichen (Ziffern) festgehalten werden; dass außerhalb einer Stellentabelle nur die Position des Zeichens anzeigt, ob Zehner oder Einer damit bezeichnet werden; dabei zur Kenntnis nehmen und sich früher oder später merken, dass Zehner links von den Einern notiert werden (Positionsprinzip);
- schließlich: die Idiotien der Zahlwortbildung in der deutschen Sprache durchschauen“ (Gaidoschik 2015, S. 173).

Dieses Vorgehen kann problematisch sein und zu einer Häufung von Inversionsfehlern führen. Van der Ven et al. (2017, S. 499) weisen bis zum Ende der Grundschulzeit Inversionsfehler nach. Im 2. Schuljahr ist die Inversionsfehlerrate am höchsten und sinkt dann zwar kontinuierlich, aber nicht auf Null, so dass auch in der 6. Klasse 1% der Kinder noch Inversionsfehler machen.

In Bezug auf (deutschsprachige) rechenschwache Kinder stellen u. a. Lorenz und Schwank (nach Schwank 2013, S. 130 f.) einen weiteren Aspekt heraus: Sie unterscheiden zwischen (i) Kindern mit unterdurchschnittlichen Fähigkeiten in Mathematik, aber noch geringerer Leistung im Schriftspracherwerb, und (ii) Kindern ohne komorbide Störung im Bereich Rechtschreiben. Die zweite Gruppe sei vergleichsweise sprachstark, habe aber ein geringes Verständnis der mathematischen Algorithmen auf Grund einer Störung des Prozesses der nichtsprachlichen Begriffsbildung; allerdings können laut Schwank diese Kinder ihre Stärke bei sprachlichen Begriffsbildungsprozessen in der Arithmetik gar nicht ausspielen, da die deutsche Sprache keine exakt zum arithmetischen Betrieb passenden Zahlwörter zur Verfügung stelle.

Dass Kinder der Gruppe (i), also solche mit Rechenstörung und LRS, vor noch größeren Herausforderungen bei der Erschließung des dezimalen Stellenwertsystems stehen, liegt auf der Hand. Die Prävalenzen zeigen, dass es sich keinesfalls um Einzelfälle handelt – häufig leiden Kinder unter Rechen- und Lese-Rechtschreibstörungen (Kap. 2.4, 2.5). Betroffen sind neben dieser Gruppe auch Kinder mit Migrationshintergrund: Entweder, weil sprachliche Defizite im Deutschen bestehen, oder aber, weil in ihren Erstsprachen (z. B. im Türkischen, Italienischen oder Spanischen) keine Abweichungen in der Reihenfolge von Sprech- und Schreibweise existieren (Gottbrath 1984, zitiert nach Padberg & Benz 2021, S. 82 f.).

Es ist zusammenfassend festzuhalten, dass die Zahleninversion negative Auswirkungen auf die Transkodierfähigkeit und die mathematischen Leistungen hat. Das bei rechenschwachen Kindern ohnehin defizitäre Arbeitsgedächtnis wird durch die inverse Zahlensprechweise zusätzlich belastet. Die jüngsten Studien mit der Zwanzigeins-App (Schmid 2023; Kuhl 2025 in Zusammenschau mit vorliegender Arbeit; Summer et al. in Vorb.) bestätigen den Effekt, dass die Verwendung der inversen Zahlensprechweise beim Transkodieren zu einer erhöhten Fehleranfälligkeit führt, während die Verwendung einer stellenwertgerechten

Sprechweise nicht nur zur Verringerung der Fehlerzahl, sondern gleichzeitig zur Reduzierung der Verarbeitungsdauer führt. Der durch die Verwendung der Zwanzigeins-App ermöglichte direkte Vergleich („within“-Vergleich, Kap. 4.3) ermöglicht den Rückschluss, dass die Fehleranfälligkeit des Transkodierens auf die nicht stellenwertgerechte deutsche Zahlsprechweise zurückzuführen ist: „Die inverse (verdrehte) Sprechweise führt zu einer signifikant erhöhten Fehlerhäufigkeit beim Transkodieren“ (Morfeld & Summer 2024, S. 1). Dieser Effekt soll im folgenden empirischen Teil in Bezug auf eine Untersuchungsgruppe rechenschwacher Kinder untersucht werden.

II EMPIRISCHER TEIL

In einer österreichischen Studie, die im Jahr 2023 mit 55 Kindern der 2. Klassenstufe durchgeführt wurde (Schmid 2023), konnte ein klarer Vorteil für eine stellenwertgerechte Sprechweise nachgewiesen werden: Beim Transkodieren eines Zahlendiktats verringerten sich sowohl die Eingabedauer als auch die Fehlerzahl signifikant bei Verwendung einer stellenwertgerechten Zahlensprechweise im Vergleich zu der im deutschen Sprachraum gebräuchlichen inversen Zahlensprechweise. Die stellenwertgerechte Sprechweise war den Kindern bis zum Zeitpunkt der Testung unbekannt; sie erhielten lediglich im Vorfeld eine kurze Einweisung.

Diese Ergebnisse führten zu der Annahme, dass auch Kinder mit Rechenschwierigkeiten von einer stellenwertgerechten Sprechweise profitieren könnten. Diese Hypothese soll in diesem zweiten Teil der Arbeit durch eine empirische Untersuchung mit 43 rechenschwachen Kindern untersucht werden. Es stellen sich folgende Fragen, die jeweils im Vergleich zur traditionellen inversen Sprechweise beantwortet werden sollen:

- 1) Wird die Eingabe von Zahlen in der Zwanzigeins-App bei stellenwertgerechter Sprechweise schneller realisiert?
- 2) Werden bei stellenwertgerechter Sprechweise weniger Fehler gemacht?
- 3) Gibt es mehr fehlerfreie Durchgänge bei stellenwertgerechter Sprechweise?

Es soll zudem überprüft werden, ob weitere Faktoren wie z. B. Alter, Geschlecht oder die Erstsprache(n) der Kinder einen modifizierenden Einfluss auf die Ergebnisse haben.

Die Arbeit von Anna Kuhl (2025) beschäftigt sich mit der Eingabedauer der Zahlen, die vorliegende Arbeit mit der Fehlerzahl und der Anzahl fehlerfreier Durchgänge. Im Folgenden werden das Studiendesign einschließlich der verwendeten Hard- und Software sowie der Rekrutierungsprozess erläutert, bevor die Erhebung der Daten und die Untersuchungsergebnisse in Kapitel 5 detailliert dargestellt werden.

Diese werden abschließend diskutiert, wobei Argumente für und gegen eine stellenwertgerechte Sprechweise sowie Reflexionen der Kinder (Kap. 6), die jeweils nach den Testungen dokumentiert wurden, berücksichtigt werden.

4 ERHEBUNGSINSTRUMENTE UND STUDIENDESIGN

Das Forschungsdesign (= Studiendesign) ist für die Aussagekraft der erzielten Ergebnisse von entscheidender Bedeutung, da es festlegt, wie eine theoretische Forschungsfrage konkret in eine empirische Untersuchung umgesetzt werden soll (Stein 2022, S. 135). Die Entwicklung des Designs der vorliegenden Studie, das auch für weitere empirische Forschungen mit der Zwanzigeins-App zum Einsatz kommt, erfolgte durch PD Dr. Peter Morfeld, dessen Forschungspraxis sowie fundierte Statistik- und Methodenkenntnisse unerlässlich für die Durchführung der empirischen Untersuchung und die Auswertung waren. Nach dem Abschluss der Erhebung wurden ihm die Daten aller getesteten Personen und Durchgänge pseudonymisiert und in adäquater Form (Baur & Blasius 2022b, S. 17) zur statistischen Auswertung überstellt. Die Analyse der Daten erfolgte mit dem Programmpaket Stata 14 (StataCorp 2015). Aufgrund einer Powerberechnung auf Basis der Vorerfahrungen durch Schmid (2023) ergab sich, dass zur Beantwortung der zentralen Forschungsfragen 1) bis 3) mindestens 40 Schulkinder getestet werden sollten.

Da bereits bei der Datenerhebung die Datenauswertung berücksichtigt und geplant werden muss (Baur & Blasius 2022b, S. 17), wurden im Vorfeld der Durchführung mit Prof. Dr. Michael von Aster und Anna Kuhl mögliche Einflussfaktoren diskutiert und gemeinsam Entscheidungen zur Datenauswahl getroffen, was zur Erhebung von Kovariablen (Kap. 4.4) sowie zu einer erweiterten Hypothesenbildung führte (Kap. 4.5). Die Interpretation und Diskussion der mittels der statistischen Auswertung gewonnenen Ergebnisse und Erkenntnisse erfolgte in einer gemeinsamen Videokonferenz am 12. Juli 2024.

4.1 EINGESETZTE HARDWARE & BERÜCKSICHTIGUNG DER UNTERSCHIEDLICHEN ANSAGEDAUER

Für die Testungen des Gesamtkollektivs, also aller 43 Kinder der zwei Studienkollektive (Kap. 4.7), kamen verschiedene Geräte zum Einsatz.

Anna Kuhl testete 4 Kinder des Studienkollektivs 1 mit einem iPhone mit Android 16.7 und Safari-Browser und alle anderen Kinder mit einem Lenovo Desktop

PCFHJVK im Tablet-Mode mit Windows 11 und Edge-Browser. Die Verfasserin führte die Tests des Studienkollektivs 2 mit Hilfe eines Tablets PEAQ PET 101-H232E mit Android 12 und Chrome-Browser bei allen Kindern durch.

Die Dauer der Ansage der diktierten Zahlen unterscheidet sich zum einen je nach Sprechweise und zum anderen in Abhängigkeit von der genutzten Hardware.

Der Sprechaufwand und damit die Ansagezeit ist bei Tv im Vergleich zu Ze leicht erhöht: Bei Tv tritt fast immer ein „und“ auf, welches bei Ze entfällt, z. B. „einundzwanzig“ vs. „zwanzigeins.“ Da es nur wenige Ausnahmen gibt, auf die dies nicht zutrifft, sind die Ansagedauern für Ze im Durchschnitt etwas kürzer als für Tv. Die inverse Sprechweise für die Zahlen 11 bis 99 benötigt im Mittel 1.7 s zusätzlich pro zehn Zahlwörter im Vergleich zur stellenwertgerechten Sprechweise.

Um zu verifizieren, wie sich die Ansagedauern von Ze und Tv auf der verwendeten Hardware exakt unterscheiden, wurde dies auf den oben aufgeführten Geräten gemessen (detaillierte Darstellung: Kuhl 2025). Die Unterschiede bei der Ansagedauer wurden in der Auswertung berücksichtigt. Zudem sicherte dieser vorausgehende Check, dass die Geräte nicht entgegen der Erwartung abweichend bei der späteren Testung der Kinder arbeiteten.

4.2 EINGESETZTE SOFTWARE UND APP-PARAMETERFESTLEGUNGEN

Zur Durchführung der Erhebung wurde die auf der Website des Vereins Zwanzigeins e. V. oder im *google play store* frei verfügbare Zwanzigeins-App auf das oben genannte Gerät installiert, so dass etwaige Performanceschwankungen des Internets keinen Einfluss auf die Studie haben konnten; die App lief zuverlässig ohne Unterbrechung. Sie steht einerseits jeder interessierten Person zum Ausprobieren (und bei Bedarf zum kostenlosen Download) zur Verfügung und dient andererseits wissenschaftlichen Zwecken, um die Zwanzigeins-Hypothese weiter zu erforschen.

Die App diktiert Zahlen, die in die Tastatur einzutippen sind; es handelt sich also um eine Transkodierungsaufgabe: Mündlich präsentierte Zahlen werden unter Verwendung der indo-arabischen Notation transkribiert. Dabei stellt die App pro Durchgang von jeweils 10 diktierten Zahlen zwischen 11 und 99 die Eingabedauer und die Zahl der bei der Eingabe unterlaufenen Fehler fest. Die Daten werden –

sofern sie nicht explizit gelöscht werden – gespeichert und stehen unter der Funktion *Statistik* für die weitere Verwendung zur Verfügung. Somit konnte die Datensammlung für die vorliegende Studie als csv-Datei exportiert, in eine Tabelle umgewandelt und mit weiteren Informationen als möglichen Einflussfaktoren (= Kovariablen, Kap. 4.4.2) in eine statistisch auswertungsfähige Form gebracht werden.

Die App verfügt über verschiedene frei wählbare Modi. Für die vorliegende Studie wurden im Vorfeld folgende Parameter festgelegt:

- 1) Aufgabentyp „Hören und Schreiben“
- 2) Level „Leicht“
- 3) Sprechweisen „traditionell verdreht“ im Wechsel mit „zehneins“
- 4) Sprechgeschwindigkeit „100%“
- 5) Darstellung „hell“

Vor der Testung jedes neuen Kindes musste darauf geachtet werden, dass die Sprechweise entsprechend der Randomisierungsliste (Anhang 3) im Menü „Hören und Schreiben“ eingestellt wurde, da die Kinder gemäß Studiendesign „hälftig über kreuz“ wahlweise entweder mit Tv oder Ze starten sollten (Kap. 4.3). Um auswertungsfähige Ergebnisse zu erhalten, war es weiterhin wichtig, während der Testung diese Sprechweise entsprechend dem Studienplan umzustellen.

Des Weiteren wurde für jedes Kind ein neues Profil angelegt. Der dreistellige Profilname stellte dabei gleichzeitig das Pseudonym für das getestete Kind dar. Pseudonyme, die mit der Ziffer 1 beginnen, sind dem Studienkollektiv 1 (Untersucherin: Anna Kuhl) zuzuordnen, also 101 bis 121. Die Pseudonyme der Kinder des Studienkollektivs 2 beginnen entsprechend mit 2, also 201 bis 221. Mit insgesamt 3 Kindern konnte entweder die Testung nicht abschließend durchgeführt werden oder die Teilnahme wurde vollständig verweigert. Da keine vollständigen Datensätze erstellt wurden, konnten diese Ergebnisse nicht in die Auswertung einbezogen werden; dies bedeutet, dass 40 statistisch verwertbare Datensätze zur Auswertung vorlagen.

Eine Umsteigeliste mit jeweils den Profil- und Klarnamen der Kinder wurde von den Untersucherinnen vorgehalten, um etwaigen Rückfragen zu Auffälligkeiten, die erst während der Auswertung deutlich hätten werden können, begegnen oder Zusatzfragen beantworten zu können. Die Übergabe zur Auswertung erfolgte, wie

oben dargelegt, pseudonymisiert, um den Schutz der persönlichen Daten der Teilnehmenden zu gewähren.

4.3 GRUNDSTRUKTUR UND STUDIENDESIGN

Die vorliegende Studie wurde als eine experimentelle Längsschnittstudie konzipiert; das Studiendesign kann als „randomized controlled panel⁴ study with cross-over“ (Younge et al. 2015), also *balanciertes randomisiertes kontrolliertes Experiment mit Crossover* bezeichnet werden und soll im Folgenden erläutert werden.

Als experimentell (Experimentalstudien) werden Studien bezeichnet, bei denen die Exposition pro Periode festgelegt wird, d. h. in unserem Fall ist die Sprechweise (= Exposition) pro Durchgang (= Periode) mit Ze (stellenwertgerecht) oder Tv (traditionell verdreht) festgelegt. Diese Sprechweisen sind also unsere unabhängigen Variablen, da sie laut Hypothese (Kap. 4.5) Einfluss auf Dauer (= Geschwindigkeit) und Fehlerhäufigkeit haben. Zusätzlich wurde der Einfluss von Kovariablen (Kap. 4.4.2) überprüft.

Längsschnitt bezieht sich bei der vorliegenden Studie darauf, dass die Kinder (= Beobachtungseinheiten) im Gegensatz zu einer Querschnittstudie mehrfach beobachtet wurden: Es fand also kein Vergleich zwischen den Beobachtungseinheiten statt („between“-Vergleich), sondern jede Beobachtungseinheit (= Kind) realisierte die verschiedenen Sprechweisen („within“-Vergleich). Die Expositionen für ein und dasselbe Kind erfolgten also mehrfach, jedes Kind realisierte jeweils 2 Perioden mit Ze und 2 Perioden mit Tv und fungiert somit als seine eigene Referenz. Daraus ergibt sich gegenüber Querschnittstudien der Vorteil, dass die individuellen Variablen, die der Beobachtungseinheit (= Kind) immanent sind, keine Störgrößen darstellen (Pforr / Schröder 2015). Konstante individuelle Variablen wie Alter, Geschlecht, der soziale Status der Eltern, der kulturelle Hintergrund, das pädagogische Konzept der Schule, die Persönlichkeit der Mathematiklehrkraft, der Grad der pädagogischen Unterstützung durch die

⁴ Panelstudien sind im Unterschied zu Zeitreihen dadurch gekennzeichnet, dass die Zahl der Perioden kleiner ist als die Zahl der Beobachtungseinheiten, was zutreffend für unsere Studie ist: Jedes der 40 Kinder (= Beobachtungseinheiten) absolvierte 4 Durchgänge (= Perioden).

Eltern, die Anzahl der Geschwister, der durchschnittliche Unterrichtsausfall, der Sitzplatz in der Klasse etc. können bei dem vorliegenden Design der Studie keinen Einfluss auf die Ergebnisse haben.

Mit jedem Kind wurden vier Durchgänge gemäß den festgelegten Parametern (Kap. 4.2) realisiert. Pro Durchgang wurden dem jeweiligen Kind zehn zufällig generierte Zahlen aus dem Bereich 11 bis 99 diktiert. Unsere unabhängige Variable – die Sprechweise – wurde dabei als vermutete Einflussgröße planmäßig variiert, um die Studie im Mittel vor Lerneffekten als Störvariable (Kap. 4.3.1) zu schützen.

Die systematische Variation „hälftig über Kreuz“ (= *cross over*) wurde wie folgt umgesetzt: Es wurden zwei unterschiedliche Sequenzen festgelegt, die randomisiert an die Kinder bzw. deren Profile vergeben wurden. Die Hälfte der Kinder bekamen per Münzwurf die erste Sequenz, die andere Hälfte die zweite Sequenz zugewiesen. Bei den der ersten Sequenz zugeordneten Kindern wurden jeweils ein Probe- und ein Messdurchgang mit Tv, im Anschluss jeweils ein Probe- und ein Messdurchgang mit Ze realisiert, d. h. die Abfolge war TvTvZeZe. Die der zweiten Sequenz zugewiesenen Kinder starteten mit Ze, d. h. die Abfolge war ZeZeTvTv. Die von den Kindern zu realisierende Sequenz wurde unter ihrem Pseudonym (Profilname) in der Randomisierungsliste (Anhang 3) festgehalten.

4.3.1 Lerneffekte als Störgrößen

Wie oben beschrieben, war es aus methodischer Sicht wichtig, unsere Studie im Mittel gegen Lerneffekte als Störgrößen zu schützen. Daher wurde zum einen die Sequenz balanciert variiert, zum anderen wurden lediglich die Messdurchgänge (2. und 4. Periode = Testdaten) in die Hauptanalyse einbezogen. Der Vergleich zwischen der Auswertung der Testdaten vs. der Auswertung aller Daten (1. - 4. Periode) belegt die Wirksamkeit dieses Designs und soll im Folgenden dargestellt werden.

Die Auswertung aller Perioden – bezogen auf die Eingabedauer – zeigt den Lerneffekt:

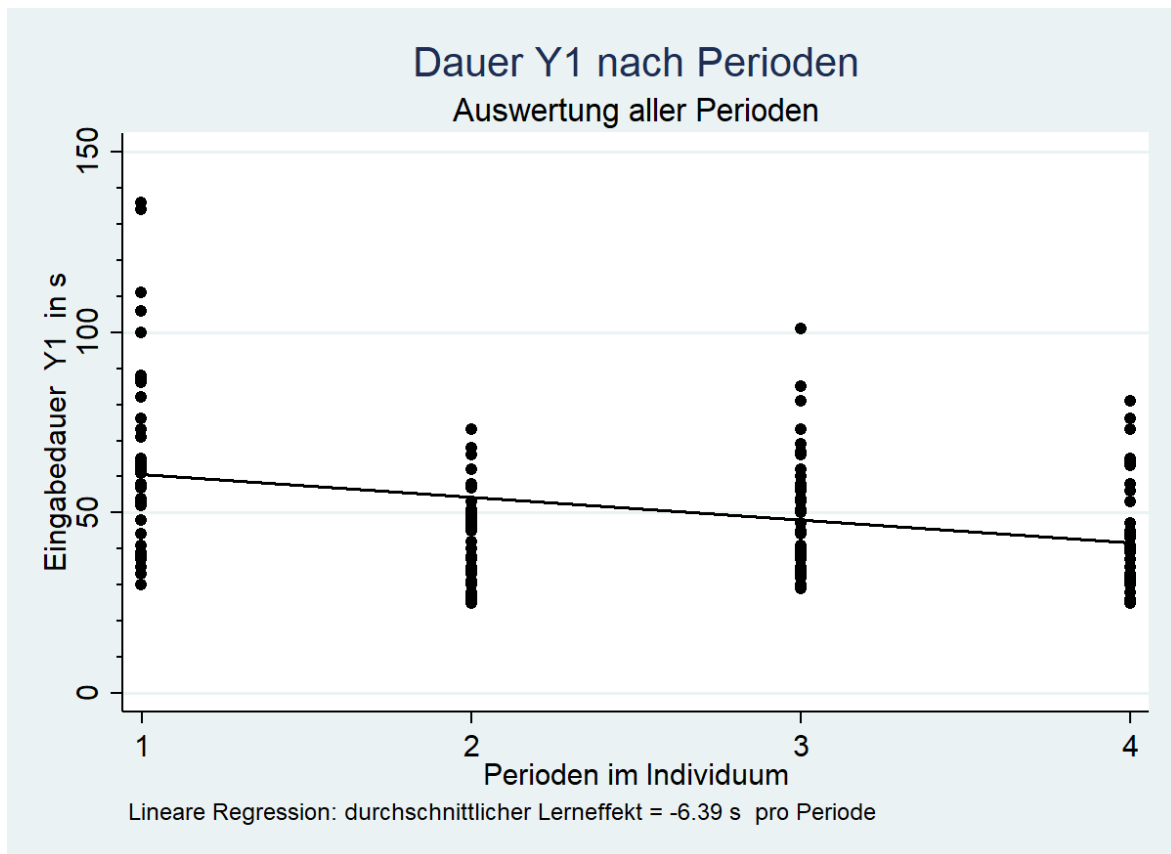


Abb. 2: Y1_period_all.tif: Dauer / Auswertung aller Perioden

Die Hauptdauervariable wird über die 4 Perioden, also Probe- und Messdurchgang, gezeigt und eine Ausgleichsgerade durch die Punkte gelegt. Man sieht, dass sich die Eingabedauer verkürzt und die Kinder im Durchschnitt um 6.4 s schneller pro Periode werden.

Beim Übergang von Periode 2 zu 3 tritt dieser Effekt nicht auf: Die Sprechweise wird nach der 2. Periode entweder von Tv zu Ze oder von Ze zu Tv gewechselt; die Kinder müssen sich also umstellen. Wird die Auswertung auf die Testdaten der 2. und 4. Perioden beschränkt, so ist kein Lerneffekt erkennbar, wie die Grafik zeigt:

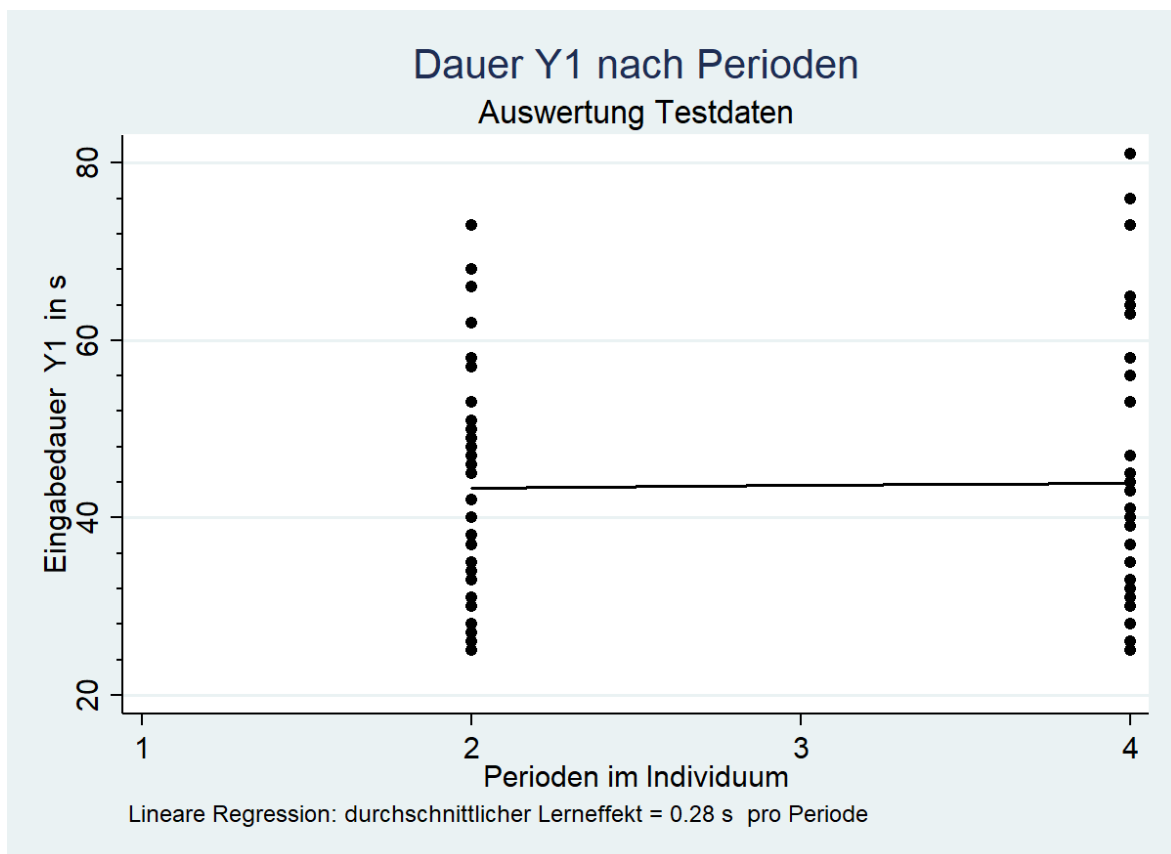


Abb. 3: Y1_period_test.tif: Dauer / Auswertung Testperioden

Dies bestätigt unser Studiendesign, nach dem die Hauptanalyse auf die Testdaten der jeweils zweiten Durchgänge beschränkt bleiben soll.

4.3.2 Der Zusammenhang von Eingabedauer und Fehlerzahl

Im Folgenden soll aufgezeigt werden, weshalb eine separate Betrachtung der Analysen für Eingabedauer und Fehlerzahl sinnvoll ist.

Die Zwanzigeins-App gibt in der Spielanleitung zu „Hören & Schreiben“ vor, dass ein Aufgabenblock (10 diktierte Zahlen = 1 Periode) erst erledigt ist, wenn alle Aufgaben korrekt eingegeben wurden. Falls eine Zahl fehlerhaft transkribiert wurde, wird sie rot unterlegt und noch einmal vorgesprochen; man hat also zwei bzw. mehrere Versuche, wenn nötig. Das Spiel wird erst nach korrekter Eingabe der jeweiligen Zahl fortgesetzt.

Vermutet werden könnte demnach, dass mit erhöhter Eingabedauer auch die Fehlerzahl steigt: Macht Kind A mehr Fehler Y2 als Kind B, so kann (tendenziell bei

vielen Kindern im Mittel) bei Kind A eine längere Dauer Y1 erwartet werden als bei Kind B.

Davon ausgehend, dass die beiden Variablen korreliert sind, soll bestimmt werden, wie hoch die Korrelation ist. Bei einer hohen Korrelation wäre eine separate Analyse der Auswertung nicht angezeigt.

Die untenstehende Grafik bildet den Zusammenhang über alle Messpunkte ab (4 Perioden pro Kind):

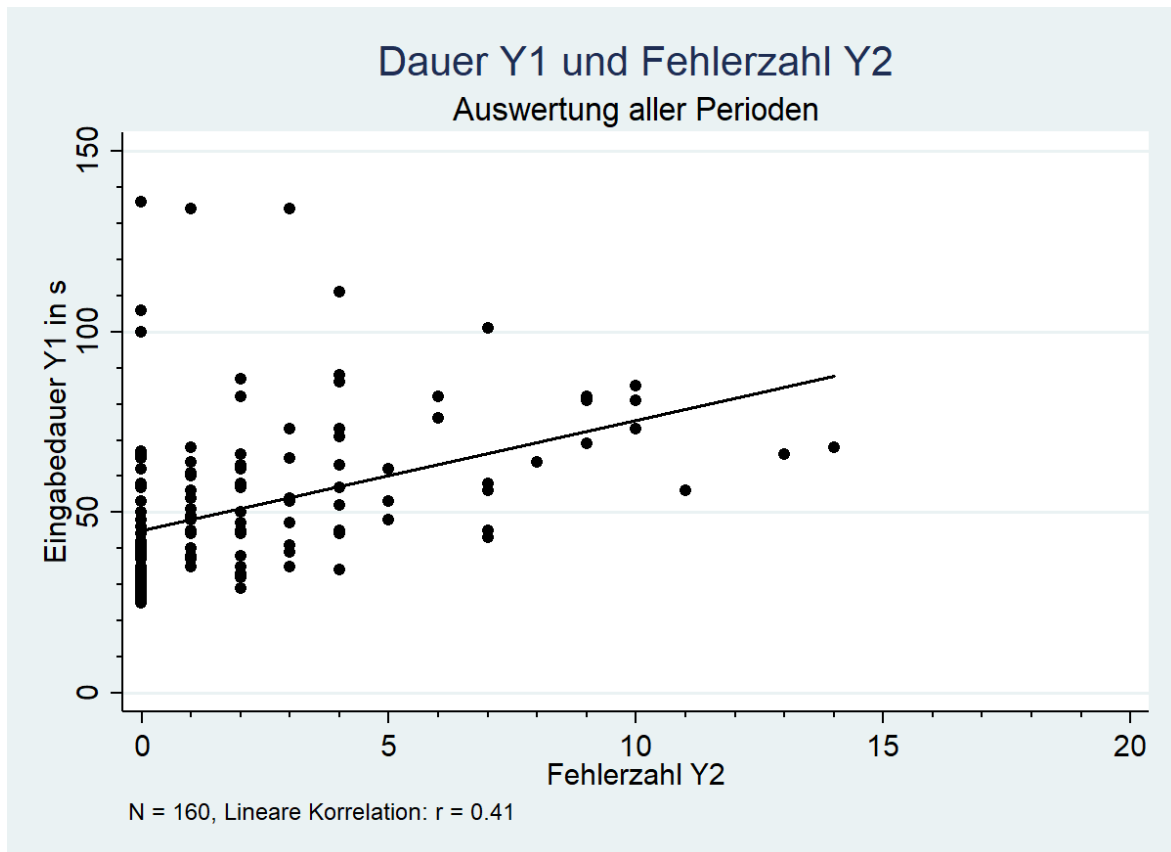


Abb. 4: Y1_Y2_all.tif: Zusammenhang Y1 und Y2 / Auswertung aller Perioden

Es wurde eine Regressionsgerade (Ausgleichsgerade) eingetragen, die den Zusammenhang von Y1 und Y2 verdeutlicht. Man sieht den Anstieg: Je mehr Fehler ein Kind gemacht hat (Y2), desto länger hat es für die Eingabe der zehn diktierten Zahlen gebraucht (Y1). Allerdings – wie in der Grafik zu erkennen – ist eine große Streuung zu verzeichnen: Bei z. B. 7 Fehlern reicht die Spanne der Zeitdauer über ca. 50 s bis 100 s. Die Korrelation zwischen Dauer und Fehlerzahl ist entsprechend auch nicht sehr ausgeprägt, der Korrelationskoeffizient ist nur 0.41, d. h. dass nur $0.41 \cdot 0.41 = 0.168 = 17\%$ der Variation von Y1 durch Y2 erklärbar ist. Die Variablen sind zwar korreliert, aber die Ausprägung ist so schwach, dass eine

getrennte Analyse sinnvoll und indiziert ist. Auch bei Beobachtung der Kinder unter der Testung konnte festgestellt werden, dass zum Teil eine lange Eingabedauer mit einer geringen Fehlerzahl korreliert war: Die konzentrierte Herangehensweise war zeitintensiv, führte aber zur korrekten Eingabe beim 1. Versuch, also zu weniger Fehlern bzw. mehr fehlerfreien Durchgängen.

Um zu überprüfen, ob es innerhalb der beiden Datenmengen von Studienkollektiv 1 (AK) oder Studienkollektiv 2 (VH) eine höhere Korrelation gibt, die durch die Zusammenführung der Daten (obige Abb. 3) ggf. nur verschleiert wurde, soll nun noch die Aufschlüsselung beider Studienkollektive untersucht werden.

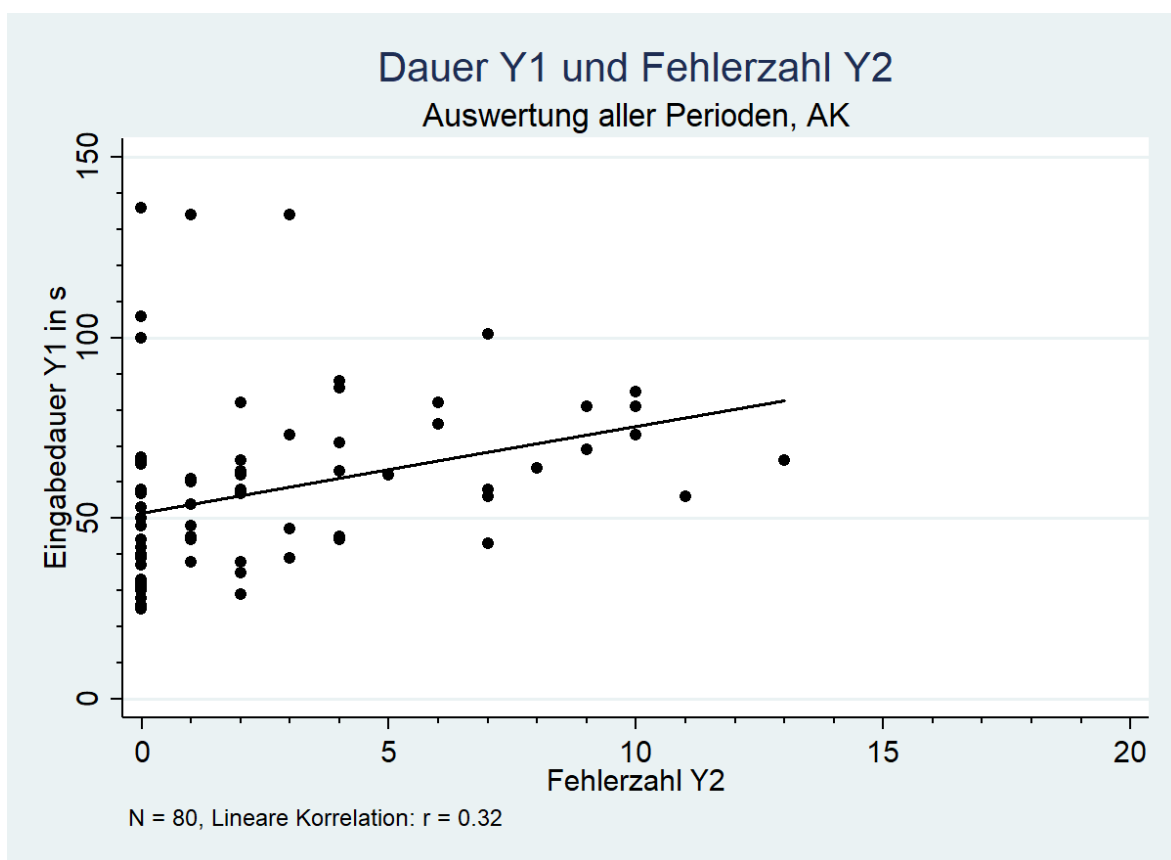


Abb. 5: Y1_Y2_AK_all.tif: Zusammenhang Y1 und Y2 / Auswertung der Studiengruppe AK

Die Korrelationen sind auch hier mit 0.32 (Abb. 4) und 0.50 (Abb. 5) gering. Auch der in Abb. 5 vergleichsweise hohe Anstieg bedeutet mit einem Korrelationskoeffizient von 0.5, dass lediglich 25% ($0.5 \cdot 0.5 = 25\%$) der Variation von Y1 durch Y2 erklärbar ist.

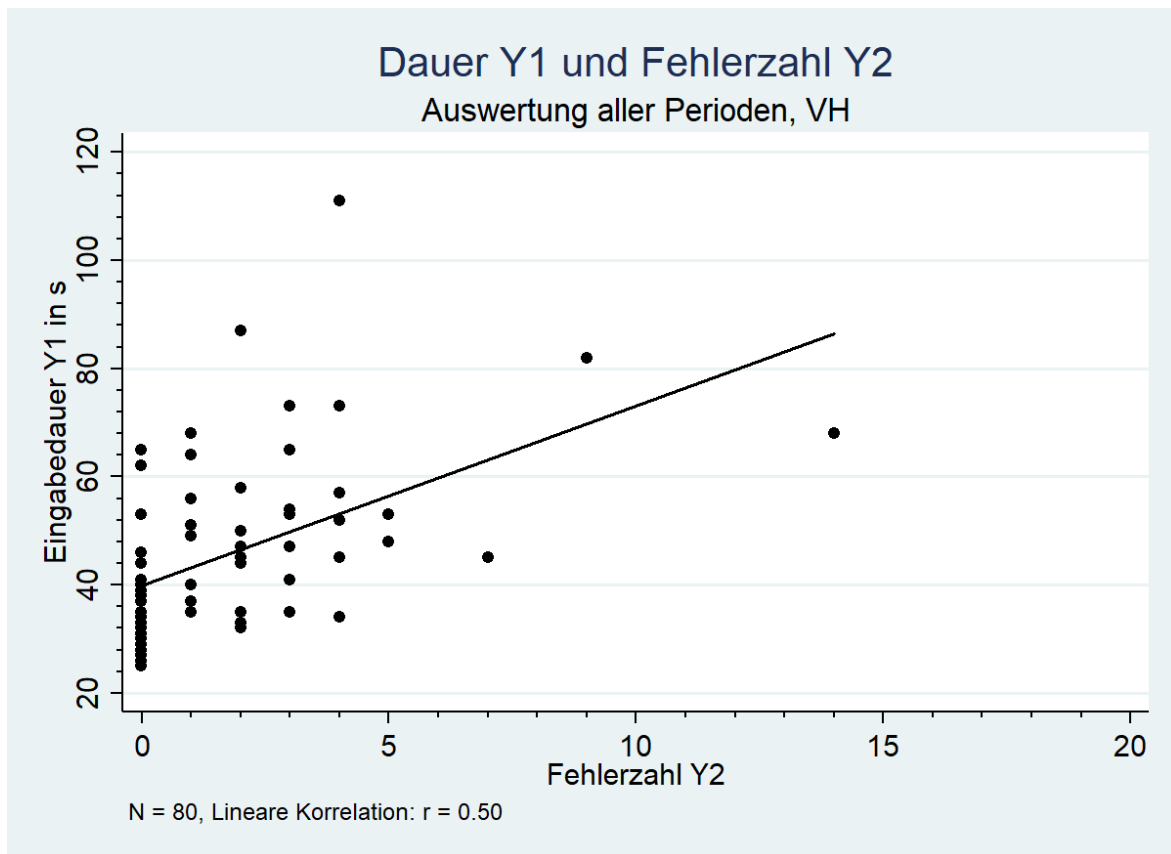


Abb. 6: Y1_Y2_VH_all.tif: Zusammenhang Y1 und Y2 / Auswertung der Studiengruppe VH

Die getrennte Betrachtung von Dauer und Fehlerzahl ist also auch nach separater Analyse der Studienkollektive angezeigt und rechtfertigt separate Auswertungen.

4.4 ERLÄUTERUNGEN ZUR DURCHFÜHRUNG

4.4.1 Responsvariablen

Zur Beantwortung der Forschungsfrage, ob sich bei rechenschwachen Kindern eine stellenwertgerechte Sprechweise (Ze) vorteilhaft gegenüber einer inversen Sprechweise (Tv) auswirkt, wurden die Zeitdauer (zwei Responsvariablen) sowie die Fehlerzahl (zwei Responsvariablen) pro Periode (Durchgang) für jeweils 10 zufällig generierte diktierter und zu transkribierende Zahlen erhoben. Zusätzlich zu den von der Zwanzigeins-App generierten Daten wurden also zwei weitere Responsvariablen ermittelt. Diese stellen sich wie folgt dar:

- 1) Als Hauptdauervariable wurde die Gesamtdauer ab Beginn der Ansage bis zum Ende der Eingabe der Zahlen durch das jeweilige Kind gemessen.

- 2) Da die Ansagedauer variiert (Kapitel 4.1), wurde zusätzlich eine Nebendauervariable ermittelt. Hierbei wurde von der Gesamtdauer der mittlere Unterschied der Ansagedauern subtrahiert.
- 3) Die vierte Responsvariable ist die Zahl der Fehler pro Durchgang.
- 4) Da auch viele fehlerfreie Durchgänge zu erwarten waren, wurde als fünfte Responsvariable die Anzahl der fehlerfreien Testungen definiert.

Die Ergebnisse zu den Fehlervariablen (Responsvariablen 3 und 4) werden in Kapitel 5 dargestellt und erläutert. Die Auswertungen zur Dauer (Responsvariablen 1 und 2) sind Gegenstand der Masterarbeit von Anna Kuhl (2025); sie werden daher in der vorliegenden Arbeit nicht detailliert dargestellt; die Ergebnisse werden jedoch als Gesamtbefund in die abschließende Diskussion aufgenommen.

4.4.2 Auswahl von Kovariablen

Wie oben erläutert, wurde während der Planungsphase diskutiert, welche Faktoren einen Einfluss darauf haben könnten, ob Ze bei rechenschwachen Kindern gegenüber Tv einen Vorteil hat.

Als Ergebnis des Austauschs wurden neben den oben beschriebenen Zielgrößen weitere Kovariablen definiert, die im Rahmen der Datenaufbereitung für jedes Profil pseudonymisiert in der App-Ausgabedatei mittels entsprechender Wertedefinition (bzw. Format, jeweils in Klammern) dokumentiert wurden. Die Ausgabedatei hält für weitere Informationen leere Spalten mit entsprechenden Variablennamen vor, von denen jedoch nicht alle als für die vorliegende Studie zielführend erschienen. Im Detail wurden neben der Eingabedauer und der Anzahl der Fehler folgende Daten als mögliche Einflussfaktoren erhoben bzw. aus hier aufgeführten Gründen nicht dokumentiert:

- 1) Studienkollektiv 1 (AK: Anna Kuhl), Studienkollektiv 2 (VH: Vivien Hartwig)
- 2) Eingesetzte Hardware (iPhone, Lenovo Desktop, Tablet)
- 3) Name der Fördereinrichtung (Lernpraxis, Zentrum für Schulische und Psychosoziale Rehabilitation der DRK-Kliniken Berlin-Westend, kurz: ZSPR, Duden-Institut)
- 4) Sequenz (TvTvZeZe, ZeZeTvTv)
- 5) Geschlecht (m, w, d für männlich, weiblich und divers, wobei letztere keine praktische Relevanz hatte)

- 6) Geburtsdatum (im üblichen 10-stelligen Format dd.mm.yyyy für Tag-Monat-Jahr. Bsp.: 15.08.2010, 15.12.2005; aus Datenschutzgründen wurde das Geburtsdatum nur monatsgenau angegeben, indem der Tag durchgängig auf den 15. gesetzt wurde)
- 7) Klassenstufe (1,2, 3, ...; - 1 für unbekannt)
- 8) Erstsprache (einzutragen in Erstsprache_1, Erstsprache_2, Erstsprache_3): Hieraus konnte abgeleitet werden, ob dem Kind eine stellenwertgerechte Sprechweise bereits vor der Testeinweisung bekannt war. Da es Kinder gibt, die z. B. in Deutsch beschult werden und gleichzeitig mit Mutter und Vater verschiedene Sprachen sprechen, ist die Erstsprache nicht immer eindeutig; daher wurden die Erstsprachenvariablen 1, 2 und 3 eingeführt.
- 9) Vorliegen eines Nachteilsausgleich (ja/nein/-; hier: - für unbekannt)
- 10) Vorliegen einer Lese-Rechtschreib-Störung (ja/nein/-): Besteht gleichzeitig eine diagnostizierte LRS? Dies galt bei „nein“ gleichzeitig als Indikator für eine Teilleistungsstörung ausschließlich in Mathematik.
- 11) Rückstellung (ja/nein/-; hier: - für unbekannt): Gemeint ist eine Rückstellung vor der regulären Einschulung, d. h. der Verbleib eines weiteren Jahres in einem Kindergarten bzw. die Wiederholung der Vorschule.
- 12) Schulwechsel (Anzahl: 0, 1, 2, ... -1; -1 für unbekannt). Es wurden alle Schulwechsel gezählt, also auch die von Grund- zu einer anderen Grundschule und z. B. zum ZSPR. In den meisten Fällen waren die Wechsel hier auf schulische Schwierigkeiten zurückzuführen, nur in zu vernachlässigenden Ausnahmefällen auf andere Gründe wie z. B. Umzug der Familie.
- 13) Wiederholung (ja/nein/-): Gemeint ist die Wiederholung einer Klassenstufe.
- 14) Diagnose (ja/nein/-): Unterteilung in Kinder mit kinderpsychiatrisch-psychotherapeutischer Diagnose bzw. Behandlung und ohne solche, falls gegeben mit ICD-10.
- 15) Besondere Belastung: Abnorme psychosoziale Umstände, bezogen auf die 5. Achse der ICD-10, wurden hier als freier Text vermerkt.

Die Angabe der Mathematik-Note des letzten Halbjahres wurde diskutiert, da sich die Frage stellte, ob leistungsschwache Kinder möglicherweise eher von Ze profitieren als leistungsstarke oder umgekehrt. Dies wurde jedoch aus verschiedenen Gründen verworfen: Ein Großteil des Gesamtkollektivs unterlag dem

Notenschutz, d. h. hier lag keine Note vor; der Nachteilsausgleich wird nicht an jeder Schule und von jeder Lehrkraft einheitlich umgesetzt, so dass die Noten nur verzerrt die Leistungen widerspiegeln, was einen Vergleich unmöglich macht; die Kinder waren zum Zeitpunkt der Testung unterschiedlich lange in Therapie – bei einigen gelingt der Transfer in den schulischen Kontext schon gut und es können dadurch gute Noten erzielt werden, bei anderen ist dies noch nicht der Fall. Die Vergleichbarkeit von Leistungsbewertungen in Form einer Mathematik-Note war für unsere Untersuchungsgruppe daher nicht aussagekräftig und wurde als Einflussvariable verworfen. Ebenso wurde auf die Angabe der Schulen und genauen Klassenbezeichnungen (z. B. 3b) aus Datenschutzgründen verzichtet, um eine Reidentifikation der Teilnehmenden auszuschließen. Anstelle der Klasse und Schule wurde lediglich die Klassenstufe angegeben.

Die Aufnahme der vorgenannten Kovariablen in die Erhebung und Auswertung führte zu einer erweiterten Hypothesenbildung, die im Folgenden aufgezeigt wird.

4.5 FORSCHUNGSFRAGE UND HYPOTHESEN

Aus den oben erläuterten Fragestellungen und Ausführungen zu den erhobenen Kovariablen ergab sich unsere Forschungsfrage:

Können rechenschwache Kinder der Grundschule bei Verwendung einer stellenwertgerechten Sprechweise diktierte Zahlen schneller und fehlerfreier transkodieren als bei der Verwendung der im deutschen Sprachraum gebräuchlichen inversen Zahlsprechweise und gibt es weitere Faktoren, die auf einen etwaigen Effekt Einfluss nehmen können?

Wir stellten folgendes Hypothesenpaar auf:

H1 Rechenschwache Kinder der Primarstufe transkodieren diktierte Zahlen bei Verwendung einer stellenwertgerechten Zahlsprechweise schneller und weniger fehleranfällig als bei der Verwendung der im deutschen Sprachraum gebräuchlichen inversen Zahlsprechweise, und zwar unabhängig von anderen möglichen Einflussfaktoren.

H0 Zwischen einer stellenwertgerechten Zahlsprechweise und der im deutschen Sprachraum gebräuchlichen inversen Zahlsprechweise besteht kein Unterschied in

Bezug auf Geschwindigkeit und Fehleranfälligkeit bei der Transkodierung von diktierten Zahlen durch rechenschwache Kinder der Primarstufe.

4.6 REKRUTIERUNG UND BESCHREIBUNG DER UNTERSUCHUNGSGRUPPE

Im Folgenden wird der Rekrutierungsprozess für die 21 Kinder der Gruppe 2 des Studienkollektivs (= Gesamtkollektiv, Untersuchungsgruppe) dargestellt. Die statistischen Auswertungen der nachfolgenden Kapitel im Teil II beziehen sich jedoch stets auf das Gesamtkollektiv, das in Kapitel 4.6.2 näher beschrieben wird.

4.6.1 Rekrutierungsprozess für das Studienkollektiv 2

Die Teilnehmenden für die Gruppe 2 des Studienkollektivs konnten mit der vorbehaltlosen Zustimmung des Leiters der Duden Institute Falkensee und Oranienburg, Mario Cordes, an diesen Standorten rekrutiert werden. Die Sorgeberechtigten erklärten mit der Unterzeichnung der Einverständniserklärung ihrerseits die Zustimmung zur Teilnahme an dieser Studie (Anhang 2). Als Maßnahme zum Erhalt des Datenschutzes (Baur / Blasius 2022a, S. 22) wurde nicht nur die Anonymität zugesichert, sondern auch einer möglichen Reidentifikation von Kindern vorgebeugt, indem auf die Erhebung von diesbezüglich kritischen Daten wie z. B. Name der Schule oder Klasse verzichtet wurde. Eine Schädigung der Untersuchungsteilnehmenden konnte auf diese Weise im Sinne der Forschungsethik (Döring / Bortz 2016, S. 86) ausgeschlossen werden.

Durch die Unterstützung von interessierten KollegInnen, die Eltern kontaktierten, über das Forschungsthema informierten und Termine koordinierten, konnten insgesamt 21 Kinder der 3. bis 6. Klassenstufen für das Studienkollektiv 2 (s. Kap. 4.6.2) rekrutiert werden. In Berlin und Brandenburg umfasst die Grundschule 6 Schuljahre. Alle in die Studie aufgenommenen Kinder besuchen also eine Grundschule, im Fall des Studienkollektivs 1 die Klassenstufen 2 bis 6.

4.6.2 Beschreibung der Untersuchungsgruppe

Das Gesamtkollektiv (= Untersuchungsgruppe) setzt sich also aus den Studienkollektiven 1 und 2 zusammen, die im Folgenden näher beschrieben werden sollen.

Alle in die Studie aufgenommenen Kinder sind von einer Rechenstörung betroffen: Sie wurden entweder von fachlicher Seite, in der Regel durch eine kinderpsychiatrische Praxis, gemäß der ICD-10 diagnostiziert oder von der zuständigen Lehrperson als „rechenschwach“ eingestuft und so in die verschiedenen lerntherapeutischen Einrichtungen aufgenommen, in denen für diese Studie Teilnehmende gewonnen werden konnten.

Da viele Kinder nicht nur unter Dyskalkulie, sondern auch unter komorbiden Störungen leiden, sind sie teilweise nicht nur in lerntherapeutischer Behandlung, sondern zusätzlich in einer kinderpsychiatrischen Praxis angebunden oder sie lernen vorübergehend am Zentrum für Schulische und Psychosoziale Rehabilitation in Berlin (ZSPR). In dieser Einrichtung sowie in einer lerntherapeutischen Praxis konnte Anna Kuhl (2025) 22 Kinder für das Studienkollektiv 1 (AK) rekrutieren, wobei es zwei Non-Responder gab: Ein Kind verweigerte die Teilnahme und mit einem weiteren konnte die Testung nur unvollständig realisiert werden, so dass kein auswertungsfähiger Datensatz vorlag. Auch ein Kind des Studienkollektivs 2 verwehrte auf Grund seiner spezifischen Diagnostik und somit nachvollziehbaren Gründen die Teilnahme, so dass insgesamt 20 Probanden dieses Studienkollektivs die App testeten.

Aus dem Gesamtkollektiv von 43 Kindern der Jahrgangsstufen 2 bis 6 konnten somit 40 vollständige Datensätze gewonnen werden. Die große Altersspanne von 8;9 bis 13;9 Jahren liegt darin begründet, dass Kinder ggf. von der Einschulung zurückgestellt wurden und / oder ein Schuljahr wiederholt haben.

Das Geschlechterverhältnis der Teilnehmenden des Studienkollektivs 1 war mit 10 Mädchen und 10 Jungen ausgeglichen. Bei den teilnehmenden Kindern der Gruppe 2 war sie hingegen mit 15 Mädchen und 5 Jungen unausgeglichen. Das Gesamtkollektiv bestand demnach aus 25 Mädchen und 15 Jungen. Da die Teilnehmenden innerhalb der Institutionen zufällig gewählt wurden, konnte eine Ursache für diese ungleiche Geschlechterverteilung nicht identifiziert werden. Es

gibt jedoch verschiedene Studien, die belegen, dass Mädchen etwas häufiger von Rechenschwäche betroffen sind als Jungen (Kap. 2.4.1).

Die Heterogenität des Gesamtkollektivs war nicht nur durch die Altersspanne und Rekrutierung in verschiedenen Einrichtungen gegeben, sondern auch durch die unterschiedlichen Bevölkerungsstrukturen der Einzugsgebiete. Die Duden-Institute Falkensee und Oranienburg liegen hinter der Berliner Stadtgrenze. Anders als in Berlin, dem hauptsächlichen Einzugsbereich von ZSPR und lerntherapeutischer Praxis (Teilnehmende der Gruppe 1), bilden in diesen Städten und anliegenden Gemeinden Kinder mit Migrationshintergrund und/oder nichtdeutscher Herkunftssprache die Ausnahme, bzw. sie besuchen aus uns nicht näher bekannten Gründen nur in Einzelfällen die genannten Duden-Institute. Die Kovariable *Erstsprachen* (Kap. 4.4.2) der teilnehmenden Kinder der Duden-Institute war daher durchgehend Deutsch, wohingegen die Teilnehmenden der Gruppe 1 unterschiedlichen Herkunftsländern und Erstsprachen zuzuordnen waren.

Wie oben bereits erläutert, leiden rechenschwache Kinder häufig unter komorbiden Störungen und / oder weiteren belastenden psychosozialen Umständen. Einzelne Kinder unserer Untersuchungsgruppe erhielten neben der Rechenstörung u. a. folgende Diagnosen nach der ICD-10:

- Emotionale Störung des Kindesalters mit Stimmungsschwankung und mangelndem Selbstwertgefühl (F93.8), Rechenstörung (F81.2)
- Kombinierte Störung schulischer Fertigkeiten (F81.3), Sonstige emotionale Störung des Kindesalters (F93.8), Undifferenzierte Somatisierungsstörung (F45.1)
- Emotionale Störung des Kindesalters mit Insuffizienzgefühl, überhöhter Anpassbarkeit und sozialer Erwünschtheit (F93.8), Aufmerksamkeits- und Hyperaktivitätsstörung (F90.0), Kombinierte Störung schulischer Fertigkeiten (F81.3)
- Generalisierte Angststörung (F41.1), Einfache Aufmerksamkeits- und Hyperaktivitätsstörung (F90.0), Kombinierte Störung schulischer Fertigkeiten (F81.3)
- Einfache Aufmerksamkeits- und Hyperaktivitätsstörung (F90.0), Nichtorganische Enuresis (F98.0), Anpassungsstörung (F43.2), Lese- und

Rechtschreibstörung (F81.0), Kombinierte umschriebene
Entwicklungsstörung (F83)

Weitere dokumentierte Diagnostiken innerhalb der Untersuchungsgruppe umfassten: Störung des Sozialverhaltens mit oppositionellem, aufsässigem Verhalten; Versagensängste; Verhaltensstörung im Kindesalter; Emotionale Störung mit Geschwisterrivalität; Bindungsstörung; Institutionelle Erziehung; Vorübergehende Ticstörung; Hyperkinetische Störung des Sozialverhaltens; Rezeptive Sprachstörung; Expressive Sprachstörung; Umschriebene Entwicklungsstörung der Fein- und Graphomotorik; Umschriebene Entwicklungsstörung der Grobmotorik; Leichte depressive Episode; Mittelgradige Depression sowie Hyperkinetische Störung des Sozialverhaltens.

Viele Kinder stehen bzw. standen auf Grund ihrer Diagnose unter Medikation.

Zu den belastenden Begleitumständen, die im Rahmen der Anamnesen erfasst wurden, gehören unter anderem eigene körperliche Beeinträchtigungen oder körperliche Beeinträchtigungen von Familienmitgliedern; psychische Störungen eines Elternteils; Disharmonien in der Familie zwischen Erwachsenen; Migration; Verlust einer liebevollen Beziehung; Streitbeziehungen mit Schülern; inadäquate oder verzerrte intrafamiliäre Kommunikation sowie traumatische Erfahrungen wie Gewalterfahrungen oder Todesfälle in der Familie.

Einige Kinder wurden vor der Einschulung ein Jahr zurückgestellt (d. h. die Einschulung erfolgte ein Jahr später als üblich) und / oder sie wiederholten einmal oder mehrmals eine Klassenstufe, so dass sie wesentlich älter als ihre MitschülerInnen sind, was die Pflege sozialer Beziehungen erschweren kann. Gleiches gilt für häufige Schulwechsel. Auch der Besuch von verschiedenen Unterstützungsangeboten wie Logopädie, Psycho-, Ergo- und Lerntherapie, der in vielen Fällen indiziert ist, schränkt die Freizeit und somit die soziale Teilhabe ein.

Auf Grund der Diagnostiken und besonderen Belastungen sind viele Kinder sozial-emotional schwer belastet. Eine therapeutische Intervention kann mitunter über Jahre notwendig sein. Andere Kinder haben gute kognitive Voraussetzungen und ein stabiles, unterstützendes und wertschätzendes Umfeld (Eltern, Lehrkräfte, TherapeutIn); hier kann eine Therapie innerhalb kurzer Zeit zu einer Abmilderung der Symptome und Verbesserung der Leistungen und Lebensumstände führen.

Vor dem Hintergrund der beschriebenen Symptomaten und Diagnostiken der rekrutierten Kinder lässt sich einordnen, warum es in Einzelfällen nicht möglich war, die Testungen durchzuführen oder zum Abschluss zu führen.

4.7 DIE DURCHFÜHRUNG DER ERHEBUNG

Die am Rekrutierungsprozess beteiligten TherapeutInnen wurden vorab eingewiesen, so dass weder dem Kind noch den Eltern im Vorfeld Informationen zur Durchführung gegeben wurden, um etwaigen Übungseffekten (durch vorheriges Trainieren einer stellenwertgerechten Zahlensprechweise und / oder Ausprobieren der App) vorzubeugen. Der Name der App *Zwanzigeins* war auch den TherapeutInnen im Vorfeld nicht bekannt. Auf diese Weise konnte sichergestellt werden, dass die Kinder zum ersten Mal während der kurzen Einweisung mit der stellenwertgerechten Zahlensprechweise konfrontiert wurden und keinerlei Vorwissen durch vorherige Briefings oder Ausprobieren bestand. Für die Einweisung vor dem Test mit der App wurde ein Gesprächsleitfaden eingesetzt (s. Anhang 1).

4.8 METHODIK DER AUSWERTUNG

Zur Auswertung wurde das Programmpaket Stata 14 eingesetzt (StataCorp 2015).

Die Ergebnisse werden im Folgenden mit Hilfe von beschreibenden Darstellungen präsentiert. In der deskriptiven Statistik (oder beschreibenden Statistik) können die Informationen im Datensatz so komprimiert werden, dass man mit ihnen schon erste Forschungsfragen beantworten kann. So werden z. B. die Informationen im Datensatz zu einer Tabelle, einer Grafik oder zu aussagekräftigen Maßzahlen verdichtet (Blasius & Baur 2022b, S. 734). Bezüglich der Fehlerzahl Y2 und den Anteilen fehlerfreier Durchgänge Y2neg umfassen die beschreibenden Darstellungen Tabellen z. B. zu Umfang, Median und Mittelwert sowie grafische Darstellungen (Bar Charts und Box-Plots).

Zur zufallskritischen Analyse wurden fixed-effect Längsschnittdatenanalysen durchgeführt: Für die Responsvariable „Fehlerzahl“ (natürliche Zahl) wurde der Stata-Befehl `xtpoisson` eingesetzt und der Befehl `xtlogit` für den Respons

"fehlerfreier Durchgang" (binäre Größe). Zur Bearbeitung von Zusatzfragestellungen (z. B. Einfluss des Geschlechts auf den Effekt) wurden entsprechende Interaktionsmodelle eingesetzt.

4.8.1 Hauptanalyse

Die jeweils zweiten Perioden (Durchgänge) der unterschiedlichen Sprechweisen wurden in der Hauptanalyse ausgewertet. Dies sind die Testdaten, um unsere Hypothesen analytisch zu testen. Die jeweils ersten Perioden wurden als Einübungsphase (Probedurchgänge) ignoriert.

4.8.2 Zusätzliche Analyse aller Daten

Die Auswertung aller Daten (d. h. aller Perioden inkl. der Einübungsphase) wurde ergänzend durchgeführt, um zu überprüfen, ob die Ergebnisse der Hauptanalyse bei Verwendung aller zur Verfügung stehenden Daten bestätigt werden. In dieser zweiten Analyse wurden alle vier Perioden unter Adjustierung für die Periode ausgewertet. Diese Adjustierung war insofern unverzichtbar, als sie der Kontrolle von Lernfortschritten während der Untersuchung diene. Eine gemeinsame Auswertung aller Perioden hat den Vorteil, dass der Umfang größer ist und keine Vorab-Selektion von Perioden stattfindet, aber den Nachteil, dass die Ergebnisse von dem erheblichen Lernfortschritt an der App überlagert sind, der – wie nachgewiesen – ausgeschlossen wird, wenn ausschließlich die Testdaten ausgewertet werden.

4.8.3 Bezeichnungen der Dateien und Graphen

Für die in den Auswertungsdateien und Dateinamen verwendeten Bezeichnungen wurden folgende Abkürzungen festgelegt: Die Eingabedauer ist Y1, die Fehlerzahl wird mit Y2 bezeichnet. Y2neg bedeutet, dass keine Fehler gemacht wurden (= Fehlerzahl negativ), d. h. diese Bezeichnung wird für die fehlerfreien Durchgänge verwendet. Die Sprechweise als unabhängige Variable wird mit X abgekürzt.

5 DARSTELLUNG UND ANALYSE DER ERGEBNISSE ZU DEN FEHLERZAHLEN

Im Folgenden werden die Ergebnisse zu den Variablen, die die Fehlerzahl beschreiben, detailliert dargestellt und erläutert.

Wie im methodischen Teil (4.8) beschrieben, liegen den Auswertungen der Hauptanalyse die Daten der jeweils zweiten Perioden (Testdaten) zugrunde; die jeweiligen Probedurchläufe wurden nicht in die Analyse einbezogen. Es gibt also für die Hauptanalyse von Y1 und Y2 jeweils 80 Beobachtungen („number of observations“ = 80) und 40 Gruppen („number of groups“ = 40), d. h. für die Hauptanalyse wurden 80 Messungen ausgewertet. Da es jedoch bei der Auswertung der Fehlerzahlen und Differenzen viele Null-Werte gibt, musste die Anzahl der auswertbaren Beobachtungen („number of obs“) angepasst werden (Kap. 5.1.2).

In der vorliegenden Arbeit umfasst die deskriptive Analyse der Responsvariablen *Fehlerzahl* (Y2) und *fehlerfreie Durchgänge* (Y2neg) die Angabe von Kennwerten wie Median- und Mittelwert sowie die grafische Darstellung mittels Bar-Charts, da dies auf Grund der vielen Null-Werte aus methodischer Sicht das aussagekräftigste Darstellungsmedium ist. Diese Null-Werte ergeben sich immer dann, wenn kein Fehler gemacht wurde, also alle 10 diktierten Zahlen von den Kindern sofort korrekt eingegeben wurden.

Die Darstellung der Differenz erfolgt jedoch mittels Box-Plot, da hier die Streuung besser veranschaulicht werden kann als bei einem Bar-Chart.

Die zufallskritischen Testungen der Differenzverteilungen wurden mittels Regressionsrechnungen durchgeführt. Die dabei zu beachtenden methodischen Besonderheiten werden an entsprechender Stelle erläutert.

5.1 ANALYSEN DER MITTLEREN FEHLERZAHL (RESPONSVARIABLE)

5.1.1 Deskriptive Analyse der Ergebnisse zu den Fehlerzahlen

Zunächst soll der Fragestellung nachgegangen werden, welchen konkreten Effekt die Expositionsänderung, also der Wechsel der Sprechweisen, auf die Fehlerzahl hat.

Variable Y2 enthält die Zahl der Fehler pro Durchgang. Wie erläutert wurde für eine komprimierte grafische Darstellung der Daten ein Bar-Chart gewählt:

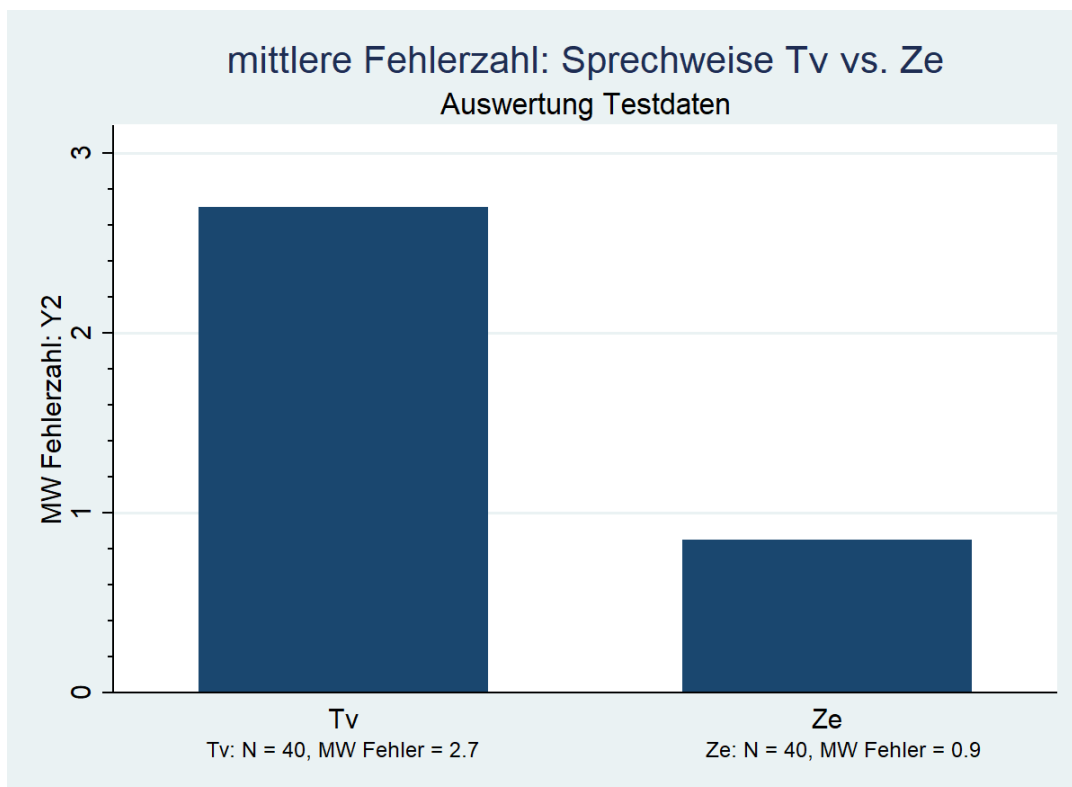


Abb. 7: Bar_Y2_test.tif: Auswertung der mittleren Fehlerzahl Sprechweise: Tv. vs. Ze

Die grafische Darstellung zeigt den eindeutigen Vorteil einer stellenwertgerechten Zahlsprechweise: Mit Tv wurden im Mittel 2.7 Fehler pro Durchgang von 10 diktierten Zahlen gemacht, bei der Verwendung von Ze nur 0.9 Fehler pro Durchgang. Es handelt sich dabei um gerundete Werte.

Für das Differenzbild bietet sich ein Box-Plot an: Der Vorteil von Ze ist auf Abbildung 8 deutlich zu erkennen.

Der Mittelwert gibt für die inverse Sprechweise (Tv) 1.9 zusätzliche Fehler an. Der Anteil der negativen Differenzen (< 0) beträgt 17.5%. Dies bedeutet, dass 17.5% der Kinder bei der Sprechweise Ze mehr Fehler hatten als bei der inversen Sprechweise.

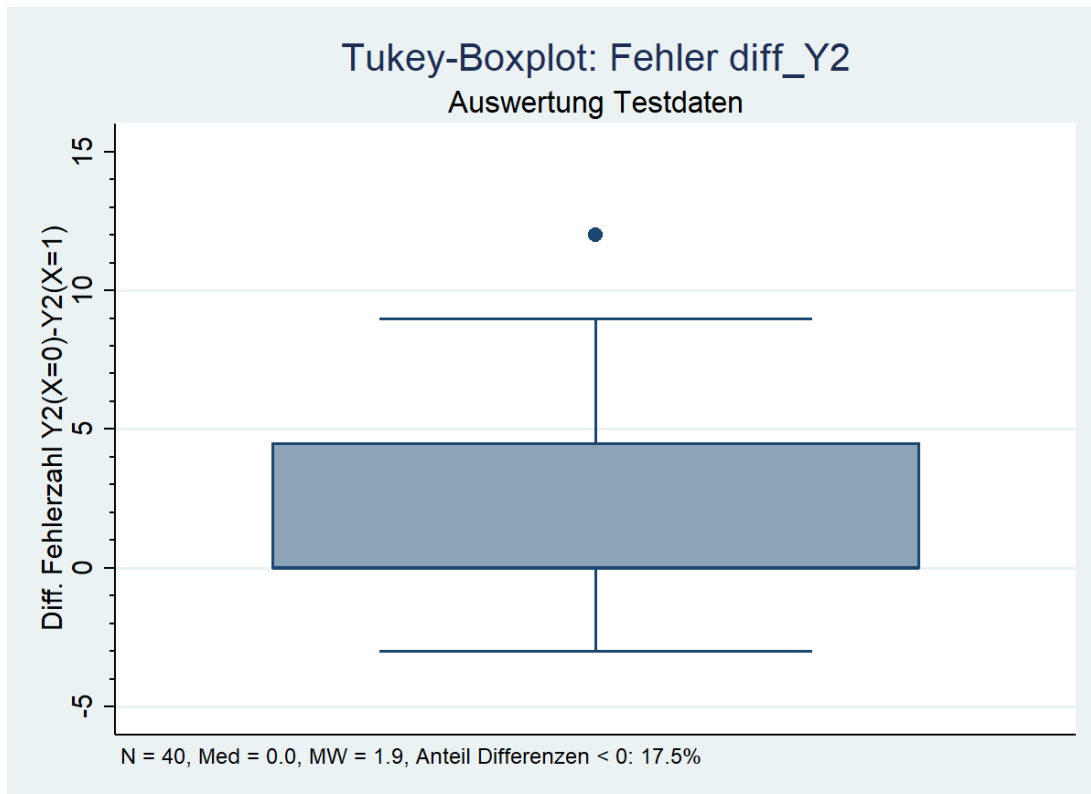


Abb. 8: Box_diff_Y2_test.tiff: Verteilung der Fehlerzahl-Differenzen

5.1.2 Zufallskritische Analyse der Ergebnisse zur mittleren Fehlerzahl Y2 (Testdaten)

Im Folgenden wird die Regressionsrechnung zu der Abbildung 7: Bar_Y2_test.tif unter Anwendung von xtpoisson („extended Poisson regression“) dargestellt, d. h. einer im Längsschnitt erweiterten („within“) Poisson-Regression zu der Zielgröße *Fehlerzahl* mit festen Effekten („fixed effects“), auch *Panel-Poisson-Regression mit festen Effekten* genannt.

Übliche Effekt-Analyse

```
-----
      Y2 |          IRR   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
      x |   .3148148   .0619082   -5.88   0.000   .2141247   .4628535
-----
```

Die Zahl der ausgewerteten Beobachtungen (Number of obs = 54) wie auch die Zahl der Kinder (Number of groups = 27) ist jeweils reduziert, d. h. statt 40 werden 27 Kinder mit je einer Messung für Ze und Tv, insgesamt also 54 Messungen, ausgewertet. Die anderen 13 Kinder haben in beiden Testungen, also bei Tv und bei Ze, keine Fehler gemacht. Das Poisson-Modell misst, wie sich die Fehlerzahl relativ ändert, also welcher Prozentsatz an Änderung in der Fehlerzahl durch den Wechsel auf Ze vorliegt. Ob sich Fehler prozentual vermehren oder verringern, kann jedoch an diesen 13 Kindern nicht festgestellt werden, da jeder beliebige Prozentsatz von Null immer Null ist. Deshalb werden diese Kinder hier nicht berücksichtigt.

IRR kennzeichnet das Verhältnis der Fehlerzahl bei stellenwertgerechter Sprechweise (Ze) zur Fehlerzahl bei traditionell verdrehter Sprechweise (Tv), d. h. das Verhältnis der in der Abbildung 7 angegebenen Fehleranteile 0.9 und 2.7 (bis auf Rundung). $IRR=0.3148$ bedeutet, dass die mittlere Fehlerzahl bei stellenwertgerechter Sprechweise nur 31.5% der mittleren Fehlerzahl bei Verwendung von der inversen Sprechweise beträgt.

Das 95%-Konfidenzintervall reicht von 21.4% bis 46.3%, also eine statistisch klare Reduktion unter 100% (= Ausgangswert bei Tv). Entsprechend ist der zweiseitige p-Wert <0.0005 , also kleiner als 0.05%. Der Effekt ist hochsignifikant.

5.1.3 Zusammenfassung der Ergebnisse zur mittleren Fehlerzahl Y2

Das Ergebnis für die mittlere Fehlerzahl Y2 ohne Berücksichtigung von Kovariablen stellt sich wie folgt dar:

Durch den Wechsel nach Ze reduziert sich die Fehlerzahl auf 31% des Ausgangswerts bei Tv. Es konnte also eine relative Reduktion der Fehlerzahl um 69% festgestellt werden. Der zweiseitige p-Wert ist kleiner als 0.05% und somit hochsignifikant, entsprechend schließt das Konfidenzintervall (0.21 bis 0.46) den Neutralwert 1 (100%) deutlich aus.

Die Verringerung der Fehlerzahl durch die Verwendung der stellenwertgerechten Sprechweise im Vergleich zur inversen Sprechweise ist somit hochsignifikant.

5.2 ANALYSEN DER FEHLERFREIEN DURCHGÄNGE (RESPONS VARIABLE)

5.2.1 Deskriptive Analyse der Ergebnisse zu den fehlerfreien Durchgängen

Da es zahlreiche fehlerfreie Durchgänge gab, verdienen diese eine gesonderte Betrachtung. Sie wurden zusätzlich ausgewertet, um der Besonderheit der Nullen (= Durchgänge ohne Fehler) gerecht zu werden.

Wir definieren die Variable $Y_{2neg} = ja$, wenn der Durchgang ohne jeden Fehler erfolgte.

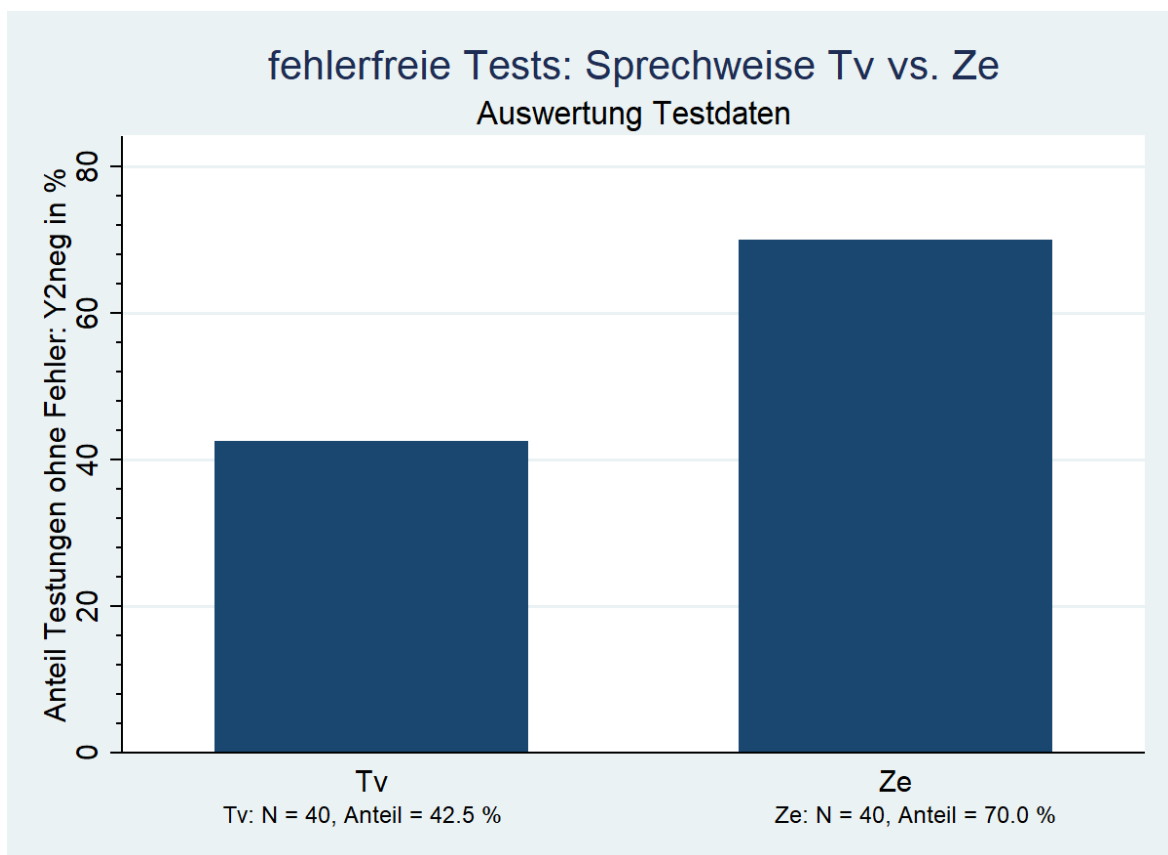


Abb. 9: Bar_Y2negp_test.tif: Anteile der fehlerfreien Testungen

Die Abbildung zeigt die Anteile fehlerfreier Durchgänge Y_{2negp} bei beiden Sprachmodi: Bei Tv sind es 42.5%, bei Ze hingegen 70%. Die Anzahl der fehlerfreien Testungen ist also bei Verwendung von Ze wesentlich erhöht. Die Zunahme der fehlerfreien Durchgänge verdeutlicht die Grafik.

Die Häufigkeit fehlerfreier Durchgänge nimmt also statistisch gesichert durch den Wechsel auf eine stellenwertgerechte Sprechweise zu.

5.2.2 Zufallskritische Analyse und Zusammenfassung der Ergebnisse zu den fehlerfreien Durchgängen

Auch dieser Unterschied lässt sich regressionsanalytisch testen. Es muss eine logistische Regression (fixed effect longitudinal) mit dem Befehl xtlogit: („extended logistic regression“) zu der Zielgröße „fehlerfreier Durchgang“ mit festen Effekten (fixed effects), auch *Panel-Logistik-Regression mit festen Effekten* genannt, angesetzt werden, da die Zielgröße nur zwei Merkmale als Ausprägung hat (fehlerfrei: ja oder nein):

Übliche Effekt-Analyse (ohne Kovariablen)

Y2neg	OR	Std. Err.	z	P> z	[95% Conf. Interval]
x	3.75	2.110243	2.35	0.019	1.24461 11.29872

Die Zahl der ausgewerteten Beobachtungen ist auch hier reduziert (Number of obs = 38), entsprechend die der Kinder (Number of groups = 19). Insgesamt konnten also 38 Messungen ausgewertet werden: 19 Kinder mit je einer Messung für Ze und Tv. Die anderen 21 Kinder zeigten keine Änderung: Bei ihnen lagen entweder in beiden Durchgängen keine Fehler vor, oder es lagen in beiden Durchgängen Fehler vor.

Hier wird das Chancenverhältnis (Odds Ratio) geschätzt, also das Verhältnis der Anzahl positiv diskordanter Paare zur Anzahl negativ diskordanter Paare aus den individuell gematchten Daten mit Tv bzw. Ze. Als positiv diskordant zählen Kinder, die bei Tv keinen fehlerfreien Durchgang hatten, aber bei Ze; als negativ diskordant Kinder, die bei Tv einen fehlerfreien Durchgang hatten, aber keinen bei Ze. Alle anderen Kinder sind konkordant. Es waren 15 Kinder positiv diskordant, 4 Kinder negativ diskordant und 21 Kinder konkordant. Mit $OR = 15/4 = 3.75$ ergibt die Analyse ein gegenüber dem Neutralwert 1 (100%) deutlich erhöhtes Chancenverhältnis. Diese Erhöhung des fehlerfreien Anteils durch Ze ist signifikant: $p = 0.019 = 1.9\% < 5\%$. Das 95%-Konfidenzintervall reicht von 1.2 bis 11.3.

5.3 MITTLERE FEHLERZAHL: EFFEKT DES ALTERS

Da zu erwarten war, dass das Alter der Kinder einen Einfluss auf die Untersuchungsergebnisse haben würde, wurde die Kovariable *Alter* erhoben. Die Ergebnisse zeigen signifikante Unterschiede und werden daher detailliert dargestellt. Wir unterscheiden in die Altersgruppen 10 Jahre und jünger (≤ 10) und älter als 10 Jahre (> 10).

5.3.1 Deskriptive Analyse der Ergebnisse zum Effekt des Alters

Abbildung 10 zeigt, dass die älteren Kinder (älter als 10 Jahre) erwartungsgemäß in beiden Sprechweisen weniger Fehler machen. Deutlich wird auch, dass beide Altersgruppen mit Ze weniger Fehler machen als mit Tv.

Die 18 Kinder, die 10 Jahre und jünger sind, machen mit Tv 3.3 und mit Ze 1.4 Fehler im Durchschnitt, wohingegen den 22 Kindern, die älter als 10 Jahre sind, mit Tv 2.2 und mit Ze lediglich 0.4 Fehler im Durchschnitt unterlaufen.

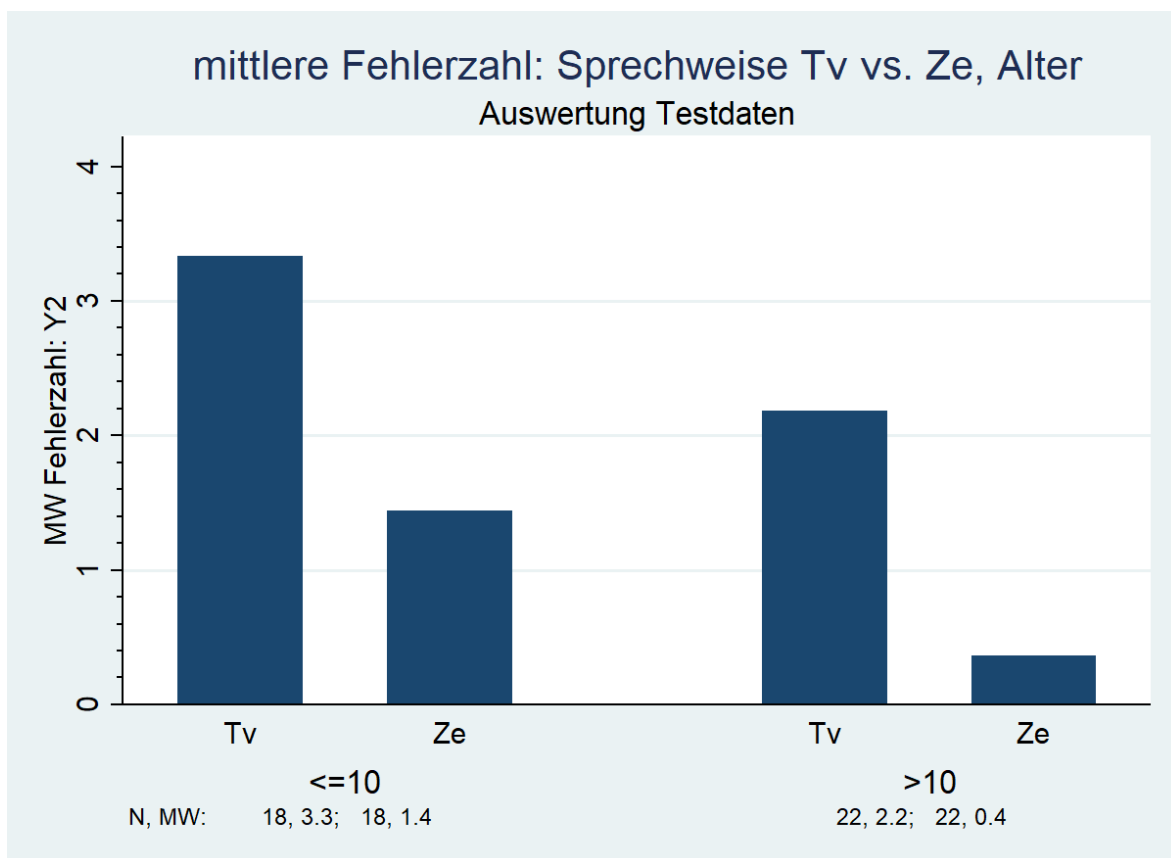


Abb. 10: Bar_Y2_Alt_kat_test.tif: Mittlere Fehlerzahl / Auswertung der Kovariable Alter

Die Box-Plots zur Differenz der Fehler zwischen Tv und Ze (Abbildung 11) zeigen den Vorteil der stellenwertgerechten Sprechweise. Beide Box-Plots liegen im für Ze günstigen Bereich.

Bei den 18 jüngeren Kinder lagen 27.8% und bei den 22 älteren Kindern nur 9.1% im für Tv günstigen Bereich. Für die älteren Kinder stellt sich im Unterschied zu den jüngeren im Median (= 0.0) kein Vorteil von Ze ein.

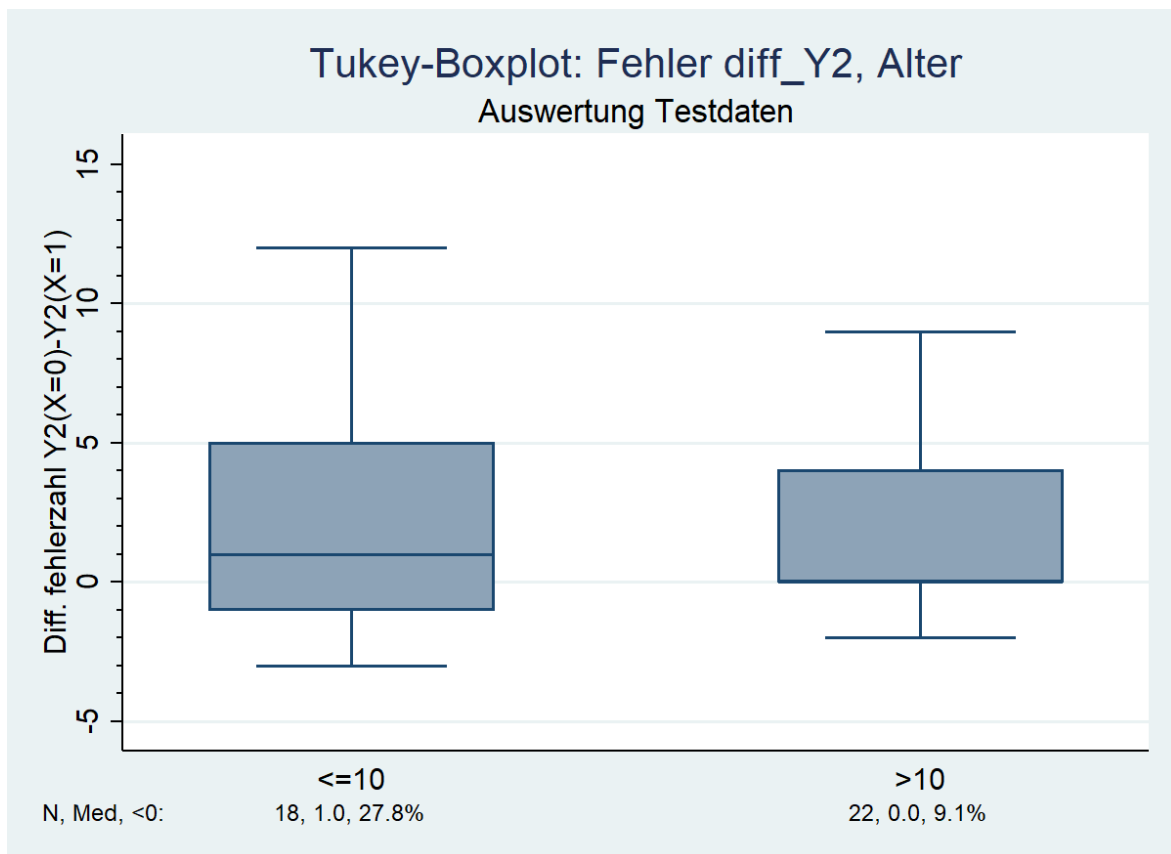


Abb. 11: Box_diff_Y2_kat_test.tif: Fehlerdifferenz / Auswertung der Kovariable Alter

Diese Box-Plots können den Vorteil von Ze aber nicht vollständig abbilden, da für beide Sprechweisen – insbesondere bei den älteren Kindern – viele Durchgänge ohne Fehler vorliegen und demnach keine Fehlerdifferenz ermittelt werden kann.

5.3.1.1 Fehlerfreie Durchgänge (Kovariable Alter)

Daher ist die Auswertung der fehlerfreien Durchgänge zusätzlich indiziert. Abbildung 12 zeigt:

a) Den älteren Kindern gelingen in beiden Sprechweisen häufiger fehlerfreie Durchgänge und

b) der Anteil der fehlerfreien Durchgänge nimmt in beiden Altersgruppen mit Ze zu.

Von den 18 jüngeren Kindern (10 Jahre oder jünger) gelangen bei inverser Sprechweise 27.8% fehlerfreie Durchgänge, bei stellenwertgerechter Sprechweise jedoch 61.1%. Von den 22 älteren Kindern (älter als 10 Jahre) gab es 54% fehlerfreie Durchgänge bei Tv und 77.3% bei Ze.

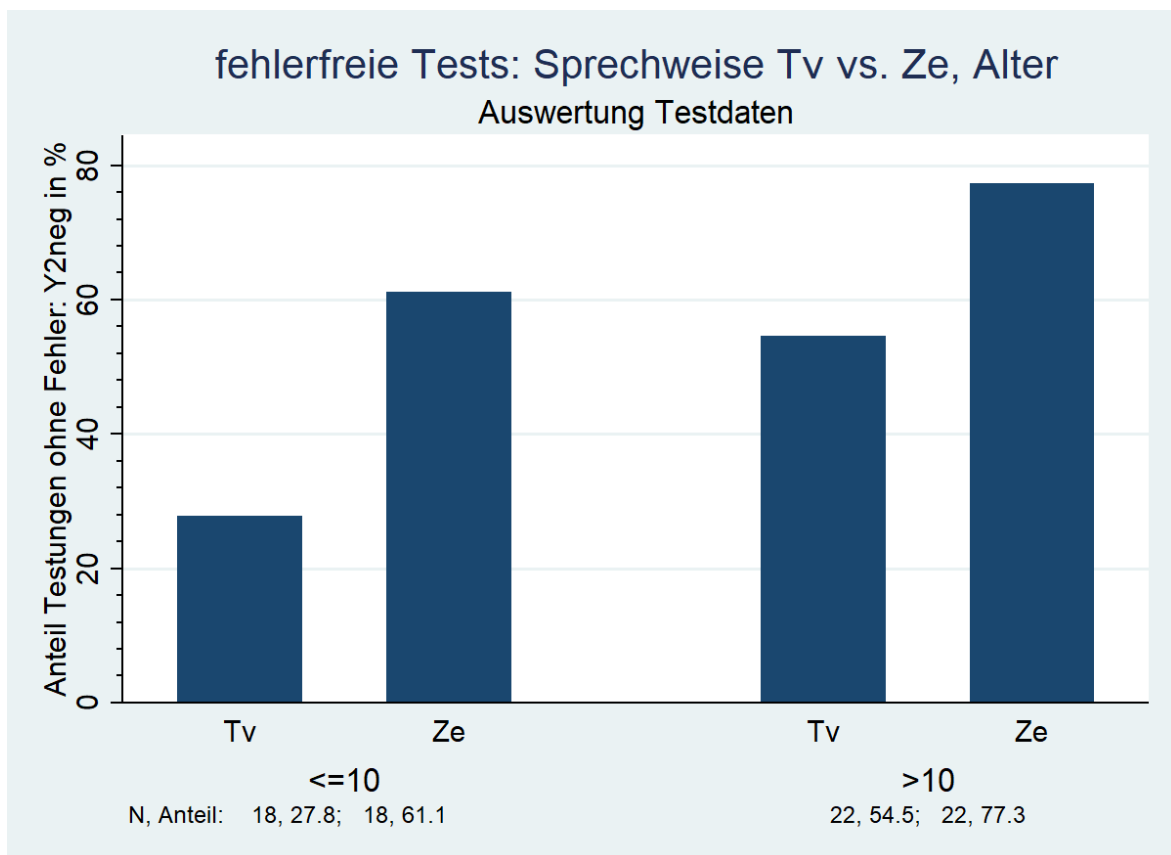


Abb. 12: Bar_Y2neg_Alt_kat_test.tif: Fehlerfreie Durchgänge / Auswertung der Kovariable Alter

5.3.2 Zufallskritische Analyse der Ergebnisse zur Kovariable Alter

Die zufallskritischen Analysen müssen für die Fehlerzahl mit xtpoisson und poisson („extended poisson“ für die Längsschnittanalyse und normale Poisson-Regression für den Vergleich der Höhe der Fehlerzahl zwischen den Altersgruppen) durchgeführt werden. Für den Anteil fehlerfreier Tests kommen xtlogit und logit (extended logistic für die Längsschnittanalyse und normale logistische Regression für den Vergleich der Häufigkeit fehlerfreier Durchgänge zwischen den Altersgruppen) zur Anwendung.

Die Koeffizienten sind Odds Ratios. Sie sind ähnlich zu relativen Risiken, d. h. der Normalwert ist 1, wenn kein Zusammenhang besteht und größer als 1, wenn der Anteil an fehlerfreien Testungen höher ist.

Die nachfolgenden Analysen beziehen sich jeweils auf die vorstehenden drei Abbildungen.

5.3.2.1 Analyse der Fehlerzahl (Kovariable Alter)

Daten zu Abb. 10: Bar_Y2_Alt_kat_test.tif

Alt_kat	Sprach_Modus		Total	
	traditionell	Verdreht zehneins		
<=10	18	18	36	
	0.00	0.00	0.00	
	1.50	0.00	1.00	
	3.33	1.44	2.39	1.44/3.33=0.43
	14.00	11.00	14.00	
>10	22	22	44	
	0.00	0.00	0.00	
	0.00	0.00	0.00	
	2.18	0.36	1.27	0.36/2.18=0.17
	9.00	2.00	9.00	0.17/0.43=0.39
Total	40	40	80	1.27/2.39=0.53
	0.00	0.00	0.00	

	1.00	0.00	0.00
	2.70	0.85	1.77
	14.00	11.00	14.00

Daten zu Abb. 11: Box_diff_Y2_Alt_kat_test.tif

Alt_kat	N(diff_Y2)	min(diff_Y2)	med(diff_Y2)	mean(diff~2)	max(diff_Y2)
<=10	18	-3.00	1.00	1.89	12.00
>10	22	-2.00	0.00	1.82	9.00
Total	40	-3.00	0.00	1.85	12.00

Y2	IRR	Std. Err.	z	P> z	[95% Conf. Interval]
X	.4333333	.1017441	-3.56	0.000	.2735049 .6865609
Xalter	.3846154	.1724182	-2.13	0.033	.1597511 .9259969

Y2	IRR	Std. Err.	z	P> z	[95% Conf. Interval]
X	.3148148	.0619082	-5.88	0.000	.2141247 .4628535
Alt_kat	.5271318	.0905149	-3.73	0.000	.3764936 .7380416
_cons	3.847556	.4527204	11.45	0.000	3.055124 4.845528

Die Analyse mit xtpoisson zeigt uns bei Xalter einen Koeffizienten von 0.38 = 38%. Die Reduktion der Fehlerzahl durch den Wechsel von Tv zu Ze ist also im Mittel kleiner bei den älteren Kindern; er beträgt nur 38% des Wertes der jüngeren. Dieser Unterschied ist signifikant: $p = 0.033 = 3.3\% < 5\%$. Beide

Altersgruppen zeigen also weniger Fehler bei Ze, aber der Effekt ist deutlich kleiner bei den älteren Kindern.

Die Poisson-Regression zeigt uns bei Alt_kat einen Koeffizienten von 0.53 = 53%. Die älteren Kinder machen also im Mittel weniger Fehler, ihre Fehlerzahl beträgt nur 53% der Fehlerzahl der jüngeren. Dieser Unterschied ist hochsignifikant, denn der p-Wert ist mit 0.000 angegeben, was $p < 0,0005$, also $p < 0,05\%$ bedeutet.

5.3.2.2 Analyse fehlerfreier Durchgänge (Kovariable Alter)

Daten zu Abb. 12: Bar_Y2negp_Alt_kat_test.tif

```

-----
      |
      |          Sprach_Modus
Alt_kat | traditionellVerdreht      zehneins      Total
-----+-----
<=10 |          18              18              36
      |          27.78          61.11          44.44
      |
>10 |          22              22              44
      |          54.55          77.27          65.91
      |
Total |          40              40              80
      |          42.50          70.00          56.25
-----

```

```

-----
      Y2neg |          OR      Std. Err.      z      P>|z|      [95% Conf. Interval]
-----+-----
      x |          4      3.162277      1.75      0.080      .849427      18.83622
Xalter | .8750001      .9852428      -0.12      0.906      .0962858      7.951588
-----

```

```

-----
      Y2neg | Odds Ratio      Std. Err.      z      P>|z|      [95% Conf. Interval]
-----+-----
      x | 3.596905      1.806801      2.55      0.011      1.343855      9.627318
-----

```

Alt_kat		2.834095	1.426718	2.07	0.039	1.056596	7.601854
_cons		.35559	.1640675	-2.24	0.025	.1439492	.8783948

Die Analyse mit xtlogit zeigt für Xalter einen p-Wert von $0.906 = 91\% > 5\%$. Es ist also kein Unterschied im Effekt des Sprechweisenwechsels auf den Anteil fehlerfreier Durchgänge zwischen jüngeren und älteren Kindern nachweisbar.

Die Auswertung mit logit zeigt bei Alt_kat einen Koeffizienten von 2.8, also eine Erhöhung gegenüber dem Normalwert von 1. Das bedeutet, dass der Anteil fehlerfreier Durchgänge unter den älteren Kindern höher ist. Diese Erhöhung ist signifikant: $p = 0.039 = 3.9\% < 5\%$.

5.3.3 Interpretation der Regressionsanalyse (Kovariable Alter)

Ältere Kinder machen signifikant weniger Fehler ($p = 3\%$) als jüngere. Es ergibt sich ein signifikant größerer relativer Vorteil (Rückgang der Fehler in %) von Ze bei den Älteren ($p = 3\%$). Der absolute Vorteil (Rückgang der Anzahl an Fehlern) ist jedoch ähnlich in beiden Altersgruppen. Beide Altersgruppen zeigen also weniger Fehler bei Ze, aber der Effekt ist bei den älteren Kindern deutlich kleiner als bei den jüngeren (Abb. 10: Bar_Y2_Alt_kat_test und Abb.11: Box_diff_Y2_Alt_kat_test).

In Bezug auf die fehlerfreien Durchgänge ergibt sich ein nicht-signifikant höherer Anteil unter den Älteren ($p = 9\%$) und ein fast identischer Vorteil von Ze in beiden Altersgruppen ($p = 90\%$), der Vorteil ist also unabhängig vom Alter (Abb. 12: Bar_Y2negp_Alt_kat_test).

Zusammengefasst ergibt die Regressionsanalyse, dass unter den älteren Kindern die Fehlerzahl signifikant niedriger und der Anteil fehlerfreier Durchgänge erhöht ist. Jedoch zeigen beide Altersgruppen – sowohl die Jüngeren als auch die Älteren – eine Verbesserung mit Ze: Es werden weniger Fehler gemacht und es gibt mehr fehlerfreie Durchgänge bei Verwendung einer stellenwertgerechten Sprechweise, wobei die Reduktion der Fehlerzahl bei den älteren statistisch signifikant weniger ausgeprägt ist.

Diese Ergebnisse waren zu erwarten; sie sprechen somit für die Methodik.

5.3.4 Ergebnisse der Auswertung zur Kovariable Klassenstufe

Die Auswertungen zu den Klassenstufen (Kovariable 7) untermauern den Befund, der sich schon in Bezug auf das Alter zeigte: In höheren Klassen ereignen sich weniger Fehler. Der Unterschied ist mit $p < 0.05\%$ hochsignifikant. Wie auf der Grafik (Abbildung 13) zu erkennen, gibt es einen ähnlichen Vorteil von Ze in allen Klassenstufen.

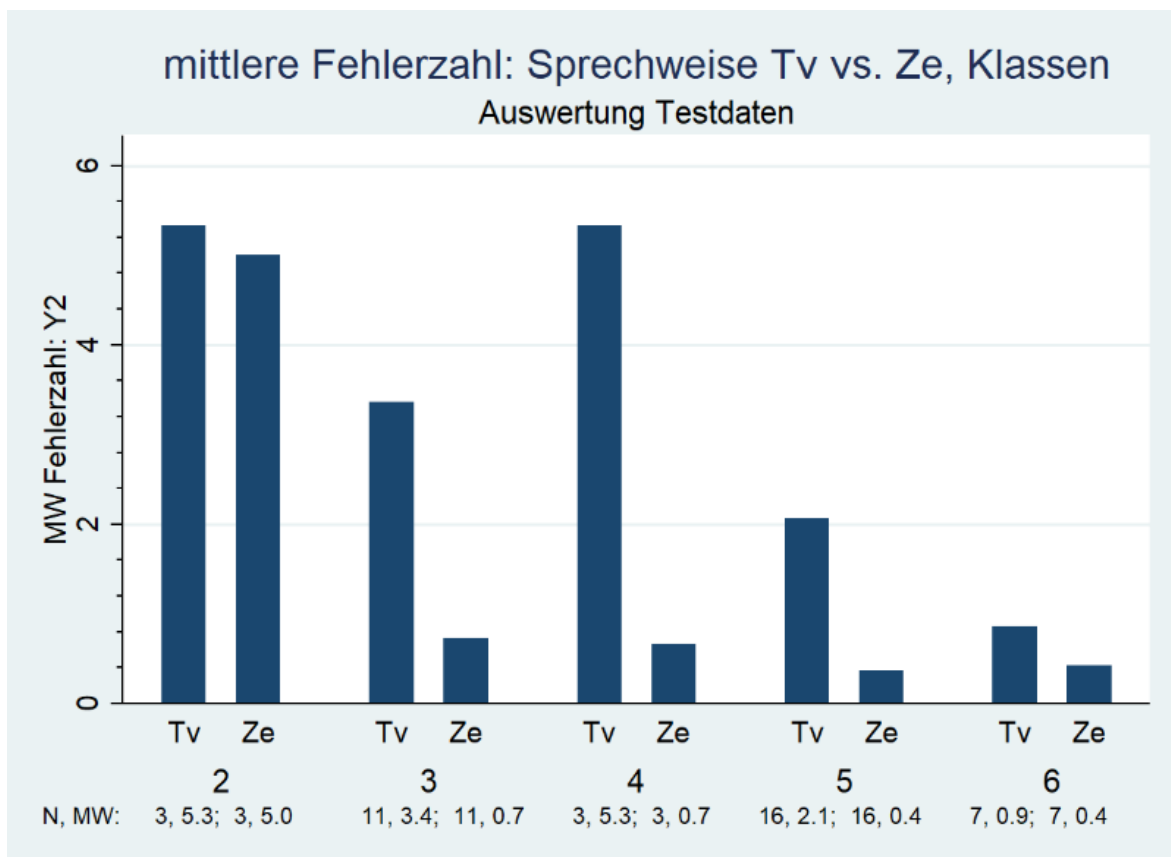


Abb. 13: Bar_Y2_Klassen_test: Mittlere Fehlerzahl / Auswertung der Kovariable Klassenstufen

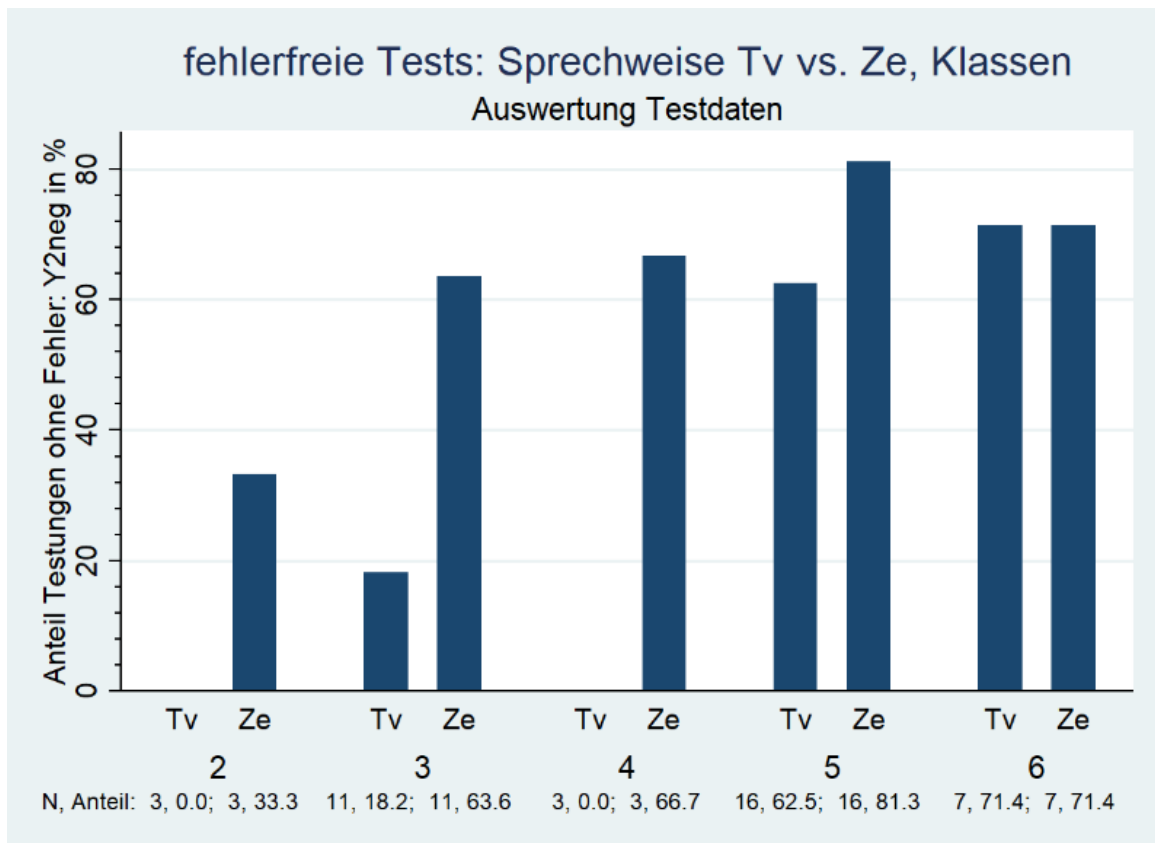


Abb. 14: Bar_Y2neg_Klassen_test: Fehlerfreie Tests / Auswertung der Kovariable Klassenstufen

Abbildung 14 zeigt, dass mit steigender Klassenstufe die Anzahl der fehlerfreien Durchgänge tendenziell steigt. Bei den 2.- und 4.-Klässlern (jeweils 3 Kinder) konnten mit inverser Sprechweise keine fehlerfreien Durchgänge dokumentiert werden. Mit stellenwertgerechter Sprechweise gelangen 33.3% fehlerfreie Durchgänge bei den 2.-Klässlern, 63.6 % bei den 3.-Klässlern, 66.7% bei den 4.- Klässlern und 81.3 % bei den 5.-Klässlern. Der Vorteil von Ze ist in der Grafik gut zu erkennen; einzig für die 6. Klassenstufe trifft dies nicht zu, hier beträgt der Anteil fehlerfreier Durchgänge bei beiden Sprechweisen 71.4 %.

Es sind keine p-Werte bestimmbar.

5.4 WEITERE BEFUNDE UNTER BERÜCKSICHTIGUNG DER KOVARIABLEN

Wie schon im Basisbefund dokumentiert, ergibt sich auch in allen Aufgliederungen nach den Kovariablen stets ein Vorteil bei Verwendung einer stellenwertgerechten Zahlensprechweise.

Dieser Vorteil von Ze zeigt sich geräteunabhängig (Kovariable 2: *Geräte*): Es können keine Unterschiede in der Fehlerzahl festgestellt werden. Auch die Anteile fehlerfreier Durchgänge sind ähnlich ($p > 80\%$), unabhängig davon, ob mit dem iPhone, dem Lenovo Desktop Tablet oder dem PEAQ Tablet getestet wurde. Für diese Testungen sind keine p-Werte bestimmbar.

Im Folgenden werden weitere ausgewählte Befunde zur Fehlerzahl (Y2) und den fehlerfreien Durchgängen (Y2neg) in Abhängigkeit von den untersuchten Kovariablen zusammengestellt.⁵

⁵ Aus methodischer Sicht ist die Bestimmung der p-Werte und Konfidenzintervalle numerisch schwierig, da die Skalen der Variablen eingeschränkt sind (Y2: es sind nur natürliche Zahlen möglich, Y2neg: es ist nur 0 oder 1 möglich). Daher müssen nicht-lineare Methoden eingesetzt werden. Es ist deshalb möglich, dass in einigen Situationen keine p-Werte bestimmbar sind.

5.4.1 Ergebnisse der Auswertung zur Kovariable Fördereinrichtung

Der Vorteil von Ze zeigt sich in allen Einrichtungen, in denen Testungen vorgenommen wurden (Lernpraxis, ZSPR, Duden-Institut).

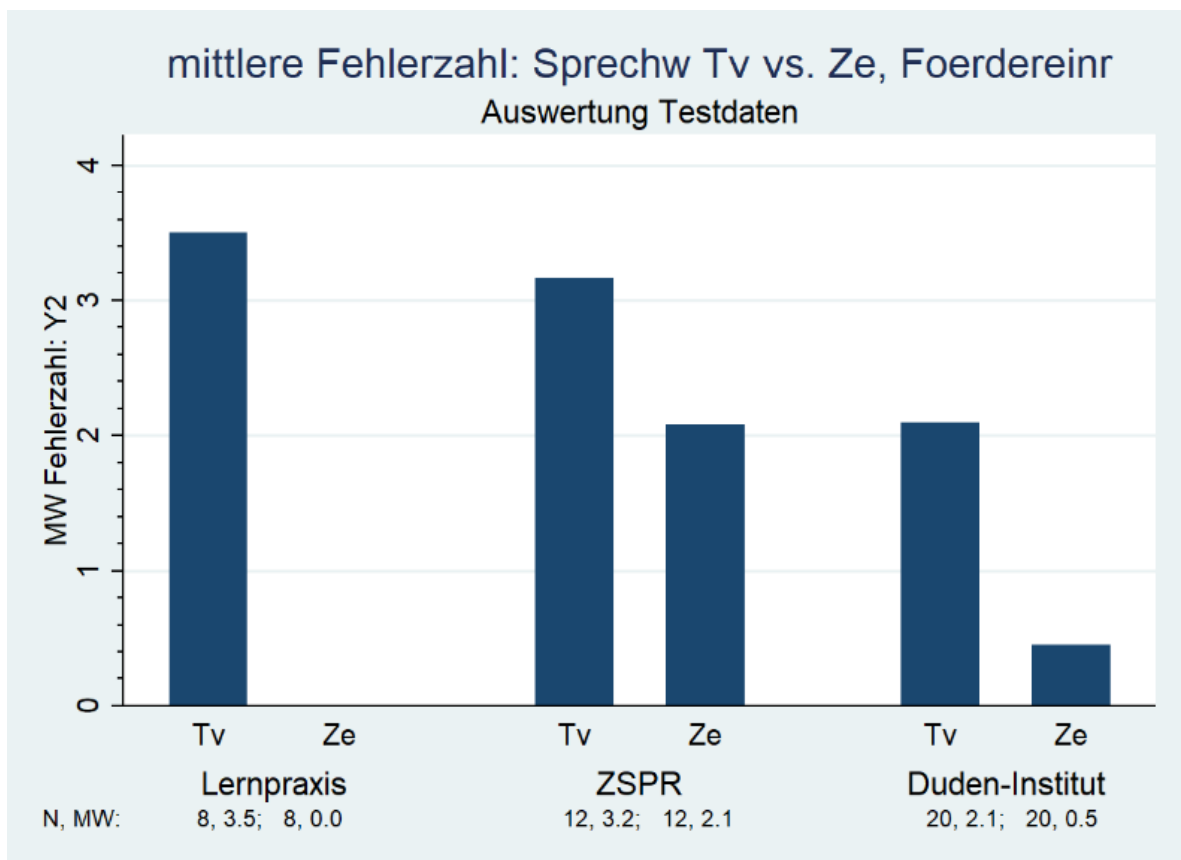


Abb. 15: Bar_Y2_Foerder-Einrichtung_test: Mittlere Fehlerzahl / Auswertung der Kovariable Fördereinrichtung

Die 8 Probanden der Lernpraxis machten bei stellenwertgerechter Sprechweise keine Fehler, bei inverser Sprechweise hingegen im Mittel 3.5 Fehler pro Durchgang. An den Duden-Instituten wurden die wenigsten Fehler gemacht (grenzwertig signifikant: $p = 6\%$). Das ZSPR zeigt den kleinsten Vorteil bei Ze (kein p -Wert bestimmbar, denn die Schätzwerte divergieren).

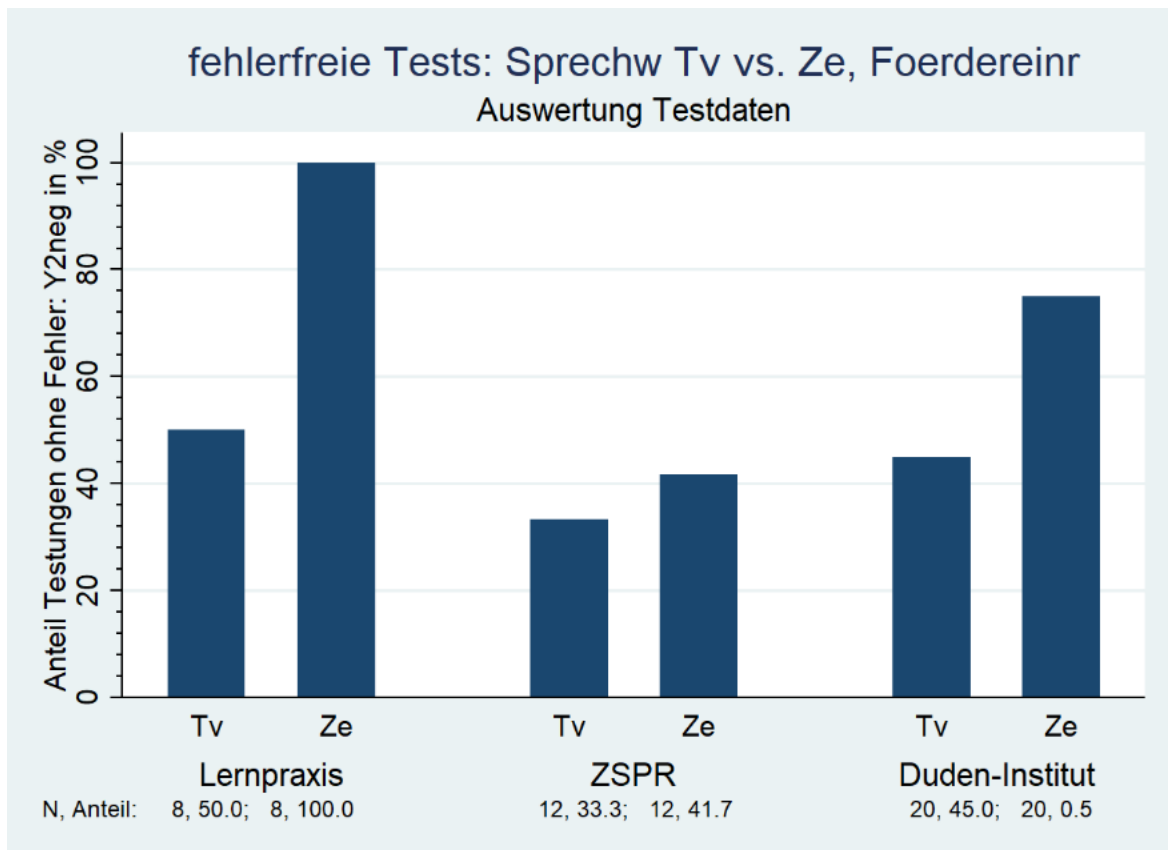


Abb. 16: Bar_Y2negp_Foerder-Einrichtung_test: Fehlerfreie Tests / Auswertung der Kovariable Fördereinrichtung

Bei allen Einrichtungen zeigen sich ähnliche Anteile fehlerfreier Durchgänge ($p > 70\%$). Der Vorteil von Ze ist in allen drei Einrichtungen gegeben (kein p-Wert bestimmbar).

5.4.2 Ergebnisse der Auswertung zur Kovariable Geschlecht

Untersucht wurde, ob sich eine Modifikation des Effekts, bedingt durch das Geschlecht, feststellen lässt.

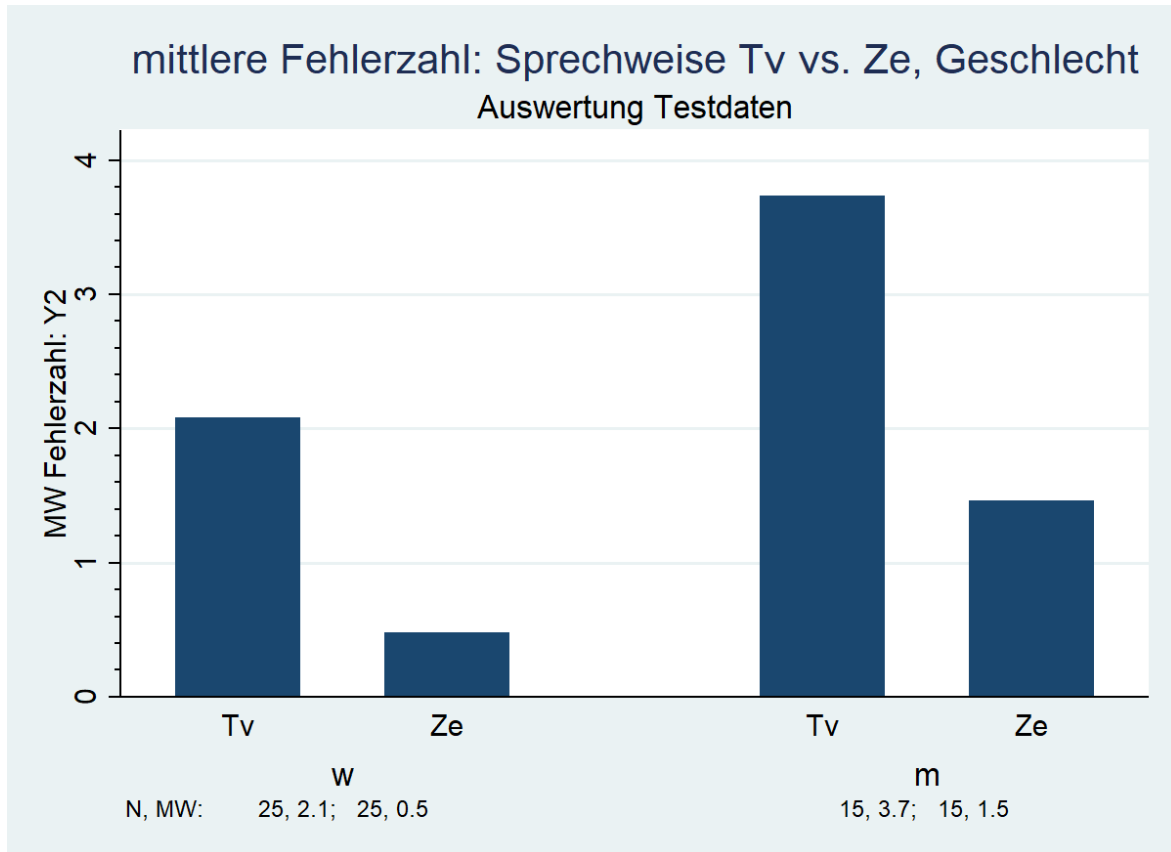


Abb. 17: Bar_Y2_Geschl_test: Mittlere Fehlerzahl / Auswertung der Kovariable Geschlecht

Die Abbildung zeigt: Mädchen machen signifikant weniger Fehler ($p = 0.2\%$): Bei Tv unterlaufen ihnen im Mittel 2.1, den Jungen 3.7 Fehler pro Durchgang. Mit Ze sind es 0.5 bei den Mädchen und 1.5 bei den Jungen.

Bei einem p-Wert von ca. 20% ergibt sich ein ähnlicher Vorteil einer stellenwertgerechten Sprechweise in beiden Geschlechtern.

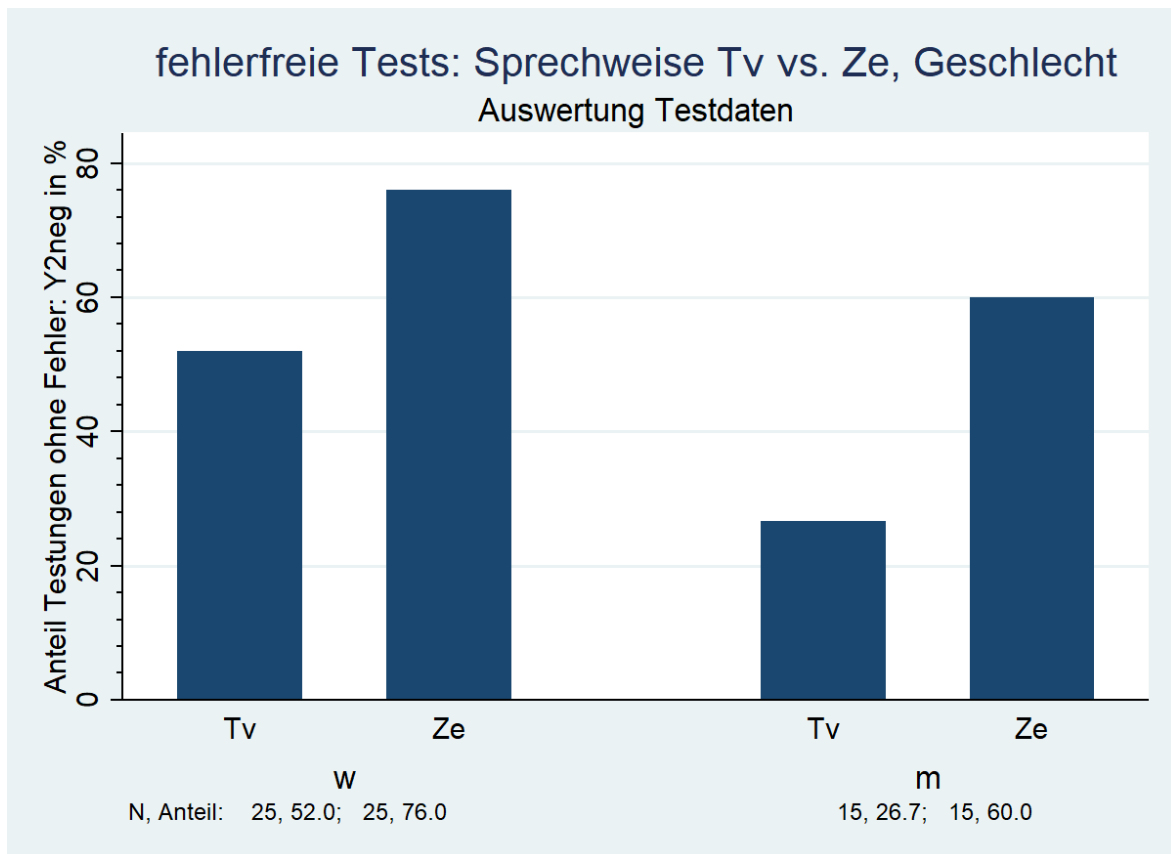


Abb. 18: Bar_Y2neg_Geschl_test: Fehlerfreie Durchgänge / Auswertung der Kovariable Geschlecht

Wie in der Abbildung zu sehen, zeigen Mädchen auch häufiger fehlerfreie Durchgänge. Eine statistische Signifikanz wurde jedoch nicht erreicht ($p = 12\%$). Bei einem zweiseitigen p -Wert von ca. 90% ergibt sich ein ähnlicher Vorteil von Ze bei Jungen und Mädchen.

5.4.3 Ergebnisse der Auswertung zur Kovariable Erstsprache

Die Kovariable *Erstsprache* wurde erhoben, um festzustellen, ob es einen Unterschied in der Fehleranzahl oder Anzahl fehlerfreier Durchgänge in Abhängigkeit von der (oder den) Erstsprache(n) des Kindes gibt (Kovariable 8). In der Abbildung 19 bedeutet Erstsprache „ja“, dass den Kindern eine stellenwertgerechte Sprechweise aus dem Gebrauch ihrer Erstsprache(n) bekannt ist (Erstsprache mit „Zwanzigeins-Struktur“). Diesen 9 Kindern unterliefen mit Tv im Durchschnitt 4.3 Fehler, mit Ze hingegen nur 2.0 Fehler. Den anderen 31 Kindern (ohne Vorkenntnisse einer stellenwertgerechten Sprechweise = Erstsprache „nein“) unterliefen mit Tv 2.2 und mit Ze lediglich 0.5 Fehler. Kindern, die durch ihre Erstsprache mit der „Zwanzigeins-Struktur“ vertraut sind, unterlaufen demnach signifikant mehr Fehler ($p = 0.1\%$).

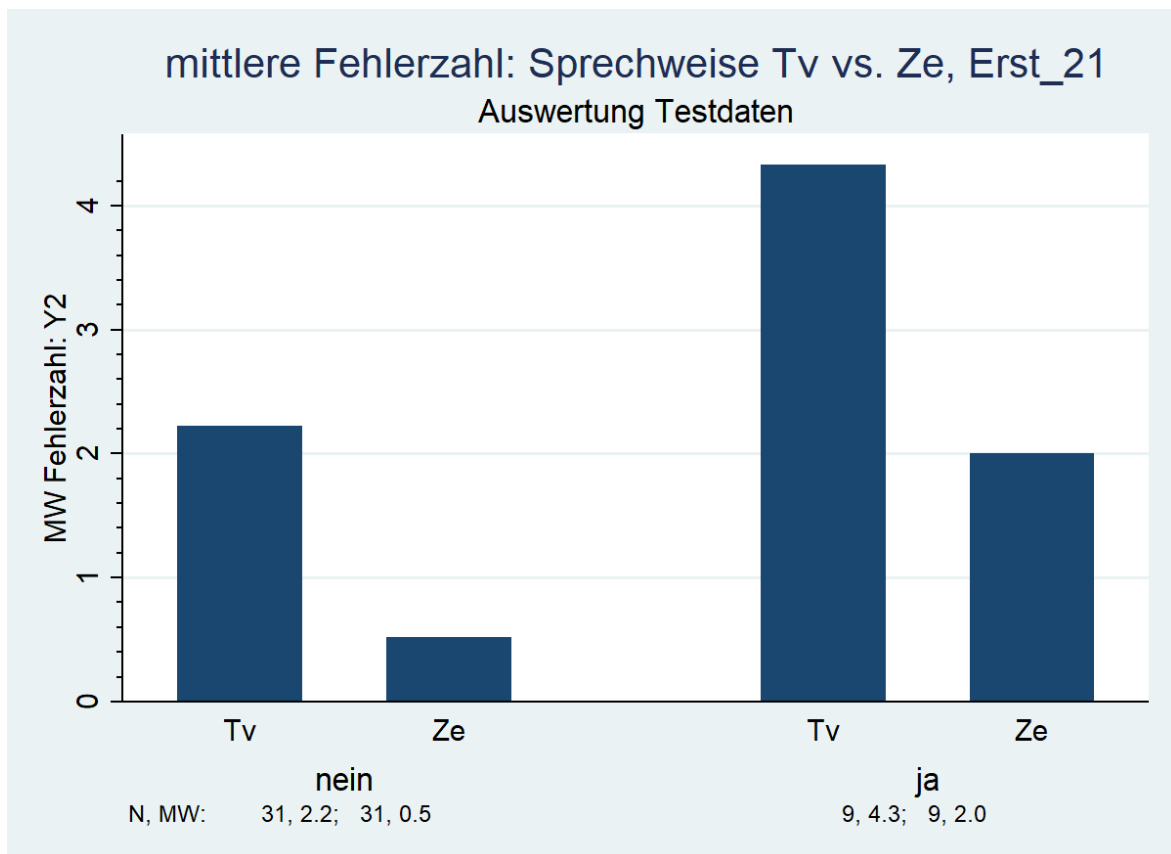


Abb. 19: Bar_Y2_Erst_21_test: Mittlere Fehlerzahl / Auswertung der Kovariable Erstsprache

Es ergibt sich jedoch ein ähnlicher absoluter Vorteil von Ze in beiden Sprachgruppen, also auch bei den Kindern ohne Erstsprache mit „Zwanzigeins-Struktur“ (relativer Vorteil: $p > 8\%$). Insofern ist der Vorteil einer

stellenwertgerechten Sprechweise bezüglich der Fehlerzahl unabhängig von der Erstsprachenstruktur.

Abbildung 20 zeigt, dass den Kindern, die Ze aus ihrer Erstsprache kennen (Erstsprache „ja“) weniger fehlerfreie Durchgänge gelangen: Der Anteil fehlerfreier Durchgänge betrug bei Tv 33.3%, bei Ze 55.6%. Die anderen Kinder (Erstsprache „nein“) erreichten bei Tv 45.2% und bei Ze 74.2%. Bei einem p-Wert von $> 50\%$ sind die Unterschiede bei Tv jedoch deutlich nicht-signifikant. Es ergibt sich zudem ein ähnlicher Vorteil der stellenwertgerechten Sprechweise in beiden Sprachgruppen ($p > 35\%$); der Vorteil von Ze ist bezüglich der Anzahl fehlerfreier Durchgänge also weitgehend unabhängig von der Erstsprachenstruktur.

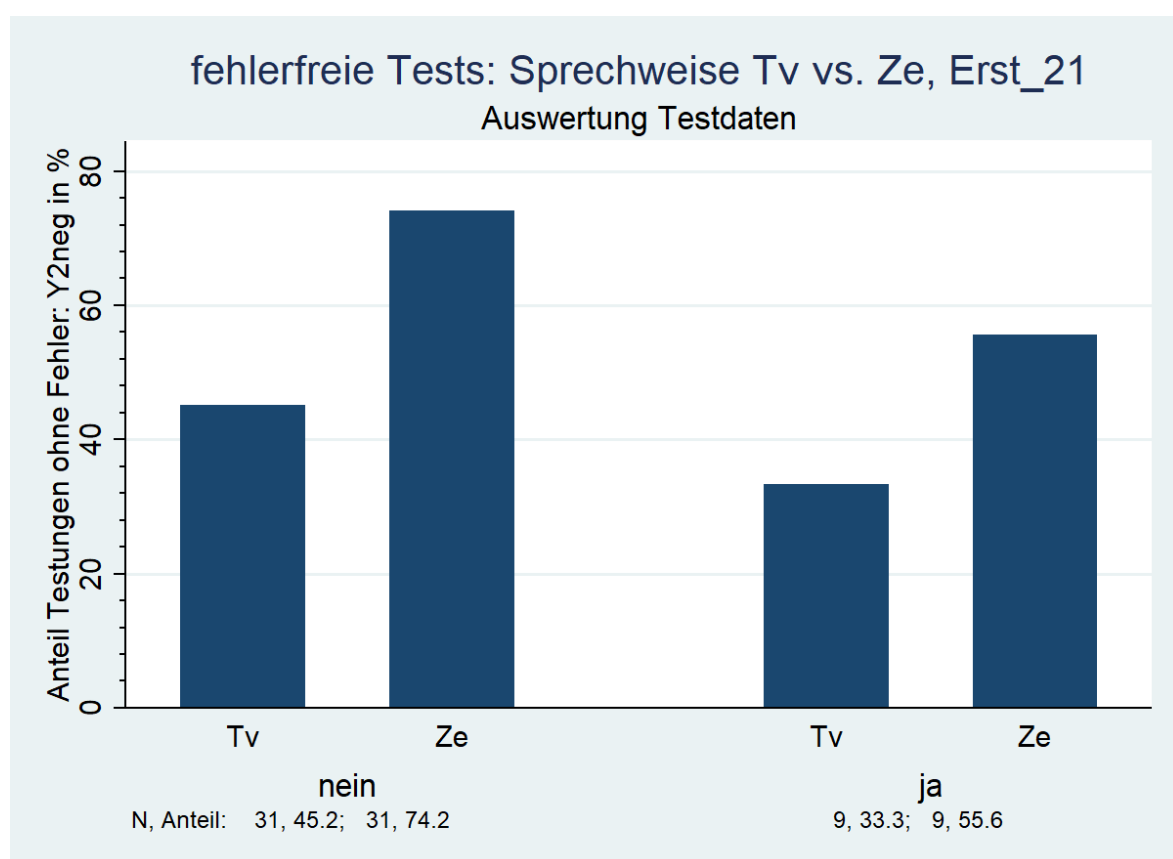


Abb. 20: Bar_Y2neg_Erst_21_test: Fehlerfreie Durchgänge / Auswertung der Kovariable Erstsprache

Zusammenfassend lässt sich festhalten, dass insbesondere Kinder von einer stellenwertgerechten Sprechweise profitieren, die eine Erstsprache mit Zwanzigeins-Struktur haben, da sie mit Hilfe von Ze ihre erhöhte Fehlerhäufigkeit bei Tv auf das Tv-Niveau der Kinder reduzieren können, deren Erstsprache keine Zwanzigeins-Struktur hat. Der Vorteil beschränkt sich jedoch nicht auf diese Kinder:

Auch Kinder, die diese Erstsprachenstruktur nicht haben, profitieren deutlich von einer stellenwertgerechten Zahlensprechweise (Reduktion der Fehler um $1 - 0.5/2.2 = 78\%$, Erhöhung der Anzahl der fehlerfreien Bearbeitungen um $74.2/45.2 - 1 = 64\%$).

5.4.4 Ergebnisse der Auswertung zur Kovariable LRS

Kinder, die zusätzlich zur Rechenschwäche unter einer Lese-Rechtschreib-Störung (Kovariable 10) leiden, machten, wie Abbildung 21 zeigt, mehr Fehler.

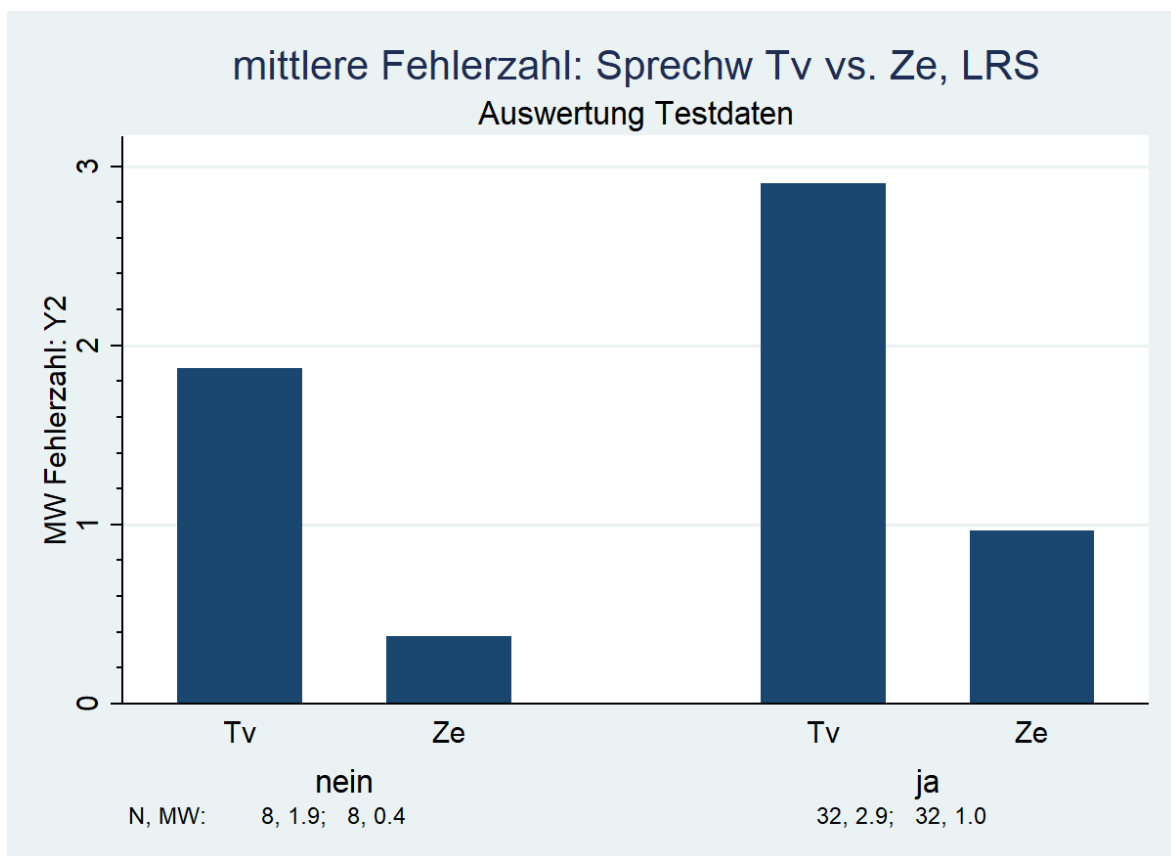


Abb. 21: Bar_Y2_LRS_test: Mittlere Fehlerzahl / Auswertung der Kovariable LRS

Den 8 Probanden ohne zusätzliche LRS-Diagnose unterliefen im Mittel mit Tv 1.9 Fehler pro Durchgang, bei Ze waren es 0.4 Fehler.

Bei den 32 Kindern mit LRS wurden im Mittel 2.9 Fehler mit Tv gemacht und 1.0 mit Ze. Die Ergebnisse weisen bei $p > 10\%$ jedoch keine statistische Signifikanz auf. Der Vorteil von Ze zeigt sich in beiden Gruppen ($p > 40\%$).

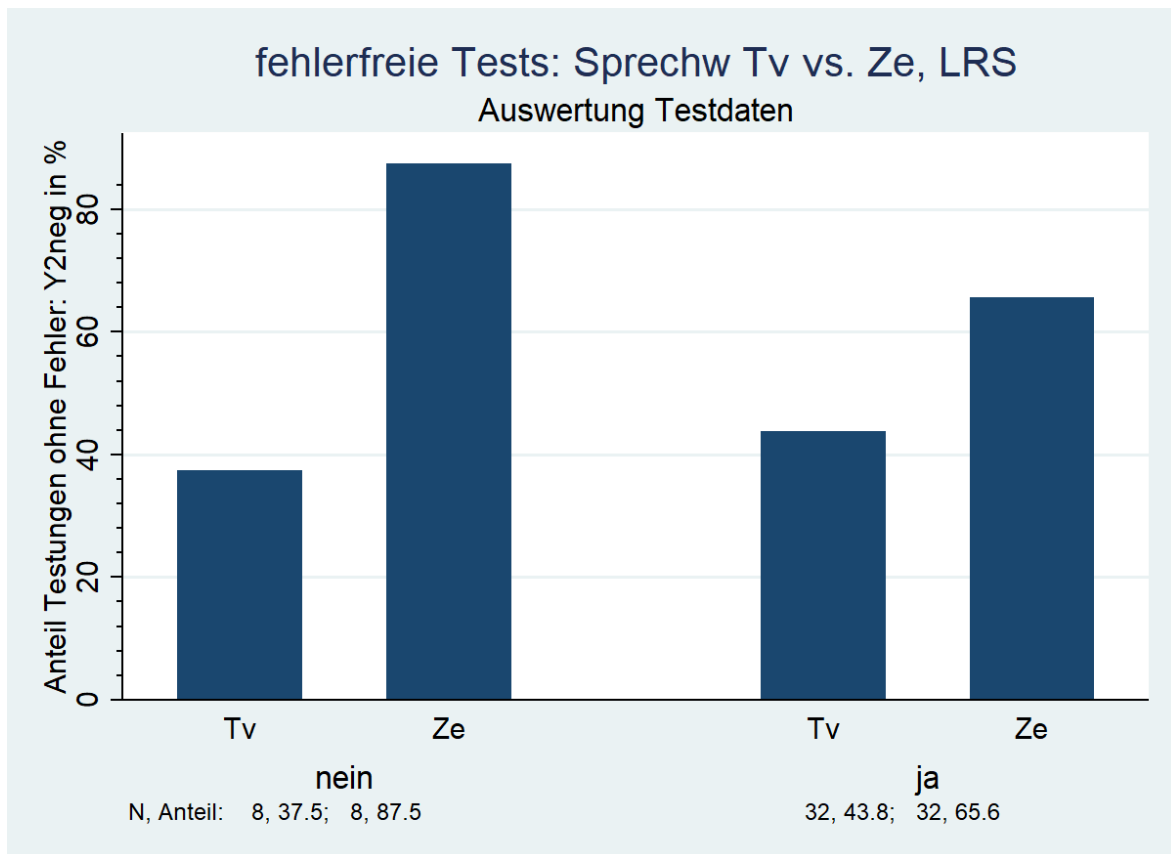


Abb. 22: Bar_Y2negp_LRS_test: Fehlerfreie Tests / Auswertung der Kovariable LRS

Der Anteil fehlerfreier Durchgänge lag bei den Kindern ohne LRS bei 37.5% mit Tv vs. 87.5% mit Ze. Sie hatten somit etwas weniger fehlerfreie Durchgänge als die Kinder mit LRS (43.8% bei Tv vs. 65.6% bei Ze). Auch hier liegt mit $p > 70\%$ keine statistische Signifikanz vor.

Ein ähnlicher Vorteil von Ze ergibt sich in beiden Gruppen, er ist etwas schwächer bei Kindern mit LRS ($p > 70\%$) und somit weitgehend unabhängig vom Vorliegen einer LRS.

5.4.5 Ergebnisse der Auswertung zur Kovariable Sequenz

Von Interesse war, ob sich ein Unterschied in der Sprechweise in Abhängigkeit vom gewählten Startmodus feststellen lässt (Kovariable 4: Sequenz).

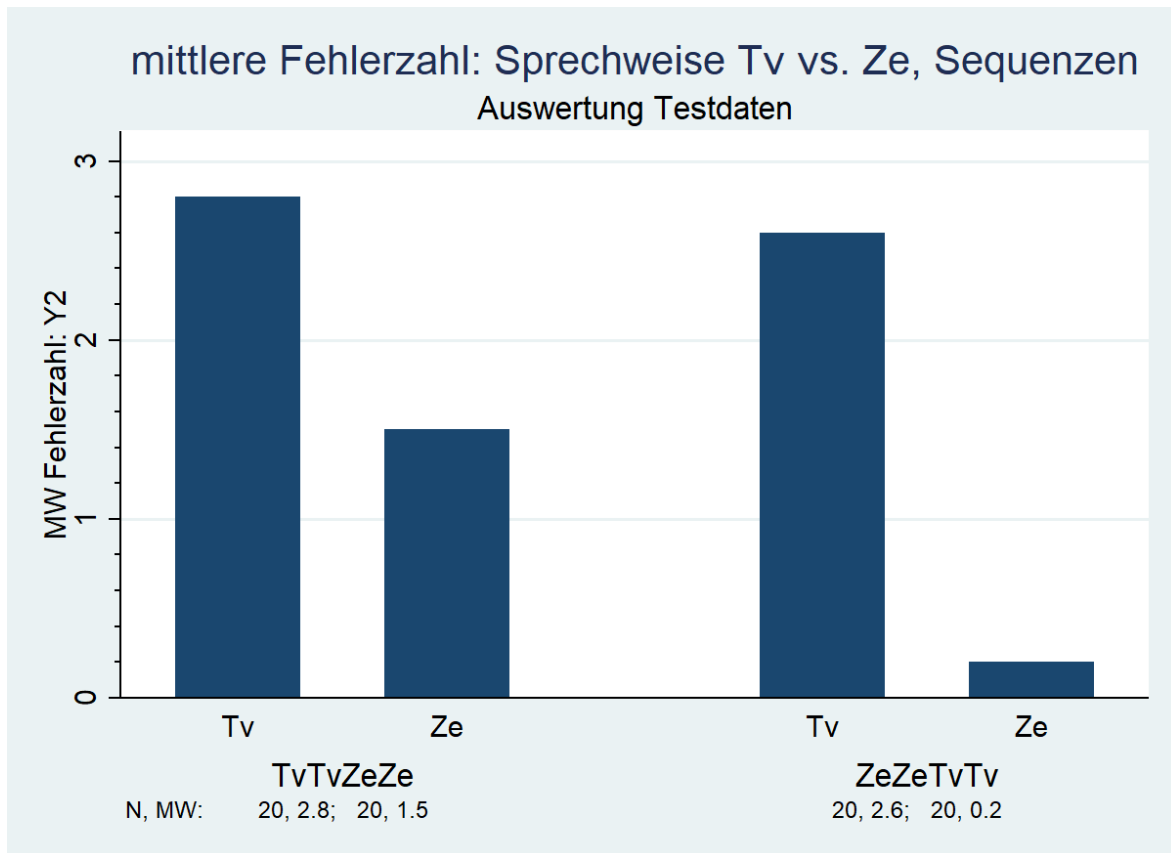


Abb. 23: Bar_Y2_sequ_test: Fehlerfreie Tests / Auswertung der Kovariable Sequenz

Für beide Sequenzen ergibt sich zwar ein Vorteil hinsichtlich einer verringerten Fehlerzahl bei Ze, aber signifikant ausgeprägter ist dieser Vorteil in der Sequenz ZeZeTvTv, also bei Start mit stellenwertgerechter Sprechweise. Der Vorteil von Ze ist also nicht homogen.

Der zugehörige zweiseitige p-Wert für einen Unterschied zwischen den Sequenzen ist 0.1%. Der Vorteil durch Ze ist somit abhängig von der Sequenz.

5.4.6 Ergebnisse der Auswertung zur Kovariable Schulwechsel

Wie bei der Sequenz zeigte sich auch bei der Auswertung zur Kovariable Schulwechsel (12) ein nicht homogener Vorteil von Ze.

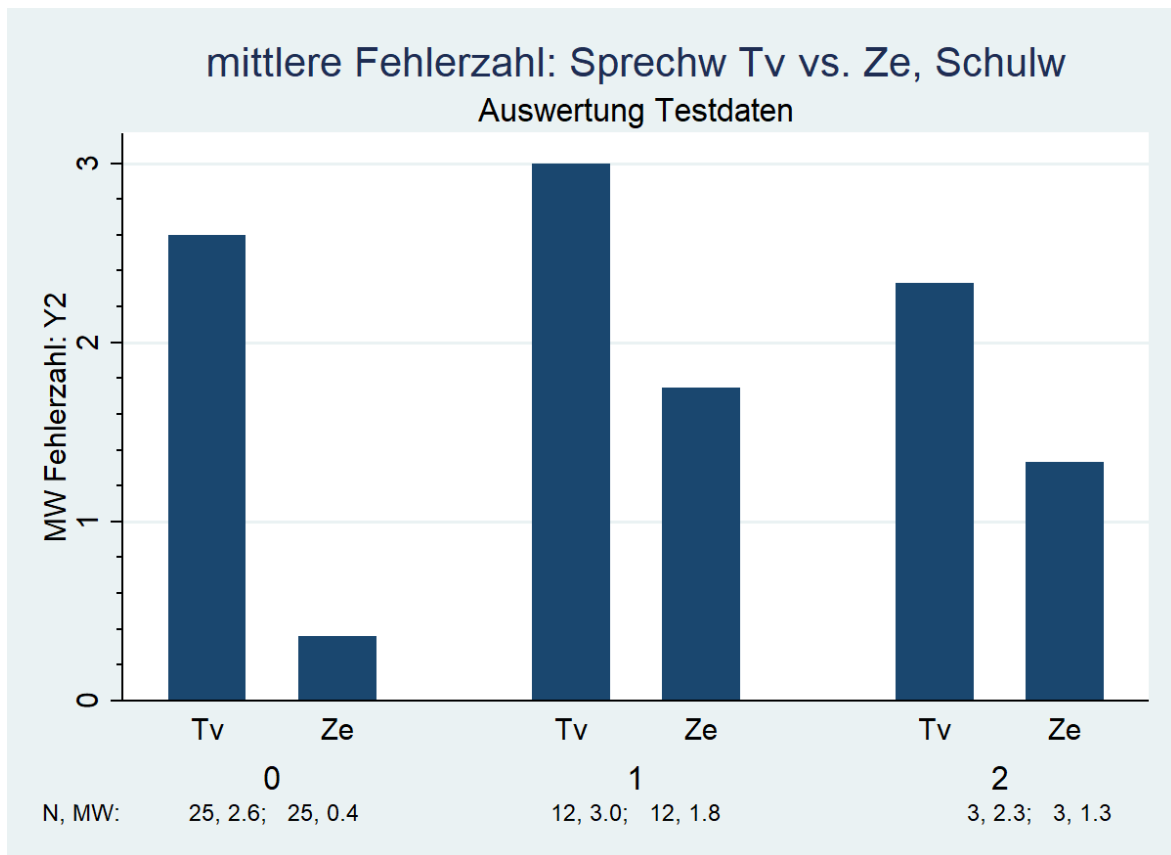


Abb. 24: Bar_Y2_Schulwechsel_test: Mittlere Fehlerzahl / Auswertung der Kovariable Schulwechsel

Der Beschreibung liegt die Unterteilung in zwei Gruppen zu Grunde: Kinder ohne Schulwechsel und Kinder mit 1 und 2 Schulwechseln (die Grafik hingegen zeigt die Aufteilung in 0, 1 oder 2 Schulwechsel). Mit und ohne Schulwechsel gibt es einen Vorteil für Y2 bei stellenwertgerechter Sprechweise. Die Fehlerzahl bei inverser Sprechweise ist in beiden Gruppen ähnlich ($p > 60\%$). Aber der Vorteil von Ze ist signifikant ausgeprägter bei Kindern ohne Schulwechsel ($p < 3\%$). Der Vorteil von Ze ist also abhängig davon, ob ein Kind die Schule gewechselt hat.

5.4.7 Interpretation der Befunde nach Auswertung der Kovariablen

Aus den Auswertungen ergibt sich zusätzlich zum Basisbefund, dass keine der untersuchten 15 Variablen den nachgewiesenen Zusammenhang so modifiziert, dass sich nicht immer ein Vorteil bei Verwendung einer stellenwertgerechten Sprechweise ergibt. Der Vorteil durch Verwendung von Ze ist also nicht auf eine spezielle Teilgruppe wie Mädchen oder Kinder, die häufige Schulwechsel erlebt haben, beschränkt, sondern übergreifend für alle Teilgruppen ermittelbar: Alle Gruppen profitieren von einer Umstellung auf eine stellenwertgerechte Zahlensprechweise. So machen z. B. Mädchen signifikant weniger Fehler (Y2) und haben mehr fehlerfreie Durchgänge (Y2neg) als Jungen, aber dennoch profitieren beide Untergruppen – Jungen und Mädchen – in beiden Variablen (Y2, Y2neg) von einer Umstellung auf Ze.

Allerdings ist zu berücksichtigen, dass sich die durchgeführte Poweranalyse (siehe Kap. 4: 40 Kinder) auf die Beantwortung der zentralen Forschungsfragen bezog und in dieser Hinsicht offensichtlich zielführend war. Aber für die Testung von Interaktionshypothesen (Kovariableneffekte) sind deutlich größere Umfänge indiziert, so dass es nicht verwundert, dass diese Testungen auch bei durchaus plausiblen Ergebnismustern, wie etwa beim Alter, nicht signifikant wurden.

Der Vorteil einer stellenwertgerechten Sprechweise ist nur in zwei Fällen nicht homogen, wie oben gezeigt wurde. Diese Ergebnisse sind qualitativ übereinstimmend mit den Befunden zu den Dauer-Variablen (Kuhl 2025). Es konnte insgesamt ein sehr überzeugender Gesamtbefund zum Vorteil einer stellenwertgerechten Zahlensprechweise ermittelt werden.

5.5 BASISBEFUND BEI AUSWERTUNG ALLER PERIODEN

Zum Abschluss soll im Sinne der Qualitätskontrolle der vorliegenden Studie (Baur & Blasius 2022a, S. 10) noch überprüft werden, ob sich der ermittelte Befund eines Vorteils der stellenwertgerechten Sprechweise nur bei einer Auswertung der Testdaten (Perioden 2 und 4) zeigt oder auch bei einer Analyse aller Perioden 1 bis 4 ermittelbar ist. Eine gemeinsame Auswertung aller Perioden hat den Vorteil, dass der Umfang größer ist und keine Vorab-Selektion von Perioden stattfindet, aber den Nachteil, dass die Ergebnisse von dem erheblichen Lernfortschritt an der App überlagert sind, der – wie nachgewiesen – bei einer Konzentration auf die Testdaten ausgeschlossen werden kann.

Wie die Abbildungen 25 und 26 zeigen, ergeben sich auch bei gemeinsamer Auswertung aller 4 Perioden einheitlich Vorteile für Ze: Die mittlere Fehlerzahl Y2 beträgt 2.8 Fehler bei Tv, dagegen nur 1.3 Fehler bei Ze bei $p < 0.05\%$ (ermittelt mit $xtpoisson$).

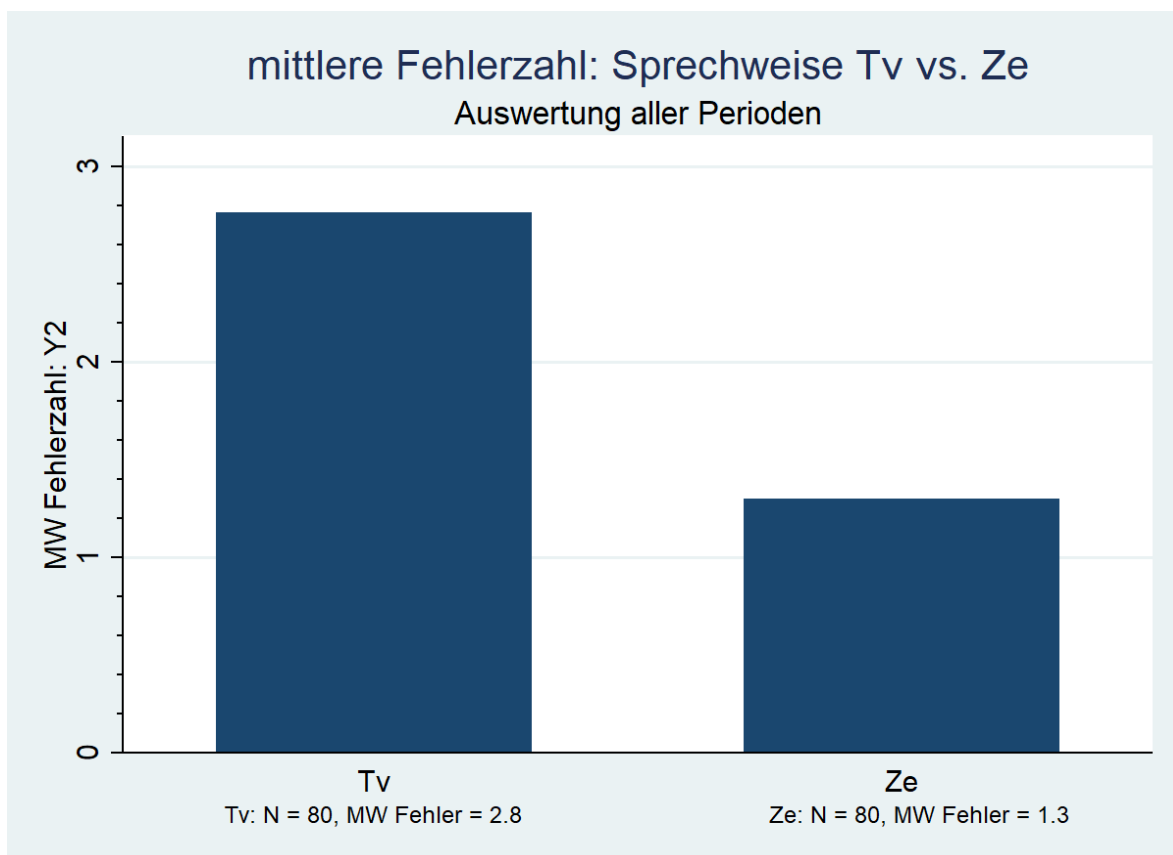


Abb. 25: Bar_Y2_all.tif: Mittlere Fehlerzahl / Auswertung aller Perioden

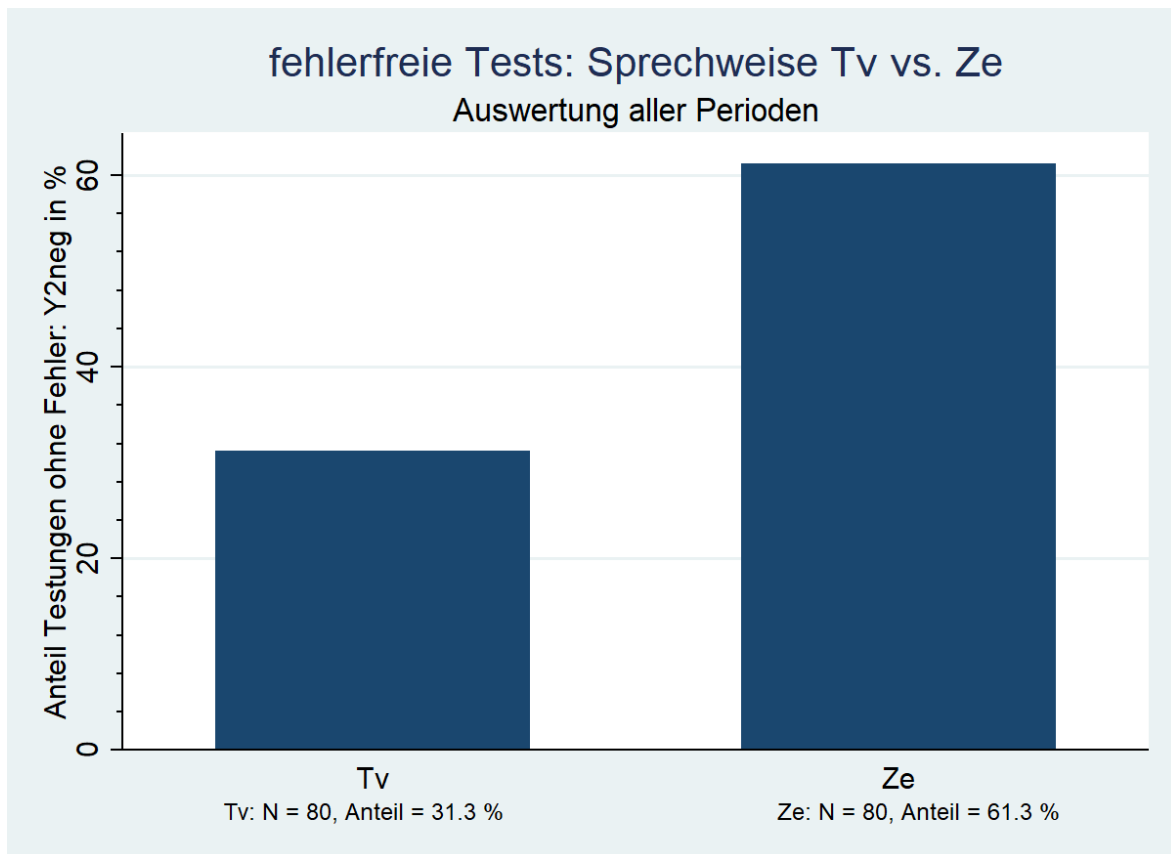


Abb. 26: Bar_Y2neg_all.tif: Fehlerfreie Tests / Auswertung aller Perioden

Der Anteil fehlerfreier Tests beträgt bei Tv 31.3% und bei Ze 61.3% bei einem zweiseitigen p-Wert unter 0.05% (ermittelt mit xtlogit).

Die Vorteile durch Ze sind auch bei Auswertung aller Perioden hochsignifikant.

5.5.1 Auswertung aller Perioden: Interpretation

Für die beiden Fehlervariablen ergeben sich, wie dargestellt, hochsignifikante Vorteile durch eine stellenwertgerechte Sprechweise. Damit ist der an den Testdaten (2. und 4. Periode) ermittelte Basisbefund auch an dem gesamten Datensatz (alle Perioden) ermittelbar, trotz der Verunklarung durch die Überlagerung mit Lernfortschritten an der App.

6 REFLEXIONEN DER KINDER

Die Kinder der Studiengruppe 1 (n = 20) zeigten sich fast durchgehend offen gegenüber der Zwanzigeins-Hypothese, so dass einige bereits während der kurzen Einweisungsphase den Sachverhalt reflektierten und kommentierten. Ihre Reflexionen wurden jeweils im Anschluss an jede einzelne Testung als Gedächtnisprotokolle notiert. Die offene Befragung (Baur & Blasius 2022, S. 20) erfolgte regulär im Anschluss an den Test, so dass die Kinder noch unter dem Eindruck ihrer neuen Erfahrung mit Ze standen und möglichst viele Gedanken äußern konnten. Suggestives Fragen wurde dabei strikt vermieden.

Einige Kinder waren sehr kommunikativ und interagierten, wie oben erwähnt, bereits während der Einweisungsphase, was unter der Prämisse eines Gesprächsangebots zugelassen wurde, um ein möglichst umfassendes Stimmungsbild zu erhalten und ggf. auch Meinungsänderungen abzubilden. Andere Kinder reagierten mit Zurückhaltung auf die Testsituation und waren weniger mitteilsam, was auf Grund der spezifischen psychischen Belastungen und Diagnosen erwartbar war.

Durch die gewohnte Umgebung im Therapieraum und – sofern von dem Kind gewünscht im Beisein der therapeutischen Fachkraft – wurde darauf hingewirkt, dass die Kinder sich wohl fühlten und frei äußern konnten. Eine vertrauensvolle Atmosphäre wurde als wichtig erachtet, um die Kinder zum Aussprechen ihrer Gedankengänge zu ermutigen. Die Testanweisung erfolgte gemäß Gesprächsleitfaden (s. Anhang 1) und umfasste ca. 2-5 Minuten, je nach Interaktion. Die Kinder hatten also nur ein sehr kurzes Zeitfenster, um die Thematik der inversen Zahlsprechweise zu verstehen und die alternative Sprechweise Ze direkt im Anschluss umzusetzen und am Tablet auszuprobieren. Es ist hervorzuheben, dass es bei keinem Kind weiterer Erläuterung oder einer ausführlicheren Einweisung als der im Leitfaden angegebenen bedurfte. Die Frage nach der Ze-Sprechweise der Zahl 14, die im Rahmen der Testanweisung gestellt wurde, beantworteten alle Kinder korrekt. Während des Tests wurden keine Hinweise gegeben.

Der Zeitpunkt der Äußerungen wurde in der nachfolgenden Dokumentation jeweils festgehalten, um den Prozess der Meinungsbildung zu dokumentieren, sofern Kinder vor oder während der Testung mit der Untersucherin interagierten. Zudem wurden die Gesprächsprotokolle um die statistischen Werte Eingabedauer (Y1) und

Fehlerzahl (Y2) ergänzt, um zu untersuchen, ob diese mit den Reflexionen der Kinder korrelierten: Es stellte sich die Frage, ob die Kinder eine kürzere Eingabedauer und eine geringere Fehlerzahl auch als „besser“, „leichter“ oder „angenehmer“ wahrnahmen. Dies war fast immer, aber nicht durchgängig der Fall.

Reflexion Mädchen, 10;9 Jahre

Nach dem Test: „Es war gleich schwer.“

Sprechweise⁶	Dauer Y1	Fehlerzahl Y2
Tv	40	0
Ze	47	3

Testergebnisse

Die Abb. zeichnet insofern ein atypisches Bild, als dass sie Tv im Vergleich zu Ze sowohl schneller (40 s vs. 47 s) als auch weniger anfällig für Fehler (0 vs. 3) zeigt. Obgleich die Testauswertung einen Vorteil für Tv suggeriert, nimmt das Kind subjektiv beide Sprechweisen als gleichrangig wahr.

Reflexion Mädchen, 11;0 Jahre

Nach dem Test: „Es ist leichter mit Ze.“

Sprechweise	Dauer Y1	Fehlerzahl Y2
Tv	40	0
Ze	33	0

Testergebnisse

⁶ Bei den Angaben zu der Hauptdauervariablen Y1 und der Anzahl der Fehler pro Durchgang Y2 handelt es sich um die Rohwerte der jeweils zweiten Perioden (Testdaten) von Tv und Ze. Die Werte sind hier aus Gründen der Übersichtlichkeit einheitlich in die Tabellen eingefügt worden: Tv in die obere und Ze in die untere Zeile. Die Testungen wurden jedoch dem Studienplan entsprechend entweder mit der Sequenz 1 (TvTvZeZe) oder der Sequenz 2 (ZeZeTvTv) realisiert.

Reflexion Mädchen, 10;1 Jahre

Nach dem Test: „Es ist gleich schwer.“

Sprechweise	Dauer Y1	Fehlerzahl Y2
Tv	50	2
Ze	35	0

Testergebnisse

Reflexion Junge, 9;5 Jahre

Nach dem Test „Ze ist viel besser, z. B. bei fünfundachtzig: Also bei achtzigfünf kapiert man es viel schneller.“

Sprechweise	Dauer Y1	Fehlerzahl Y2
Tv	58	2
Ze	46	0

Testergebnisse

Reflexion Junge, 12;0 Jahre

Nach dem Test: „Ze ist leichter; es ist auch besser, wenn man am PC ist. Ich musste mich damals umgewöhnen, weil ich habe in der Schule immer erst die Einer und dann die Zehner geschrieben, aber am Computer muss ich das ja andersrum eingeben, deshalb ist es auch besser, wenn man es so spricht.“

Sprechweise	Dauer Y1	Fehlerzahl Y2
Tv	31	0
Ze	25	0

Testergebnisse

Reflexion Mädchen, 11;0 Jahre

Nach dem Test: „Ze ist leichter.“

Sprechweise	Dauer Y1	Fehlerzahl Y2
Tv	31	0
Ze	27	0

Testergebnisse

Reflexion Mädchen, 10;8 Jahre

Vor dem Test: „Das kann ich mir nicht vorstellen, dass Ze leichter geht.“

Nach dem Test: „Ze ist eigentlich doch viel angenehmer. Es ist leichter.“

Sprechweise	Dauer Y1	Fehlerzahl Y2
Tv	68	1
Ze	41	0

Testergebnisse

Reflexion Profil Junge, 10;1 Jahre

Nach dem Test: „Ze ist viel einfacher, es gefällt mir besser.“

Sprechweise	Dauer Y1	Fehlerzahl Y2
Tv	44	2
Ze	34	0

Testergebnisse

Reflexion Mädchen, 11;8 Jahre

Vor dem Test: „Das kann ich mir nicht vorstellen, es ist umständlich.“

Nach dem Test: „Ze gefällt mir sehr gut. Es ist einfacher.“

Das Kind erklärte im Anschluss an die Testung dem Vater ausführlich, worum es ging und dass es ja im Englischen auch so sei; es sei auch viel logischer.

Es kann sich Ze grundsätzlich nicht nur in der Schule, sondern auch im Alltag, z. B. beim Einkaufen vorstellen, gibt jedoch zu bedenken, dass die Uhrzeit mit Ze auszudrücken schwieriger wäre: „Meine Freunde würden es bestimmt komisch finden, wenn ich z. B. sagen würde: Wir treffen uns um zehnsieben Uhr.“

Sprechweise	Dauer Y1	Fehlerzahl Y2
Tv	45	4
Ze	32	0

Testergebnisse

Reflexion Profil 210: Mädchen, 11;5 Jahre

Vor dem Test: „Ah, interessante Idee!“

Nach dem Test: „Ze gefällt mir besser. Es ist bei Tv verwirrend, weil ich es ja andersrum schreibe.“

Sprechweise	Dauer Y1	Fehlerzahl Y2
Tv	33	0
Ze	30	0

Testergebnisse

Reflexion Mädchen, 11:9 Jahre

Nach dem Test: „Ze geht schneller. Das gefällt mir besser.“

Das Mädchen überlegt, ob das Aussprechen von Ze länger dauert im Vergleich zu Tv und erprobt einige Zahlen, z. B.: dreißigfünf vs. fünfunddreißig: also 3 Silben vs. 4 Silben, um dann festzustellen: „Siebzehn hat nur 2 Silben und zehn-sie-ben 3. Das ist länger. Aber Ze gefällt mir trotzdem besser; die meisten Zahlen sind auch kürzer, also wenn man sie spricht.“

Sprechweise	Dauer Y1	Fehlerzahl Y2
Tv	48	5
Ze	26	0

Testergebnisse

Reflexion Mädchen, 11;10 Jahre

Während der Einweisung: „Ah, so wie die Zahlen im Englischen sind. Das ist auch viel logischer.“

Nach dem Test: „Ze ist definitiv besser.“

Sprechweise	Dauer Y1	Fehlerzahl Y2
Tv	37	1
Ze	28	0

Testergebnisse

Reflexion Mädchen, 11;3 Jahre

Während der Einweisung: „Das ist doch länger!“

Das Kind zählt die Silben einiger selbstgewählter Zahlen jeweils in Tv und Ze; um festzustellen: „Ah doch nicht. Es kam mir nur umständlich vor, weil man es dann auf einmal anders sprechen muss.“

Während der Testung (Sequenz TvTvZeZe): „Ich werde immer schneller!“

Nach dem Test: „Ze ist eigentlich besser. Warum spricht man das eigentlich andersrum? Ich schreibe es auch immer so, wie ich es spreche, und dann dauert es länger.“

Untersucherin: „Das ist eine gute Frage. Weißt du, wie man es im Englischen ausspricht?“

„*Twentyone*. Ja stimmt. Dann ist Ze viel besser. Man könnte es doch im Deutschen auch so machen.“

Sprechweise	Dauer Y1	Fehlerzahl Y2
Tv	38	0
Ze	28	0

Testergebnisse 213

Reflexion Junge, 12;5 Jahre

Beginnt den Test mit Ze: „Das ist leicht.“

Nach dem Test: „Das Zweite (Tv) war verwirrend, das Erste (Ze) war einfacher.“

Sprechweise	Dauer Y1	Fehlerzahl Y2
Tv	35	3
Ze	33	2

Testergebnisse

Reflexion Mädchen, 10;7 Jahre alt

Nach dem Test: „Beides ist leicht und schwer.“

Die Auswertung zeigt zwar, vor allem hinsichtlich der Eingabedauer, einen eindeutigen Vorteil für Ze (s. Abb.); dieser wird jedoch von dem Kind subjektiv nicht wahrgenommen.

Sprechweise	Dauer Y1	Fehlerzahl Y2
Tv	51	1
Ze	39	0

Testergebnisse

Reflexion Mädchen, 9;6 Jahre

Nach dem Test: „Ze ist besser. Es geht einfacher.“

Das Kind empfindet Ze als einfacher, obwohl die Eingabedauer Y1 bei Tv niedriger war als bei Ze; allerdings stieg auch die Fehlerzahl mit 7 bei Tv vs. 1 bei Ze (s. Abb.). Weiterhin fällt auf, dass die erste Periode („Probendurchgang“) bei Ze mit 44 s Eingabedauer (Y1 = 44) gleichzeitig die schnellste Periode des Kindes war, die zudem fehlerfrei gelang (Y2neg = 0). Dies macht die subjektive Wahrnehmung des Kindes zugunsten von Ze nachvollziehbar.

Sprechweise	Dauer Y1	Fehlerzahl Y2
Tv	45	7
Ze	49	1

Testergebnisse

Reflexion Mädchen, 13;9 Jahre

Während der Einweisung: (Untersucherin): „Du würdest mit der neuen Sprechweise z. B. sagen: zwanzigeins, dreißig – das bleibt so – sechzigfünf...“

Kind: „... und 11 [wurde gerade an die Tafel geschrieben] wäre dann zehneins!“

Nach dem Test: „Im ersten Durchgang von Ze habe ich noch kurz überlegt: Wird siebzigzwei so wie 70 und 2 geschrieben, also dass man die Null mitschreibt? Aber dann war es eigentlich leicht. Also leichter als das andere.“

Sprechweise	Dauer Y1	Fehlerzahl Y2
Tv	38	0
Ze	30	0

Testergebnisse

Reflexion Junge, 11;11 Jahre

Nach dem Test: „Ze ist leichter. Ich war auch schneller!“

Sprechweise	Dauer Y1	Fehlerzahl Y2
Tv	44	0
Ze	35	1

Testergebnisse

Reflexion Mädchen, 10 ;0 Jahre

Während des Tests: (Kommuniziert während der Durchführung): „Kann ich auch mittel und schwer probieren?“

Nach dem Test: „Es geht leichter so wie du mir es gezeigt hast (Ze). Ich habe auch früher immer links und rechts verwechselt und deswegen war es schwer, also beim Rechnen.“ Dass das Mädchen deutlich von Ze profitiert, zeigt die Tabelle:

Sprechweise	Dauer Y1	Fehlerzahl Y2
Tv	68	14
Ze	47	2

Testergebnisse

Reflexion Mädchen, 10;11 Jahre

Nach dem Test: „Ze ist eigentlich angenehmer.“

Sprechweise	Dauer Y1	Fehlerzahl Y2
Tv	35	0
Ze	34	0

Testergebnisse

7 ZUSAMMENFASSUNG UND DISKUSSION

Aus den Ergebnissen der vorliegenden Arbeit ergibt sich ein sehr überzeugender Gesamtbefund: Es konnte ein direkter Zusammenhang zwischen der Nutzung einer stellenwertgerechten Zahlensprechweise und schnellerem sowie weniger fehleranfälligerem Lösen von Transkodieraufgaben nachgewiesen werden. Die Befunde zu den Dauer- und Fehlervariablen sind dabei qualitativ übereinstimmend: Siehe Anna Kuhls Arbeit (2025) mit den detaillierten Auswertungen und Ergebnissen zu den Dauervariablen (Responsvariablen 1 und 2). Im Vergleich zu der gebräuchlichen inversen Zahlensprechweise lösen rechenschwache Kinder Transkodieraufgaben schneller mit einer stellenwertgerechten Sprechweise, sie machen weniger Fehler und es gibt mehr fehlerfreie Durchgänge. Die eingangs erörterten Forschungsfragen können daher wie folgt beantwortet werden:

- 1) Die Eingabe von diktierten Zahlen in der Zwanzigeins-App wird von rechenschwachen Kindern bei stellenwertgerechter Sprechweise schneller realisiert.
- 2) Rechenschwache Kinder machen bei stellenwertgerechter Sprechweise weniger Fehler.
- 3) Es gibt mehr fehlerfreie Durchgänge bei stellenwertgerechter Sprechweise.

Die Auswertungen zu den erhobenen Kovariablen ergaben, dass alle Untergruppen (z. B. Jungen und Mädchen, ältere und jüngere Kinder, solche mit der Erstsprache Deutsch oder anderen Erstsprachen) von einer Umstellung auf eine stellenwertgerechte Zahlensprechweise profitieren.

Auf der Basis unserer Befunde kann daher unsere Hypothese H1 bestätigt werden:

H1 Rechenschwache Kinder der Primarstufe transkodieren diktierte Zahlen bei Verwendung einer stellenwertgerechten Zahlensprechweise schneller und weniger fehleranfällig als bei der Verwendung der im deutschen Sprachraum gebräuchlichen inversen Zahlensprechweise, unabhängig von anderen möglichen Einflussfaktoren.

Die Ergebnisse der vorliegenden Arbeit lassen darauf schließen, dass den Kindern tendenziell die stellenwertgerechte Zahlensprechweise leichter fällt, was diese in unseren Feedbackgesprächen mehrheitlich bestätigten. Auch frühere Unterrichtsversuche, u. a. von Eckstein (2020a) und Kimmeskamp (2008),

bezeugen die Zustimmung vieler Kinder zu einer stellenwertgerechten Sprechweise. Die Argumentation von Schuppener gegen eine Reform der Zahlensprechweise ist vor dem Hintergrund der mit der Zwanzigeins-App durchgeführten Forschungen nicht haltbar. Dass ein möglicher Effekt der stellenwertgerechten Zahlensprechweise „Verwirrung bei den Lernenden“ sei (Morfeld & Schuppener 2024, S. 23), steht im Widerspruch zu den Ergebnissen dieser Untersuchung; wäre dies der Fall, müsste man bei Verwendung einer stellenwertgerechten Sprechweise eine höhere Fehlerzahl und / oder Eingabedauer annehmen als bei Verwendung der inversen (gebräuchlichen) Zahlensprechweise. Dies ist, wie in vorliegender Arbeit in Zusammenarbeit mit Kuhl (2025) sowie bei Schmid (2023) gezeigt werden konnte, nicht zutreffend. Auch die Aussagen der Kinder widersprechen dem Postulat der „Verwirrung.“ Wenngleich die Stichprobe (n=20) keine verallgemeinernden Aussagen erlaubt, ist das Meinungsbild der Kinder zugunsten einer stellenwertgerechten Sprechweise eindeutig: In ihrer subjektiven Wahrnehmung ist diese „einfacher“, „angenehmer“, „besser“ oder „leichter“ (17 Probanden) oder es gibt für sie keinen wahrnehmbaren Unterschied zwischen den verschiedenen Zahlensprechweisen (3 Probanden).

Bemerkenswert war, dass einige Kinder spontan weitere Aspekte der Verwendung einer stellenwertgerechten Zahlensprechweise äußerten: So wurde ein Bezug zur englischen Sprache hergestellt oder untersucht, ob die Zahlwörter in der stellenwertgerechten Sprechweise mehr oder weniger Silben haben als mit der inversen Sprechweise. Ein Kind überlegte, ob man die Uhrzeit ausdrücken könnte, ein anderes argumentierte, dass die Verwendung stellenwertgerechten Aussprechens einen Vorteil für das Schreiben am PC mit sich bringen würde. Auch Eckstein (2020b, S. 7) bestätigt aus seiner Förderpraxis, dass die neue Sprechweise (Ze) Fragen aufwirft, „denen zusammen mit den Lernenden nachgegangen werden sollte: Warum sind die Zahlwörter der deutschen Sprache so, wie sie sind? Wer hat die Zahlwörter der deutschen Sprache eigentlich erfunden? Wie alt sind sie? Gibt es das Problem der verdrehten Zahlwörter auch in anderen Sprachen? Gibt es in allen Sprachen Zahlwörter für große und für ganz große Zahlen?“

Das Reflektieren und die inhaltliche Auseinandersetzung mit der Thematik der inversen Zahlensprechweise könnte in jedem Fall nicht nur bereichernd für jede Unterrichts- oder Förderstunde sein, sondern auch einen grundlegenden Beitrag

zum Verständnis des Stellenwertsystems leisten. Dies ist deshalb so wichtig, weil „Missverständnisse und Verständnislücken im Bereich des Dezimalsystems weitreichende Konsequenzen auf den arithmetischen Kompetenzaufbau haben und einen Kernbereich anhaltender Lernschwierigkeiten im Mathematikunterricht bilden“ (Gaidoschik 2015, S. 173).

Gerade rechenschwache Kinder und Jugendliche sollten befähigt werden, Vertrauen in die eigenen mathematischen Kompetenzen zu erwerben, was nur über lernbezogene Erfolgserlebnisse geschehen kann: „Da zwischen spezifischen Lernschwächen und zugehörigen niedrigen Selbstbewertungen eine hohe Koppelung besteht, ist der Aufbau lernbezogener selbstwertdienlicher Überzeugungen von großer Bedeutung, um einem ungünstigen Teufelskreislauf negativer Selbsteinschätzungen, geringer Lernmotivation und ungünstiger Leistungsbereitschaft insgesamt vorzubeugen“ (Schuchardt et al. 2015, S. 523). Aus therapeutischer Sicht und vor dem Hintergrund der komplexen Schwierigkeiten sowie defizitärer Arbeitsgedächtniskapazitäten, mit denen sich rechenschwache Kinder häufig konfrontiert sehen, sollte jede unnötige Hürde vermieden werden, um diese lernbezogenen Erfolgserlebnisse zu ermöglichen. Umgekehrt kann jede Hürde, die überwunden werden muss, für lern- und rechenschwache Kinder zu einer erneuten Bestätigung des eigenen Versagens führen und zur Entstehung negativer Kreisläufe beitragen (Betz & Breuninger 1998). Wesentliche Faktoren für Schwierigkeiten beim Rechnenlernen sind Arbeitsgedächtnisprobleme und Störungen der basalen Zahlenrepräsentation (van Eimeren & Ansari 2009, S. 31). Eine stellenwertgerechte Zahlensprechweise entlastet das Arbeitsgedächtnis und erleichtert die Herausbildung einer mentalen Zahlraumvorstellung, das Verständnis des Stellenwertsystems und letztlich den Erwerb arithmetischer Kompetenzen.

Unter Einbeziehung der statistischen Untersuchungsergebnisse und der Reflexionen der Kinder ergibt sich ein Gesamtbild, das die Argumente für eine Verwendung einer stellenwertgerechte Zahlensprechweise untermauert. Vor dem Hintergrund der globalen Bedeutung scheint es angemessen, die Erfindung und Verbreitung des dezimalen Stellenwertsystems, die historischen Zusammenhänge sowie länderübergreifende Aspekte wie z. B. die Zahlensprechweisen in anderen Ländern und Möglichkeiten der Reform wie in Tschechien und Norwegen in der Ausbildung von Fachpersonal (Lehrkräfte und TherapeutInnen) zu thematisieren und in den Rahmenlehrplänen zu verankern. Die Möglichkeit, die

stellenwertgerechte Zahlensprechweise mit der Zwanzigeins-App auszuprobieren und sich mit ihr auseinanderzusetzen, könnte maßgeblich zum Verständnis des Stellenwertsystems beitragen. In diesem Sinn wäre eine größere Beachtung der Zwanzigeins-Hypothese in der Grundschuldidaktik angemessen und im Rahmen einer Förderung für rechenschwache Kinder indiziert. Es bedarf der Sensibilisierung von pädagogischem und therapeutischem Personal, von Eltern, Betroffenen und Interessierten sowie der Information über die in dieser Arbeit dargestellten Zusammenhänge von Zahlensprechweise, Transkodieren von Zahlen, Arbeitsgedächtnis und Mathematikleistungen besonders in Bezug auf rechenschwache Kinder.

Anders als bei der Rechtschreibreform von 1996, mit deren erheblichen Aufwand eine mögliche Reformierung der Zahlensprechweise Schuppener vergleicht (Morfeld & Schuppener 2024, S. 23), betreffen die Reformbemühungen, so wie sie in dem Positionspapier von Zwanzigeins e. V. (2024a) dargestellt sind, lediglich den mündlichen Sprachgebrauch, so dass zum einen der Aufwand deutlich geringer gehalten ist und zum anderen – bei einer breiten Aufklärung über die Vorteile hinsichtlich des Erwarbs arithmetischer Kompetenzen – mit einer hohen Akzeptanz in der Gesellschaft zu rechnen ist. Die Zulassung einer stellenwertgerechten Zahlensprechweise wäre ein erster möglicher Schritt und würde auch die Pädagogen und Therapeuten unterstützen, die diese bereits in ihre Lernangebote integrieren.

Ausblick

Obwohl die Auswertungen der bisher vorliegenden Studien den signifikanten Vorteil der stellenwertgerechten Zahlensprechweise bereits hinreichend belegen, können weitere Untersuchungen aufschlussreich sein, um z. B. die Unterschiede in den Sprechweisen auf höheren Levels der Zwanzigeins-App, beim Lösen von Kopfrechenaufgaben oder beim Rechnen mit der App nach einer Übungsstunde zur stellenwertgerechten Zahlensprechweise festzustellen. Übergreifend hochsignifikante Ergebnisse wären auf Grund der vorliegenden Auswertung zum Einfluss des Alters erwartbar, wenn nur rechenschwache Kinder der Klassenstufen 1 - 4 getestet würden oder solche, die erst kürzlich als rechenschwach aufgefallen sind bzw. diagnostiziert wurden und noch keine therapeutische Intervention

begonnen haben. Sinnvoll wären ergänzend Forschungsarbeiten zum Effekt der Zahlensprechweise auf Kinder und Erwachsene mit Hochbegabung oder mit Problemen der Rechts-Links-Orientierung. Darüber hinaus wären Studien mit funktioneller Bildgebung (fMRT) zur Überprüfung der Hypothesen zur Arbeitsgedächtnisbelastung bzw. -entlastung wünschenswert.

Danksagung

Mein besonderer Dank gilt Prof. Dr. Michael G. von Aster für die außergewöhnliche und fürsorgliche Betreuung sowie die stete Inspiration.

PD Dr. Peter Morfeld danke ich für sein unermüdliches Engagement, die vertiefenden Einblicke in die Welt der Statistik und sein stets offenes Ohr.

Meiner Mit-Masterin Anna Kuhl: Danke für den immer anregenden Austausch.

Allen teilnehmenden Kindern möchte ich für ihre Offenheit und das Teilen ihrer Gedanken danken. Nur ihr Mitwirken ermöglichte dieses Forschungsprojekt.

Die Arbeit hätte ohne die freundliche Genehmigung zur Durchführung der Studie von Herrn Cordes, Leiter der Duden-Institute Oranienburg und Falkensee, nicht entstehen können. Das Engagement meiner KollegInnen Ingrid, Margit, Martina und Michael bei der Ansprache der Kinder, Kontaktierung der Eltern und Koordination der Termine war unerlässlich – vielen Dank dafür! Allen Eltern gilt mein Dank für das Vertrauen, das sie mir mit der Einwilligung in die Teilnahme ihrer Kinder geschenkt haben.

Bettina half mit wertvollen Tipps, vor allem aber danke ich ihr für die anregenden Gespräche.

Nicht zuletzt: Meiner Familie danke ich für ihre Geduld und die fortwährende Unterstützung.

Gender Disclaimer

In dieser Arbeit wird teilweise ein generisches Maskulinum für Personen- und Berufsbezeichnungen verwendet. Sofern die Aussagen es erfordern, sind immer alle Geschlechtsidentitäten mitgemeint. Die Verwendung der maskulinen Form dient ausschließlich der besseren Lesbarkeit und beinhaltet keine Wertung.

8 LITERATURVERZEICHNIS

- Alloway, T., Gathercole, S., Kirkwood, H. & Elliott, J. (2009). The Cognitive and Behavioral Characteristics of Children With Low Working Memory. *Child development*, 80 (2), 606–621.
- Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften e. V. (AWMF). *S3-Leitlinie: Diagnostik und Behandlung der Rechenstörung*. https://www.bvl-legasthenie.de/images/static/pdfs/Leitlinien/S3-Leitlinie_Rechenstrung_Langfassung.pdf. Aufruf am 24.09.2024.
- Ashcraft, M. H. & Krause, J. A. (2007). Working memory, math performance, and math anxiety. In: *Psychonomic Bulletin & Review*, 14(2), 243-248.
- Aster, M. von & Shalev, R. S. (2007). Number development and developmental dyscalculia. *Developmental Medicine & Child Neurology*, 49, 868–873.
- Aster, M. von & Lorenz, J. H., Hrsg. (2013). *Rechenstörungen bei Kindern. Neurowissenschaft, Psychologie, Pädagogik*. 2. überarb. und erw. Auflage. Göttingen: Vandenhoeck & Ruprecht.
- Aster, M. von. (2013a). Wie kommen Zahlen in den Kopf? In: M. von Aster & J. H. Lorenz (Hrsg.), *Rechenstörungen bei Kindern. Neurowissenschaft, Psychologie, Pädagogik*, 15–38. 2. überarb. und erw. Auflage. Göttingen: Vandenhoeck & Ruprecht.
- Aster, M. von, Bzufka, M. W., Neumärker, K.-J. (2013b) Diagnostik von Rechenstörungen. In: M. von Aster & J. H. Lorenz (Hrsg.), *Rechenstörungen bei Kindern. Neurowissenschaft, Psychologie, Pädagogik*, 79–91. 2. überarb. und erw. Auflage. Göttingen: Vandenhoeck & Ruprecht.
- Aster, M. von, Kucian, K., Schweiter, M. et al. (2013c) Rechenstörungen im Kindesalter. *Monatsschrift Kinderheilkunde* 153: 614–622.
- Aster, M. von, Weinhold-Zulauf, M. & Horn, R. (2013d). *ZAREKI-R: Neuropsychologische Testbatterie für Zahlenverarbeitung und Rechnen bei Kindern – Revidierte Fassung*. 4., aktualisierte Auflage. Frankfurt: Pearson Clinical.
- Aster, M. von (2017). Dyskalkulie. Wenn Kinder nicht rechnen lernen. In: L. Huck & A. Schulz (Hrsg.), *Lerntherapie und inklusive Schule (Lerntherapie in Theorie und Praxis)*, 107–119. Berlin: Dudenverlag.
- Aster, M. von, Kaufmann, L., McCaskey, U. & Kucian, K. (2021) Rechenstörungen im Kindes- und Jugendalter. In: J. Fegert et al. (Hrsg.), *Psychiatrie und Psychotherapie des Kindes- und Jugendalters*. Berlin, Heidelberg: Springer Reference Medizin. https://doi.org/10.1007/978-3-662-49289-5_120-1

- Aster, M. von (2023). *Legasthenie und Dyskalkulie. Schulisches Scheitern ist kein Schicksal. Interview mit S. Vieth-Entus*. Tagesspiegel vom 25.07.2023. <https://www.tagesspiegel.de/berlin/legasthenie-und-dyskalkulie-schulisches-scheitern-ist-kein-schicksal-10171231.html>. Aufruf am 13.12.2024.
- Baddeley, A. (2012). Working Memory: Theories, Models, and Controversies. In: *Annual Review of Psychology* 63, 1–29. <https://www.annualreviews.org/doi/abs/10.1146/annurev-psych-120710-100422>. Aufruf am 12.09.2024.
- Barrow, J. D. (2005). *Ein Himmel voller Zahlen – Auf den Spuren mathematischer Wahrheit*. 4. Aufl. Reinbeck: Rowohlt.
- Baur, N. & Blasius, J., Hrsg. (2022a). *Handbuch Methoden der empirischen Sozialforschung*. Wiesbaden: Springer Fachmedien. <https://doi-10org-15ww5t0cp03a9.erf.sbb.spk-berlin.de/10.1007/978-3-658-37985-8>
- Baur, N. & Blasius, J. (2022b). Multivariate Datenanalyse. In: N. Baur & J. Blasius (Hrsg.) *Handbuch Methoden der empirischen Sozialforschung*, 997-1016. Wiesbaden: Springer Fachmedien.
- Bender, F., Brandelik, K., Jeske, K., Lipka, M., Löffler, C., Mannhaupt, G., Naumann, C. L., Nolte, M., Ricken, G., Rosin, H., Scheerer-Neumann, G., von Aster, M. von Orloff., M. (2017). Die integrative Lerntherapie. Therapieform zur Behandlung von Lernstörungen. *Lernen und Lernstörungen* 6(2): 65-73.
- Benz, Ch., Grüßing, M., Lorenz, J. H., Reiss, K., Selter, Ch. & Wollring, B. (2017). *Frühe mathematische Bildung - Ziele und Gelingensbedingungen für den Elementar- und Primarbereich. Wissenschaftliche Untersuchungen zur Arbeit der Stiftung "Haus der kleinen Forscher" 8*. Opladen, Berlin, Toronto: Verlag Barbara Budrich. DOI: 10.25656/01:12932
- Bellos, A. (2010). *Alex's Adventures in Numberland*. London: Bloomsbury.
- Betz, D. & Breuninger, H. (1998). *Teufelskreis Lernstörungen: Theoretische Grundlegung und Standardprogramm*. 1. Auflage 1982. München, Weinheim: Beltz.
- Bolin, J. E. (2022). *Regression Analysis in R: A Comprehensive View for the Social Sciences*. Boca Raton: Chapman and Hall / CRC. <https://doi-10org-1gx5abjwm0068.erf.sbb.spk-berlin.de/10.1201/9780429295843>
- Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM), https://www.bfarm.de/DE/Kodiersysteme/Klassifikationen/ICD/ICD-10-GM/_node.html. Aufruf am 24.09.2024.
- Dehaene, S. (1992). Varieties of Numerical Abilities. *Cognition*, 44 (1–2), 1-42, [https://doi.org/10.1016/0010-0277\(92\)90049-N](https://doi.org/10.1016/0010-0277(92)90049-N).

- Dehaene, S. & Cohen, L. (1995). Towards an Anatomical and Functional Model of Number Processing. *Math Cognition* 1995. 83-120.
https://www.unicog.org/publications/DehaeneCohen_TripleCodeModelNumberProcessing_MathCognition1995.pdf
- Dehaene, S., Piazza, M., Pinel, P. & Cohen, L. (2003). Three parietal circuits for number processing. *Cognitive Neuropsychology*, 20(3), 487–506.
<https://doi.org/10.1080/02643290244000239>
- Delazer, M., Handl, P., Kaufmann, L., Pixner, S. (2013). Wie Kinder rechnen lernen und was ihnen dabei hilft. Eine kognitiv-neuropsychologische Perspektive. In: M. von Aster & J. H. Lorenz (Hrsg.), *Rechenstörungen bei Kindern. Neurowissenschaft, Psychologie, Pädagogik*. 231–249. 2. überarb. und erw. Auflage. Göttingen: Vandenhoeck & Ruprecht.
- Dilling, H., Mombour, W. & Schmidt, M. H. (2013). *Internationale Klassifikation psychischer Störungen ICD-10 Kapitel V (F) Klinisch-diagnostische Leitlinien (9. Aufl.)*. Bern: Huber.
- Döring, N. & Bortz, J. unter Mitarbeit von Pöschl, S. (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften*. Berlin, Heidelberg: Springer-Verlag. DOI 10.1007/978-3-642-41089-5_3
- Dorsch, F., Wirtz, M. A. & Strohmer, J. (2017). *Dorsch - Lexikon der Psychologie*. 18., überarb. Auflage. Bern: Hogrefe.
- Eckstein, B. (2020a). *Verdrehte Zahlwörter. Trick zehnsieben hilft!* Wuppertal: Eigenverlag Berthold Eckstein.
- Eckstein, B. (2020b). *Für die Website des Vereins Zwanzigeins e. V. / Text für Erwachsene*. [https://zwanzigeins.jetzt/downloads/Eckstein_2020-Text%20f%C3%BCr%20Erwachsene%20\(Auszug\).pdf](https://zwanzigeins.jetzt/downloads/Eckstein_2020-Text%20f%C3%BCr%20Erwachsene%20(Auszug).pdf). Aufruf am 18.12.2024.
- Eichler, K.-P. (2006). *Lernbereich mathematisches Denken im Elementarbereich*. <https://www.mathematikus.de/lehren-und-lernen/lerntherapie>. Aufruf am 12.06.2024.
- Falkai, P., Wittchen, H.-U., Döpfner, M., Gaebel, W., Maier, W., Rief, W., et al. (2015). *Diagnostisches und Statistisches Manual Psychischer Störungen DSM-5®*. Göttingen: Hogrefe.
- Fischbach, A. et al. (2013). Prävalenz von Lernschwächen und Lernstörungen: Zur Bedeutung der Diagnosekriterien. *Lernen und Lernstörungen*, 2(2), 65–76, doi: 10.1024/2235-0977/a000035
- Fritz, A., Ricken, G. & Schmidt, S., Hrsg. (2009). *Handbuch Rechenschwäche: Lernwege, Schwierigkeiten und Hilfen bei Dyskalkulie*. Weinheim und Basel: Beltz.

- Fuchs, L. S., Compton, D. L., Fuchs, D., Paulsen, K., Bryant, J. D. & Hamlett, C. L. (2005). The prevention, identification, and cognitive determinants of math difficulty. *Journal of Educational Psychology* 97, 493–513.
- Gaidoschik, M. (2015). Einige Fragen zur Didaktik der Erarbeitung des „Hunderterraums“. *Journal für Mathematik-Didaktik*, 35 (1), 163–190.
- Gerritzen, L., Hrsg. (2008). *Zwanzigeins – für die unverdrehte Zahlensprechweise*. Bochum: Brockmeyer. <https://zwanzigeins.jetzt/infos/literatur>. Aufruf am 08.04.2024.
- Gerster, H. D. & Schultz, R. (1998). *Schwierigkeiten beim Erwerb mathematischer Konzepte im Anfangsunterricht. Bericht zum Forschungsprojekt Rechenschwäche – Erkennen, Beheben, Vorbeugen*. Freiburg im Breisgau: PH Freiburg. Erw. Auflage 2004.
- Gialluisi, A., Andlauer, T. F. M., Mirza-Schreiber, N. et al. (2021). Genome-wide association study reveals new insights into the heritability and genetic correlates of developmental dyslexia. *Molecular Psychiatry* 26, 3004–3017.
- Greco, A. (1995). The concept of representation in psychology (Vol. 4). In: A. Greco (Hrsg.): *Cognitive Systems*, 247-256.
- Griepenburg, C. & Schuchardt, K. (2019). Psychoedukation bei Lernstörungen: Relevanz und Umsetzung. *Lernen und Lernstörungen* 8 (1), 33–45.
- Grond, U., Schweiter, M. & Aster, M. von (2013). Neuropsychologie numerischer Repräsentationen. In: M. von Aster & J. H. Lorenz (Hrsg.), *Rechenstörungen bei Kindern. Neurowissenschaft, Psychologie, Pädagogik*, 39-57. 2. überarb. und erw. Auflage. Göttingen: Vandenhoeck & Ruprecht.
- Haberstroh, S. & Schulte-Körne, G. (2019). Clinical practice guideline: The diagnosis and treatment of dyscalculia. *Deutsches Ärzteblatt International* 2019; 116: 107–114. DOI: 10.3238/arztebl.2019.0107
- Haffner, J., Baro, K., Parzer, P. & Resch, F. (2005). *Heidelberger Rechentest (HRT 1–4). Erfassung mathematischer Basiskompetenzen im Grundschulalter*. Göttingen: Hogrefe.
- Hasselhorn, M. (2022). Lernstörungen: ein unvermeidbares Schicksal? *Zeitschrift für pädagogische Psychologie* 36 (1-2), 1–17.
- Helmreich, I., Zuber, J., Pixner, S., Kaufmann, L., Nürk, H. C., Möller, K. (2011). Language Effects on Children’s Mental Number Line: How cross-cultural differences in number word systems affect spatial mappings of numbers in a non-verbal task. *Journal of Cross-Cultural Psychology*, 598–613. doi: 10.1177/0022022111406026
- Hopko, D. R., Ashcraft, M. H., Gute, J., Ruggiero, K. J., Lewis, C. (1998). Mathematics Anxiety and Working Memory: Support for the Existence of a

- Deficient Inhibition Mechanism. *Journal of Anxiety Disorders*, 12(4), 1998, 343–355. [https://doi.org/10.1016/S0887-6185\(98\)00019-X](https://doi.org/10.1016/S0887-6185(98)00019-X)
- Ifrah, G. (1991). *Die Universalgeschichte der Zahlen*, 2. Aufl. Frankfurt/New York: Campus.
- Imbo, I., van den Bulcke, C., de Brauwer, J., Fias, W. (2014). Sixty-four or four-and-sixty? The influence of language and working memory on children's number transcoding. *Frontiers in Psychology | Developmental Psychology* 5: 1-10. URL: doi: 10.3389/fpsyg.2014.00313
- Käser, T. & Aster, M. von (2013). Computerbasierte Lernprogramme für Kinder mit Rechenschwäche. In: M. von Aster & J. H. Lorenz (Hrsg.), *Rechenstörungen bei Kindern. Neurowissenschaft, Psychologie, Pädagogik*, 259–275. 2. überarb. und erw. Auflage. Göttingen: Vandenhoeck & Ruprecht.
- Kaufmann, L., Mazzocco, M., Dowker, A., Aster, M. von, Goebel, S., Grabner, R., Henik, A., Jordan, N., Karmiloff-Smith, A., Kucian, K., Rubinsten, O., Szucs, D., Shalev, R. & Nürk, H.-C. (2013). Dyscalculia from a developmental and differential perspective. *Frontiers in Psychology Vol. 4, Art. 516*, 1–5. DOI=10.3389/fpsyg.2013.00516
- Kaufmann, S. & Wessolowski, S. (2021). *Rechenstörungen: Diagnose und Förderbausteine*. 8. Auflage. Hannover. Klett / Kallmeyer.
- Kimmeskamp, P. (2008). Kinder mögen „Zwanzigeins.“ In: L. Gerritzen (Hrsg.), *Zwanzigeins: Für die unverdrehte Zahlensprechweise; Fakten, Argumente und Meinungen*, 59–61. Brockmeyer.
- Klein, E., Bahnmüller, J., Mann, A., Pixner, S., Kaufmann, L., Nürk, H. C., Möller, K. (2013). Language influences on numerical development — Inversion effects on multi-digit number processing. *Frontiers in Psychology | Developmental Psychology* 4,1-6. doi:10.3389/fpsyg.2013.00480
- Krinzinger, H. & Günther, T. (2013). Lesen, Schreiben, Rechnen – gibt es Unterschiede zwischen den Geschlechtern? *Lernen und Lernstörungen* 2(1), 35–49.
- Kucian, K. & Aster, M. von (2013). Dem Gehirn beim Rechnen zuschauen. Ergebnisse der zerebralen Bildgebung. In: M. von Aster & J. H. Lorenz (Hrsg.), *Rechenstörungen bei Kindern. Neurowissenschaft, Psychologie, Pädagogik*, 59–78. 2. überarb. und erw. Auflage. Göttingen: Vandenhoeck & Ruprecht.
- Kucian, K. (2017). Mathematikangst: Welche Effekte hat Rechentraining auf das emotionale Erleben und auf das Gehirn? *Lernen und Lernstörungen* 6(4), 221.

- Kucian, K. (2022). Einstellungen und Überzeugungen von Grundschullehrer_innen über räumliches Denken und Mathematik. *Lernen und Lernstörungen* 11(4), 239.
- Kucian, K. (2023). Der Einfluss von Angst und Lebensqualität auf die mathematische Leistung von Kindern mit Dyskalkulie: The Impact of Anxiety and Life Quality on the Mathematical Performance of Dyscalculic Middle School Children. *Lernen und Lernstörungen* 12(4), 229.
- Kuhl, A. (2025). *Zwanzigeins vs. einundzwanzig: Profitieren rechenschwache Kinder in Bezug auf die Dauer von einer stellenwertgerechten Zahlensprechweise? Eine empirische Untersuchung*. Masterarbeit in Integrative Lerntherapie. Pädagogische Hochschule Schwäbisch Gmünd / Zentrum für Wissenstransfer.
- Landerl, K. & Moll, K. (2010). Comorbidity of learning disorders: prevalence and familial transmission. *Journal of Child Psychology and Psychiatry: Vol. 51, No. 3 (2010-03)*, 287-294.
- Lang, P. J. (1968). Fear reduction and fear behavior: Problems in treating a construct. In: J. M. Schlien (Hrsg.), *Research in psychotherapy, Vol. III*. Washington, DC: American Psychological Association.
- Lewis, K. E. & Fisher M. B. (2016). Taking stock of 40 years of research on mathematical learning disability: methodological issues and future directions. *Journal for Research in Mathematics Education* 2016, Vol. 47, No. 4 (July 2016), 338-371.
- Lorenz, J. H. (2013). Grundlagen der Förderung und Therapie. Wege und Irrwege. In: M. von Aster & J. H. Lorenz (Hrsg.), *Rechenstörungen bei Kindern. Neurowissenschaft, Psychologie, Pädagogik*, 181–194. 2. überarb. und erw. Auflage. Göttingen: Vandenhoeck & Ruprecht.
- Lorenz, J. H. (2017). Einige Anmerkungen zur Repräsentation von Wissen über Zahlen. *Journal für Mathematik-Didaktik* 38, 125–139.
- Luttenberger, S., Wimmer, S. & Paechter, M. (2018). Spotlight on math anxiety. *Psychology Research and Behavior Management* 11, 311–322.
- Ma, M., Li, D. & Zhang, L. (2021). Longitudinal prediction of children's math anxiety from parent-child relationships, *Learning and Individual Differences*, Vol. 88, 102016. <https://doi.org/10.1016/j.lindif.2021.102016>
- Mähler, C. (2017). Diskrepanzkriterium. In: F. Dorsch, M. A. Wirtz & J. Strohmmer (Hrsg.). *Dorsch - Lexikon der Psychologie*. 18. überarb. Auflage. <https://dorsch.hogrefe.com/stichwort/diskrepanzkriterium>
- Maloney, E. A., Ramirez, G., Gunderson, E. A., Levine, S. C. & Beilock, S. L. (2015). Intergenerational Effects of Parents' Math Anxiety on Children's Math

- Achievement and Anxiety. *Psychological Science*, 26(9), 1480–1488.
<https://doi.org/10.1177/0956797615592630>
- Mazzocco, M. M. & Myers, G. F. (2003). Complexities in identifying and defining mathematics learning disability in the primary school-age years. *Ann. Dyslexia* 53, 218–253. doi: 10.1007/s11881-003-0011-7
- Möller, K., Pixner, S., Zuber, J., Kaufmann, L. & Nürk, H.-C. (2011). Early place-value understanding as a precursor for later arithmetic performance – a longitudinal study on numerical development. *Research in Developmental Disabilities* (32): 1837–1851. doi: 10.1016/j.ridd.2011.03.012
- Morfeld, P. & Gerritzen, L. (2024). *Stellenwertsystem und Ziffernrechnen als immaterielles Kulturerbe der Menschheit: ein Unterrichtsgegenstand für die Mathematik?* <https://zwanzigeins.jetzt/aktivitaeten/projekte/stellenwertsystem>
- Morfeld, P. & Schuppener, G. (2024). Lieber zwanzigeins als einundzwanzig? Eine Pro-Contra-Rubrik. *ZEIT SPRACHEN, Deutsch perfekt* 5/24, 22–23.
- Morfeld, P. & Summer, A. (2024). Zwanzigeins. *Lernen und Lernstörungen* 13(1), 1–3.
- Nelson, J. M. & Harwood, H. (2011). Learning Disabilities and Anxiety: A Meta-Analysis. *Journal of Learning Disabilities*, 44(1), 3-17.
<https://doi.org/10.1177/0022219409359939>
- Ng, S. S. N. & Rao, N. (2010). Chinese Number Words, Culture, and Mathematics Learning. *Review of Educational Research*, 80(2), 180–206.
- Nürk, H.-C., Olsen, N. & Willmes, K. (2005). “Better teach your children Japanese number words: how transparent number structure helps number acquisition,” in *Poster presented at the 23rd Workshop of Cognitive Neuropsychology* (Bressanone).
- Opp, K.-D. (2010). Kausalität als Gegenstand der Sozialwissenschaften und der multivariaten Statistik. In: C. Wolf & H. Best (Hrsg.), *Handbuch der sozialwissenschaftlichen Datenanalyse*. VS Verlag für Sozialwissenschaften.
- Padberg, F. & Benz, C. (2021). *Didaktik der Arithmetik*. 5. überarb. Auflage. Berlin: Springer.
- Pennac, D. (2009). *Schulkummer*. Köln: Kiepenheuer & Witsch. (Frz. Original: *Chagrin d'école*).
- Pfarr, K. & Schröder, J. (2016). *Why Panel Surveys? GESIS Survey Guidelines*. Mannheim, Germany. GESIS – Leibniz Institute for the Social Sciences. doi: 10.15465/gesis-sg_en_008

- Pinheiro-Chagas, P., Daitch, A., Parvizi, J., Dehaene, S. (2018). Brain Mechanisms of Arithmetic: A Crucial Role for Ventral Temporal Cortex. *Journal of Cognitive Neuroscience*; Vol. 30, No. 12 (2018-12), 1757–1772.
- Pixner, S., Zuber, J., Heřmanová, V., Kaufmann, L., Nürk, H.-C. & Möller, K. (2011). One language, two number-word systems and many problems: Numerical cognition in the Czech language. *Research in Developmental Disabilities*, 32(6), 2683–2689. <https://doi.org/10.1016/j.ridd.2011.06.004>
- Power, R. J. D. & Dal Martello, M. F. (1990). The dictation of Italian Numerals Lang. *Cognit. Processes*, 5, 237-254.
- Ramirez, G., Shaw, S. T. & Maloney, E. A. (2018). Math Anxiety: Past Research, Promising Interventions, and a New Interpretation Framework. *Educational Psychologist*, 53(3), 145–164.
- Roland, P. E. & Friberg, L. (1985). Localization of cortical areas activated by thinking. *Journal of Neurophysiology*, 53(5), 1219–1243. <https://doi.org/10.1152/jn.1985.53.5.1219>
- Saga, M., Rkhaila, A., Oubaha, D. & Ounine, K. (2022). The impact of anxiety and life quality on the mathematical performance of dyscalculic middle school children. *Applied Neuropsychology: Child*, 1–9. <https://doi.org/10.1080/21622965.2022.2105146>
- Schmid, S. (2023). Zwanzigeins – Eine empirisch-quantitative Untersuchung zur Zahleninversion in der zweiten Schulstufe. Masterarbeit in Erziehungswissenschaft, Primarstufe. Kirchliche Pädagogische Hochschule Wien / Krems. <https://zwanzigeins.jetzt/aktivitaeten/projekte/unterrichtsversuche>
- Schmidt, S. (2009). Arithmetische Kenntnisse am Schulanfang. In: A. Fritz, G. Ricken, S. Schmidt (Hrsg.), *Handbuch Rechenschwäche: Lernwege, Schwierigkeiten und Hilfen bei Dyskalkulie*, 77–99. Weinheim und Basel: Beltz.
- Schneider, W., Hasselhorn, M. & Bengel, J. (2008). *Handbuch der pädagogischen Psychologie*. Göttingen, Bern, Wien, Paris, Oxford, Prag, Toronto, Cambridge, Amsterdam, Kopenhagen: Hogrefe.
- Schipper, W., Ebeling A. & Dröge R. (2022). *Handbuch für den Mathematikunterricht: 2. Schuljahr*. Braunschweig: Westermann.
- Schuchardt, K., Brandenburg, J., Fischbach, A., Büttner, G., Grube, D., Mähler, C., Hasselhorn, M. (2015). Die Entwicklung des akademischen Selbstkonzeptes bei Grundschulkindern mit Lernschwierigkeiten. *Zeitschrift für Erziehungswissenschaft* 18, 513–526.

- Schulte-Körne, G. (2016). Mental health problems in a school setting in children and adolescents. *Deutsches Ärzteblatt International* 113, 183–90.
- Schuppener, G. (2014). *Warum 21 einundzwanzig heißt: Die höheren Einerzahlwörter im Deutschen. Geschichte ihrer Bildung und Reformideen*. Wien: Praesens Verlag.
- Schwank, I. (2013). Die Schwierigkeiten des Dazu-Denkens. In: M. von Aster & J. H. Lorenz (Hrsg.), *Rechenstörungen bei Kindern. Neurowissenschaft, Psychologie, Pädagogik*, 93–138. 2. überarb. und erw. Auflage. Göttingen: Vandenhoeck & Ruprecht.
- Shalev, R. S., Auerbach, J., Manor, O., Gross-Tsur, V. (2000). Developmental dyscalculia: prevalence and prognosis. *European Child & Adolescent Psychiatry*, Vol. 9, No. S2 (2000-6), II/58–II/64.
- Shalev, R., Manor, O., Kerem, B., Ayali, M., Badichi, N., Friedlander, Y. Y. F. et al. (2001). Developmental dyscalculia is a familial learning disability. *Journal of Learning Disabilities*, 34, 59–65.
- Statistisches Bundesamt (Destatis): Allgemeinbildende und berufliche Schulen: Schüler/innen nach Schularten.
www.destatis.de/DE/ZahlenFakten/GesellschaftStaat/BildungForschungKultur/Schulen/Tabellen/AllgemeinBildendeBeruflicheSchulenSchulartenSchueler.html. Aufruf am 03.11.2024.
- Stein, P. (2022). Forschungsdesigns für die quantitative Sozialforschung. In: N. Baur & J. Blasius (Hrsg.), *Handbuch Methoden der empirischen Sozialforschung*, 135–151. Wiesbaden: Springer Fachmedien.
- Stern, E. (2005). Kognitive Entwicklungspsychologie mathematischen Denkens. In: M. von Aster & J. H. Lorenz (Hrsg.), *Rechenstörungen bei Kindern. Neurowissenschaft, Psychologie, Pädagogik*, 15–38. 2. überarb. und erw. Auflage. Göttingen: Vandenhoeck & Ruprecht.
- Sorvo, R., Koponen, T., Viholainen, H., Aro, T., Räikkönen, E., Peura, P., Dowker, A., Aro, M. (2017). Math anxiety and its relationship with basic arithmetic skills among primary school children. *British journal of educational psychology*, 87(3), 309. <https://doi.org/10.1111/bjep.12151>
- Van der Ven, S. H. G., Klaiber, J. D. & van der Maas, H. L. J. (2017). Four and twenty blackbirds: how transcoding ability mediates the relationship between visuospatial working memory and math in a language with inversion. *Educational Psychology* 37(4): 487–505.
- Van Eimeren, L. & Ansari, D. (2009). Rechenschwäche – eine neurokognitive Perspektive. In: A. Fritz, G. Ricken, S. Schmidt (Hrsg.), *Handbuch*

Rechenschwäche: Lernwege, Schwierigkeiten und Hilfen bei Dyskalkulie.
25–33. Weinheim und Basel: Beltz.

- Vannebo, K.-I. (2008). Die norwegische Zahlsprechreform von 1951. In: L. Gerritzen (Hrsg.), *Zwanzigeins: Für die unverdrehte Zahlensprechweise; Fakten, Argumente und Meinungen*, 95–104. Bochum: Brockmeyer.
<https://zwanzigeins.jetzt/infos/literatur>. Aufruf am 20.04.2024.
- Vloet, T. D. & Romanos, M. (2021). Angststörungen – von der ICD-10 zur ICD-11. *Zeitschrift für Kinder- und Jugendpsychiatrie und Psychotherapie* 49(6). 429–435, Nov. 2021. doi: 10.1024/1422-4917/a000768
- Voigt, J. (2008). *Die Zahlensprache in Norwegen*. In: L. Gerritzen (Hrsg.). *Zwanzigeins - für die unverdrehte Zahlensprechweise*. Bochum: Brockmeyer.
<https://zwanzigeins.jetzt/infos/literatur>. Aufruf am 02.04.2024
- Wagner, J., Arifoski, A. & Kucian, K. (2025). Unterscheiden sich Mädchen und Jungen im Interesse und der geschlechterspezifischen Leistungseinschätzung in Mathematik? *Lernen und Lernstörungen*, 14(1), 15–25. <https://doi.org/10.1024/2235-0977/a000468>
- Weißhaupt, S. & Peucker, S. (2009). Entwicklung arithmetischen Vorwissens. In: A. Fritz, G. Ricken, S. Schmidt (Hrsg.), *Handbuch Rechenschwäche: Lernwege, Schwierigkeiten und Hilfen bei Dyskalkulie*, 52-76. Weinheim und Basel: Beltz.
- Wendt, H., Bos, W., Selter, C., Köller, O., Schwippert, K., Kasper, D. (2016). *TIMSS 2015. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich*. Münster und New York: Waxmann. <https://zwanzigeins.jetzt/aktivitaeten/projekte/timss>
- World Health Organization. <https://www.who.int/>. Aufruf am 11.02.2025.
- Xie, F., Xin, Z., Chen, X. et al. (2019). Gender Difference of Chinese High School Students' Math Anxiety: The Effects of Self-Esteem, Test Anxiety and General Anxiety. *Sex Roles* 81: 235–244. <https://doi.org/10.1007/s11199-018-0982-9>
- Younge, J. O., Kouwenhoven-Pasmooij, T. A., Freak-Poli, R., Roos-Hesselink J. W. & Hunink M. G. M. (2015). Randomized study designs for lifestyle interventions: a tutorial. *International Journal of Epidemiology*: 1–14. doi: 10.1093/ije/dyv183
- Zuber, J., Pixner, S., Möller, K. & Nürk, H.-C. (2009). On the language specificity of basic number processing: Transcoding in a language with inversion and its relation to working memory capacity. *Journal of experimental child psychology*, 102 (1), 60–77.

Zwanzigeins e. V. (2017). *Immaterielles Kulturerbe 2017*.
<https://zwanzigeins.jetzt/aktivitaeten/projekte/immaterielles-kulturerbe-2017>.
Aufruf am 12.09.2024.

Zwanzigeins e. V. (2024a). *Positionspapier des Vereins Zwanzigeins e. V. – Vorschlag zur Reform der Zahlensprechweise im Deutschen*.
https://zwanzigeins.jetzt/downloads/Positionspapier_Zahlensprechweise_Zwanzigeins%20eV_03.10.2024.pdf. Aufruf am 07.12.2024.

Zwanzigeins e. V. (2024b). *Deutsch und Niederländisch – Auswirkungen der irregulären Zahlensprechweise*. <https://zwanzigeins.jetzt/infos/deutsch-und-niederlaendisch>. Aufruf am 12.09.2024.

Zwanzigeins e. V. (2024c). *Walisisch (Kymrisch), Tamil, Chinesisch vs. Englisch*.
<https://zwanzigeins.jetzt/infos/walisisch-tamil-chinesisch-vs-englisch>. Aufruf am 13.09.2024.

9 PROGRAMM & APP

StataCorp. (2015). *Stata Statistical Software: Release 14*. College Station, USA: StataCorp LP.

Zwanzigeins-App. <https://zwanzigeins.jetzt/app/>

10 ANHÄNGE

ANHANG 1 **GESPRÄCHSLEITFADEN EINWEISUNG**

Untersucherin (= Verfasserin): „Dein(e) Therapeut(in) hat dir ja erzählt, dass wir eine wissenschaftliche Studie machen. Schön, dass du mitmachst! Es ist ein Spiel am Tablet und es geht um die verdrehte Zahlsprechweise im Deutschen. Ich zeige dir mal, was ich meine.

[Erstellung eines Tafelbilds]

Wenn ich hier [*Name des Kindes*] schreibe, in welche Richtung liest du das, also von wo nach wo?

Kind: Von links nach rechts. [Die Kinder zeigten es manchmal auch: „Von da nach da.“]

Untersucherin: Sehr gut. [Am Tafelbild wurde ein Pfeil von links nach rechts über der Schrift eingezeichnet]. Wenn du jetzt die Zahl hier [21] liest, wie ist es da?

Kind: Einundzwanzig. Da liest man es andersrum, von rechts nach links.

Untersucherin: Genau. [Am Tafelbild wurde ein Pfeil von rechts nach links über der Zahl eingezeichnet]. Und da gibt es Forscher, die vermuten, dass es an dieser umgekehrten Sprechweise, liegen könnte, dass manchen Kindern das Rechnen schwerfällt. Sie denken, dass es einfacher wäre, wenn man die Zahlen auch von links nach rechts aussprechen würde [Pfeilrichtung wurde bei 21 entsprechend abgeändert], so, wie du die normale Schrift liest.

Also man würde zum Beispiel zwanzig-eins sagen. Oder [46 wurde im Tafelbild auf einem Zahlenstrahl ergänzt] vierzig-sechs. Die [30 wurde im Tafelbild auf einem Zahlenstrahl ergänzt] dreißig würde bleiben. Und das [11 wurde im Tafelbild auf einem Zahlenstrahl ergänzt] wäre zehn-eins. Und das hier [14 wurde im Tafelbild auf einem Zahlenstrahl ergänzt] – was wäre das?

Kind: Zehn-vier.

Untersucherin: Genau. Hast du Lust, das mal am Tablet auszuprobieren? Wir machen immer einen Probedurchlauf zum Ausprobieren und einen, bei dem auch

die Zeit und die Fehleranzahl gemessen werden. Das ist aber nur für die Statistik und dass du es für dich selbst vergleichen kannst, also es ist kein Wettbewerb.

Und dann würde ich danach gern wissen, wie es für dich war, also ob das eine oder das andere leichter für dich ist oder ob es gleich ist.

Das Tablet spricht dir jetzt die Zahlen vor, einmal so, wie du es kennst, und einmal so, wie ich es dir gerade erklärt habe und du sollst die Zahlen eingeben, das sieht dann so ähnlich aus wie eine Handytastatur. Bist du bereit?

Dann geht es los.

Dauer: ca. 2-5 Minuten, je nach Interaktion.

Nach dem Test:

Untersucherin: Wie war es für dich?

Hat dir eine Sprechweise besser gefallen oder war leichter als die andere?

ANHANG 2 EINVERSTÄNDNIS ZUR TEILNAHME AN DER STUDIE

Liebe Eltern,

ich bin Kollegin der Lerntherapeutin / des Lerntherapeuten Ihres Kindes am Duden Institut und würde im Rahmen einer Studie gern ein ca. 3-minütiges Konzentrationsspiel am Tablet mit Ihrem Kind durchführen, in dem es um die verdrehte Zahlensprechweise in der deutschen Sprache geht (Sprechweise für 21: einundzwanzig).

Die Auswertung erfolgt anonymisiert; über die Ergebnisse informieren wir Sie gern nach dem Abschluss der Studie.

Wir würden uns freuen, wenn Sie uns unterstützen und die Einwilligung dazu geben, vor oder nach einer Lerntherapieeinheit Ihr Kind an der Studie teilnehmen zu lassen.

Mit freundlichen Grüßen

.....

Untersucherin

Ich stimme der Teilnahme meines Kindes an o.g. Studie zu.

.....

Datum, Unterschrift Sorgeberechtigte/r

ANHANG 3 UMSTEIGELISTE UND RANDOMISIERTE SEQUENZEN (VORLAGE)

Klarname des Kindes		Profilname des Kindes	Randomisierte Sequenz per Münze <small>Angegeben ist der Startmodus</small>	Anmerkung
		201	Tv	
		202	Ze	
		203	Tv	
		204	Ze	
		205	Tv	
		206	Ze	
		207	Tv	
		208	Ze	
		209	Tv	
		210	Ze	
		211	Tv	
		212	Ze	
		213	Tv	
		214	Ze	
		215	Tv	
		216	Ze	
		217	Tv	
		218	Ze	
		219	Tv	
		220	Ze	
		221	Tv	

Abkürzungen:

Tv: traditionell verdreht / realisierte Sequenz: TvTvZeZe

Ze: zehneins / realisierte Sequenz: ZeZeTvTv

Kopf: Startmodus Tv

Zahl: Startmodus Ze

Eigenständigkeitserklärung

für Abschlussarbeiten

für schriftliche Modulprüfungsarbeiten

Name, Vorname:	Hartwig, Vivien
Matrikelnummer:	
Studiengang:	M.A. Integrative Lerntherapie
Semester:	8
1. Prüfer/in:	Prof. Dr. Michael G. von Aster
2. Prüfer/in:	PD Dr. Peter Morfeld

Hiermit versichere ich, dass ich die vorliegende Arbeit mit dem Titel

Zwanzigeins vs. einundzwanzig: Profitieren rechenschwache Kinder in Bezug auf die Fehlerzahl von r

stellenwertgerechten Zahlensprechweise?

Eine empirische Untersuchung

selbstständig und ohne unzulässige fremde Hilfe verfasst habe. Ich habe keine anderen als die angegebenen Quellen und Hilfsmittel benutzt und alle wörtlich oder sinngemäß übernommenen Textstellen als solche kenntlich gemacht. Die Arbeit hat in gleicher Form noch keiner Prüfungsbehörde vorgelegen. Ebenfalls erkläre ich, dass ich noch keine Bachelor- oder Masterthesis in der gleichen oder einer vergleichbaren Studienrichtung endgültig nicht bestanden habe oder mich in einem laufenden Prüfungsverfahren befinde. Die Versicherung gilt auch für beigefügte Zeichnungen, Skizzen oder graphische Darstellungen.

Sollten in der Prüfungsstellung Chatbots oder andere Formen künstlicher Intelligenz (KI) ausdrücklich als Hilfsmittel erlaubt worden sein, versichere ich

- verstanden zu haben, dass diese generativen Sprachmodelle in ihrer Funktionsweise technisch auf dem Wahrscheinlichkeitsprinzip basieren und somit auch wissenschaftlich nicht korrekte Antworten vorliegen können
- alle Abschnitte kenntlich gemacht zu haben, für die Textmaterial aus Chatbots verwendet wurden
- in einer vollständigen Liste im Anschluss an das Literaturverzeichnis alle genutzten Tools, den Zeitpunkt des Zugriffs sowie den von mir jeweils eingegebenen Ausgangstext bzw. die Anfangsphase („prompt“) dokumentiert zu haben
- verstanden zu haben, dass die Qualität der Antwort eines KI-Textmodells stark von der Qualität des „prompts“ abhängt und sich hierin auch die persönliche Leistung und Eigenständigkeit meiner Arbeit zeigen kann.

Dem ausgedruckten Text habe ich einen Datenträger mit der digitalisierten Version der Arbeit beigefügt.

Durch meine Unterschrift bestätige ich, dass die Datenschutzverordnung (EU-DSGVO) vom 27.04.2016 – insbesondere bei personenbezogenen Daten – in der vorliegenden Arbeit eingehalten wurde.

Ort, Datum

Unterschrift