

Automatisierung im Lesen und Schreiben

Studienergebnisse zu den hierarchieniedrigen schriftsprachlichen Kompetenzen im 4. Schuljahr

Eva Odersky¹ , Angelika Speck-Hamdan², Maximilian Stark² 

¹Lehrstuhl für Grundschulpädagogik und Grundschuldidaktik, KU Eichstätt-Ingolstadt, Deutschland

²Lehrstuhl für Grundschulpädagogik und Grundschuldidaktik, Ludwig-Maximilians-Universität München, Deutschland

Zusammenfassung: *Hintergrund:* Basale schriftsprachliche Kompetenzen sind grundlegend für schulische und gesellschaftliche Teilhabe. Während die Lesekompetenz seit Jahren gut erforscht wird, besteht hinsichtlich der Schreibkompetenz, insbesondere die Flüssigkeit des Schreibens betreffend, ein gewisser Nachholbedarf. Die Komplexität des Schreibprozesses macht es unmöglich, einfache Wirkzusammenhänge zu beschreiben, doch gelingt es in (Strukturgleichungs-)Modellen zunehmend, unterschiedliche Faktoren und Bedingungen für Schreibflüssigkeit (*transcription* und *translation*) herauszukristallisieren. *Methoden:* Die vorliegende explorative Studie mit 336 Viertklässler_innen vertieft diese Befunde durch prozessorientierte Daten zur Automatisierung des Handschreibens und prüft Zusammenhänge zur Rechtschreib- und Leseleistung. *Ergebnisse:* Bei der Analyse der hierarchieniedrigen *transcription*-Prozesse (Hand- und Rechtschreiben) wird sichtbar, dass sich die Ergebnisse insbesondere in Bezug auf das Geschlecht unterscheiden. Mädchen, die automatisiert handschreiben, machen weniger Rechtschreibfehler, bei den Jungen ist ein solcher Zusammenhang nicht feststellbar. Auch Automatisierung des Handschreibens und Leseleistung korrelieren nur bei den Mädchen signifikant. Für alle Kinder kann dagegen eine starke Korrelation der Rechtschreib- mit der Leseleistung gezeigt werden. *Diskussion:* Unter den untersuchten Indikatoren für die Automatisierung des Handschreibens erweist sich die benötigte Schreibzeit als bester Prädiktor für Rechtschreib- und Leseleistung.

Schlüsselwörter: Schreibflüssigkeit, Automatisierung, Handschreiben, Rechtschreiben, Lesen

Automated Reading and Writing: Results on Correlations in Written Language Competencies in Grade 4

Abstract: *Background:* Basic literacy skills are crucial for participation in school and society. While reading skills have been well researched for years, there is a certain amount of catching up to do when it comes to writing skills, especially writing fluency. While the complexity of the writing process precludes explanations based on simple causal relations, (structural equation) models increasingly enable researchers to determine different factors and conditions affecting writing fluency (*transcription* and *translation*). *Methods:* The present explorative study with 336 fourth graders deepens these findings with process-oriented data on handwriting automaticity and examines correlations with spelling and reading performance. *Results:* When analyzing the *transcription* processes (*handwriting* and *spelling*), it becomes apparent that the results differ especially with respect to gender. For girls, higher handwriting automaticity is associated with fewer spelling errors, but no such correlation can be found for boys. Likewise, handwriting automaticity and reading performance only correlate significantly for girls. A strong correlation of spelling performance with reading performance was observed for all children. *Discussion:* Out of the possible indicators of handwriting automaticity examined in this study, writing time proved to be the best predictor of spelling and reading performance.

Keywords: writing fluency, handwriting, spelling, reading, gender specifics

Einleitung

Das Hand- bzw. Tastaturschreiben zählt ebenso wie das Rechtschreiben zu den hierarchieniedrigen Schreibprozessen (vgl. Sturm, Nänny & Wyss, 2017). Hierarchieniedrige Prozesse zeichnen sich dadurch aus, dass sie bei entsprechender Expertise weitgehend automatisiert ablaufen und hierarchiehöhere Prozesse kaum beeinträchtigen. Erfahrene Schreiberinnen und Schreiber können sich daher auf die kognitiv anspruchsvolleren Prozesse des Formulierens, Planens und

Revidierens besser konzentrieren. In diesem Kontext werden die exekutiven Funktionen, speziell das Arbeitsgedächtnis mit seinen limitierten Ressourcen, als eine Art Nadelöhr für den Schreibprozess betrachtet (Olive, 2014; Kellogg, 2022). Je reibungsloser, sprich automatisierter, die hierarchieniedrigen Teilprozesse ablaufen, desto mehr Ressourcen sind für die hierarchiehöheren Teilprozesse frei. Diese beiden hierarchieniedrigen Prozessen konkurrieren allerdings auch untereinander, wobei dieser Konflikt mit zunehmender Expertise abnimmt (Torrance & Galbraith, 2006; Odersky, 2018).

Der Grad der erreichten Automatisierung, der seinerseits stark von Übung abhängt, beeinflusst die Textqualität, wie Studien nahelegen, die sowohl Textumfang als auch Textqualität in ihrer Beziehung zu den basalen Schreibfähigkeiten (handschriftliche Flüssigkeit sowie Rechtschreiben) untersucht haben (vgl. z. B. Meta-Analyse von Kent & Wanzenek, 2016). Demnach können hierarchieniedrige Schreibfähigkeiten als Prädiktor für hierarchiehöhere Schreibkompetenzen gelten (Sturm et al., 2017), und zwar in unteren Jahrgangsstufen deutlicher als in höheren, was sich mit der Abhängigkeit von Erfahrung und Übung gut erklären lässt. Aktuell konnte die Studie von Skar et al. (2022) mit insgesamt fast 5000 Grundschulkindern einen signifikanten Zusammenhang zwischen der *handwriting fluency* und der *writing quality* nachweisen. Ein besseres Verständnis dieser Zusammenhänge erlaubt daher nicht nur Einsichten in den Schreibprozess. Es trägt auch dazu bei, den Zusammenhang zwischen den exekutiven Funktionen und komplexen intellektuellen Handlungen (wie dem Schreiben) besser zu verstehen und stellt damit eine wichtige Grundlage für die Entwicklung effektiven Schreibunterrichts dar (Kellogg, 2022).

Theoretische Einbettung

Auf die Komplexität des Schreibens als *full-time cognitive overload* hatten schon 1980 Hayes und Flower im Zusammenhang ihres bis heute viel zitierten, grundlegenden Schreibmodells hingewiesen. Die Darstellung eines Zusammenspiels von kognitiven, sprachlichen und motorischen Teilprozessen beim Schreiben ist keineswegs trivial. Das Bild einer linearen Abfolge von der Idee zum fertigen Text griffe zu kurz. Die beteiligten Prozesse greifen nicht nur ineinander, sie erfordern Rückkoppelungen und Vorausplanungen, und das nicht nur nacheinander, sondern auch gleichzeitig, vorausschauend und rückblickend, wie beispielsweise Olive (2014) in seinem Kaskadenmodell veranschaulicht. Diese Komplexität des Schreibens wird auch in der Gegenüberstellung der *Simple View of Writing* mit einer *Not-so-simple View of Writing* durch Berninger und Winn (2006) deutlich, wobei sich das „nicht so einfache“ Modell des Schreibens dadurch auszeichnet, dass es zusätzlich zur Transkription (Handschriften/Tastaturschreiben, Rechtschreiben) und zur Textproduktion (Formulieren) die exekutiven Funktionen als Teilprozesse des Schreibens einführte, insbesondere das Arbeitsgedächtnis, die Aufmerksamkeit und die inhibitorische Kontrolle.

Im weiterentwickelten *Direct and Indirect Effects Model of Writing (DIEW)* wird neben kognitiven Funktionen auch die mündliche Sprachverwendung als Einflusskomponente berücksichtigt und später noch um den Einfluss der Lesefähigkeit auf das Schreiben erweitert (Kim & Schatschneider, 2017; Kim & Graham, 2022). Bereits Kent und Wanzenek

wiesen in ihrer Meta-Analyse (2016) auf die Bedeutung des Lesens für das Schreiben hin. Sowohl Kim und Graham (2022) wie auch Linnemann et al. (2022) können in ihren Strukturgleichungsmodellen ebenfalls die Bedeutung exekutiver Funktionen und vielfache und vielschichtige Zusammenhänge zwischen dem Lesen und Schreiben aufzeigen.

In einer Längsschnittstudie stellten Kim, Petcher, Wanzenek und Al Otaiba (2018) einen signifikanten Zusammenhang zwischen der Entwicklung der basalen Lesefähigkeit und dem Rechtschreiben – beides hierarchieniedrige schriftsprachliche Fähigkeiten – fest, wobei der Einfluss einseitig vom Lesen auf das Rechtschreiben gegeben war, nicht aber vom Rechtschreiben auf das Lesen: „... reading-writing relations are stronger at the lexical level than at the discourse level and may be a unidirectional one from reading to writing at least between Grades 3 and 6“ (Kim et al., 2018, S. 1591).

Diese Ergebnisse lassen sich sehr gut einordnen in das *Interactive Dynamic Literacy Model* (Kim, 2020), das den Zusammenhang zwischen Lese- und Schreibkompetenzen im Erwerbsprozess berücksichtigt und in einen größeren Kontext der Sprachverwendung stellt. Ausgehend von der Tatsache, dass Sprache einerseits mündlich über Ohr und Mund (aural und oral) und andererseits schriftlich auch über Auge und Hand (Lesen und Schreiben) genutzt wird, dabei Schriftlichkeit Mündlichkeit voraussetzt, wird in diesem Modell auch die mündliche Sprache einbezogen und die Bedeutung der allgemeinen Kognition bzw. der exekutiven Funktionen als basale Voraussetzung gesehen. Sowohl auf Ebene der grundlegenden Lese- und Schreibfähigkeiten (*Foundational Literacy*) als auch auf Ebene der höheren Lese- und Schreibfähigkeiten (*Discourse-Level Literacy*) nimmt Kim eine Wechselwirkung zwischen Lesen und Schreiben an, wobei beide Modi nicht nur auf allgemeinen kognitiven und sprachlichen Fähigkeiten aufbauen, sondern dabei auch interagieren, sich beeinflussen, sich gegenseitig verstärken und gemeinsam entwickeln (Kim, 2020).

Empirisch konnten Kim und Graham (2022) entsprechend dem erweiterten *DIEW*-Modell einen signifikanten Einfluss aller untersuchten sprachlichen und kognitiven Fähigkeiten auf die Textqualität zeigen; dazu gehörten Leseverstehen, Handschreibflüssigkeit, mündliche Ausdrucksfähigkeit, Rechtschreiben, Wortschatz, grammatikalisches Wissen, Aufmerksamkeit, Perspektivenübernahme, Schlussfolgerung, Arbeitsgedächtnis und Selbstmonitoring-Fähigkeiten. Das Leseverstehen korrelierte eng mit der Textqualität (*writing quality*), nicht jedoch mit der Textmenge (*writing productivity*) und der Korrektheit des Textes (*correctness in writing*). Für alle drei Bereiche allerdings spielte die Handschreibflüssigkeit neben der Aufmerksamkeitskontrolle eine entscheidende Rolle.

Auch andere Studien belegen eine Korrelation von Lesen bzw. Rechtschreiben mit der Handschreibflüssigkeit,

d.h. schwächere Rechtschreib- bzw. Leseleistungen und ein niedrigerer Grad an Handschreibflüssigkeit scheinen miteinander zusammenzuhängen (Hurschler Lichtsteiner, Saxer Geiger & Wicki, 2010; Kent & Wanzek, 2016). Als entscheidende Größe für die Handschreibflüssigkeit gilt der Grad der Automatisierung, welche in den erwähnten Studien in der Regel (nicht bei Hurschler Lichtsteiner et al., 2010) pauschal über die Schreibzeit erfasst wird.

Ausgehend von diesen Erkenntnissen lohnt es sich daher, die Einflussgrößen auf das Schreiben gerade im Erwerbsprozess in einem größeren Kontext zu sehen und dabei einzelne Einflüsse näher zu beleuchten. Sprachverständnis und Weltwissen etwa dürften ebenso eine Rolle spielen wie Gedächtnis und Konzentrationsfähigkeit. Nicht zu vernachlässigen sind bei allen Lernprozessen die Effekte der Lernumgebung, die sich z. B. mit der Schulklassenzugehörigkeit überprüfen lassen, aber auch mit den subjektiven Lerntheorien der Lehrpersonen oder den Vorgaben eines bestimmten Curriculums (vgl. Diskussion zur Ausgangsschrift und zur Ein- bzw. Zweiphasigkeit des Erwerbsprozesses Odersky, Speck-Hamdan, Mahrhofer-Bernt & Marquardt, 2021, S. 78f).

Immer wieder wird in den genannten schriftsprachlichen Fähigkeiten ein Vorsprung der Mädchen aufgezeigt. Im Bereich Orthografie lag der Kompetenzvorsprung der Mädchen z. B. in der IQB-Studie am Ende der 4. Jahrgangsstufe 2016 mit $d = 0.33$ (Stanat et al. 2022, S. 129) und 2021 mit $d = 0.29$ (Stanat et al., 2022, S. 131) im kleinen bis mittleren Bereich (Gignac & Szodorai, 2016). Damit ist die Orthografie im Fach Deutsch – untersucht wurden außerdem die Bereiche Zuhören und Lesen – der Bereich mit dem größten Kompetenzunterschied zwischen Mädchen und Jungen (Stanat et al., 2022, S. 131). Die Metaanalyse von Cordeiro, Castro und Limpo (2018) bestätigt in allen untersuchten Teilprozessen des Schreibens eine Überlegenheit der Mädchen und vermutet als einen Grund für die besseren Leistungen der Mädchen beim Texteverfassen ihre Überlegenheit im Handschreiben: „Overall, findings confirmed the gender difference typically found in writing and indicated that potential explanatory variables for it may be handwriting and self-efficacy“ (Cordeiro et al., 2018, S. 1). Diese Vermutung wird in den oben zitierten elaborierten neuen Schreibmodelle abgebildet, die die Komplexität des Schreibprozesses und die Bedeutung der Automatisierung hierarchieniedriger Prozesse betonen.

Ableitung der Forschungsfragen

Es gilt die theoretische Annahme zu prüfen, nach der die Automatisierung der Graphomotorik Kapazitäten freisetzt, die dann für andere Teilprozesse des Schreibens zur Verfügung stehen und deren Qualität positiv beeinflussen,

und zwar spezifiziert am Beispiel des Rechtschreibens. Da orthografische und graphomotorische Prozesse in der Ausführung jeder Schreibhandlung am Ende stehen und allein durch diese zeitliche Nähe eng verknüpft sind, stellt sich die Frage, inwiefern diese beiden hierarchieniedrigen, möglicherweise konkurrierenden schriftsprachlichen Kompetenzen korrelieren und ob eventuelle Effekte auch bei Kontrolle von Schulklassenzugehörigkeit, Geschlecht und anderen Einflüssen zu beobachten sind.

Der für die Textqualität nachgewiesene Gewinn durch eine bessere Handschreibflüssigkeit wirkt auf Grundlage des Forschungsstandes darüber hinaus die Frage auf, ob generell die hierarchiehöheren schriftsprachlichen Kompetenzen, also auch das Leseverständnis, von einer besseren Automatisierung der Graphomotorik profitieren.

Folgenden Forschungsfragen wollen wir daher nachgehen:

1. Gibt es nach Berücksichtigung der Zugehörigkeit der Schüler_innen zu ihren Klassen einen Zusammenhang zwischen der Automatisierung des Handschreibens und der Rechtschreibleistung (F1a)? Welche Rolle spielen dabei geschlechtsbezogene Unterschiede (F1b)?

Unabhängig vom Verhältnis der beiden hierarchieniedrigen Teilprozesse des Schreibens (Handschreiben und Rechtschreiben) zueinander wird ein positiver Zusammenhang zwischen diesen Fertigkeiten und der Lesefähigkeit vermutet:

2. Gibt es nach Berücksichtigung der Zugehörigkeit der Schüler_innen zu ihren Klassen einen Zusammenhang zwischen den Leistungen im Rechtschreiben und Lesen (F2a) bzw. zwischen der Automatisierung des Handschreibens und der Leseleistung (F2b)? Welche Rolle spielen jeweils geschlechtsbezogene Unterschiede (F2c und 2d)?

Methoden

Um diese Forschungsfragen zu beantworten, wurde eine quantitativ-empirische Studie durchgeführt, in die sowohl standardisierte Leistungstests zum Rechtschreiben und Lesen als auch digitale Aufnahmen zur validen, prozessorientierten Erhebung der Automatisierung des Handschreibens integriert waren.

Beschreibung der Stichprobe

Die Erhebung wurde in 23 vierten Klassen an zwölf Schulen durchgeführt. Die Schulen wurden so ausgewählt, dass sie sich in ihrer Struktur unterscheiden. Es waren ein- und

mehrzügige Schulen im ländlichen Raum genauso beteiligt wie mehrzügige Schulen in unterschiedlichen Stadtvierteln einer Großstadt (> 1 Million) und fünf- von Jahrgang 1 bis 4-jahrgangsgemischte Klassen. Die Teilnahme an der Studie war für Schulen, Lehrer_innen und Kinder (bei Vorliegen der Einwilligung der Erziehungsberechtigten) freiwillig. Die Zahl der teilnehmenden Viertklässler_innen pro Schule und pro Klasse variiert aus diesen Gründen insbesondere aufgrund der kleinen Lerngruppen des 4. Schuljahrs in der Jahrgangsmischung, und liegt im Mittel bei 14 Kindern pro Klasse.

Für die Leistungserhebung wurde in einem Zeitraum von acht Wochen in allen 23 Klassen der standardisierte Schulleistungstest KLASSE 4 im Klassenverband von der Versuchsleiterin durchgeführt. Anschließend (und immer am gleichen Tag) wurde das Handschreiben der Schüler_innen einzeln in einem separaten Raum der Schule digital aufgezeichnet; diese Aufzeichnung dauerte pro Kind circa 15 Minuten und wurde von zwei Versuchsleiter_innen parallel durchgeführt. Die Erhebungen erfolgten anonym; personenbezogen wurden das Alter (Jahre; Monate), das Geschlecht, die Händigkeit und die (Mehr-)Sprachigkeit erfasst.

Insgesamt konnten von $N = 336$ Kindern sowohl Werte zur Rechtschreib- und Leseleistung als auch zur Automatisierung der Handschrift erhoben werden. Die Kinder waren im Durchschnitt $M = 10;3$ Jahre alt ($SD = 5$ Monate), 52% der Kinder ($n = 174$) weiblich und 48% ($n = 162$) männlich, 92% ($n = 308$) rechts- und 8% ($n = 28$) links-händig. Bei 68% ($n = 228$) der Kinder wurde zuhause nur deutsch gesprochen, 32% ($n = 107$) hatten einen mehrsprachigen familiären Hintergrund.

Alle 336 Kinder besuchten die 4. Jahrgangsstufe einer deutschsprachigen Grundschule in Bayern und hatten zu Beginn des 1. Schuljahrs eine – im bayerischen Lehrplan vorgeschriebene – Druckschrift als unverbundene Erstschrift erlernt und im 2. Schuljahr die Vereinfachte Ausgangsschrift als verbundene Schrift.

Instrumente für die Erhebung der Rechtschreib- und Leseleistung

In den 23 beteiligten Klassen wurde der Allgemeine Schulleistungstest „KLASSE 4. Kombiniertes Leistungsinventar zur allgemeinen Schulleistung und für Schullaufbahneempfehlungen in der vierten Klasse“ (Lenhard, Hasselhorn & Schneider, 2011) eingesetzt, da er die standardisierte Durchführung mehrerer Untertests zu verschiedenen Lernbereichen unter objektiven Bedingungen und vertretbarem Zeitaufwand (ca. 60 Minuten) ermöglicht. Zudem weist er insbesondere für die hier relevanten Lernbereiche eine gute und stabile Reliabilität auf (Lenhard et al., 2011, S. 23):

- Das im Rahmen von KLASSE 4 durchgeführte Diktat ist dem DERET 3–4+ entnommen (Cronbachs $\alpha = 0.93$).
- Der im Rahmen von KLASSE 4 durchgeführte Untertest zum Leseverständnis ist dem ELFE 1–6 entnommen (Cronbachs $\alpha = 0.80$).

Durchführung und Auswertung des KLASSE-4-Tests erfolgten gemäß Testmanual.

Instrument für die Erhebung der Automatisierung des Handschreibens

Um den Grad der Automatisierung der Schriften zu bestimmen, schrieb jedes Kind mit einem Wacom Inking Pen, der kabellos und im Handling einem Kugelschreiber vergleichbar ist, auf handelsübliches liniertes Schreibpapier, das auf einem Wacom Intuos Tablet auflag. Dieses Grafik-Tablet war mit einem Laptop verbunden, über das eine digitale Aufnahme des Schreibens gesteuert und gespeichert wurde. Für diese Aufnahmen und die spätere Analyse der Schriftkennwerte wurde das Programm CS-Win 2012 verwendet (Marquardt, 2014).

Jedes Kind absolvierte 16 Schreibaufgaben: kurze Aufnahmen zur Testung der Schreibmotorik (Kritzeln, Kringeln, Schleifen), repetitives Schreiben von Wörtern und Schreiben von Sätzen in unterschiedlichen Schreibmodi: Diktat, Abschreiben und Freies Schreiben. Für die hier präsentierte Studie sind zwei dieser Schreibmodi relevant und berücksichtigt:

- Diktat: „Die Kinder schreiben auf Papier.“ (Anhören des gesamten Satzes)
- Abschreiben: „Löwen und Tiger sind gefährliche Raubkatzen.“ (Präsentation des gesamten Satzes)

Die Schreibmodi „Diktat“ bzw. „Abschreiben“ verlangen den Kindern ganz spezifische Fertigkeiten ab, und zwar gerade im Hinblick auf die für Automatisierungsprozesse relevanten zugrunde liegenden Arbeitsgedächtnisleistungen. Um einen Wiederholungseffekt zu vermeiden, unterscheiden sich die beiden Sätze für die Schreibmodi „Diktat“ und „Abschreiben“. In diesen beiden Aufnahmen wurden – anders als beim Freien Schreiben – von allen Schülern_innen aber jeweils die identischen Inhalte geschrieben, so dass auch die für das Schreiben benötigte Zeit verglichen werden kann.

Diese Schreibzeit wird in so genannten produktorientierten Verfahren zur Bestimmung der Automatisierung bzw. Flüssigkeit des Handschreibens als einziges Kriterium verwendet, da sie niedrigschwellig – durch Stoppen der fürs Schreiben benötigten Zeit – erhoben werden kann. Darüber hinaus können in so genannten prozessorientierten Verfahren mit Programmen wie dem hier verwendeten

CSWin zahlreiche kinematische Parameter erhoben und das Schreiben insgesamt einer detaillierten qualitativen und quantitativen Analyse zugänglich gemacht werden (Odersky et al., 2021).

Berücksichtigte Werte für die Beschreibung der Automatisierung des Handschreibens

Folgende drei Werte sind für die Bestimmung der Automatisierung einer Schrift von besonderer Bedeutung und wurden in der hier vorgestellten Studie berücksichtigt (vgl. inhaltlich und zu allen Werten Marquardt, 2011):

- **Schreibzeit:** Die effektiv für das Schreiben benötigte Zeit wird in der digitalen Aufnahme sehr präzise zwischen dem Aufsetzen des Stiftes für den ersten Buchstaben des zu schreibenden Satzes bis zum Abheben nach dem letzten Buchstaben in Millisekunden erfasst.
- **Frequenz:** Die Schreibfrequenz erfasst die Schreibgeschwindigkeit anhand der Anzahl der Auf- und Abstriche in der Bewegungsausführung, sozusagen die Grundfrequenz der Handbewegungen. Sie wird in Hertz (Hz) gemessen. Je höher die Frequenz beim Schreiben ist, desto flüssiger und automatisierter ist das Schreiben. Erwachsene schreiben in einer Frequenz von etwa fünf Hertz, in einer Untersuchung der normalen Schreibleistung von 91 Testpersonen zum Beispiel mit $M = 4.67$ Hz ($SD = 0.71$ Hz) (Marquardt, 2011, S. 387).
- **NIV:** In einem Geschwindigkeitsprofil wird die Anzahl der Geschwindigkeitswechsel (*Number of Inversions in Velocity*) pro Bewegungseinheit berechnet. Je niedriger diese Anzahl, desto automatisierter erfolgt die Schreibeinheit. Der Optimalwert ist 1.0, wenn pro Bewegungseinheit nur ein Geschwindigkeitswechsel erfolgt, also ein Wechsel zwischen Beschleunigung und Bremsung, z. B. bei der Ausführung eines Strichs oder Bogens. Je stärker das Schreiben (visuell) kontrolliert ist, je häufiger also innerhalb der Bewegungsausführung Kontrollvorgänge – und damit Geschwindigkeitswechsel – ausgeführt werden, desto höher ist der NIV und desto weniger automatisiert ist die Schrift. Erwachsene erreichen einen NIV von etwa 1.1, die oben erwähnten 91 Testpersonen beispielsweise $M = 1.13$ ($SD = 0.11$) (Marquardt, 2011, S. 387).

Statistische Methoden

Die Zusammenhänge zwischen den relevanten Variablen wurden als lineare Modelle berechnet. Da der Fokus der Forschungsfragen auf interindividuellen Unterschieden liegt, musste die Clusterung der Kinder in Schulklassen berücksichtigt werden. Die drei Automatisierungsparameter NIV, Frequenz und Schreibzeit wiesen in Bezug auf die

Clustervariable „Klasse“ Intraklassen-Korrelationen zwischen .07 und .14 auf. Ab einem Intraklassen-Koeffizienten von .10 kann die Clusterung der Daten nicht mehr als trivial betrachtet werden (Lee, 2000). Deshalb wurde ein gemischtes Modell mit zufälligen Effekten für die Clusterzugehörigkeit der Kinder gewählt.

Durch ein schrittweises Vorgehen wurde bei jeder Fragestellung entschieden, welche zufälligen Parameter (zufälliger *Intercept* oder zufälliger *Intercept* und zufällige Steigung) in das Modell einfließen sollten, wobei als Entscheidungsgrundlage das *Akaike Information Criterion* oder AIC zum Einsatz kam (Müller, Scealy & Welsh, 2013; Yamashita, Yamashita & Kamimura, 2007).

Die Regressionskoeffizienten werden durchgehend standardisiert berichtet. Um das Ausmaß der berechneten Zusammenhänge einstuft zu können, wurde das marginale Pseudo- R^2 nach Nakagawa, Johnson & Schielzeth (2017) herangezogen. Dieser Kennwert gibt an, wie viel der Varianz der abhängigen Variablen durch die festen (nicht jedoch die zufälligen) Prädiktoren erklärt wird. Nach Gignac & Szodorai (2016) in Erweiterung der klassischen Richtwerte von Cohen (1988) lässt sich bei einer Studie wie dieser eine Effektstärke von R^2 bzw. Pseudo- R^2_m von .02 als klein, .04 als moderat und .09 als groß betrachten.

Für alle Datenanalysen kam RStudio (RCore Team, 2013; RStudio Team, 2023) zum Einsatz. Die gemischten Modelle wurden mit dem lme4-Paket für R (Bates, Mächler, Bolker & Walker, 2015) berechnet und die statistischen Grafiken wurden mit dem ggplot2-Paket für R (Wickham, 2016) erstellt.

Um die komplexeren Fragestellungen zu geschlechtsbezogenen Unterschieden zu beantworten, wurden zunächst Modelle berechnet, in die das Geschlecht als Dummy-Variable aufgenommen wurde. Zusätzlich wurde ein Interaktionsterm für das Geschlecht mit dem kontinuierlichen Prädiktor hinzugefügt (Cohen & Cohen, 2003). Anschließend wurden zusätzlich getrennte gemischte Modelle berechnet, um die unterschiedlichen Muster bei Mädchen und bei Jungen zu illustrieren.

Die beiden berücksichtigten Schreibmodi (Diktat und Abschreiben) wurden für jede Forschungsfrage getrennt betrachtet, da sich nicht nur die beiden Sätze inhaltlich und in ihrer Länge unterscheiden (um bewusst eine Wiederholung zu vermeiden), sondern insbesondere auch der Modus des Schreibens mit seinen spezifischen Anforderungen.

Ergebnisse

Aufgrund der Vielzahl von Zusammenhängen, die überprüft wurden, steht in der folgenden Darstellung der Ergebnisse nicht die Signifikanz einzelner Zusammenhänge im Vorder-

Tabelle 1. Pearson-Korrelationen der Variablen; Mittelwert, Standardabweichung und Spanne auf der Diagonalen

	DERET	ELFE	NIV (Diktat)	Frequenz (Diktat)	Schreibzeit (Diktat)	NIV (Abschreiben)	Frequenz (Abschreiben)	Schreibzeit (Abschreiben)
DERET	M = 15.25, SD = 10.62, [0.00; 58.00]	-0.57**	0.11*	-0.13*	0.16**	0.16**	-0.18**	0.27**
ELFE	M = 15.25, SD = 10.62, [0.00; 20.00]		-0.13*	0.10	-0.17**	-0.16**	0.15**	-0.24**
NIV (Diktat)			M = 1.28, SD = 0.32, [1.00; 4.02]	-0.71**	0.71**	0.87**	-0.63**	0.59**
Frequenz (Diktat)				M = 3.33, SD = 0.68, [1.29; 5.42]	-0.83**	-0.67**	0.86**	-0.71**
Schreibzeit (Diktat)					M = 20.44, SD = 4.82, [11.24; 41.25]	0.68**	-0.73**	0.81**
NIV (Abschreiben)						M = 1.40, SD = 0.39, [1.00; 4.55]	-0.75**	0.65**
Frequenz (Abschreiben)							M = 3.19, SD = 0.72, [1.27; 4.92]	-0.77**
Schreibzeit (Abschreiben)								M = 33.77, SD = 8.12, [19.02; 60.00]

Anmerkungen: Grundlage für DERET: Zahl der Rechtschreibfehler im Rechtschreibtest DERET 3–4 +; Grundlage für ELFE: Punktzahl im Lesetest ELFE 1–6; Frequenz in Hertz, Schreibzeit in Sekunden; M = Mittelwert; SD = Standardabweichung; * $p < .05$, ** $p < .01$.

grund; stattdessen liegt der Fokus auf den übergeordneten Mustern, die sich aus den Ergebnissen ablesen lassen.

Deskriptive Statistik

Zunächst wird in deskriptiven Statistiken ein Überblick über die erfassten und im folgenden berücksichtigten Werte für die Gesamtstichprobe und geschlechtsspezifisch gegeben, wobei die Werte wie folgt zu interpretieren sind:

- DERET: Die im DERET 3–4 + erfasste Rechtschreibleistung wird in den Rohwerten (Fehlerzahl) abgebildet, das heißt: Bessere Leistung entspricht einem *niedrigeren* Wert.
- ELFE: Die im ELFE 1–6 erfasste Leseleistung wird über die erreichten Punktwerte abgebildet, das heißt: Bessere Leistung entspricht einem *höheren* Wert.

- NIV: Je niedriger die Anzahl der Geschwindigkeitswechsel, desto automatisierter erfolgt die Schreibbewegung, optimal ist 1.0, das heißt: Bessere Leistung entspricht einem *niedrigeren* Wert.
- Frequenz: Je höher die Schreibfrequenz ist, desto flüssiger und automatisierter erfolgt die Schreibbewegung, das heißt: Bessere Leistung entspricht einem *höheren* Wert.
- Schreibzeit: Je geringer die für das Schreiben benötigte Zeit ist, desto flüssiger ist das Schreiben, das heißt: Bessere Leistung entspricht einem *niedrigeren* Wert.

Die durchschnittliche Fehlerzahl der gesamten Stichprobe im Rechtschreibtest DERET ($M = 15.25$) entspricht gemäß der Normierung des KLASSE-4-Leistungsinventars einem Prozentrang von 60 und einem T-Wert von 52, der durchschnittliche Punktwert der gesamten Stichprobe im

Tabelle 2. Deskriptive Statistiken nach Geschlecht

Mädchen					
Variable	<i>M</i>	<i>SD</i>	min	max	Fehlend
DERET	13.83	9.02	1.00	44.00	0.57 %
ELFE	15.51	3.58	6.00	20.00	0.57 %
Diktat					
NIV	1.20	0.22	1.00	2.48	0.00 %
Frequenz	3.41	0.66	1.57	4.92	0.00 %
Schreibzeit	19.87	4.41	11.24	37.57	0.00 %
Abschreiben					
NIV	1.29	0.29	1.00	3.15	0.00 %
Frequenz	3.33	0.70	1.43	4.92	0.00 %
Schreibzeit	32.31	7.37	19.02	59.92	0.00 %
Jungen					
Variable	<i>M</i>	<i>SD</i>	min	max	Fehlend
DERET	16.79	11.96	0.00	58.00	1.23 %
ELFE	14.48	3.95	0.00	20.00	0.00 %
Diktat					
NIV	1.36	0.38	1.00	4.02	0.00 %
Frequenz	3.25	0.70	1.29	5.42	0.00 %
Schreibzeit	21.06	5.18	11.55	41.25	0.00 %
Abschreiben					
NIV	1.51	0.44	1.03	4.55	0.00 %
Frequenz	3.04	0.73	1.27	4.72	0.00 %
Schreibzeit	35.35	8.61	19.32	60.00	0.00 %

Anmerkungen: *M* = Mittelwert, *SD* = Standardabweichung; Werte für DERET: Zahl der Rechtschreibfehler im Rechtschreibtest DERET 3–4+; Werte für ELFE: Punktzahl im Lesetest ELFE 1–6; Frequenz in Hertz, Schreibzeit in Sekunden.

Lesetest ELFE ($M = 15.01$) einem Prozentrang von 59 und ebenfalls einem T-Wert von 52 (Lenhard et al., 2011, S. 29f.).

Das Handschreiben erfolgt im Diktat eines Satzes („Die Kinder schreiben auf Papier.“) automatisierter als beim Abschreiben eines Satzes („Löwen und Tiger sind gefährliche Raubkatzen.“). Der NIV ist im Diktat niedriger ($M = 1.28$) als beim Abschreiben ($M = 1.40$), die Frequenz im Diktat höher ($M = 3.33$) als beim Abschreiben ($M = 3.19$). Beide Werte stehen für besser automatisiertes Schreiben im Diktat und können evtl. durch die beim Abschreiben notwendigen Lese- und Kontrollvorgänge erklärt werden. Die Schreibzeit kann aufgrund der inhaltlich und in der Länge unterschiedlichen Sätze nicht verglichen werden.

Die drei Automatisierungs-Parameter (NIV, Frequenz und Schreibzeit) korrelierten für den gleichen Schreibmodus (Diktat bzw. Abschreiben) sehr hoch miteinander (Betrag zwischen .65–.85; siehe Tabelle 1 für alle Korrelationen); noch stärker war die Korrelation der drei Automatisierungs-Parameter mit dem jeweils gleichen Parameter im anderen Schreibmodus (Betrag .81–.87). Die negativen Zusammenhänge erklären sich daraus, dass niedrigerer NIV-Wert und niedrigere Schreibzeit, aber höhere Frequenz jeweils Indikatoren für ein höheres Niveau der Automatisierung sind.

Der NIV-Indikator zeigte kleine bis annähernd moderate (Gignac & Szodorai, 2016) Korrelationen mit den Ergebnissen des Rechtschreib- und des Lesetests (Betrag .11–.16). Bei der Frequenz lagen kleine bis moderate Korrelationen vor (Betrag .10–.18), bei der Schreibzeit moderate bis große Korrelationen mit dem Rechtschreib- und Lesetest (Betrag .17–.27).

In allen Werten erreichten die Mädchen bessere Ergebnisse als die Jungen: Weniger Fehler im Rechtschreibtest, mehr Punkte im Lesetest und eine bessere Automatisierung des Handschreibens (niedrigerer NIV, höhere Frequenz, weniger benötigte Schreibzeit; siehe Tab. 2 für alle Korrelationen, zusätzlich visualisiert durch Boxplots in den Abbildungen E1 bis E3 in den elektronischen Supplementen ESM1, ESM2 und ESM3).

Fehlende Werte

Aufgrund des Testaufbaus liegen keine fehlenden Werte für die Automatisierungs-Indikatoren NIV, Frequenz und Schreibzeit vor. Für das Geschlecht der Kinder fehlen ebenfalls keine Angaben, während für die Ergebnisse des Rechtschreib- und Lesetests je weniger als 1% der Werte fehlen. Diese fehlenden Werte sind für die folgenden Analysen nicht bedeutsam und liegen deutlich unter dem von Bennett (2001) vorgeschlagenen Richtwert von 10%. Eine visuelle Begutachtung der Struktur der fehlenden Daten zeigt zudem, dass sie zufällig fehlen (*Missing at Random*). Daher reicht fallweiser Ausschluss für den Umgang mit fehlenden Daten aus (Howell, 2007).

Inferenzstatistik

Beim schrittweisen Aufbau der gemischten Modelle ergab sich in allen Fällen aus dem AIC, dass das Modell mit zufälligem *Intercept*, aber keiner zufälligen Steigung überlegen ist. Daher sind alle folgenden Modelle des Zusammenhangs gemischte lineare Modelle mit zufälligem *Intercept* für die Clustervariable „Klasse“.

Forschungsfrage 1a: Handschreiben und Rechtschreiben

Gibt es nach Berücksichtigung der Zugehörigkeit der Schüler_innen zu ihren Klassen einen Zusammenhang zwischen der Automatisierung des Handschreibens und der Rechtschreibleistung?

Zunächst wurde überprüft, inwiefern die Automatisierungs-Parameter mit der Anzahl der Rechtschreibfehler im DERET zusammenhängen, nachdem die Clusterung in Klassen berücksichtigt wurde.

- NIV und DERET: Für den beim Diktat gemessenen NIV-Wert zeigt sich kein signifikanter Zusammenhang mit

Tabelle 3. Zusammenhänge der Automatisierungs-Parameter mit der Anzahl der Fehler im DERET-Rechtschreibtest im gemischten Modell

	df	F	p	β	Pseudo- R^2_m
NIV (Diktat)	1, 332.96	2.36	.13	0.09	< .01
NIV (Abschreiben)	1, 332.94	6.76**	< .01	0.01	.02
Frequenz (Diktat)	1, 324.71	3.54	.06	-0.10	.01
Frequenz (Abschreiben)	1, 328.95	8.74**	< .01	-0.16	.03
Schreibzeit (Diktat)	1, 332.76	6.55*	.01	0.01	.02
Schreibzeit (Abschreiben)	1, 332.97	21.85**	< .01	0.03	.06

Anmerkungen: df wird als „Freiheitsgrade des Zählers, Freiheitsgrade des Nenners“ angegeben; β ist der standardisierte Regressionskoeffizient; * $p < .05$, ** $p < .01$.

dem DERET-Ergebnis. Beim Abschreiben liegt dagegen ein kleiner, aber signifikanter Zusammenhang vor (siehe Tabelle 3).

- Frequenz und DERET: Es besteht kein signifikanter Zusammenhang zwischen der beim Diktat erfassten Frequenz und dem DERET-Ergebnis. Die Frequenz beim Abschreiben hängt dagegen im kleinen bis mittleren Maß mit dem DERET-Ergebnis zusammen (siehe Tab. 3).
- Schreibzeit und DERET: Das DERET-Ergebnis hängt signifikant mit der Schreibzeit in beiden Schreibmodi (Diktat und Abschreiben) zusammen. Für das Diktat war der Zusammenhang klein, für das Abschreiben moderat bis groß (siehe Tabelle 3).

Zusammengefasst zeigte sich eine stärkere Assoziation des DERET-Ergebnisses mit den Automatisierungs-Indikatoren beim Abschreiben eines Satzes als mit den Automatisierungs-Indikatoren beim Schreiben eines diktierten Satzes. Die benötigte Schreibzeit war dabei in beiden Schreibmodi ein wichtigerer Prädiktor der Rechtschreibleistung als NIV oder Frequenz.

Forschungsfrage 1b: Handschreiben und Rechtschreiben geschlechtsbezogen

Welche Rolle spielen dabei geschlechtsspezifische Unterschiede?

Anschließend gingen wir auf die Geschlechterunterschiede in Bezug auf die Automatisierung beim Hand-

schreiben und DERET-Ergebnisse ein und untersuchten mögliche Zusammenhänge in einem gemischten Modell (alle Werte des gemischten Modells siehe Tabelle E1 im ESM4):

- NIV und DERET: Zunächst wurde in einem kombinierten Modell überprüft, ob NIV und Geschlecht die Anzahl der Fehler im DERET-Test vorhersagen, wenn die Cluster-Variable Klasse berücksichtigt wird. Hier gibt es im Diktat keine signifikanten Zusammenhänge mit den Prädiktoren oder ihrem Interaktionsterm. Beim Abschreiben wird der Haupteffekt von NIV signifikant, $F(1, 332.29) = 5.57, p = .02, \beta = 0.05$, aber nicht der Haupteffekt des Geschlechts oder der Interaktionseffekt. Die Effektstärke für die Modelle ist klein bis mittelgroß.
- Frequenz und DERET: Im gemeinsamen Modell wird beim Diktat der Haupteffekt des Geschlechts signifikant, $F(1, 331.41) = 5.28, p = .02, \beta = -0.25$, aber nicht der der Frequenz oder der Interaktionsterm. Beim Abschreiben sind die Frequenz, $F(1, 329.77) = 6.41, p = .01, \beta = -0.11$, und das Geschlecht, $F(1, 331.70) = 3.92, p = .05, \beta = -0.21$, signifikante Prädiktoren, nicht jedoch die Interaktion. Für beide Modelle zeigt sich insgesamt eine kleine bis mittlere Effektstärke.
- Schreibzeit und DERET: Für das gemeinsame Modell von Geschlecht und Schreibzeit im Diktat werden nur die Haupteffekte als Prädiktoren für den DERET signifikant, Schreibzeit $F(1, 332.99) = 5.78, p = .02, \beta = 0.06$, und Geschlecht $F(1, 331.42) = 4.90, p = .03, \beta = -0.24$, nicht jedoch die Interaktion; im Modell für das Ab-

Tabelle 4. Zusammenhänge der Automatisierungs-Parameter mit der Anzahl der Fehler im DERET-Rechtschreibtest im gemischten Modell, getrennt nach Geschlecht

	df	F	p	β	Pseudo- R^2_m
Mädchen					
NIV (Diktat)	1, 167.90	4.51*	.04	0.19	.02
NIV (Abschreiben)	1, 170.35	8.33**	< .01	0.24	.04
Frequenz (Diktat)	1, 172.84	4.91*	.03	-0.15	.03
Frequenz (Abschreiben)	1, 172.94	6.93**	< .01	-0.17	.04
Schreibzeit (Diktat)	1, 170.96	9.23**	< .01	0.21	.05
Schreibzeit (Abschreiben)	1, 172.47	24.71**	< .01	0.33	.12
Jungen					
NIV (Diktat)	1, 155.32	0.09	.77	0.02	< .01
NIV (Abschreiben)	1, 156.02	0.89	.35	0.08	< .01
Frequenz (Diktat)	1, 126.26	0.56	.45	-0.07	< .01
Frequenz (Abschreiben)	1, 150.77	2.35	.13	-0.14	.01
Schreibzeit (Diktat)	1, 145.12	0.93	.34	0.08	< .01
Schreibzeit (Abschreiben)	1, 154.50	4.74*	.03	.18	.03

Anmerkungen: df wird als „Freiheitsgrade des Zählers, Freiheitsgrade des Nenners“ angegeben; β ist der standardisierte Regressionskoeffizient; * $p < .05$, ** $p < .01$.

schreiben wird nur die Schreibzeit signifikant, $F(1, 332.96) = 19.64, p < 0.01, \beta = 0.16$, nicht jedoch das Geschlecht oder die Interaktion. Beide Modelle zeigen moderate Effektstärken.

Eine getrennte Modellierung der Zusammenhänge von Handschrift-Parametern und Testergebnissen zeigt allerdings ein klares Muster:

- NIV und DERET: Bei einem nach Geschlechtern getrennten Modell zeigt sich bei den Mädchen sowohl für NIV im Schreibmodus Diktat als auch für NIV beim Abschreiben ein signifikanter, kleiner bis mittlerer, Zusammenhang mit dem DERET-Ergebnis. Bei den Jungen ist dagegen kein Zusammenhang zu sehen (siehe Tabelle 4).
- Frequenz und DERET: Bei der nach Geschlechtern getrennten Modellierung sagt die Frequenz sowohl im Diktat als auch beim Abschreiben die Rechtschreibleistung der Mädchen voraus; die Effektstärke ist klein bis moderat. Bei den Jungen dagegen hängt die Frequenz für keinen Schreibmodus mit der Rechtschreibleistung zusammen (siehe Tab. 4).
- Schreibzeit und DERET: Trennt man hier nach Geschlechtern, zeigt sich bei den Mädchen ein moderater signifikanter Zusammenhang der Schreibzeit im Schreibmodus Diktat mit dem DERET-Ergebnis. Für die Schreibzeit beim Abschreiben ist der Zusammenhang sogar groß. Bei den Jungen dagegen ist kein Zusammenhang der Schreibzeit im Diktat mit dem DERET-Ergebnis feststellbar, für die Schreibzeit beim Abschreiben ist der Zusammenhang immerhin klein bis moderat (siehe Tabelle 4).

Zusammenfassung zu Forschungsfrage 1: Handschreiben und Rechtschreiben

Bildet man größere Modelle, die das Geschlecht als zusätzlichen Prädiktor beinhalten, sind die Ergebnisse gemischt. Geschlecht und Automatisierungs-Parameter des Hand-

schreibens sind manchmal signifikante Prädiktoren für die DERET-Ergebnisse, manchmal nicht.

Klarer wird das Bild mit getrennten Modellen für beide Geschlechter. Hier ist durchgehend zu sehen, dass bei den Mädchen nennenswerte Zusammenhänge zwischen Automatisierungs-Parametern und DERET-Ergebnissen bestehen, bei den Jungen dagegen keine oder nur kleine Zusammenhänge. Zudem hängen die Automatisierungs-Parameter beim Abschreiben durchgehend stärker mit dem DERET-Ergebnis zusammen als im Schreibmodus Diktat.

Insgesamt lässt sich unter den drei hier berücksichtigten Parametern für die Automatisierung des Handschreibens (NIV, Frequenz, Schreibzeit) die Rechtschreibleistung am besten aus der Schreibzeit vorhersagen.

Forschungsfrage 2a: Rechtschreiben und Lesen

Gibt es nach Berücksichtigung der Zugehörigkeit der Schüler_innen zu ihren Klassen einen Zusammenhang zwischen den Leistungen im Rechtschreiben und Lesen?

Die Fehlerzahl beim DERET-Rechtschreibtest und die Punktzahl im ELFE-Lesetest korrelieren nicht nur deutlich miteinander ($r = -.57$), sie hängen auch nach einer Berücksichtigung der Clusterung der Kinder in Klassen noch stark zusammen: $F(1, 332.70) = 157.23, p < .01, \beta = -0.57, \text{Pseudo-}R^2_m = .32$.

Forschungsfrage 2b: Handschreiben und Lesen

Gibt es nach Berücksichtigung der Zugehörigkeit der Schüler_innen zu ihren Klassen einen Zusammenhang zwischen der Automatisierung des Handschreibens und der Leseleistung?

- NIV und ELFE: Der NIV im Schreibmodus Diktat hängt nicht signifikant mit der Leseleistung zusammen, beim

Tabelle 5. Zusammenhänge der Automatisierungs-Parameter mit der Punktzahl im ELFE-Lesetest im gemischten Modell

	df	F	p	β	Pseudo- R^2_m
NIV (Diktat)	1, 334.54	3.74	.05	-0.11	.01
NIV (Abschreiben)	1, 334.94	6.38*	.01	-0.14	.02
Frequenz (Diktat)	1, 324.66	1.68	.20	< 0.01	< .01
Frequenz (Abschreiben)	1, 329.66	5.36	.02	0.01	.02
Schreibzeit (Diktat)	1, 333.66	7.87**	< .01	-0.15	.02
Schreibzeit (Abschreiben)	1, 334.99	18.55**	< .01	-0.23	.05

Anmerkungen: df wird als „Freiheitsgrade des Zählers, Freiheitsgrade des Nenners“ angegeben; β ist der standardisierte Regressionskoeffizient; * $p < .05$, ** $p < .01$.

Abschreiben gibt es einen kleinen signifikanten Zusammenhang (siehe Tab. 5).

- Frequenz und ELFE: Erneut ist ein signifikanter Zusammenhang nur beim Abschreiben, aber nicht im Diktat zu beobachten (Tab. 5).
- Schreibzeit und ELFE: Die Schreibzeit im Schreibmodus Diktat hängt in geringem Maß mit der Leseleistung zusammen, im Abschreiben ist er erneut stärker ausgeprägt und moderat (Tab. 5).

Forschungsfrage 2c: Rechtschreiben und Lesen geschlechtsbezogen

Welche Rolle spielen dabei geschlechtsspezifische Unterschiede?

In einem Modell mit der ELFE-Punktzahl und dem Geschlecht als Prädiktoren für die Zahl der Fehler im DERET-Test ist nur die ELFE-Punktzahl ein signifikanter Prädiktor, $F(1, 332.56) = 149.19, p < .01, \beta = -0.60$; beim Geschlecht lässt sich kein signifikanter Zusammenhang beobachten, $F(1, 332.94) = 1.91, p = .17, \beta = -0.13$, ebenso wenig bei der Interaktion, $F(1, 332.99) = 1.14, p = .29, \beta = 0.10$. Das Modell hat eine große Effektstärke mit $\text{Pseudo-}R^2_m = .33$.

Auch bei getrennter Betrachtung der Geschlechter bestätigt sich, dass kein Effekt des Geschlechts feststellbar ist: Sowohl bei den Mädchen mit $F(1, 169.31) = 79.86, p < .01, \beta = -0.51$ $\text{Pseudo-}R^2_m = .32$ als auch bei den Jungen mit $F(1, 159.88) = 73.22, p < .01, \beta = -0.61$ $\text{Pseudo-}R^2_m = .32$ sagt

die ELFE-Punktzahl die Fehlerzahl im DERET vorher und der Zusammenhang ist bei beiden Geschlechtern gleichermaßen äußerst groß. Demnach sind keine geschlechtsbezogenen Unterschiede feststellbar.

Forschungsfrage 2d: Handschreiben und Lesen geschlechtsbezogen

Welche Rolle spielen dabei geschlechtsspezifische Unterschiede?

Anschließend gingen wir auch hier auf die Geschlechterunterschiede in Bezug auf die Automatisierung beim Handschreiben und Ergebnisse im ELFE-Lesetest ein und untersuchten mögliche Zusammenhänge zunächst in einem gemischten Modell (alle Werte in Tabelle E2 im ESM4):

- NIV und ELFE: Beim gemeinsamen Modell von NIV und Geschlecht als Prädiktoren für die Leseleistung im ELFE-Test gab es im Diktat keine signifikanten Haupt- oder Interaktionseffekte. Das Modell hat eine kleine bis mittlere Effektstärke. Für das Abschreiben sind dagegen der Haupteffekt für NIV, $F(1, 334.48) = 6.25, p = .01, \beta = -0.03$, und der Interaktionsterm signifikant, $F(1, 327.95) = 4.15, p = .04, \beta = -0.24$, nicht jedoch der Haupteffekt des Geschlechts. Die Effektstärke ist moderat.
- Frequenz und ELFE: Für einen Zusammenhang zwischen der Frequenz und dem Geschlecht mit der ELFE-Punktzahl gibt es im gemeinsamen Modell einen sig-

Tabelle 6. Zusammenhänge der Automatisierungs-Parameter mit der Punktzahl im ELFE-Lesetest im gemischten Modell, getrennt nach Geschlecht

	df	F	p	β	$\text{Pseudo-}R^2_m$
Mädchen					
NIV (Diktat)	1, 165.05	5.23*	.02	-0.23	.03
NIV (Abschreiben)	1, 167.09	10.85**	< .01	-0.29	.06
Frequenz (Diktat)	1, 172.97	4.19*	.04	0.15	.02
Frequenz (Abschreiben)	1, 172.89	6.13*	.01	0.18	.03
Schreibzeit (Diktat)	1, 169.02	8.70**	< .01	-0.22	.05
Schreibzeit (Abschreiben)	1, 171.50	15.13**	< .01	-0.29	.08
Jungen					
NIV (Diktat)	1, 155.65	0.26	.61	-0.04	< .01
NIV (Abschreiben)	1, 159.22	0.33	.57	-0.04	< .01
Frequenz (Diktat)	1, 139.83	0.08	.77	-0.02	< .01
Frequenz (Abschreiben)	1, 159.22	0.36	.55	0.05	< .01
Schreibzeit (Diktat)	1, 147.59	1.15	.29	-0.08	< .01
Schreibzeit (Abschreiben)	1, 159.57	4.77*	.03	-0.17	

Anmerkungen: df wird als „Freiheitsgrade des Zählers, Freiheitsgrade des Nenners“ angegeben; β ist der standardisierte Regressionskoeffizient; * $p < .05$, ** $p < .01$.

nifkanten Zusammenhang nur mit dem Geschlecht, sowohl im Diktat, $F(1, 332.94) = 5.45, p = .02, \beta = 0.25$, als auch beim Abschreiben, $F(1, 333.48) = 4.25, p = .04, \beta = 0.22$. In beiden Fällen hat das Modell eine kleine bis moderate Effektstärke.

- Schreibzeit und ELFE: Zuletzt wurden auch für den Zusammenhang von Schreibzeit und Geschlecht mit dem ELFE-Ergebnis gemeinsame Modelle eingesetzt, eines für die Schreibzeit im Schreibmodus Diktat und eines für die Schreibzeit beim Abschreiben. Die Schreibzeit im Diktat ist ein signifikanter Prädiktor, $F(1, 334.68) = 7.19, p < .01, \beta = -0.07$, wie auch das Geschlecht, $F(1, 330.30) = 4.63, p = .03, \beta = 0.23$, die Interaktion von Schreibzeit und Geschlecht jedoch nicht. Die Effektstärke des Modells ist moderat. Für das Abschreiben ist nur die Schreibzeit signifikant, $F(1, 334.94) = 16.34, p < .01, \beta = -0.16$, bei gleichermaßen moderater Effektstärke.

Auch in diesem Fall ergeben sich interessante Erkenntnisse aus dem Berechnen getrennter Modelle für Jungen und Mädchen:

- NIV und ELFE: Bei getrennten Modellen zeigt sich für die Mädchen im Diktat ein kleiner bis moderater Zusammenhang, beim Abschreiben ist der Zusammenhang moderat. Für die Jungen lässt sich dagegen erneut kein Zusammenhang feststellen (siehe Tab. 6).
- Frequenz und ELFE: Bei getrennter Betrachtung der Geschlechter zeigt sich ein Zusammenhang von Frequenz und Leseleistung nur bei Mädchen. Im Diktat ist der Zusammenhang klein, beim Abschreiben klein bis moderat. Bei den Jungen zeigten sich keine Zusammenhänge (Tab. 6).
- Schreibzeit und ELFE: Beim Berechnen getrennter Modelle für beide Geschlechter ist die Schreibzeit bei den Mädchen ein signifikanter Prädiktor. Im Schreibmodus Diktat ein mittlerer, beim Abschreiben ein großer Effekt. Für die Jungen ergibt sich im Diktat kein Zusammenhang, beim Abschreiben ist ein kleiner bis moderater Zusammenhang feststellbar (Tab. 6).

Zusammenfassung zu Forschungsfrage 2: Handschreiben / Rechtschreiben und Lesen

Für den Lesetest zeigen sich vergleichbare Ergebnisse wie für den Rechtschreibtest: Die Assoziation zwischen Lesetest-Ergebnis und Automatisierungs-Parametern ist für die Mädchen größer als für die Jungen, bei denen sie klein oder gar nicht vorhanden ist. Erneut ist die Schreibzeit der stärkste Prädiktor. In diesem Fall ist zudem der Zusammenhang bei NIV stärker ausgeprägt als bei der Frequenz. Zuletzt zeigt sich erneut, dass die Automatisierungs-Indikatoren, die beim Abschreiben gemessen wurden, enger

mit den Lesetest-Ergebnissen zusammenhängen als die beim Diktat.

Diskussion

Insgesamt bestätigen diese Ergebnisse, dass es sich lohnt, mehrere Einflussfaktoren in den Blick zu nehmen. Geschlechterübergreifend besteht ein Zusammenhang zwischen den Lese- und Rechtschreibleistungen, wobei die Rolle der Handschreibflüssigkeit eine spezifische zu sein scheint. Während die Automatisierungs-Indikatoren bei Mädchen, die dieselbe Schulklasse besuchen, mit den Ergebnissen des Lese- und des Rechtschreibtests zusammenhängen, ergibt sich bei den Jungen nur ein geringer oder gar kein Zusammenhang. Da auch die Werte der Automatisierungs-Indikatoren in dieselbe Richtung weisen (also bessere Werte der Mädchen), kann angenommen werden, dass ein höherer Grad an Automatisierung im Handschreiben mit besseren Leistungen im Rechtschreiben und im Lesen zusammenhängt. Warum aber Mädchen zu besser automatisierten Kompetenzen gelangen und sich damit Vorteile auch für die höheren Prozesse des Texteverfassens sichern, bleibt an dieser Stelle offen.

Auch das Alter der Kinder ist hier in Rechnung zu stellen. Am Ende der vierten Jahrgangsstufe befinden sie sich noch in einer frühen Phase der Konsolidierung schriftsprachlicher Kompetenzen.

Außerdem lassen NIV, Frequenz und Schreibzeit beim Abschreiben bessere Rückschlüsse auf Lese- und Rechtschreibleistung zu als im Schreibmodus „Diktat“. Das mag damit erklärt werden, dass beim Diktat einige weitere kognitive Prozesse erforderlich sind. Beim Abschreiben sind zwar Kontrollvorgänge mit der Vorlage nötig, die Zusammenhänge mit der Leseleistung wahrscheinlicher machen, aber das Nachdenken über Rechtschreibregeln entfällt. Das Arbeitsgedächtnis ist durch die Vorgabe des zu schreibenden Satzes entlastet.

Zuletzt ist der wohl beste Indikator, wenn man Lese- oder Rechtschreibleistung aus den Automatisierungs-Indikatoren vorhersagen will, die Schreibzeit. Die prozessorientierten Parameter NIV und Frequenz zeigen einen geringeren Zusammenhang mit der Lese- und Rechtschreibleistung. NIV und Frequenz analysieren gewissermaßen die Bewegung an sich, während die Schreibzeit das Ergebnis der (gelungenen) Bewegungsausführung markiert. Insofern bestätigt dieses Ergebnis ebenfalls die Bedeutung der Automatisierung des Handschreibens aufgrund der damit einhergehenden Entlastung des Arbeitsgedächtnisses. Interessanterweise bezieht sich dies nicht nur auf die Rechtschreib-, sondern – zunächst deutlich weniger klar interpretierbar – auch auf die Leseleistung. Es ist also nach Inter-

pretation unserer Ergebnisse davon auszugehen, dass sich die hierarchieniedrigen Prozesse beider Modi, des Lesens und des Schreibens, interdependent zueinander verhalten, was der Annahme des *Interactive Dynamic Literacy Model* (Kim, 2020) entspricht.

Limitationen

Limitierend für unsere Studie ist in erster Linie die Beschränkung auf die Untersuchung hierarchieniedriger schriftsprachlicher Prozesse. Wir können weder etwas über hierarchiehöhere Prozesse aussagen, da wir die Textqualität nicht erhoben haben, noch über das Textverständnis beim Lesen ganzer Texte. Auch die vielen weiteren, das Schreiben insgesamt beeinflussenden Faktoren (z. B. auf Ebene der exekutiven Funktionen) haben wir nicht erhoben, doch könnten sie sicher weitere Aufschlüsse geben, wie die eingangs zitierten Schreibmodelle von Kim und Graham (2022) und Linnemann et al. (2022) zeigen.

Auf der statistischen Seite ist zu beachten, dass die vorliegenden Ergebnisse mit einem explorativen Ansatz gefunden wurden, wobei verschiedene Modelle verwendet wurden, um Zusammenhänge aufzuklären. Hier war das Ziel nicht die Suche nach einzelnen signifikanten Zusammenhängen, sondern die Darstellung von Mustern. Die Ergebnisse sollten dennoch in zukünftigen Untersuchungen konfirmatorisch repliziert werden. Hinzu kommt, dass in diesen Analysen das Konstrukt der Automatisierung nicht als übergeordnete Variable abgebildet werden konnte, sondern durch drei getrennte Indikatoren (NIV, Frequenz und Schreibzeit) approximiert wurde.

Zuletzt bleibt hinzuzufügen, dass als mögliche Einflussfaktoren in dieser Studie die Automatisierungs-Kennwerte der Schriften der Kinder, ihr Geschlecht und ihre Klassenzugehörigkeit berücksichtigt wurden, nicht jedoch andere denkbare Einflussvariablen wie *beliefs* der Lehrkräfte, sozioökonomischer Status, Mehrsprachigkeit oder andere Persönlichkeitsmerkmale (z. B. Skar et al., 2022; Cordeiro et al., 2018).

Relevanz für die Praxis

Für die Praxis bedeuten unsere Ergebnisse einerseits, dass auch ohne umfangreiche technische Ausstattung über die Schreibzeit recht gut auf die Flüssigkeit von Handschriften geschlossen werden kann. Andererseits scheint es geboten, „langames“ Schreiben in der Schule nicht einfach als gegeben hinzunehmen, sondern gezielt die Schreibmotorik zu fördern. Denn wie jede motorische Herausforderung können auch Schreibbewegungen nur durch häufige Wiederholung gefördert und in immer besser automatisierte Be-

wegungsroutinen überführt werden. Die Beschleunigung des Schreibtempos wiederum korreliert, wie hier gezeigt werden konnte, mit der Lese- und Rechtschreibleistung, ein Zusammenhang, der natürlich für die Leistungen in allen Lernbereichen, in denen geschrieben wird, relevant ist.

Elektronische Supplemente (ESM)

Die elektronischen Supplemente sind mit der Online-Version dieses Artikels verfügbar unter <https://doi.org/10.1024/2235-0977/a000437>.

ESM1. Abbildung E1. NIV.

ESM2. Abbildung E2. Frequenz.

ESM3. Abbildung E3. Schreibzeit.

ESM4. Tabellen E1 und E2.

Literatur

- Bates, D., Mächler, M., Bolker, B & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1).
- Bennett, D. A. (2001). How can I deal with missing data in my study? *Australian and New Zealand Journal of Public Health*, 25(5), 464 – 469.
- Berninger, V. W. & Winn, W. D. (2006) Implications of Advancements in Brain Research and Technology for Writing Development, Writing Instruction, and Educational Evolution. In C. A. MacArthur, S. Graham, J Fitzgerald (Eds.), *Handbook of writing research* (S. 96 – 114) New York: Guilford Press.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: L. Erlbaum Associates.
- Cohen, J. & Cohen, P. (2003). *Applied multiple regression / correlation analysis for the behavioral sciences* (2nd ed.). New York, London: Psychology Press.
- Cordeiro, C., Castro, S. L. & Limpo, T. (2018). Examining Potential Sources of Gender Differences in Writing: The Role of Handwriting Fluency and Self-Efficacy Beliefs. *Written Communication*, 35(4), 349 – 410.
- Gignac, G. E. & Szodorai, E. T. (2016). Effect size guidelines for individual differences researchers. *Personality and Individual Differences*, 102, 74 – 78.
- Hayes, J. R. & Flower, L. S. (1980). Identifying the organization of writing processes. In L. Gregg & E. R. Steinberg (Eds.), *Cognitive processes in writing* (S. 3 – 30). Hillsdale, NJ: Lawrence Erlbaum.
- Howell, D. C. (2007). The treatment of missing data. In W. Outhwaite & S. Turner (Eds.), *The SAGE handbook of social science methodology* (S. 212 – 226). London: SAGE Publications.
- Hurschler Lichtsteiner, S., Saxer Geiger, A. & Wicki, W. (2010). *Schreibmotorische Leistungen im Primarschulalter in Abhängigkeit vom unterrichteten Schrifttyp*. Luzern: Pädagogische Hochschule Zentralschweiz.
- Kellogg, R. T. (2022). Executive functions and writing [Book Review of: Executive functions and writing by T. Limpo and T. Olive, Eds.]. *Journal of Writing Research*, 13(3), 473 – 479.
- Kent, S. C. & Wanzek, J. (2016). The Relationship Between Component Skills and Writing Quality and Production Across Develop-

- mental Levels: A Meta-Analysis of the Last 25 Years. *Review of Educational Research* 86(2), 570 – 601.
- Kim, Y.-S. G. & Schatschneider, C. (2017). Expanding the developmental models of writing: A direct and indirect effects model of developmental writing (DIEW). *Journal of Educational Psychology*, 109(1), 35 – 50.
- Kim, Y.-S.G., Petcher, Y., Wanzek, J. & Al Otaiba, S. (2018). Relations between reading and writing: a longitudinal examination from grades 3 to 6. *Reading and Writing*, 31, 1591 – 1618.
- Kim, Y.-S. G. (2020). Interactive dynamic literacy model: An integrative theoretical framework for reading and writing relations. In R. Alves, T. Limpo, & M. Joshi (Eds.), *Reading-writing connections: Towards integrative literacy science* (S. 11 – 34). Netherlands: Springer.
- Kim, Y.-S. G. & Graham, S. (2022). Expanding the Direct and Indirect Effects Model of Writing (DIEW): Reading-Writing Relations, and Dynamic Relations as a Function of Measurement / Dimensions of Written Composition. *Journal of Educational Psychology*, 114(82), 215 – 238.
- Lee, V.E. (2000). Using hierarchical linear modeling to study social contexts: The case of school effects. *Educational Psychologist*, 35(2), 125 – 141.
- Lenhard, W., Hasselhorn, M. & Schneider, W. (2011). KLASSE 4. Kombiniertes Leistungsinventar zur allgemeinen Schulleistung und für Schullaufbahneempfehlungen in der vierten Klasse. Manual. Göttingen: Hogrefe.
- Linnemann, M., Stephany, S., Lemke, V., Bulut, N., Haider, H., Roth, H. J., & Becker-Mrotzek, M. (2022). The dimensionality of writing and reading fluency and its impact on comprehension and composition. *Journal of Writing Research*, 14(2), 185 – 227.
- Marquardt, C. (2011). Schreibanalyse. In: Dennis A. N. (Hrsg.): *Handfunktionsstörungen in der Neurologie. Klinik und Rehabilitation* (S. 380 – 394). Berlin, Heidelberg, New York, NY: Springer.
- Marquardt, C. (2014). *CSWin Bedienungshandbuch Version 2012. Computerunterstützte Analyse des Bewegungsablaufs beim Schreiben*. München: MedCom.
- Müller, S., Scealy, J. L. & Welsh, A. H. (2013). Model Selection in Linear Mixed Models. *Statistical Science*, 28(2).
- Nakagawa, S., Johnson, P. C. D. & Schielzeth, H. (2017). The coefficient of determination R^2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *Journal of The Royal Society Interface*, 14(134), 20170213.
- Odersky, E. (2018). *Handschrift und Automatisierung des Handschreibens. Eine Evaluation von Kinderschriften im 4. Schuljahr*. Berlin: J.B.Metzler.
- Odersky, E., Speck-Hamdan, A., Mahrhofer-Bernt, C. & Marquardt, C. (2021) Handschriften und Handschreiben in der Schule – ein Forschungsbericht. *Didaktik Deutsch, Heft 51*, 78 – 92
- Olive, T. (2014). Toward a parallel and cascading model of the writing system. A review of research on writing processes coordination. *Journal of Writing Research* 6 (2), 173 – 194.
- RCore Team. (2013). *R: A language and environment for statistical computing*. [Software] R Foundation for Statistical Computing.
- RStudio Team. (2023). *RStudio: Integrated development for R* (2023.03.0) [[Software]]. RStudio.
- Skar, G. B., Lei, P.-W., Graham, S., Aasen, A. J., Johansen, M. B. & Kvistad, A. H. (2022). Handwriting fluency and the quality of primary grade students' writing. *Reading and Writing* 35, 509 – 538.
- Stanat, P., Schipolowski, S., Schneider, R., Sachse, K.A., Weirich, S. & Henschel, S. (Hrsg.) (2022). *IQB-Bildungstrend 2021. Kompetenzen in den Fächern Deutsch und Mathematik am Ende der 4. Jahrgangsstufe im dritten Ländervergleich*. Münster, New York: Waxmann
- Sturm, A., Nänny, R. & Wyss, S. (2017). Entwicklung hierarchieniedriger Schreibprozesse. In M. Philipp (Hrsg.): *Handbuch Schriftspracherwerb und weiterführendes Lesen und Schreiben*. (S. 84 – 104) Weinheim: Beltz Juventa.
- Torrance, M., Galbraith, D. (2006). The processing demands of writing. In C. A. MacArthur, S. Graham. J. Fitzgerald (ed.): *Handbook of writing research* (S. 67 – 82). New York: Guilford Press.
- Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis* (2. Aufl.). Dordrecht, Heidelberg, London, New York: Springer International Publishing.
- Yamashita, T., Yamashita, K. & Kamimura, R. (2007). A stepwise AIC method for variable selection in linear regression. *Communications in Statistics – Theory and Methods*, 36(13), 2395 – 2403.



Historie

Manuskript eingereicht: 02.08.2023
 Manuskript angenommen: 06.02.2024
 Onlineveröffentlichung: 12.03.2024

Förderung

Open-Access-Veröffentlichung ermöglicht durch die Universitätsbibliothek der Katholischen Universität Eichstätt-Ingolstadt.

ORCID

Eva Odersky
 <https://orcid.org/0009-0008-7827-0525>
 Maximilian Stark
 <https://orcid.org/0000-0003-3055-842X>



Dr. Eva Odersky

Lehrstuhl für Grundschulpädagogik und Grundschuldidaktik
 Katholische Universität Eichstätt-Ingolstadt
 Ostenstraße 26
 85072 Eichstätt
 Deutschland
eva.odersky@ku.de