Fakultät Kulturwissenschaft
Institut für Diversitätsstudien
Technische Universität Dortmund
Erstgutachterin: Prof. Dr. Barbara Mertins
Zweitgutachterin: Katrin Odermann

Masterarbeit

Zur Erlangung des akademischen Grades Master of Education
Im Studiengang Lehramt an Grundschulen

Lesekompetenz ist mehrdimensional

Eine empirische Eye-Tracking-gestützte Analyse von Leseflüssigkeit und Textverstehen von Zweitklässler*innen unter Berücksichtigung zentraler Einflussfaktoren

Vorgelegt von
Maja Sternemann
[redacted]

[11.09.2025]

Das Projekt „BeeMEHR" wird im Rahmen des Programms „Integration durch Bildung" durch das Bundesministerium für Bildung und Forschung und die Europäische Union über den Europäischen Sozialfonds Plus (ESF Plus) gefördert.

Danksagung

Mein berufliches Ziel war von Anfang an klar: Ich möchte Grundschullehrerin werden. Die vorliegende Masterarbeit stellte dabei die letzte Etappe auf diesem Weg dar. Es war eine Etappe, die nicht immer einfach und mit zahlreichen Herausforderungen verbunden war. Umso mehr gilt mein aufrichtiger Dank allen, die mich während der Anfertigung dieser Arbeit sowie im Verlauf meines gesamten Studiums unterstützt, begleitet und ermutigt haben.

Zunächst möchte ich meiner Erstprüferin, Frau Prof. Dr. Barbara Mertins, meinen herzlichen Dank aussprechen. Ich schätze nicht nur ihre fachliche Begleitung während der Entstehung dieser Arbeit sehr, sondern auch die Möglichkeit im Rahmen dieser Forschung mit dem Eye-Tracker arbeiten zu dürfen. Besonders dankbar bin ich ihr für ihre wohlwollende Unterstützung, das entgegengebrachte Vertrauen sowie ihr überzeugtes Engagement in diesem Forschungsfeld – eine Haltung, die sich spürbar auf das gesamte Team der Psycholinguistik übertragen hat. Ihre konstruktiven Rückmeldungen und ihre ermutigende, unterstützende Art haben mir geholfen, die Arbeit mit Klarheit und Selbstvertrauen weiterzuentwickeln.

Ein ebenso herzlicher Dank gilt meiner Zweitprüferin Katrin Odermann. Das Seminar bei ihr hat in mir das Interesse für das Thema geweckt und maßgeblich dazu beigetragen, dass ich meine Masterarbeit zu diesem Thema schreiben wollte. Seit der Entscheidung, die Masterarbeit im Team der Psycholinguistik zu verfassen, hat sie mich durch alle Phasen begleitet. Katrin hatte stets ein offenes Ohr, unter der Woche aber auch am Wochenende oder in ihrem Urlaub, und ich konnte mich jederzeit mit Fragen an sie wenden. Diese Unterstützung und ihre beruhigenden Worte waren für mich von unschätzbarem Wert.

Mein aufrichtiger Dank gilt auch dem gesamten Team der Psycholinguistik. Ich wurde nicht nur herzlich aufgenommen, sondern auch mit zahlreichen wertvollen Anregungen und Ideen zur Entwicklung meiner Studie und der Ausarbeitung der Arbeit unterstützt. Besonders danken möchte ich Vincent Schröder, der mich bei der statistischen Auswertung unterstützt hat, sowie Renate Delucchi Danhier. Renate stand mir nicht nur bei der Erstellung des Eye-Tracking-Materials und der Auswertung zur Seite, sondern hat mir mit ihrer stets positiven Art und guten Laune auch regelmäßig ein Lächeln ins Gesicht gezaubert.

Ein großes Dankeschön gilt auch der Schulleitung sowie den Klassenlehrerinnen, den beteiligten Eltern und vor allem den Kindern, ohne die eine solche Studie nicht möglich gewesen wäre. Die Schule stellte mir nicht nur einen eigenen Raum zur Verfügung, sondern ermöglichte mir auch flexibel zu unterschiedlichen Zeiten, während des Unterrichts, zu testen – dafür bin ich sehr dankbar.

Mein tief empfundener Dank gilt vor allem meiner Familie und meinen Freunden. Mein Bruder Lasse und meine Eltern standen mir während des gesamten Prozesses mit bedingungsloser Unterstützung zur Seite – sei es mit aufmunternden Worten, beständiger Motivation oder in jenen Momenten, in denen ich besonders viel Halt brauchte. Ein besonderer Dank gilt auch meinen Freundinnen und Freunden – allen voran Melina. Mit ihnen konnte ich über jede noch so kleine Idee sprechen und mit ihren humorvollen Kommentaren haben meine Freundinnen und Freunde stets meine Stimmung aufgehellt.

Recklinghausen, den 11.09.2025

Inhaltsverzeichnis

I. ABBILDUNGSVERZEICHNIS I

II. TABELLENVERZEICHNIS I

1 EINLEITUNG 6

2 THEORETISCHER RAHMEN 9

2.1 LESEN 9

2.1.1 Prozesse 10

2.1.2 Leseflüssigkeit und Textverstehen 12

2.1.3 Determinanten 17

2.2 ZENTRALE EINFLUSSFAKTOREN AUF DIE LESEKOMPETENZ 21

2.2.1 Wortschatz 21

2.2.2 Sprachlicher Input 23

2.2.3 Sozioökonomischer Status 25

2.3 DIAGNOSTISCHE VERFAHREN ZUR ERFASSUNG DER LESEKOMPETENZ 25

2.4 EYE-TRACKING 28

2.4.1 Grundlegende Funktionsweise des Auges 29

2.4.2 Methodik des Eye-Trackings 31

2.4.3 Analyseverfahren der Augenbewegung 32

2.4.4 Forschungsstand in Bezug auf die Lesekompetenz 34

3 EMPIRISCHER TEIL 41

3.1 FRAGESTELLUNG UND HYPOTHESEN 41

3.2 OPERATIONALISIERUNG 44

3.3 STICHPROBE 46

3.4 METHODEN 47

3.4.1 Fragebogen 48

3.4.2 Peabody Picture Vocabulary Test 50

3.4.3 Eye-Tracking 51

3.5 UNTERSUCHUNGSMATERIAL 52

3.6 DURCHFÜHRUNG 57

3.7 DATENAUSWERTUNG 59

3.8 DARSTELLUNG DER ERGEBNISSE 60

3.8.1 Deskriptive Auswertung der Ergebnisse 61

3.8.2 Auswertung der gemischten Modelle 64

3.9 DISKUSSION 67

3.9.1 Interpretation der deskriptiven Ergebnisse 67

3.9.2 Interpretation der GLMM-Ergebnisse – Textverstehen 69

3.9.3 Interpretation der LMM-Ergebnisse – Leseflüssigkeit 72

3.9.4 Gesamtdiskussion *74*
3.9.5 Limitationen und Ausblick auf zukünftige Studien *77*

4 FAZIT UND AUSBLICK 80

5 LITERATURVERZEICHNIS I

6 ANHANG A VII

6.1 TEXTE ZUM BAUERNHOF VII
6.1.1 Schwierigkeitsgrad 1 ..... *VII*
6.1.2 Schwierigkeitsgrad 2 ..... *VII*
6.2 TEXTE ZUM SOMMERTAG VIII
6.2.1 Schwierigkeitsgrad 1 ..... *VIII*
6.2.2 Schwierigkeitsgrad 2 ..... *VIII*

7 EIDESSTATTLICHE VERSICHERUNG IX

I. Abbildungsverzeichnis

Abbildung 1: Mehrebenenmodell des Lesens (Rosebrock & Nix, 2025, S. 15) 13
Abbildung 2: Theoretische Struktur der Lesekompetenz in PISA (Hochstadt et al., 2022, S. 144) 17
Abbildung 3: Modell des Zusammenspiels familiärer und psychologischer Einflüsse auf das Leseverstehen (Schaffner, 2009, S. 36) 18
Abbildung 4: verwendeter Remote Eye-Tracker 32
Abbildung 5: Heat Map & Gaze Plots (Deilen, 2022, S. 277) 33
Abbildung 6: Lesetext - Bauernhof geringerer Schwierigkeitsgrad 54
Abbildung 7: Lesetext - Bauernhof höherer Schwierigkeitsgrad 54
Abbildung 8: Aufnahme der Blickbewegung mit Gaze Plots 59
Abbildung 9: Verteilung des T-Wertes; des sprachlichen Inputs; des sozioökonomischen Status ; des sozioökonomischen Status zwischen bilingualen und monolingualen Kindern 61
Abbildung 10: Einteilung der Kinder in die Normbereiche des PPVTs auf Basis der T-Werte 62
Abbildung 11: Verteilung der Gesamtlesezeit & Der Lesezeit des ersten Durchganges 63

II. Tabellenverzeichnis

Tabelle 1: Operationalisierung der Zielvariablen und Einflussfaktoren 46
Tabelle 2: Zusammensetzung der Stichprobe 47
Tabelle 3: Zusammenfassung des Online-Elternfragebogens 48
Tabelle 4: Skala des sozioökonomischen Status nach Lampert et al. (2018) 49
Tabelle 5: Wort-, Zeichen- und Frequenzwerte der Texte 53
Tabelle 6: Fragen mit den korrekten Antworten für jeden Text 56
Tabelle 7: Übersicht der deskriptiven Werte; SD = standard derivation 64
Tabelle 8: Zuordnung der p-Werte zu den Signifikanzniveaus 65
Tabelle 9: Generalized linear mixed model: Textverstehen – fixed effects; SD = standard derivation 66
Tabelle 10: Linear mixed model - Leseflüssigkeit; df = degress of freedom 67

1 Einleitung

„Reading comprehension is a critical skill that forms the foundation for academic success and lifelong learning“ (Mihret & Joshi, 2025, S. 1).

Nicht nur für den Berufserfolg, sondern auch für das lebenslange Lernen und die Bewältigung alltäglicher Aufgaben stellt die Lesefähigkeit eine essentielle Schlüsselkompetenz dar. Sie bildet somit die Grundlage für schulischen sowie akademischen Erfolg, für die Arbeit in vielen Berufszweigen, das alltägliche Leben und die gesellschaftliche Teilhabe (Lenhard, 2024). Dabei bezeichnet der Begriff der Lesekompetenz die Fähigkeit, geschriebene Grapheme, Wörter, Sätze und Texte nicht nur technisch zu entziffern, sondern auch ihre Bedeutung zu erfassen und zu interpretieren. Die scheinbar einfache Fähigkeit des Lesens setzt sich dabei aus komplexen, zeitgleich ablaufenden Teilprozessen zusammen (Lenhard, 2024). Hierzu zählen die grundlegenden Basiskompetenzen des sicheren Erfassens sprachlicher Einheiten sowie der Herstellung lokaler Kohärenz (Jesch, 2020; Lenhard, 2024). Eine automatisierte Worterkennung, die präzise Dekodierung, eine angemessene Lesegeschwindigkeit und die prosodische Sicherheit, zusammengefasst unter dem Begriff Leseflüssigkeit, entlasten das Arbeitsgedächtnis und schaffen Raum für weiterführende Verstehensleistungen (Jesch, 2020; Lenhard, 2024; Rosebrock, 2022). Aufbauend auf diesen Prozessen setzen Lesestrategien an, die auf die globale Kohärenz abzielen: Lesende verknüpfen Textinhalte mit ihrem Vorwissen, extrahieren Bedeutungen und integrieren sie zu umfassenden mentalen Modellen (Lenhard, 2024).

Zahlreiche empirische Untersuchungen aus der Leseforschung belegen, dass individuelle, soziale und kognitive Variablen die Entwicklung der Lesekompetenz sowohl fördern als auch hemmen können (Schaffner, 2009). Aus diesen Erkenntnissen wurden unterschiedliche theoretische Modelle zu den Einflussfaktoren auf die Lesekompetenz abgeleitet (Cromley & Azevedo, 2007; Jenkins, 1979; Schaffner, 2009). Das in dieser Arbeit herangezogene Modell nach Schaffner (2009) fasst die zuvor genannten Einflussfaktoren zusammen und unterscheidet dabei zwischen personengebundenen Merkmalen wie beispielsweise kognitive Fähigkeiten, motivationalen Dispositionen und familiären Determinanten, insbesondere den sozioökonomischen Status (Schaffner, 2009).

Die Lesekompetenz von Schülerinnen und Schülern wird in vielen Studien untersucht, jedoch bleiben Zweitklässlerinnen und Zweitklässler oft unberücksichtigt. Nationale Vergleichsarbeiten (VERA) und die internationale Grundschul-Lese-Untersuchung (IGLU) erfassen die Lesekompetenz in der Regel erst ab der dritten Klasse (McElvany et al., 2023; Tarkian et al., 2019). Zur Analyse der beim Lesen ablaufenden Prozesse hat sich in den letzten

Jahren die Eye-Tracking-Methodik etabliert. Durch eine hochauflösende, simultane Erfassung visueller Aufmerksamkeits- und kognitiver Verarbeitungsprozesse während des Lesens ermöglicht diese Methode, Aussagen über Leseprozesse zu treffen (Holmqvist & Andersson, 2017). Hierfür werden Augenbewegungsparameter wie Fixationen, Fixationsdauern, Sakkadenlängen und Regressionen als Indikatoren herangezogen, um die Leseflüssigkeit zu quantifizieren (Krause & Ritter, 2022; Rayner, 1998). Allerdings setzen bisherige Eye-Tracking-Studien den Schwerpunkt auf eben diese Kennwerte der Leseflüssigkeit und vernachlässigen dabei die Dimension des Textverstehens weitgehend (u.a. Häikiö et al., 2009; Hautala et al., 2024; Krause & Ritter, 2022; Rayner, 1998; Wertli et al., 2023). Vor diesem Hintergrund adressiert die vorliegende Untersuchung ein bislang unzureichend erforschtes Desiderat, indem sie Leseflüssigkeit und Textverstehen gleichermaßen integriert und zentrale Einflussfaktoren am Ende der zweiten Jahrgangsstufe systematisch betrachtet.

Zur Beantwortung der Forschungsfragen wurden verschiedene Methoden gewählt. Die Erhebung individueller und familiärer Einflussfaktoren erfolgte über einen digitalisierten Elternfragebogen. Zur Bestimmung des rezeptiven Wortschatzes wurde der Peabody Picture Vocabulary Test (PPVT) nach Lenhard und Lenhard (2022) durchgeführt. Die Analyse kindlicher Leseprozesse im Sinne der Leseflüssigkeit erfolgt mit Hilfe der Eye-Tracking-Methodik, indem Kennwerte wie beispielsweise Fixationsanzahlen und -dauern erfasst wurden. Die Textverstehensleistung der Kinder wurde anhand von Verständnisfragen zu den gelesenen Texten erhoben.

Der Aufbau der vorliegenden Masterarbeit verbindet einen umfassenden theoretischen Rahmen mit einem darauf aufbauenden empirischen Teil und schafft so eine Verbindung zwischen Theorie und Praxis. Im ersten Abschnitt wird das Lesen als kognitiver Prozess systematisch erläutert. Dabei stehen Leseflüssigkeit und Textverstehen im Zentrum, bevor wichtige Determinanten der Lesekompetenz zusammengeführt werden. Anschließend werden zentrale Einflussfaktoren wie der Umfang des Wortschatzes, der sprachliche Input und der sozioökonomische Status differenziert betrachtet, um ihre Wirkmechanismen und Wechselwirkungen zu erläutern. Darauf aufbauend folgt ein Vergleich etablierter diagnostischer Verfahren, der die Auswahl der im empirischen Teil eingesetzten Testinstrumente begründet. Den Abschluss des Theorie-Kapitels bildet eine Einführung in die Eye-Tracking-Methodik. Nach dem Einblick in das visuelle System werden technische Grundlagen und Analyseverfahren vorgestellt und der aktuelle Forschungsstand zur Blickbewegungsmessung beim Lesen skizziert. Der empirische Teil knüpft direkt an diese theoretischen Vorüberlegungen an. Zunächst werden Forschungsfragen und Hypothesen formuliert, die aus den Zusammenhängen zwischen Leseflüssigkeit, Textverstehen und ihren Prädiktoren abgeleitet sind. Die theoretischen Konstrukte werden in konkrete Messgrößen

überführt und die Stichprobe dieser Studie beschrieben. Anschließend werden die eingesetzten Instrumente, ein Fragebogen, der PPVT zur Wortschatzmessung und das Eye-Tracking zur Erfassung der Augenbewegung, erläutert und ihre Anwendung im Datenerhebungsablauf dargestellt. Die Auswertung erfolgt mithilfe deskriptiver Statistiken und gemischter Modelle, wobei Textverstehen und Leseflüssigkeit getrennt analysiert werden. In der abschließenden Diskussion werden die Ergebnisse im Hinblick auf die Hypothesen reflektiert und in einen integrativen Kontext gestellt. Methodische wie inhaltliche Limitationen sowie praktische Implikationen werden aufgezeigt. Abgeschlossen wird die Arbeit durch ein zusammenfassendes Fazit, das die Theorie und empirische Befunde integriert sowie einen Ausblick auf weiterführende Untersuchungen und praktische Implikationen gibt.[1]

[1] Ausgewählte Materialien sind in einem externen Anhang dokumentiert und werden im Text durch den Verweis auf *Anhang B* erschlossen.

2 Theoretischer Rahmen

In diesem Kapitel werden die theoretischen Grundlagen der vorliegenden Studie systematisch erläutert. Zunächst wird der Leseprozess als ein komplexer, kognitiv-linguistischer Vorgang beschrieben, der von der Wahrnehmung orthografischer Einheiten über die semantische Verarbeitung einzelner Wörter bis hin zur Gestaltung und Integration globaler Textstrukturen reicht (Lenhard, 2024). Dabei erfolgt eine differenzierte Betrachtung der Teilkompetenz der Leseflüssigkeit und des Textverstehens. Auf der Grundlage des Modells nach Schaffner (2009) werden zentrale Einflussgrößen des Kompetenzerwerbs beim Lesen identifiziert und in ihren wechselseitigen Beziehungen diskutiert. Der rezeptive Wortschatz fungiert als unmittelbarer Einflussfaktor für automatisierte Dekodierungsprozesse und stellt somit die Grundlage für eine effiziente Verarbeitung sprachlicher Informationen dar (Lenhard & Artelt, 2009). Der sprachliche Input, definiert als Umfang und Qualität der in der Umwelt angebotenen mündlichen und geschriebenen Sprache, determiniert maßgeblich den Wortschatzaufbau und wirkt damit indirekt auf die Lesekompetenz (Aktas, 2020; Ehl & Grosche, 2020). Der sozioökonomische Status der Familie beeinflusst strukturell den Zugang zu bildungsrelevanten Ressourcen sowie familiären Sprachangeboten und beeinflusst dadurch den Input in qualitativer und quantitativer Hinsicht (Calvo & Bialystok, 2014). Diese Variablen bilden das theoretische Gerüst, anhand dessen im empirischen Teil der Studie ihr Einfluss überprüft wird. Im Anschluss daran wird ein Überblick über standardisierte Verfahren zur Erhebung von Lesekompetenz bei Kindern gegeben. Abschließend erfolgt eine detaillierte Darstellung der Eye-Tracking-Methode. Hierbei werden die theoretischen Annahmen zur Blickbewegungsanalyse erläutert und ihre Eignung für die Untersuchung dynamischer Leseprozesse begründet, um die forschungsmethodische Fundierung der vorliegenden Studie abzuschließen.

2.1 Lesen

Die Lesekompetenz ist eine grundlegende Voraussetzung nicht nur für die schulische und akademische Laufbahn, sondern auch für die Ausführung jeden Berufs und das alltägliche Leben (Lenhard, 2024). Sie ist eine essentielle Grundlage für „eine Teilhabe am gesellschaftlichen, kulturellen und politischen Leben“ (Lenhard, 2024, S. 12). Durch digitale Medien oder auch das Lesen diskontinuierlicher Texte gewinnt sie zunehmend an Bedeutung, wodurch auch die Entschlüsselungen und die inhaltliche Rekonstruktion von Texten weiterhin an Wichtigkeit gewinnen (Lenhard, 2024).

In der Literatur finden sich für diese Kompetenz unterschiedliche Begriffe wie *Leseverstehen* oder *Lesefähigkeit*, die häufig als Synonym genutzt werden (Rupp, 2022). Unter Lesekompetenz ist die Fähigkeit zu verstehen, Grapheme, Wörter, Sätze und Texte flüssig zu

lesen oder vorzulesen und deren Inhalte zu verstehen (Rupp, 2022). Lesen ist somit weit mehr als das bloße Dekodieren von Texten. Es ist eine komplexe Schlüsselkompetenz, die für das schulische Lernen essentiell ist (Hurrelmann, 2003 hier nach Rupp, 2022). Dabei handelt es sich stets um einen Prozess. Im weitesten Sinne handelt es sich „um den Prozess der Aufnahme von Geschriebenem, seine Verarbeitung und seine Umsetzung in eigenen sozialen Handlungen“ (Rupp, 2022, S. 112).

2.1.1 Prozesse

Während des Lesens spielen sich zahlreiche Prozesse gleichzeitig ab und greifen ineinander (Jesch, 2020; Lenhard, 2024). In der Wissenschaft existieren viele verschiedene Modelle, die unterschiedliche Teilprozesse in den Fokus rücken[2]. Einen Konsens gibt es jedoch über die wesentlichen Prozesse, die während des Lesens ablaufen (Lenhard, 2024). Lenhard (2024) und Jesch (2020) gliedern die Prozesse des Leseverständnisses in zwei Hierarchieebenen: die niedrige und hohe Hierarchieebene. In dieser Arbeit wird das Modell nach Jesch (2020) detailliert erläutert und mit Hilfe des Modells nach Lenhard (2024) ergänzt.

Unter hierarchieniedrigen Prozessen sind grundlegenden Fähigkeiten der Graphem-, Wort- und Satzerkennung sowie der lokalen Kohärenzbildung zu verstehen (Jesch, 2020). Als einen wichtigen Prozess führt Jesch (2020) die Leseflüssigkeit an, die sich in vier Dimensionen untergliedern lässt. Auf der ersten Ebene der Graphem-, Wort- und Satzerkennung nennt sie die Dekodiergenauigkeit (Jesch, 2020). Hierbei handelt es sich um das sichere Dekodieren einzelner Wörter. Die zweite Dimension stellt die Automatisierung dieses Prozesses dar. Eine automatisierte Graphem-, Wort- und Satzerkennung führt dazu, dass eine kognitive Entlastung beim Textverstehen ermöglicht wird (Jesch, 2020). Als weitere Dimension nennt Jesch (2020) die Lesegeschwindigkeit, die eine wichtige Rolle einnimmt. Handelt es sich um ein langsames Lesen kann es dazu führen, dass das Arbeitsgedächtnis überfordert wird, da vor der Herstellung einer lokalen Kohärenz, das zuvor Gelesene bereits vergessen wird (Jesch, 2020). Die letzte Dimension stellt die Vorlesefähigkeit dar, bei der es sich um die Intonation, Wortbetonung und das phrasierte Lesen, das zu einer angemessenen Vortragsweise führt, handelt (Jesch, 2020). Auch Lenhard (2024) greift die Dekodierung von Wörtern auf. Er gliedert diesen Prozess der Dekodierung in Unterkompetenzen, die auch bei Jesch (2020) teilweise zu finden sind:

- Identifizierung der Schriftzeichen unabhängig von dem Schriftsatz, der Farbe und der Schriftgröße,

[2] u.a. Mehrdimensionales Modell des Lesens (Schründer-Lenzen, 2004); Situationsmodell (Lenhard, 2024); Mehrebenenmodell des Lesens (Rosebrock & Nix, 2025).

- Kenntnis über die Lautzuordnung zu den Schriftzeichen sowie das Beherrschen von Ausnahmen,
- Zuordnung der Laute zu einzelnen Graphemen eines Wortes,
- gleichzeitiges Ablegen im Arbeitsgedächtnis in der richtigen phonologischen Reihenfolge,
- Erkennen der Silbengrenzen,
- Finden eines Wortes im Gedächtnis, das möglichst der Lautfolge im Arbeitsgedächtnis entspricht,
- kein Vergessen der bereits gelesenen Wörter des Satzes.

Aufgrund dessen, dass all diese Prozesse zeitgleich ablaufen, dient eine Automatisierung der einzelnen Prozesse dazu, die hohen kognitiven Anforderungen sicher zu bewältigen (Lenhard, 2024). Zu den hierarchieniedrigen Prozessen zählt zudem die Fähigkeit mit Hilfe von Kohäsionsmitteln[3], Sätze oder Satzteile in Bezug zu setzen und Beziehungen zwischen Wörtern zu erkennen (Lenhard, 2024). Diese Fähigkeit, eine syntaktische Analyse durchzuführen und die Struktur eines Satzes zu entschlüsseln, wird lokale Kohärenzbildung genannt (Lenhard, 2024).

Im Gegensatz zu den niedrigen Hierarchieebenen handelt es sich in den hohen Ebenen um Prozesse, die bei der Erarbeitung und dem Verständnis größerer Texteinheiten relevant sind (Lenhard, 2024). Diese gelten laut Lenhard (2024) nicht als anspruchsvoller als die hierarchieniedrigen Prozesse, wirken jedoch komplizierter, da komplexeres und längeres Material verarbeitet und so ein höheres Abstraktionsniveau erreicht wird (Lenhard, 2024). Lesestrategien dienen der globalen Kohärenzbildung, das heißt dem übergreifenden Verständnis inhaltlicher Zusammenhänge (Jesch, 2020). Hierzu verknüpfen die Lesenden semantische Inhalte mit ihrem vorhandenen Wissen (Jesch, 2020). Auch die gängige Theorie Model of Discourse Comprehension (Kintsch & van Dijk, 1978 zitiert nach Lenhard, 2024), die das Textverstehen in zwei übergeordnete Prozesse untergliedert, greift diese Verknüpfung auf. So beinhaltet der Konstruktionsprozess das Extrahieren und Interpretieren von kleinen Bedeutungseinheiten anhand der Verknüpfung verschiedener Textinhalte mit bereichsspezifischem Vorwissen (Lenhard, 2024). So werden textliche Schlussfolgerungen in einem Netzwerk mit Wissen aus dem Langzeitgedächtnis angereichert (Lenhard, 2024). Durch die Sortierung des in dem Konstruktionsprozess entstandenen Netzes entsteht im Integrationsprozess eine Bedeutungsrepräsentation des Textes (Lenhard, 2024). Dieses

[3] Kohäsionsmittel dienen der Herstellung von Zusammenhang auf grammatischer und textueller Ebene, indem sie semantische Relationen zwischen einzelnen Textelementen explizit markieren. Dazu zählen insbesondere Konnektoren wie Konjunktionen und Präpositionen, die kausale oder temporale Verknüpfungen innerhalb eines Textes ermöglichen. (Hoffmann-Erz, 2023).

sogenannte Situationsmodell umfasst über den reinen Text hinaus weitere Assoziationen und Schlussfolgerung (Lenhard, 2024). Jesch (2020) unterteilt diese Prozesse in drei kognitive Lesestrategien: Elaborationsstrategie, Ordnungsstrategie und Wiederholungsstrategie (Jesch, 2020). Die Elaborationsstrategie ist dem Konstruktionsprozess gleichzusetzen. Durch eine inhaltliche Klärung unter Berücksichtigung des Weltwissens erfolgt eine Ausarbeitung des gelesenen Textes (Jesch, 2020). Die Ordnungsstrategie spiegelt den Integrationsprozess wider. Nur die Wiederholungsstrategie fügt Jesch (2020) den Prozessen der Lesekompetenz hinzu. Um sich Inhalte besser merken zu können und ein tiefgreifendes Verständnis des Gelesenen aufzubauen, nutzen einige Kinder oder auch Erwachsene die Strategie des wiederholten Lesens (Jesch, 2020).

Während diese Modelle die hierarchische Struktur der Leseprozesse differenziert darstellen, wird in der kognitionspsychologischen Forschung zusätzlich betont, dass verschiedene Gedächtnissysteme eine zentrale Rolle bei der Verarbeitung von Texten spielen (Baddeley, 2000; Kintsch, 1998). Das sensorische Gedächtnis ermöglicht die kurzfristige Speicherung visueller Reize wie Buchstaben und Wörtern und bildet somit die Grundlage für die initiale Wahrnehmung beim Lesen (Kintsch, 1998). Das Arbeitsgedächtnis ist wesentlich für die Verarbeitung und Integration von Informationen über Satzgrenzen hinweg und unterstützt das Herstellen semantischer Zusammenhänge (Kintsch, 1998; Lenhard, 2024). Schließlich stellt das Langzeitgedächtnis das notwendige Hintergrundwissen bereit, etwa Wortbedeutungen, grammatikalische Regeln und Weltwissen, das für ein tiefes Textverstehen unabdingbar ist (Kintsch, 1998). Diese Gedächtnissysteme wirken nicht isoliert, sondern greifen dynamisch ineinander und bilden die kognitive Grundlage für die Prozesse auf niedriger und hoher Hierarchieebene.

2.1.2 Leseflüssigkeit und Textverstehen

Wie bereits dargelegt, stellt Lesen einen kognitiv konstruktiven Prozess dar, der eine Vielzahl simultaner Teilaktivitäten integriert (Rosebrock & Nix, 2025). Im Rahmen des Mehrebenenmodells des Lesens differenzieren Rosebrock und Nix (2025) drei zentrale Dimensionen: die soziale Ebene, die Subjektebene und die Prozessebene, die jeweils spezifische Einflussfaktoren auf die Entwicklung von Lesekompetenz abbilden (siehe Abbildung 1). Im Folgenden steht die Prozessebene im Fokus, da sie die kognitiven Mechanismen der Sprachverarbeitung beim Lesen in den Mittelpunkt stellt.

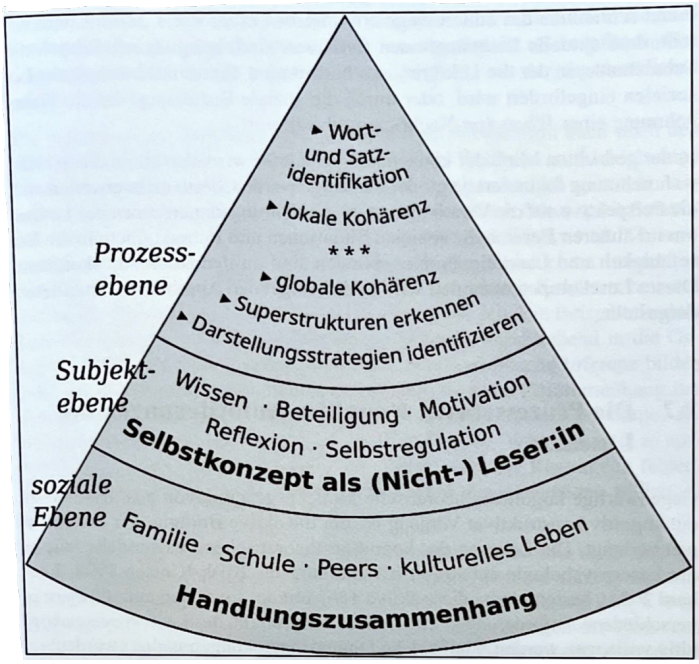

Abbildung 1: Mehrebenenmodell des Lesens (Rosebrock & Nix, 2025, S. 15)

Innerhalb dieser Prozessebene des Lesens kann zwischen den Kompetenzen der Leseflüssigkeit und des Textverstehens unterschieden werden. Diese Prozesse finden zwar stets zeitgleich statt, dennoch bildet die Leseflüssigkeit eine grundlegende Voraussetzung für das erfolgreiche Verstehen schriftlicher Texte (Rosebrock, 2022).

Die Leseflüssigkeit lässt sich dem Prozess der Wort- und Satzidentifikation zuordnen (vgl. Abbildung 1) und beschreibt die grundlegende Fähigkeit, sprachliche Einheiten schnell, korrekt und bedeutungshaltig zu erfassen und zu verarbeiten (Rosebrock, 2022). Dabei handelt es sich nicht um einen rein mechanischen Vorgang, sondern um einen komplexen kognitiven Prozess, der verschiedene Teilfähigkeiten integriert und in enger Wechselwirkung mit höheren Verstehensleistungen steht (Rosebrock & Nix, 2025). Die Verarbeitung eines Wortes erfolgt dabei bidirektional: Der Bottom-up-Prozess umfasst die visuelle Wahrnehmung der Schriftzeichen, während der Top-down-Prozess auf gespeichertes Wissen zurückgreift, das semantische, orthografische und kontextuelle Informationen einbezieht (Rosebrock & Nix, 2025). Diese Interaktion ermöglicht eine ganzheitliche Worterkennung, die über die sequenzielle Entzifferung einzelner Grapheme hinausgeht (Rosebrock & Nix, 2025).

Wie bereits in Kapitel 2.1.1 erläutert, spielen insbesondere automatisierte Prozesse der Worterkennung und Dekodierung eine zentrale Rolle. Diese Prozesse ermöglichen es Leserinnen und Lesern, sich beim Lesen nicht auf die Entschlüsselung einzelner Wörter zu konzentrieren, sondern ihre kognitiven Ressourcen auf die Verarbeitung des Textes zu richten (Rosebrock, 2022). Eine gut ausgeprägte Leseflüssigkeit entlastet somit

das Arbeitsgedächtnis und schafft die Voraussetzung für ein vertieftes Verstehen des Textes (Jesch, 2020; Lenhard, 2024; Rosebrock & Nix, 2025).

In der aktuellen Forschung wird Leseflüssigkeit typischerweise in vier miteinander verknüpfte Komponenten unterteilt. Die erste Komponente ist die Automatisierung, die sich auf die Fähigkeit bezieht, Wörter schnell und ohne bewusste Anstrengung zu erkennen und zu dekodieren. Automatisierung ist ein zentrales Merkmal effizienter Leserinnen und Leser und wird in der Forschung als entscheidender Prädiktor für Lesekompetenz betrachtet (Rosebrock, 2022). Lenhard (2024) hebt hervor, dass automatisiertes Worterkennen nicht nur die Lesegeschwindigkeit erhöht, sondern auch die Grundlage für das Verstehen längerer und komplexerer Texte bildet. Die zweite Komponente ist die Genauigkeit, also die Fähigkeit, Wörter korrekt zu lesen und dabei möglichst wenige Lesefehler zu machen (Rosebrock, 2022). Eine hohe Genauigkeit ist essentiell für die semantische Verarbeitung, da fehlerhafte Dekodierung zu Missverständnissen auf Wort- und Satzebene führen kann und somit das Textverstehen beeinträchtigt. Die dritte Komponente betrifft die Lesegeschwindigkeit, die sich dem natürlichen Sprechtempo annähert (Rosebrock, 2022). Eine angemessene Geschwindigkeit beim Lesen ist ein Indikator für die Effizienz der Dekodierprozesse und steht in engem Zusammenhang mit der Fähigkeit, syntaktische und semantische Zusammenhänge im Text zu erkennen. Langsames, stockendes Lesen hingegen kann die Kohärenzbildung erschweren und das Verstehen komplexer Inhalte behindern (Jesch, 2020; Lenhard, 2024; Rosebrock, 2022). Die vierte und letzte Komponente stellt das sogenannte prosodische Sequenzieren dar. Darunter wir die Fähigkeit verstanden, Bedeutungszusammenhänge zwischen Wörtern und Wortgruppen korrekt zu erfassen (Rosebrock, 2022). Prosodisches Lesen trägt dazu bei, syntaktische Strukturen zu verdeutlichen und semantische Relationen im Text zu erschließen, was wiederum das globale Textverständnis unterstützt (Rosebrock, 2022).

Zusammenfassend lässt sich festhalten, dass Leseflüssigkeit weit mehr ist als die Fähigkeit, Wörter korrekt auszusprechen und zu lesen. Sie ist ein Prozess, der automatisierte Dekodierung, Fehlerfreiheit, angemessene Geschwindigkeit und prosodisches Verständnis miteinander verbindet.

Dieser hierarchieniedrige Prozess der Leseflüssigkeit nimmt Einfluss auf den zweiten hierarchieniedrigen Prozess der lokalen Kohärenzbildung (Rosebrock & Nix, 2025). Dieser Prozess ist jedoch bereits dem Textverstehen zuzuordnen. Während dieses Prozesses werden Zusammenhänge zwischen Wörtern und innerhalb eines Satzes auf Ebene der Syntax und Semantik hergestellt (Rosebrock & Nix, 2025). So können einfache Verknüpfungen und Schlussfolgerungen getroffen werden. Dabei sind insbesondere die Fähigkeit zur

syntaktischen Analyse sowie die semantische Integration von Bedeutungsträgern entscheidend, um lokale Kohärenz herzustellen. Diese Prozesse erfordern eine simultane Aktivierung und Koordination sprachlicher und kognitiver Ressourcen. Ist jedoch das Dekodieren von Wörtern noch nicht automatisiert, werden Kinder bereits an diesem Punkt auf Schwierigkeiten treffen, da ihre kognitiven Ressourcen noch für das Dekodieren aufgebracht werden, sodass keine Ressourcen für Schlussfolgerungen zur Verfügung stehen (Rosebrock & Nix, 2025).

Weiterführend muss über die Satzgrenze hinaus eine globale Kohärenz gebildet werden (Rosebrock & Nix, 2025). Dieser Prozess beschreibt, dass beim Lesen eine gesamte Vorstellung über den Inhalt entworfen wird (Rosebrock & Nix, 2025). Diese Gesamtvorstellung bildet sich in Denkprozessen aus der Verarbeitung des Textes auf semantischer Ebene sowie aus eigenen Erfahrungen über das jeweilige Thema (Rosebrock & Nix, 2025). Die globale Kohärenzbildung stellt somit einen integrativen Prozess dar, bei dem Textinformationen mit Vorwissen und Erwartungen verknüpft werden (Jesch, 2020). Leserinnen und Leser konstruieren dabei eine kohärente mentale Repräsentation des Textinhalts, die über die bloße Aneinanderreihung von Sätzen hinausgeht und eine tiefere Bedeutungsebene erschließt (Rosebrock & Nix, 2025; Jesch, 2020).

Auch das Erkennen von Superstrukturen und das Identifizieren von Darstellungsstrategien stellen zentrale kognitive Prozesse im Rahmen des Textverstehens dar (Rosebrock & Nix, 2025). Superstrukturen bezeichnen das Wissen über unterschiedliche Textsorten und deren charakteristische Gestaltungselemente, wie etwa Aufbau, sprachliche Mittel oder typische Argumentationsmuster (Rosebrock & Nix, 2025). Die Kenntnis solcher Superstrukturen ermöglicht es Leserinnen und Lesern, unabhängig von der bereits etablierten globalen Kohärenz, Hypothesen über den weiteren Textverlauf zu formulieren oder vorhandene Informationen gezielt zu interpretieren (Rosebrock & Nix, 2025). Dabei fungieren Superstrukturen als mentale Schemata, die den Leseprozess strukturieren und antizipatorisches Verstehen fördern (Rosebrock & Nix, 2025). Auch die Kenntnis über Darstellungsstrategien, also über die Intentionen eines Textes, wie etwa das Informieren, Argumentieren, Appellieren oder Unterhalten, beeinflusst die kognitiven Prozesse während des Lesens maßgeblich. Kinder gehen mit bestimmten Erwartungen an einen Text heran, die durch solche Strategien geprägt sind und die Art und Weise steuern, wie Informationen aufgenommen und verarbeitet werden (Rosebrock & Nix, 2025; Jesch, 2020).

All diese Prozesse, von der Worterkennung über lokale und globale Kohärenzbildung bis hin zur Nutzung von Superstrukturen und Darstellungsstrategien, laufen simultan ab (Rosebrock & Nix, 2025). Dabei ist jedoch zu beachten, dass eine nicht ausreichende Leseflüssigkeit

Auswirkungen auf die Ausführung weiterführender kognitiver Prozesse hat (Rosebrock & Nix, 2025). Insbesondere die Bildung globaler Kohärenz wird dadurch erheblich beeinträchtigt. Denn wenn basale dekodierende Prozesse nicht automatisiert sind, beanspruchen sie einen Großteil der kognitiven Ressourcen, die für höhere Verstehensleistungen wie Inferenzerzeugung, Strukturierung des Textes oder metakognitive Überwachung notwendig wären (Rosebrock & Nix, 2025). In solchen Fällen bleibt das Textverstehen auf einer oberflächlichen Ebene stehen, da die Integration semantischer, syntaktischer und pragmatischer Informationen nicht vollständig erfolgen kann (Lenhard, 2024).

Auch groß angelegte Schulleistungsstudien wie das Programme for International Student Assessment (PISA) entwickeln Kompetenzmodelle zur Erfassung von Lesekompetenz (vgl. Abbildung 2). Diese Modelle fokussieren jedoch primär auf das Textverstehen und bilden damit nur einen Teilbereich des komplexen Leseprozesses ab. Die Leseflüssigkeit als basale Voraussetzung für höhere Verstehensleistungen bleibt weitgehend unberücksichtigt, wodurch ein reduktionistisches Verständnis von Lesekompetenz als rein messbare Kompetenz gefördert wird (Hochstadt et al., 2022). Darüber hinaus adressiert das PISA-Kompetenzmodell die zugrunde liegenden kognitiven Verarbeitungsprozesse nicht explizit, sodass zentrale Aspekte der Sprachverarbeitung, wie die Interaktion von Dekodierung, Arbeitsgedächtnis und semantischer Integration, im Modellrahmen unterrepräsentiert bleiben.

Das Modell der PISA-Studie differenziert zwei komplementäre Verarbeitungsebenen, um systematisch zu beschreiben, wie Leserinnen und Leser Texte verstehen und interpretieren. Im textinternen Prozess erfolgt die Analyse zunächst auf globaler Ebene, indem der Text als Ganzes betrachtet wird und so ein grundlegendes Verständnis des Inhaltes entsteht (vgl. Abbildung 2). Anschließend werden auf lokaler Ebene einzelne Segmente selektiv untersucht, um unabhängige Einzelinformationen wie Fakten oder Schlüsselbegriffe zu extrahieren. Die Rekonstruktion von Kohärenzrelationen zwischen diesen Einzelelementen ermöglicht darüber hinaus erste inhaltliche Interpretationen. Der wissensgestützte Prozess bezieht externes Vorwissen in das Textverstehen ein. Dabei aktivieren Leserinnen und Leser eigene Erfahrungen und fachliches Hintergrundwissen, um die inhaltliche Plausibilität zu prüfen und den textinternen Aufbau vor dem Hintergrund gattungsspezifischer Strukturen zu reflektieren.

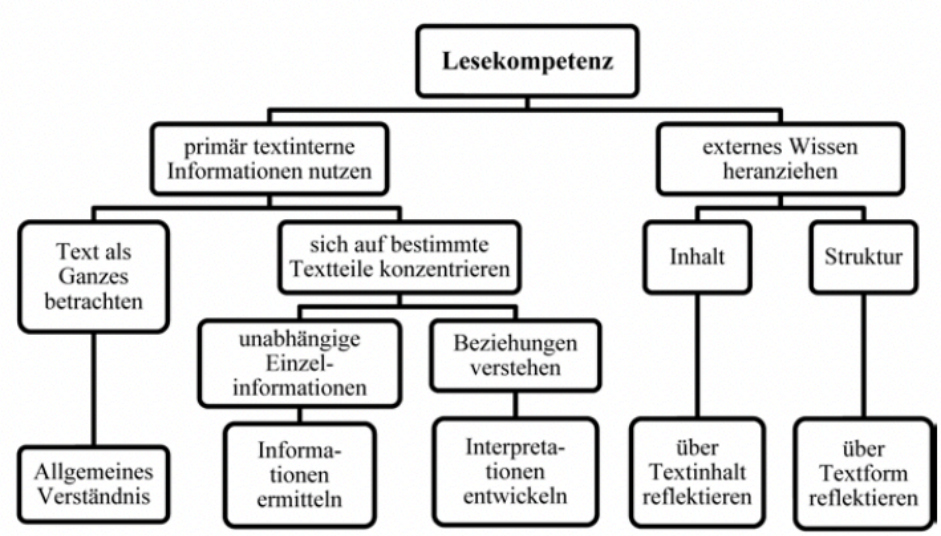

Abbildung 2: Theoretische Struktur der Lesekompetenz in PISA (Hochstadt et al., 2022, S. 144)

In beiden Modellen wird deutlich, dass Textverstehen über das reine Extrahieren explizit genannter Informationen hinausgeht. Es umfasst komplexe kognitive Prozesse wie die kritische Reflexion, die semantische Integration sowie die aktive Konstruktion von Bedeutung, wobei auch evaluative Bewertungen des Textinhalts vorgenommen werden. Lesen erweist sich damit als ein vielschichtiger, kognitiv anspruchsvoller Vorgang, der nicht nur das Verstehen auf verschiedenen Ebenen einschließt, sondern ebenso die Leseflüssigkeit im Sinne automatisierter Dekodierungsprozesse voraussetzt. Insofern stellt Lesen eine erlernbare Kompetenz dar, deren Entwicklung die koordinierte Ausbildung verschiedener Teilprozesse erfordert.

2.1.3 Determinanten

In der Grundschule zeigen sich teils deutliche Leistungsunterschiede in der Lesekompetenz von Schülerinnen und Schülern (Knickenberg, 2018). Diese Unterschiede lassen sich auf verschiedene Einflussfaktoren zurückführen, die auch als Determinanten des Leseverständnisses bezeichnet werden (Lenhard, 2024). Zur Beschreibung dieser Determinanten wurden verschiedene Modelle entwickelt, die jeweils unterschiedliche Schwerpunkte setzen[4].

So legt das Modell von Schaffner (2009) den Fokus auf familiäre und individuelle Einflussfaktoren sowie auf spezifische Leserinnen- und Lesermerkmale, die das Leseverstehen beeinflussen. Im weiteren Verlauf wird das „Modell des Zusammenspiels

[4] Tetraedermodell (Jenkins, 1979); Direct and Inferential Mediation Model (Cromley & Azevedo, 2007); Modell des Zusammenspiels familiärer und psychologischer Einflüsse auf das Leseverstehen (Schaffner, 2009). Die hier aufgeführten Modelle sind exemplarisch und stellen keine vollständige Übersicht dar.

familiärer und psychologischer Einflüsse auf das Leseverstehen“ von Schaffner (2009, S. 36) näher betrachtet. Die Entscheidung für dieses Modell basiert auf seinem umfassenden Ansatz, der sowohl familiäre als auch psychologische Determinanten integriert. Da zudem einige dieser Determinanten Gegenstand der vorliegenden Studie sind, ist dieses Modell besonders geeignet, um Einflussfaktoren auf die Lesekompetenz zu beschreiben.

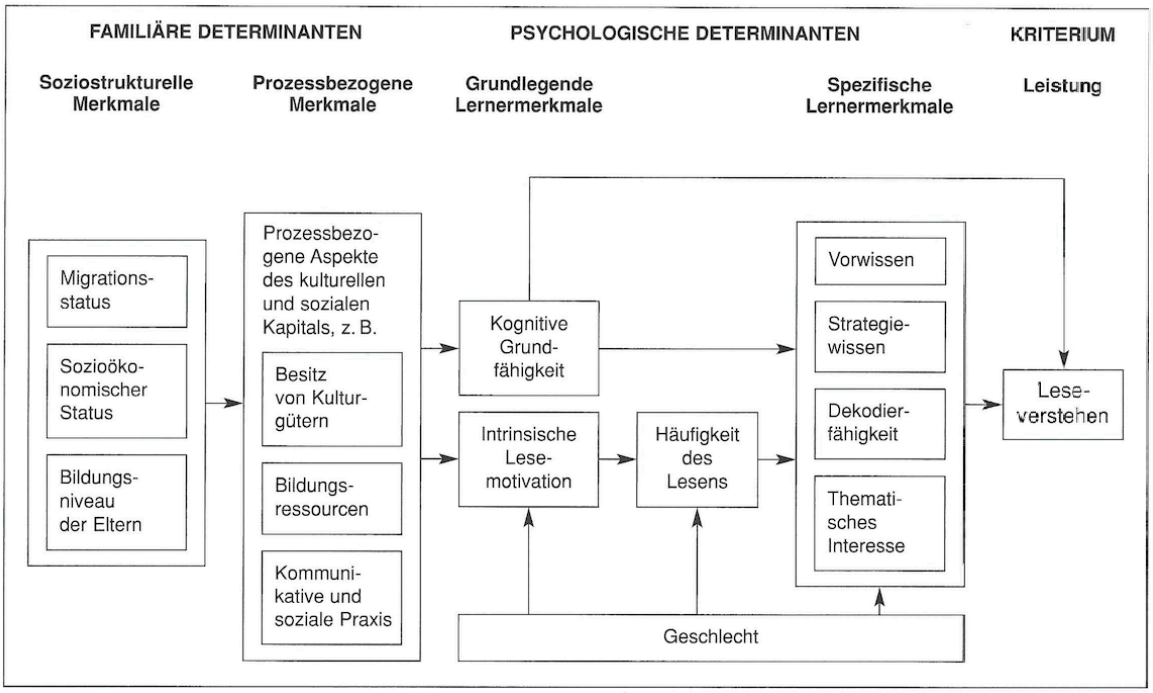

Abbildung 3: Modell des Zusammenspiels familiärer und psychologischer Einflüsse auf das Leseverstehen (Schaffner, 2009, S. 36)

Schaffner (2009) unterscheidet in ihrem Modell zwei wesentliche Einflussfaktoren auf das Leseverstehen: psychologische und familiäre Determinanten. Die psychologischen Determinanten lassen sich in grundlegende und spezifische Merkmale der Lernenden unterteilen. Unter den grundlegenden Merkmalen sind die kognitiven Grundfähigkeiten sowie die Lesemotivation, die sich aus verschiedenen Arten der Motivationen zusammensetzt, zusammengefasst. Die kognitive Grundfähigkeit stellt ein Synonym der Intelligenz, die genetisch bedingt aber auch von Umwelteinflüssen abhängig ist, dar. Die Intelligenz nimmt sowohl direkt als auch indirekt Einfluss auf das Leseverstehen, wobei die indirekte Wirkung bisher eher wenig untersucht wurde. Direkten Einfluss nimmt die Intelligenz, indem die „Verknüpfung von Propositionen oder die Inferenzbildung erleichtert“ (Schaffner, 2009, S. 24) wird. Im Bereich der Motivation ist festzuhalten, dass die intrinsische Motivation zu einer „höheren Verstehensleistung“ (Schaffner, 2009, S. 32) führt als die extrinsische Motivation. Schaffner (2009) greift zwei Arten der Motivation beim Lesen auf, die sowohl extrinsisch als auch intrinsisch motiviert sein können. Zum einen nennt sie die habituelle Lesemotivation, die durch ein wiederholtes Auftreten von Leseabsichten gekennzeichnet ist (Schaffner, 2009).

Zum anderen verweist sie auf die aktuelle Lesemotivation, die sich als situativ bedingte Absicht zum Lesen beschreiben lässt (Schaffner, 2009). Diese lässt sich sowohl dem Bereich der extrinsischen als auch intrinsischen Motivation einordnen. Die extrinsische Motivation ist insbesondere in Leistungssituationen zu finden, wohingegen die intrinsische Motivation durch das Lesen selbst entsteht (Schaffner, 2009). Im Hinblick auf die verschiedenen Komponenten der intrinsischen und extrinsischen Motivation zeigen Untersuchungen wie von Schaffner und Schiefele (2007, zitiert nach Schaffner, 2009), „dass die intrinsische Motivation positiv mit dem Gebrauch anspruchsvoller Strategien der Textverarbeitung zusammenhängt“ (Schiefele, 2007, zitiert nach Schaffner, 2009, S: 32). Somit wird deutlich, dass sowohl die habituelle als auch die aktuelle Lesemotivation im Bereich der intrinsischen Motivation positive Auswirkungen auf die Lesekompetenz haben kann (Schaffner, 2009).

In Anbetracht der spezifischen Merkmale werden vier kognitive und motivationale Kategorien unterschieden (Schaffner, 2009). Unter die kognitiven Merkmale fallen das Vorwissen, das Strategiewissen sowie die Dekodierfähigkeit, wohingegen das thematische Interesse den motivationalen Merkmalen zugeordnet werden kann. Die Dekodierfähigkeit wurde bereits in den Kapiteln 2.1.1 und 2.1.2 als Prozess des Lesens erläutert. Es handelt sich um das reine Entschlüsseln von Graphemen, Wörtern und Sätzen. Das Dekodieren gilt als eine „notwendige Voraussetzung für das Textverstehen“ (Schaffner, 2009, S. 25). Schaffner (2009) stellt heraus, dass diese Fähigkeit einen positiven Einfluss auf die Lesekompetenz nimmt, da durch eine Automatisierung des Dekodierens von Wörtern und Sätzen, kognitive Ressourcen für die anspruchsvollen Prozesse genutzt werden können (Schaffner, 2009). Das Vorwissen umfasst eine Vielzahl an Wissensarten, die beim Lesen für das Textverständnis herangezogen werden (Schaffner, 2009). Unter dieses Vorwissen fallen jedoch nicht nur Wortwissen oder das Wissen über Merkmale eines Textes, sondern auch das Allgemeinwissen, das zum Verstehen des Gelesenen beiträgt (Schaffner, 2009). Insbesondere das thematische Vorwissen wirkt sich positiv auf das Textverstehen aus, da anhand dieses Wissens „Kohärenzlücken“ (Schaffner, 2009, S. 26) geschlossen und Schlussfolgerungen über den Text hinaus getroffen werden können. Das dritte kognitive Merkmale ist in dem Strategiewissen zu finden. Hierbei handelt es sich, wie der Name bereits zeigt, um das Wissen über bestimmte Strategien und Techniken, die beim „verstehenden Umgang mit Texten sowie über deren zielangemessenen Einsatz“ (Schaffner, 2009, S. 27) unterstützen. Untersuchungen zeigen, dass sich anhand von gezielten Übungen zum metakognitiven Strategiewissen das Verstehen eines Textes verbessert (Souvignier & Mokhlesgerami, 2006, zitiert nach Schaffner, 2009). Das thematische Interesse wird ebenfalls in die spezifischen Merkmale eingeordnet, wobei es sich jedoch um einen motivationalen Faktor handelt (Schaffner, 2009). Dieses Interesse drückt einen Zusammenhang zwischen der oder dem Lesenden und dem Gegenstand beziehungsweise

dem Textinhalt aus (Schaffner, 2009). Laut Schaffner (2009) liegen ihnen „auf der Ebene der kognitiven Repräsentation […] Verknüpfungen zwischen Gegenständen […] und gefühls- oder wertbezogenen Attributen […] zugrunde“ (Schaffner, 2009, S. 28). Das heißt, dass thematische Gegenstände mit Gefühlen und Werten verknüpft sind, sodass durch das thematische Interesse die Verstehensleistung positiv beeinflusst werden kann. Als letzter Aspekt der psychologischen Determinanten wird die Lesehäufigkeit genannt (Schaffner, 2009). Die Lesehäufigkeit stellt einen „quantitativen Aspekt“ (Schaffner, 2009, S. 32) dar, der die zeitliche Häufigkeit des Lesens beschreibt. Empirische Studien zeigen, dass die Lesehäufigkeit positiv mit der Lesekompetenz korreliert. Durch häufiges Lesen werden die Dekodierfähigkeit, grundlegende sprachliche Fähigkeiten sowie das Strategiewissen gefördert (Schaffner, 2009).

Die familiären Determinanten untergliedert Schaffner in ihrem Modell in die soziostrukturellen und die prozessbezogenen Merkmale (Schaffner, 2009). Die soziostrukturellen Merkmale gehen auf den Migrationsstatus, den sozioökonomischen Status sowie auf das Bildungsniveau der Eltern ein (Schaffner, 2009). Untersuchungen zeigen, dass die soziostrukturellen Merkmale einen indirekten Einfluss auf die Lesekompetenz von Schülerinnen und Schülern haben. Der Einfluss auf die Lesekompetenz wird vor allem durch das kulturelle und soziale Kapital des Elternhauses beeinflusst. Unter dem sozialen und kulturellen Kapital wird der Besitz von Kulturgütern, Bildungsressourcen und die kommunikative sowie soziale Praxis gefasst (siehe Abbildung 3). Unter dieses Kapital fallen beispielsweise der Besitz von Büchern, die Häufigkeit des Vorlesens oder auch die Teilnahme an kulturellen Aktivitäten wie Museumsbesuchen. Die kommunikative und soziale Praxis umfasst die Quantität und Qualität der familiären Gespräche oder auch eine sprachlich anregende Kommunikation beim Vorlesen. Untersuchungen zeigen, dass insbesondere die prozessbezogenen Merkmale Vorläuferfähigkeiten beeinflussen, die für die Lesekompetenz relevant sind. Darunter fallen die zuvor erwähnten Bereiche des Vorwissens und der Motivation am Lesen, aber auch die sprachlichen Kompetenzen und das phonologische Bewusstsein (Jansen & Marx, 1999, zitiert nach Schaffner, 2009). Hurrelmann (2004, zitiert nach Schaffner, 2009) stellte heraus, dass insbesondere in sozial schwächere Familien die frühkindliche Förderung weniger erfolgt als in sozial besser gestellten Familien.

In ihrem Modell geht Schaffner (2009) nicht auf den Wortschatz ein. Dieser stellt jedoch „ […] einen wichtigen Einflussfaktor für das Leseverstehen dar“ (Lenhard & Artelt, 2009, S. 13). Laut McElvany und Schneider (2009) stehen die „Wortschatz- und Textverständnisentwicklung“ (McElvany & Schneider, 2009, S. 160) in einer wechselseitigen Beziehung. Wie auch bei einer ausgeprägten Dekodierfähigkeit (siehe Kapitel 2.1.1) kann durch einen sicheren und schnellen Zugriff auf das mentale Lexikon das Arbeitsgedächtnis entlastet werden, sodass höhere kognitive Prozesse stattfinden können (McElvany & Schneider, 2009). Abgesehen von dem

Einfluss des Wortschatzes auf die Lesekompetenz zeigt sich, dass bei einer höheren Lesekompetenz unbekannte Wörter durch den Kontext erschlossen werden können und so der Wortschatz erweitert wird (Lenhard & Artelt, 2009).

Die erläuterten Determinanten tragen zur Erklärung bei, wovon die Lesekompetenz abhängt und warum sie selbst innerhalb derselben Altersstufe eine große Heterogenität aufweist (Schaffner, 2009).

2.2 Zentrale Einflussfaktoren auf die Lesekompetenz

Die Entwicklung der Lesekompetenz wird durch eine Vielzahl an Faktoren beeinflusst, die sowohl individuelle als auch kontextuelle Dimensionen umfassen. Im vorliegenden Kapitel werden zentrale Einflussgrößen dargestellt, die für das Verständnis der zugrunde liegenden Studie von besonderer Relevanz sind. Im Fokus stehen dabei die Konzepte des Wortschatzes, des sprachlichen Inputs sowie des sozioökonomischen Status.

2.2.1 Wortschatz

Der Wortschatz stellt ein Element der Lexikologie dar, die als „Wissenschaft von Wort und Wortschatz einer Sprache“ (Wanzeck, 2010, S. 11) betitelt wird. Nach Römer und Sturm (2021) beinhaltet der Wortschatz alle Wörter einer Sprache, die sich im mentalen Lexikon befinden. Dieses zentrale sprachliche Wissenssystem im menschlichen Gehirn speichert sämtliche Wörter und ordnet sie so an, dass sie innerhalb von Sekundenbruchteilen abgerufen werden können. Jeder Eintrag ist mehrdimensional strukturiert, sodass eine effiziente Verarbeitung beim Sprechen und Verstehen möglich ist (Aitchison, 1997).

In der Psycholinguistik wird im mentalen Lexikon zwischen der Lemma- und der Lexem-Ebene unterschieden. Die Lemma-Ebene umfasst die semantischen und syntaktischen Informationen eines Wortes, wohingegen die Lexem-Ebene die morphologischen und phonologischen Strukturen beinhaltet (Levelt, 2001). So strukturiert die Lemma-Ebene die Wörter nach thematischen Feldern, wodurch starke Verbindungen zwischen ähnlichen Begriffen entstehen. Die Ordnung auf der Lexem-Ebene findet auf Grundlage der Lautstruktur statt, sodass ähnlich klingende Wörter nah beieinander angeordnet sind. Nach Levelt (2001) fungiert die Lemma-Selektion als zentraler Prozess der Wortfindung. Sie wählt zunächst das passende semantische Konzept aus und legt grammatische Parameter fest. Anschließend werden auf der Lexem-Ebene die phonologischen Abfolgen abgerufen und an die artikulatorische Planung weitergeleitet, ein Schritt, den Levelt (2001) Form-Encoding nennt und der entscheidend für eine flüssige Artikulation ist.

Laut Aitchison (1997) erfüllt das mentale Lexikon drei zentrale Funktionen: Produktion, Wahrnehmung und Gedächtnis. Im Bereich der Produktion bilden die Semantik und Syntax den Ausgangspunkt für das Formulieren von Äußerungen. Lautstrukturen sind der Ausgangspunkt für das Erkennen und Verstehen gesprochener Sprache. Eine Speicherung in und schnelle Abrufbarkeit von Wörtern aus dem Gedächtnis ermöglicht die strukturierte Organisation von Wörtern. Dabei ist das mentale Lexikon kein statisches Wörterbuch. Die Verbindungen zwischen Einträgen passen sich dynamisch an Häufigkeit und Kontextnutzung an. So lässt sich erklären, warum man manchmal die Bedeutung eines Wortes kennt, aber nicht seine exakte Lautform und umgekehrt (Aitchison, 1997).

In der Linguistik wird zwischen Sprachproduktion und Sprachverstehen unterschieden. Nach Levelt (2001) gliedert sich die Produktion in vier Ebenen. Beginnend mit der Konzeptualisierung wird eine Auswahl sowie eine Strukturierung der kommunikativen Intention getroffen. Anschließend wird in der Lemma-Selektion das korrekte Wort hinsichtlich seiner semantischen Bedeutung und grammatischer Merkmale aus dem mentalen Lexikon abgerufen. Im Bereich des lexikalischen Zugriffs werden irrelevante Pfade der Auswahl unterdrückt, sodass das richtige Wort ausgewählt werden kann. Im Bereich des Form-Encodings werden aus dem mentalen Lexikon die lautlichen Bestandteile eines Wortes und die phonologische Abfolge abgerufen. Auf Basis der Phonemabfolge werden motorische Programme erstellt, die die Bewegungen von Lippen, Zunge, Kiefer und Kehlkopf steuern. Diese Abläufe werden koordiniert, um flüssige Übergänge zwischen den Lauten zu gewährleisten. Erst nach Abschluss dieser Prozesse erfolgt die physische Ausführung der Sprachbewegung. Aitchison (1997) betont in ihrem Netzwerkmodell, dass bei Produktion und Verstehen auf Grundlage der kommunikativen Absicht zahlreiche nahegelegene Einträge im mentalen Lexikon aktiviert werden, aus denen das passende Lexem selektiert wird.

Zudem unterscheidet die Linguistik zwischen dem Verstehens- und dem Mitteilungswortschatz (Leimbrink, 2015). Der rezeptive Wortschatz, auch Verstehenswortschatz genannt, bezeichnet die Lexeme, die Lernende durch Hören und Lesen aufnehmen und verstehen können (Leimbrink, 2015; Römer & Sturm, 2021). Nation (2017) zeigt, dass extensives Hören und Lesen vor allem die semantische und orthografische Verarbeitungstiefe steigern, während intensives Zuhören beispielsweise bei Diktaten und intensives Lesen wie beim Lesen von Lückentexten die phonologische und morphologische Verankerung neuer Wörter fundieren. Der produktive Wortschatz umfasst die Wörter, die Lernende aktiv beim Sprechen und Schreiben einsetzen. Nation (2017) betont, dass aufgabenbasierte Sprechanlässe und Schreibübungen mit einer Rückmeldung zentrale Rollen spielen, um den Abruf neuer Vokabeln zu festigen und in komplexen Satzstrukturen zu verankern. Der rezeptive Wortschatz weist in allen Altersstufen eine weitaus höhere Anzahl an Wörtern als der produktive Wortschatz auf,

der im erwachsenen Alter etwa 10.000 bis 20.000 Wörter umfasst (Leimbrink, 2015). Der rezeptive Wortschatz hingegen liegt zwischen 40.000 und 100.000 Lexemen (Leimbrink, 2015; Römer & Sturm, 2021). Der Erwerb des Wortschatzes bildet einen lebenslangen Prozess ab (Leimbrink, 2015). Bereits in der Grundschule weisen Kinder einen Wortschatz von 5.000 Wörtern im produktiven und 15.000 Wörtern im rezeptiven Wortschatz auf (Leimbrink, 2015). Dieser Wert unterscheidet sich jedoch stark durch die familiären Bedingungen wie dem sprachlichen Input oder dem sozioökonomischen Status der Kinder (Leimbrink, 2015). In verschiedenen Studien zeigt sich, dass Kinder mit einem qualitativ und quantitativ höheren Input ihren Wortschatz schneller erweitern als Kinder mit geringerem Maß an Input (Hoff & Naigles, 2002; Tamis-LeMonda et al., 2001).

2.2.2 Sprachlicher Input

Der sprachliche Input ist eine essentielle Variable für den Spracherwerb von Kindern. Ohne Kontakt zur Sprache, ist ein Erwerb nicht möglich (Pearson et al., 1997). Empirische Befunde belegen darüber hinaus, dass sowohl die Qualität als auch die Quantität des sprachlichen Inputs maßgeblich zur Entwicklung des rezeptiven Wortschatzes beitragen. Kinder, die über einen reichhaltigen und differenzierten sprachlichen Input verfügen, zeigen in der Regel einen signifikant umfangreicheren Wortschatz als Kinder mit eingeschränktem Zugang zu sprachlichen Anregungen (Aktas, 2020; Ehl & Grosche, 2020).

Beim sprachlichen Input wird zwischen der Quantität und der Qualität des Sprachkontaktes unterschieden (MacLeod & Demers, 2023; Masek et al., 2021). Die Quantität des Inputs kann als Umfang des Kontaktes mit der Sprache definiert werden (MacLeod & Demers, 2023). Die Quantität des sprachlichen Inputs wird in empirischen Studien typischerweise anhand der Dauer und Häufigkeit sprachlicher Exposition erfasst, die Kinder im Alltag erfahren (Hoff & Naigles, 2002). Zur objektiven Messung kommen unter anderem technische Systeme wie das Language Environment Analysis (LENA)-System zum Einsatz (Masek et al., 2021). Dieses ermöglicht eine automatisierte Erhebung zentraler Parameter wie die Anzahl gesprochener Wörter, die Häufigkeit von Sprecherwechseln sowie die Gesamtdauer sprachlicher Interaktionen (Masek et al., 2021). Ergänzend werden häufig standardisierte Fragebögen eingesetzt, um subjektive Einschätzungen der sprachlichen Umgebung zu erfassen. Dabei kann beispielsweise erhoben werden, wie viele Bücher regelmäßig vorgelesen werden oder wie viel Zeit durchschnittlich für sprachfördernde Aktivitäten wie Vorlesen oder das Hören von Hörmedien aufgewendet wird.

Jedoch ist nicht nur die Häufigkeit des Sprachkontaktes wichtig, sondern insbesondere die Qualität des erhaltenen Inputs (Masek et al., 2021). So nimmt die Qualität des sprachlichen Inputs größeren Einfluss auf die Qualität der Sprache von Kindern als die Quantität des Inputs

(Richardson, 2023). Frühere Studien zur Qualität sprachlichen Inputs differenzierten primär zwischen einer interaktiven und einer linguistischen Dimension. Rowe und Snow (2020) erweiterten dieses Modell um eine dritte, konzeptionelle Dimension, die die kognitive Tiefe sprachlicher Interaktionen berücksichtigt. Die interaktive Dimension umfasst Merkmale wie dialogische Gesprächsführung, gemeinsame Diskussionen und die Berücksichtigung kindlicher Interessen, wodurch eine aktive Beteiligung des Kindes am Gespräch gefördert wird (Rowe & Snow, 2020). Die linguistische Dimension bezieht sich auf die sprachliche Komplexität des Inputs, einschließlich eines elaborierten und vielfältigen Wortschatzes sowie syntaktisch anspruchsvoller Satzstrukturen (Rowe & Snow, 2020). Die konzeptionelle Dimension beschreibt, in welchem Ausmaß Sprache dazu beiträgt, abstrakte Konzepte, kausale Zusammenhänge und kognitiv herausfordernde Inhalte zu vermitteln (Rowe & Snow, 2020). Sie hebt sich damit von rein kontextgebundenem oder oberflächlichem Sprachgebrauch ab und betont die Bedeutung von dekontextualisierter Sprache für die Förderung höherer Denkprozesse im Spracherwerb (Rowe & Snow, 2020). Auch MacLeod und Demers (2023) teilen den Sprach-Input in unterschiedliche Bereiche ein, bei denen es sich unter anderem um die an die Kinder gewandte Sprache, die Eltern-Kind-Interaktion und die Darstellung der Sprache handelt. Die an die Kinder gewandte Sprache umfasst die linguistische Komplexität und den Inhalt der Sprache, mit der die Eltern mit den Kindern sprechen. Die Eltern-Kind-Interaktion richtet sich an sprachliche Strategien, die von den Eltern in jeglichen Interaktionsprozessen genutzt werden. Die Darstellung der Sprache umfasst hierbei beispielsweise den bereichernden Sprachgebrauch im Zusammenhang mit dem Geschichtenerzählen, nicht nur in Bezug auf das Lesen, sondern auch das Erzählen oder das Hören von Liedern, in denen Geschichten erzählt werden.

Richardson (2023) betont die zentrale Bedeutung der lebensweltlichen Umgebung für die sprachliche Entwicklung von Kindern und hebt hervor, dass diese einen maßgeblichen Einfluss auf die Qualität des kindlichen Sprachinputs ausübt. In diesem Zusammenhang untersuchen auch Hoff et al. (2024) die Kontextabhängigkeit sprachlicher Interaktionen im Alltag. Ihre Studie basiert auf einer Stichprobe von 155 Eltern-Kind-Paaren im Alter von drei bis fünf Jahren und analysiert die sprachliche Qualität in drei alltagsrelevanten Interaktionskontexten: dem gemeinsamen Lesen, dem dialogischen Erinnern an vergangene Erlebnisse sowie dem Spielen mit Spielzeug (Hoff et al., 2024). Die Ergebnisse verdeutlichen, dass die Qualität des sprachlichen Inputs signifikant zwischen den untersuchten Kontexten variiert (Hoff et al., 2024). Insbesondere in den Situationen des gemeinsamen Lesens und des autobiografischen Erinnerns zeigen Eltern eine höhere sprachliche Elaboriertheit, während die sprachliche Qualität im Kontext des freien Spiels mit Spielzeug vergleichsweise geringer ausfällt (Hoff et

al., 2024). Diese Befunde unterstreichen die Relevanz kontextueller Faktoren für die sprachliche Anregung im frühen Kindesalter.

2.2.3 Sozioökonomischer Status

Der sozioökonomische Status zählt als ein zentraler Einflussfaktor auf die Bildung von Kindern (Becker, 2017). Der sozioökonomische Status bezeichnet ein mehrdimensionales Konzept zur Beschreibung der sozialen Lage von Individuen oder Familien innerhalb einer Gesellschaft. Dieser wird in unterschiedlichen Disziplinen wie der Bildungsforschung genutzt. Der sozioökonomische Status wird nicht ausschließlich aus einem Merkmal, sondern aus einer Vielzahl an Indikatoren gebildet (Ditton & Maaz, 2015). Darunter fallen beispielsweise der Schulabschluss, die Ausbildung oder auch das Einkommen der Erziehungsberechtigten, die den Alltag des Individuum bestimmen (Ditton & Maaz, 2015). Eine Vielzahl an Forschungsergebnissen zeigt, dass der sozioökonomische Status die Entwicklung von Kindern signifikant beeinflusst (Calvo & Bialystok, 2014). So erreichen laut Calvo und Bialystok (2014) Kinder aus Familien mit einem höheren sozioökonomischen Status bessere Ergebnisse bei kognitiven Messungen als Kinder aus Familien mit geringerem sozioökonomischen Status. Daraus resultiert, dass der sozioökonomische Status einen signifikanten Einfluss auf den Bildungserfolg von Kindern nimmt (Calvo & Bialystok, 2014).

Um den sozioökonomischen Status zu ermitteln, entwickelten Lampert et al. (2018) einen umfassenden Fragebogen, der die Bereiche Bildung, Beruf und Einkommen umfasst. Dieser wird in Kapitel 3.4.1 detailliert erläutert, da er die Grundlage des Elternfragebogens der vorliegenden Studie darstellt.

2.3 Diagnostische Verfahren zur Erfassung der Lesekompetenz

Zur Erhebung der Lesekompetenz stehen verschiedene diagnostische Verfahren zur Verfügung. Im Folgenden werden die standardisierten Verfahren des Stolperwörter-Lesetests, der internationalen Schulleistungsstudien PISA und IGLU sowie der länderübergreifenden Vergleichsarbeiten (VERA) dargestellt. Ergänzen wird ein psycholinguistisches Verfahren, das Eye-Tracking, erläutert.

Der Stolperwörter-Lesetest ist ein ökonomisches Gruppenverfahren zur Diagnose zentraler Teilkompetenzen des Lesens, das bereits ab der ersten Jahrgangsstufe eingesetzt werden kann (Backhaus, 2005). In der zweiten Klasse lesen die Schülerinnen und Schüler innerhalb von sechs Minuten 60 Sätze und markieren jeweils das semantisch oder syntaktisch inkongruente Wort (z. B. *„Mein Vater trinkt gerne schwarzen grünen Kaffee“* – *„grünen“* ist das Stolperwort) (Metze, 2024). Der Test erfasst Lesegenauigkeit, Lesegeschwindigkeit und Satzverständnis (Metze, 2024). Die Auswertung erfolgt anhand der Anzahl korrekt

identifizierter Stolperwörter pro Minute und wird mit altersbezogenen Normwerten verglichen, die eine differenzierte Einordnung der individuellen Leseleistung ermöglichen (Backhaus, 2005; Metze, 2024).

Die internationale Schulleistungsstudie PISA wird im dreijährigen Rhythmus durchgeführt und richtet sich an 15-jährige Schülerinnen und Schüler (Tarelli & Bos, 2013). Ziel ist die Erfassung grundlegender Kompetenzen, die für gesellschaftliche Teilhabe und lebenslanges Lernen erforderlich sind (Heine et al., 2023). Im Bereich Lesen werden drei zentrale Teilbereiche differenziert: Informationsentnahme, Interpretation und Bewertung von Texten. Die Schülerinnen und Schüler bearbeiten etwa 29 kurze Lesetexte, die in der Regel eine Seite nicht überschreiten. Zu jedem Text sind durchschnittlich fünf Aufgaben zu lösen (Jude et al., 2013). Die Aufgabenformate umfassen sowohl Multiple-Choice-Fragen als auch offene Antwortformate. Die Gesamtbearbeitungszeit beträgt rund zwei Stunden (Jude et al., 2013). Die Aufgaben sind im Rahmen eines „Multi-Matrix-Designs“ (Jude et al., 2013, S. 204) verteilt, sodass jede Testperson nur eine Teilmenge der Gesamtaufgaben bearbeitet. Die Auswertung erfolgt anhand standardisierter Bewertungskriterien. Offene Antworten werden nach festgelegten Kriterien beurteilt (Jude et al., 2013). Die Ergebnisse werden auf einer sechsstufigen Kompetenzskala dargestellt, die differenziert beschreibt, welche kognitiven Anforderungen Schülerinnen und Schüler in den jeweiligen Altersstufen bewältigen können (Jude et al., 2013).

Die internationale Grundschul-Lese-Untersuchung (IGLU/PIRLS) erfasst die Lesekompetenz von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe (Lorenz et al., 2023). Ziel ist ein international vergleichender Überblick über die Leistungsfähigkeit von Bildungssystemen sowie über strukturelle Bedingungen und Einflussfaktoren des Kompetenzerwerbs. Zur Erhebung werden standardisierte Lesetests eingesetzt, die sowohl informierende als auch erzählende Texte enthalten. Jeder Text ist kindgerecht gestaltet und wird durch zwölf bis 17 Aufgaben ergänzt (Lorenz et al., 2023). Die Aufgabenformate reichen von Multiple-Choice-Fragen bis zu offenen Antwortformaten. Die getesteten Verstehensprozesse umfassen das Auffinden expliziter Informationen, das Ziehen einfacher und komplexer Schlussfolgerungen sowie die Bewertung von Inhalt und Sprache (Lorenz et al., 2023). Die Auswertung basiert auf der „*Item Response*“ (Lorenz et al., 2023, S. 47) Theorie. Die Schwierigkeit der Aufgaben und die Fähigkeiten der Schülerinnen und Schüler werden auf einer gemeinsamen Skala abgebildet. Zur Schätzung individueller Kompetenzen wird der *Plausible Values*-Ansatz (Lorenz et al., 2023, S. 47) verwendet, bei dem für jede Person mehrere plausible Kompetenzwerte berechnet werden. Ergänzend werden durch Fragebögen Hintergrundinformationen zu individuellen, familiären, schulischen und gesellschaftlichen Bedingungen erhoben (Lorenz et al., 2023).

VERA wird in den Jahrgangsstufen drei und acht durchgeführt und dient der standardisierten Erfassung von Schülerkompetenzen (Groß Ophoff et al., 2008; Tarkian et al., 2019). Ziel ist die Überprüfung der Bildungsstandarderreichung sowie die Qualitätssicherung auf Klassen-, Schul- und Landesebene. Im Kompetenzbereich Lesen steht das Textverstehen im Mittelpunkt (Tarkian et al., 2019). Die Aufgaben prüfen das gezielte Entnehmen expliziter Informationen, das Herstellen lokaler und globaler inhaltlicher Zusammenhänge sowie die Reflexion über sprachliche Gestaltung und Textinhalte (Tarkian et al., 2019). Zur Erfassung dieser Kompetenzen kommen verschiedene Aufgabenformate zum Einsatz. Dazu zählen Multiple-Choice-Aufgaben, bei denen aus vorgegebenen Antwortmöglichkeiten die richtige ausgewählt werden muss, sowie Kurzantworten, die eine knappe, freie Formulierung erfordern (Tarkian et al., 2019). Ergänzt wird das Spektrum durch Zuordnungsaufgaben, bei denen Aussagen oder Begriffe passenden Textstellen zugeordnet werden, sowie durch Lückentexte, in denen fehlende Wörter oder Informationen ergänzt werden müssen (Tarkian et al., 2019). Die Bearbeitungszeit beträgt etwa 45 Minuten. Der Zeitrahmen ist so gestaltet, dass auch Schülerinnen und Schüler mit geringerer Leseflüssigkeit die Aufgaben bewältigen können (Tarkian et al., 2019). Zwar besteht keine strikte Zeitbegrenzung, jedoch wird empfohlen, den vorgegebenen Rahmen einzuhalten. Die Aufgaben basieren auf altersgerechten und thematisch vielfältigen Texten. Eingesetzt werden Sachtexte (z. B. Tierporträts, Erklärtexte), literarische Texte (z. B. Erzählungen, Märchen) und Gebrauchstexte (z. B. Fahrpläne, Anleitungen) (Groß Ophoff et al., 2008). Diese Vielfalt ermöglicht eine differenzierte Erfassung des Textverstehens und berücksichtigt die unterschiedlichen Anforderungen, die verschiedene Textarten an die Leserinnen und Leser stellen.

Zur Analyse kognitiver Prozesse beim Lesen wird in zahlreichen empirischen Studien die Methode des Eye-Trackings eingesetzt. Dieses Verfahren erlaubt eine detaillierte Erfassung der Augenbewegungen von Leserinnen und Lesern und bietet damit einen unmittelbaren Zugang zu Prozessen der Informationsverarbeitung während des Lesevorgangs (Hessels et al., 2024). Die Technologie basiert auf der kontinuierlichen Registrierung von Fixationen und Sakkaden, also jenen Bewegungsmustern, mit denen das Auge Textinhalte aufnimmt und verarbeitet (Holmqvist & Andersson, 2017). Durch die Auswertung dieser Blickbewegungen lassen sich Rückschlüsse auf verschiedene Teilaspekte der Lesekompetenz ziehen, insbesondere auf die Leseflüssigkeit. Eye-Tracking ermöglicht somit eine differenzierte Analyse individueller Leseprozesse, die über klassische Testverfahren hinausgeht.

Im Bereich der zweiten Klasse wird ausschließlich der Stolperwörter-Lesetest durchgeführt, während groß angelegte Studien zur Erhebung von Lesekompetenzen wie PISA, VERA oder IGLU in der Regel erst ab dem dritten Schuljahr einsetzen (Lorenz et al., 2023; McElvany et al., 2023; Tarkian et al., 2019). Auf Grundlage curricularer Vorgaben ist es dennoch möglich,

Aussagen über die erwarteten Lesefähigkeiten am Ende des zweiten Schuljahres zu treffen. Die expliziten curricularen Vorgaben zur Lesekompetenz am Ende des zweiten Schuljahres sind auf wenige, klar definierte Teilbereiche beschränkt. Demnach sollen Schülerinnen und Schüler in der Lage sein, häufig wiederkehrende Morpheme (etwa Präfixe wie *ver-*, *vor-* oder *nach-* sowie bekannte Stamm-Morpheme) und Silben auf einen Blick zu erkennen. Darüber hinaus wird erwartet, dass sie vertraute Wörter automatisiert identifizieren und einfache Texte durch die Erkennung einzelner Wörter und Sätze lesen können (Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen [MSB NRW], 2021). Weitere Kompetenzen im Bereich des Lesens, wie etwa die Herstellung globaler Kohärenz, die Nutzung von Vor- und Weltwissen sowie der Einsatz von Strategien zur Texterschließung und Vertiefung werden im Curriculum nicht spezifisch dem zweiten Schuljahr zugeordnet. Vielmehr sind diese Ziele über den gesamten Zeitraum der vier Grundschuljahre hinweg formuliert und stellen langfristige Entwicklungsaufgaben dar (MSB NRW, 2021).

Daher lässt sich festhalten, dass am Ende der zweiten Klasse vor allem basale technische Lesefähigkeiten wie das Dekodieren im Vordergrund stehen, während ein umfassendes verstehendes Lesen, wie es in späteren Schuljahren gefordert wird, erst im Verlauf der weiteren Grundschulzeit systematisch aufgebaut wird. Aktuell existieren keine standardisierten Testverfahren, die eine differenzierte Erhebung der Lesekompetenz speziell in dieser frühen Klassenstufe ermöglichen. Die gängigen diagnostischen Instrumente setzen meist erst in späteren Schuljahren an und fokussieren primär auf das Leseverständnis. In diesem Kontext eröffnet die Eye-Tracking-Methode ein innovatives diagnostisches Potenzial: Sie erlaubt eine objektive und feingliedrige Erfassung von Leseprozessen bereits bei Kindern im frühen Grundschulalter. Im folgenden Abschnitt wird die Methodik des Eye-Tracking-Verfahrens detailliert erläutert und hinsichtlich ihrer Anwendung im Kontext der Lesediagnostik bei Grundschulkindern näher betrachtet.

2.4 Eye-Tracking

„Die Eye-Tracking Technologie hat sich in der wissenschaftlichen Forschung zunehmend verbreitet und bietet einzigartige Einsichten in […] kognitive Prozesse“ (Übersetzung der Verfasserin, Hessels et al., 2024, S. 15). Hierbei handelt es sich um eine Methode zur Aufnahme von Augenbewegungen, um Aufschluss über kognitive Prozesse zu gewinnen (Deilen, 2022; Rakoczi, 2012). Es zählt zu den sogenannten Online-Methoden, die sich durch ihren Zugriff auf kognitive Vorgänge auszeichnen (Mertins, 2016). Im Gegensatz dazu erfassen Offline-Methoden lediglich die Ergebnisse kognitiver Prozesse in zeitlich verzögerter Form (Mertins, 2016). Sie liefern keine Echtzeitdaten, sondern basieren auf nachgelagerten Reaktionen wie etwa schriftlichen Fragebögen, Objektbenennungen oder

Entscheidungsaufgaben (Mertins, 2016). Darüber hinaus werden in der Forschung auch echte Online-Methoden unterschieden, die nicht nur kognitive, sondern auch neuronale Prozesse direkt und synchron erfassen können (Mertins, 2016). Verfahren wie das Elektroenzephalogramm (EEG) ermöglichen eine hochauflösende zeitliche Analyse der neuronalen Aktivität während der Sprachverarbeitung (Mertins, 2016).

Der Eye-Tracking-Methodik liegen zwei zentrale Annahmen zu Grunde, „die Unmittelbarkeitshypothese“ (Rakoczi, 2012, S. 87) sowie „die Auge-Geist-Hypothese“ (Rakoczi, 2012, S. 87). Die zuerst genannte Hypothese legt zu Grunde, dass während der Aufnahme von Inhalten bereits kognitive Auswertungsprozesse stattfinden (Rakoczi, 2012, S. 87). Die Auge-Geist-Hypothese zielt darauf ab, dass der Blickfokus ausschließlich in einer begrenzten Zeit auf Objekten liegt. Das heißt, besteht kein Interesse mehr oder ist das Objekt kognitiv ausgewertet, wandert der Blick weiter (Rakoczi, 2012). Diese beiden Hypothesen ermöglichen es, Aussagen mit Hilfe der Eye-Tracking-Messung über die Dauer der kognitiven Auswertung zu gewinnen.

2.4.1 Grundlegende Funktionsweise des Auges

Der Eye-Tracker ist in der Lage, die Blickbewegung eines Menschen aufzunehmen. Um seine Funktionsweise nachvollziehen zu können, wird zuerst die visuelle Wahrnehmung durch das Auge erläutert.

Beim Sehvorgang tritt Licht durch die Pupille, die das Bild umkehrt und auf die Retina, die hintere Seite des Augapfels, projiziert (Holmqvist & Andersson, 2017). Diese, bestehend aus sogenannten lichtempfindlichen Zapfen und Stäbchen, wandelt das Licht in elektrische Signale um und leitet diese in den visuellen Kortex weiter, wo diese Signale weiterverarbeitet werden (Holmqvist & Andersson, 2017). Ausschließlich ein kleiner Bereich der Netzhaut, die die höchste Sinneszelldichte aufweist, ermöglicht das scharfe Sehen, die sogenannte Fovea (Holmqvist & Andersson, 2017; Ilg & Thier, 2006). Aus diesem Grund muss beim Sehen das Auge so ausgerichtet werden, dass das Licht vom Graphem oder auch Wort genau auf die Fovea fällt, um dieses lesen und verarbeiten zu können. Ein Objekt oder auch Wort wird so lange von der Fovea fokussiert, bis der kognitive Erkennungsprozess abgeschlossen ist (Müller, 2013). Um die Blickbewegung zu steuern und auf bestimmte Punkte zu richten werden drei Muskeln benutzt, die für die horizontale, vertikale und rollende Bewegung der Augen verantwortlich sind (Holmqvist & Andersson, 2017). Der Bereich des scharfen Sehens der Fovea umfasst nach Holmqvist und Andersson (2017) circa 2°, wohingegen der parafoveale Bereich circa 5° umfasst, in dem die Sehschärfe abnimmt, Objekte dennoch unscharf erkannt werden können (Deilen, 2022). Dieser Bereich um die Fovea ist asymmetrisch, sodass sich die Lesespanne beziehungsweise Buchstabenidentitätsspanne beim Lesen unterscheidet.

Rechtseitig vom fixierten Bereich können bis zu 15 weitere Grapheme erkannt werden, wohingegen linksseitig nur circa vier Grapheme erkannt werden können (Zihl, 2006).

Für die Messung von Augenbewegungen sind die Parameter Sakkaden und Fixationen von Relevanz und werden im Folgenden erläutert. Das Lesen ist kein kontinuierlicher Aufnahmeprozess, sondern ein komplexer Vorgang der Informationsverarbeitung, der von ruckartigen Augenbewegungen begleitet wird (Radach et al., 2012). Diese ruckartigen Sprünge zwischen Textstellen beziehungsweise die Augenbewegung zwischen zwei Fixationen werden Sakkaden genannt (Deilen, 2022; Holmqvist & Andersson, 2017). Während einer Sakkade wird die Aufnahme an Informationen unterdrückt (Hessels et al., 2024). Diese stellt die schnellste Körperbewegung dar, die ein Mensch durchführen kann und umfasst zwischen 20 und 80 Millisekunden. Die durchschnittliche Sakkadenlänge bei geübten Leserinnen und Lesern beträgt sieben bis neun Buchstaben, schwächere Lesende weisen kürzere Sprünge zwischen den Fixationspunkten auf (Müller, 2013). Unterschieden wird zwischen progressiven und regressiven Sakkaden (Radach et al., 2012). Die progressiven Sakkaden beschreiben die Augenbewegung in die Leserichtung, das heißt vorwärts (Deilen, 2022). Regressive Sakkaden, die auch Regressionen genannt werden, sind als Sakkaden definiert, die sich entgegengesetzt der Leserichtung bewegen (Holmqvist & Andersson, 2017). Diese Regressionen können innerhalb eines Wortes oder auch zwischen Satzbestandteilen erfolgen und deuten somit entweder auf das Wort- oder Satzverstehen hin (Holmqvist & Andersson, 2017). Zwischen den einzelnen Sakkaden finden die Fixationen statt. Es handelt sich um einen kurzen Stillstand des Auges, in dem die Fovea einen spezifischen Punkt fokussiert (Logothetis, 2006). Während einer Fixation gehen kognitive Prozesse vonstatten, in denen die visuelle Informationsverarbeitung stattfindet (Deilen, 2022). Die gesamte Verweildauer auf einer Fixation wird Fixationsdauer (fixation duration) genannt. Dabei werden sowohl unmittelbar aufeinanderfolgende Fixationen als auch solche berücksichtigt, die im Rahmen von Regressionen erfolgen. Sie dient als Maß für die gesamte visuelle Aufmerksamkeit, die einem Wort bei der Textverarbeitung zuteil wird. Im Gegensatz dazu erfasst die First Fixation Duration (FFD) ausschließlich die Dauer der initialen Fixation auf das betreffende Wort (Deilen, 2022). Sie gilt als Indikator für die unmittelbare Verarbeitung beim erstmaligen Lesen. Die kognitive Verarbeitung während einer Fixation kann zwischen wenigen Millisekunden und über einer Sekunde vonstatten gehen (Holmqvist & Andersson, 2017; Negi & Mitra, 2020). Geübte Leserinnen und Leser weisen eine durchschnittliche Verweildauer von 200 bis 250 Millisekunden auf (Müller, 2013). In der zweiten Klasse zeigen Kinder eine durchschnittliche Fixationsdauer von 306 Millisekunden (Starr & Rayner, 2001, zitiert nach Müller, 2013).

Die Fixationsanzahl, -dauer, Sakkaden sowie Regressionen können Aufschluss über die kognitive Informationsverarbeitungsprozesse bieten (Holmqvist & Andersson, 2017). Somit können mit Hilfe der Messung der Blick- beziehungsweise Augenbewegung Aussagen über den Ablauf kognitiver Verstehens- und visueller Erkundungsprozesse getroffen werden (Müller, 2013).

2.4.2 Methodik des Eye-Trackings

Die Methodik des Eye-Tracking umfasst verschiedene Verfahren zur Erfassung von Augenbewegungen, die sich in Bezug auf „Genauigkeit, […] Auflösung, […] Dynamik und [die] Art der aufgezeichneten Augendaten“ (Rakoczi, 2012, S. 88) unterscheiden. Grundsätzlich lassen sich drei zentrale Ansätze unterscheiden: die „elektrookulografischen Messmethode“ (Duchowski, 2003, zitiert nach Rakoczi, 2012, S. 89), die Messung mittels Kontaktlinsen sowie die „reflexionsbasierte[n] Messtechniken“ (Duchowski, 2003, zitiert nach Rakoczi, 2012, S. 89). Im Rahmen dieser Studie wird der Fokus auf die reflexionsbasierten Messtechniken gelegt, da dieser Ansatz für die Durchführung dieser Untersuchung gewählt wurde. Bei dieser Messtechnik werden die Augen der Probandinnen und Probanden mithilfe von Infrarotlicht beleuchtet. Anhand von Reflexionspunkten lässt sich die Blickorientierung berechnen. Die genauen Reflexionsbereiche variieren je nach Hersteller des Eye-Tracking-Systems und umfassen zumeist die Pupille sowie die Lichtreflexion der äußeren Cornea (Hornhaut). Es wird Infrarotlicht verwendet, um die Lichtreflexion des natürlichen Lichtes nicht zu erfassen (Müller, 2013). Das Verfahren gilt als nicht-invasiv, da die Augenbewegung rein optisch von außen erfasst wird, ohne dass ein direkter physischer Kontakt mit den Augen oder dem Körper erforderlich ist.

Ein weiterer Unterschied zwischen den verschiedenen Systemen des Eye-Tracking besteht in der Intrusivität und Mobilität (Rakoczi, 2012). Die Intrusivität bezeichnet hierbei, inwieweit das Gerät die Bewegungsfreiheit oder das natürliche Verhalten der Testperson einschränkt (Müri et al., 2019). Intrusive Eye-Tracking-Verfahren fanden vor allem im 20. Jahrhundert Anwendung. Erst durch die Entwicklung einer nicht-intrusiven Kamera durch Judd im Jahr 1905 wurde die Erhebung von Blickdaten für Probandinnen und Probanden deutlich komfortabler (Müri et al., 2019). Über die Intrusivität hinaus unterscheiden sich die Systeme in ihrer Mobilität. Das heißt, inwieweit sie ein festes oder portables System darstellen (Rakoczi, 2012). Mobile Geräte sind tragbar und können den Standort wechseln, wohingegen stationäre Systeme an einen Ort gebunden sind (Rakoczi, 2012). In dieser Studie wird der sogenannte Remote Eye-Tracker genutzt, der ein nicht-intrusives sowie mobiles Verfahren ermöglicht (Rakoczi, 2012). Hierbei ist der Eye-Tracker entweder in den Computer eingebunden oder kann mit dem Computer verbunden werden (siehe Abbildung 4). Beide Möglichkeiten bieten

ein berührungsfreies Verfahren, bei dem sich der Kopf in einer bestimmten Distanz von circa 60 Zentimeter befinden sollte, damit Messungen erfolgreich durchgeführt werden können.

Auch die Qualität von Daten kann sich beim Eye-Tracking unterscheiden, da die Aufnahme der Augenbewegung von individuellen und technischen Merkmalen abhängt. Im Bereich der

Abbildung 4: verwendeter Remote Eye-Tracker (SMI RED250 mobile; SensoMotoric Instruments)

individuellen Merkmale der Testpersonen kann das Tragen von Brillen, die Augenfarbe oder auch aufgetragene Wimperntusche die Daten beeinflussen (Holmqvist & Andersson, 2017). Im technischen Bereich nehmen die zeitliche sowie räumliche Auflösung des Gerätes eine entscheidende Rolle ein (Müller, 2013). Die sogenannte Abtastrate (sampling rate) oder Abtastfrequenz (sampling frequency) gibt in Hertz die Anzahl an Messpunkten pro Sekunde an, die vom Gerät erfasst werden können (Holmqvist & Andersson, 2017; Müller, 2013). Eye-Tracker, die für die Forschung genutzt werden, weisen eine Abtastrate zwischen 25 und 2000 Hertz auf (Holmqvist & Andersson, 2017; Müller, 2013). Das Problem von Geräten mit niedrigen Abtastfrequenzen besteht darin, dass sie zeitlich sehr schnelle Blickbewegungen nicht erfassen können (Holmqvist & Andersson, 2017). Die Messgenauigkeit der räumlichen Auflösung wird maßgeblich von der Lichtreflexion- und Pupillenmessung beeinflusst (Müller, 2013).

2.4.3 Analyseverfahren der Augenbewegung

Die Analyse der Augenbewegungen kann sowohl mit Hilfe von Diagrammen oder Tabellen als auch mittels Grafiken visualisiert werden (Deilen, 2022). Zu den gängigsten Visualisierungen mittels Grafiken zählen die sogenannten „Heatmaps“ (Rakoczi, 2012, S. 90), „Gaze Plots“ (Rakoczi, 2012, S. 90) und „Gaze Replays“ (Deilen, 2022, S. 276). Hierbei kann

zwischen dynamischen und statischen Visualisierungen unterschieden werden (Rakoczi, 2012). Sowohl die Heatmaps als auch die Gaze Plots lassen sich der statischen Visualisierung zuordnen, wohingegen die Gaze Replays eine dynamische Darstellung zeigen (Deilen, 2022).

Die Heatmaps (siehe Abbildung 5a)) bilden mit Hilfe von Farben Informationen über die Verweildauer auf einzelnen Elementen ab (Rakoczi, 2012). Die Länge der Verweildauer wird über eine Farbskala angezeigt, wobei Blau geringe und Rot beziehungsweise Orange hohe Werte markieren. Das heißt, dass Positionen, die lange oder wiederholt betrachtet werden, eine höhere Farbintensität aufweisen (Deilen, 2022). Das Potential dieser Darstellungsweise besteht darin, Stellen des Textes oder auch Bildes zu ermitteln, denen eine hohe Aufmerksamkeit durch die Testperson zuteil wird (Deilen, 2022).

Die Gaze Plots (siehe Abbildung 5b)) hingegen bilden die Augenbewegung mittels Linien und Kreise ab (Deilen, 2022; Rakoczi, 2012). Hierbei werden also nicht nur die Fixationen und die Fixationsdauer, sondern auch die Sakkaden der Probandinnen und Probanden erfasst (Deilen, 2022; Rakoczi, 2012). Die Kreise zeigen hierbei die Fixationen und die Fixationsdauer an, da die Kreise mit zunehmender Betrachtung größer werden (Deilen, 2022; Rakoczi, 2012). Anhand von Linien werden die Sakkaden dargestellt und durch das Video können Regressionen im Text veranschaulicht werden (Deilen, 2022; Rakoczi, 2012).

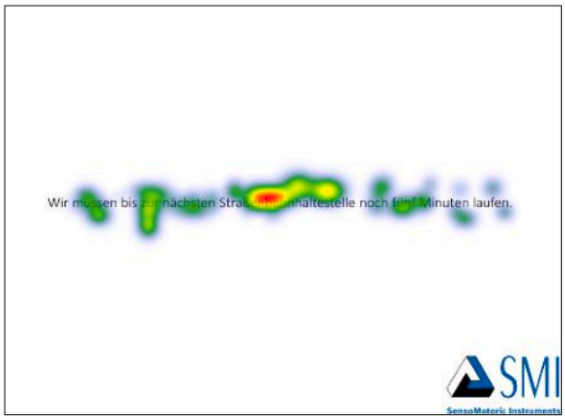

Abbildung 5a): Heat Map (Deilen, 2022, S. 277)

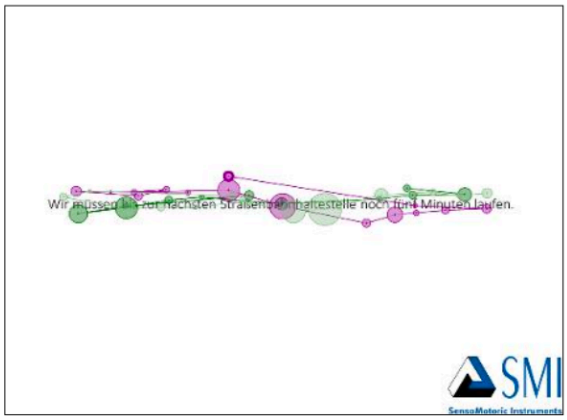

Abbildung 5b): Gaze Plots (Deilen, 2022, S. 277)

Im Gegensatz zu den Heatmaps und den Gaze Plots handelt es sich bei den Gaze Replays um dynamische Veranschaulichungen (Deilen, 2022). Hier wird zur Darstellung der Blickbewegung ein Film gezeigt. Die Blickbewegung wird anhand eines Punktes dargestellt, der über den Bildschirm wandert. Je nach Fixationsdauer nimmt auch hier der Punkt in seiner Größe zu (Deilen, 2022).

2.4.4 Forschungsstand in Bezug auf die Lesekompetenz

In der Literatur existieren zahlreiche Studien zur Lesekompetenz sowie zur Anwendung der Eye-Tracking-Methode in unterschiedlichen Kontexten. Dennoch bleibt der spezifische Einsatz des Eye-Trackings zur Analyse und Bewertung der Lesekompetenz, insbesondere bei Kindern, bislang ein vergleichsweise wenig beachtetes Feld (Radach et al., 2012; Solheim & Uppstad, 2011). Um der Zielsetzung dieser Arbeit gerecht zu werden, nämlich mittels Eye-Tracking Einblicke in die zugrunde liegenden Prozesse und Determinanten der Lesekompetenz zu gewinnen, konzentriert sich die Darstellung im Folgenden ausschließlich auf Studien, die diese Methodik im direkten Zusammenhang mit der Lesekompetenz einsetzten. Reine Studien zur Lesekompetenz oder zur Eye-Tracking-Methode ohne deren konkrete Verknüpfung bleiben unberücksichtigt. Im Folgenden wird eine Auswahl an aktuellen Studien dargestellt.

Zunächst werden die Studien von de la Peña (2024) und Mézière et al. (2023) beschrieben, die das Potenzial von Eye-Tracking in der Leseforschung hervorheben. Diese Studien zeigen auf, inwiefern sich die Eye-Tracking-Methode als alternative Herangehensweise zur Erfassung von Lesekompetenz gegenüber standardisierten Tests eignet. Mézière et al. (2023) stellten einen Vergleich zwischen drei verbreiteten Lesetests und dem Eye-Tracking an. Betrachtet wurden das „York Assessment of Reading for Comprehension“ (Mézière et al., 2023, S. 425), der „Gray Oral Reading Test“ (Mézière et al., 2023, S. 425) und der „Wide Range Achievement Test“ (Mézière et al., 2023, S. 425). Hierbei ergaben sich bei allen drei Tests unterschiedliche Ergebnisse über das Textverstehen der Probandinnen und Probanden. Mézière et al. (2023) stellten heraus, dass unterschiedliche Aufgabenformate Einfluss auf das Ergebnis von Untersuchungen der Lesekompetenz nehmen. Sie heben hervor, dass sich das Eye-Tracking dazu eignet, die kognitiven Prozesse direkt anhand der Augenbewegung von Schülerinnen und Schülern zu messen, im Gegensatz zu Lesetests, die je nach Aufgabenformat die Ergebnisse beeinflussen (Mézière et al., 2023). Auch de la Peña (2024) hebt in ihrer Untersuchung mit 67 Probandinnen und Probanden hervor, dass die Nutzung von Eye-Tracking-Systemen das Leseverständnis besser darstellen kann, als es beispielsweise Vokabel- und Verarbeitungstests tun (de-la-Peña, 2024). Gerade im Hinblick auf die Augenbewegungsmessung zeigt de la Peña (2024), dass das implizite Leseverständnis mit weniger Fixationen und kürzerer Fixationsdauer korreliert, was in standardisierten Tests nicht zu erfassen ist (de-la-Peña, 2024).

Eine essentielle Eye-Tracking-Studie im Bereich der Grundschule wurde bereits 2009 von Häikiö et al. durchgeführt. Anhand von 80 Schülerinnen und Schülern der zweiten, vierten und sechsten Klasse wurde untersucht, wie viele Grapheme in unterschiedlichen

Entwicklungsstadien identifiziert werden können und in welchem Maß die Buchstabenidentitätsspanne mit der Lesefähigkeit zusammenhängt (Häikiö et al., 2009). Diese Eye-Tracking Studie ergab, dass die Buchstabenidentitätsspanne mit steigendem Alter erhöht wird und von circa fünf Buchstaben in der zweiten Klasse auf circa neun Buchstaben wächst (Häikiö et al., 2009). Während der Grundschulzeit erreicht diese bei schnellen Leserinnen und Lesern bereits in der vierten Klasse beziehungsweise im Alter von zwölf Jahren das Niveau von Erwachsenen (Häikiö et al., 2009). Im Bereich der Lesefähigkeit ist zu erwähnen, dass schnellere Leserinnen und Leser ihren peripheren Sichtbereich effizienter nutzen. Dadurch weisen diese längere Sakkaden und weniger Fixationen auf als Leserinnen und Leser, die den peripheren Sichtbereich weniger effizient nutzen können (Häikiö et al., 2009).

De Simone et al. (2023) gehen über die Grapheme hinaus und untersuchten, welche Bedeutung Silben und Morpheme beim stillen Lesen haben. An der Studie nahmen 40 deutschsprachige Muttersprachlerinnen und Muttersprachler im Alter zwischen 20 und 54 Jahren teil. Ziel der Untersuchung war es, die Verarbeitung von Silben und Morphemen miteinander zu vergleichen. Die Ergebnisse zeigten, dass geübte deutsche Leserinnen und Leser hinsichtlich ihrer Lesezeit stärker durch die Silbenstruktur als durch die Morphemstruktur beeinflusst werden. Als mögliche Ursache nennen die Autorinnen und Autoren ein ausgeprägtes Silbenbewusstsein, das insbesondere im Anfangsunterricht, etwa durch das Silbenklatschen, gezielt gefördert wird (De Simone et al., 2023).

Einige aktuelle Studien untersuchen über die bereits erläuterten Aspekte hinaus Einflussfaktoren auf die Lesekompetenz (Krause & Ritter, 2022; Rettig & Schiefele, 2023; Wertli et al., 2023). Krause und Ritter (2022) betrachten den Einfluss des Sprachhintergrundes sowie des Alters auf die Lesekompetenz, wohingegen Rettig und Schiefele (2023) den Einfluss der Lesemotivation betrachten. Wertli et al. (2023) betrachten die Augenbewegung von Kindern im frühen Grundschulalter in Hinblick auf das Alter, das Geschlecht sowie den Sprachhintergrund.

Wertli et al. (2023) untersuchten die Augenbewegungen von 118 Kindern zwischen sieben und zwölf Jahren. Hier ist zu erwähnen, dass sich jedoch ausschließlich sechs dieser Kinder in der zweiten Klasse befunden haben (Wertli et al., 2023). Im Bereich des Alters stellte die Studie heraus, dass sich die Augenbewegung beim Lesen mit zunehmenden Alter deutlich ändert. So weisen jüngere Kinder, zu denen Rayner (1998) auch schwächere Leserinnen und Leser zählt, mehr Sakkaden, Fixationen und eine längere Fixationsdauer auf. Auch die Gesamtlesezeit umfasst einen größeren Zeitraum und es zeigen sich mehr Lesefehler auf (Wertli et al., 2023). Ältere Kinder und auch erfahrene Leserinnen und Leser weisen weniger Sakkaden und Fixationen auf. Zudem sei die Gesamtlesezeit sowie die Fehlerquote geringer (Rayner, 1998;

Wertli et al., 2023). In Anbetracht der Mehrsprachigkeit zeigt sich, dass in der ersten Klasse nicht muttersprachlich deutsche Kinder einen langsameren Leseprozess mit mehr Sakkaden und Fixationen aufweisen (Wertli et al., 2023). Dieser Unterschied ist jedoch in der fünften Klasse nicht mehr festzustellen (Wertli et al., 2023). Laut Wertli et al. (2023) nehmen sowohl das Geschlecht als auch die besuchte Schulform keinen signifikanten Einfluss auf die in Kapitel 2.4 erläuterten Parameter.

Die Stichprobe der Studie von Krause und Ritter (2022) umfasste insgesamt 83 monolinguale und bilinguale Schülerinnen und Schüler mit den Sprachhintergründen Russisch und Deutsch. Davon waren 25 Kinder russisch-deutsch bilingual, 27 Kinder monolingual deutschsprachig und 31 Kinder monolingual russischsprachig. Alle drei Gruppen wurden zusätzlich in zwei Alterskohorten unterteilt (neun bis zehn Jahre sowie 15–16 Jahre), sodass sich insgesamt sechs Untersuchungsgruppen ergaben. Die bilingualen Teilnehmenden kamen spätestens ab dem dritten Lebensjahr mit der russischen Sprache in Kontakt, während der Erwerb des Deutschen im Durchschnitt mit 1;6 Jahren in der jüngeren und mit 3;0 Jahren in der älteren Kohorte begann. Alle bilingualen Schülerinnen und Schüler verfügten über die Fähigkeit, sowohl in kyrillischer als auch in lateinischer Schrift zu lesen. Der Beginn der Lesesozialisation in deutscher Sprache erfolgte üblicherweise mit dem Schuleintritt im Alter von etwa sechs Jahren, wohingegen die russische Schriftsprache in der jüngeren Gruppe tendenziell etwas früher, in der älteren Gruppe hingegen deutlich später erworben wurde. Ein Großteil der bilingualen Kinder besuchte darüber hinaus außerschulische Angebote zum Erwerb und zur Festigung der russischen Sprach- und Schriftkompetenz. Die monolingual deutschsprachigen Schülerinnen und Schüler begannen das Lesen im Durchschnitt ebenfalls mit dem Schuleintritt (ca. sechs Jahre), wobei individuelle Unterschiede bestanden. Die monolingual russischsprachigen Teilnehmenden zeigten demgegenüber frühere Leseaktivitäten und erwarben Lesefähigkeiten im Alter von etwa vier bis fünf Jahren. In der Studie zeigten sich einige Unterschiede zwischen den monolingualen und bilingualen Kindern, zwischen lautem und leisem Lesen sowie in den beiden Altersgruppen. Wie bereits Studien zuvor schon festgestellt haben, nimmt das Alter einen Einfluss auf das Leseverhalten. Mit zunehmenden Alter nimmt die Anzahl der Fixationen sowie die Fixationsdauer ab. Mit Hinblick auf den Sprachhintergrund wird deutlich, dass bilinguale Lesende eine höhere Variation beim Leseverhalten aufzeigen als monolinguale Lesende. Im Bereich des lauten und leisen Lesens zeigen monolinguale Leserinnen und Leser größere individuelle Unterschieden, wohingegen bilinguale Leserinnen und Leser in beiden Lesemodi ähnliche Verhaltensmuster aufweisen. Anzumerken ist jedoch, dass bilingual russische Kinder beim Lesen auf Deutsch mehr Fixationen sowie eine signifikant höhere Fixationsdauer zeigten (Krause & Ritter, 2022).

Rettig und Schiefele (2023) untersuchten mit Hilfe des Eye-Trackings den Einfluss der Lesemotivation auf die Leseeffizienz anhand einer Langzeitstudie mit 131 Kinder. Es handelt sich um monolingual deutsche Kinder, die die Klassen eins bis drei besuchten (Rettig & Schiefele, 2023). An drei verschiedenen Zeitpunkten fand die Untersuchung statt, wobei der dritte Zeitpunkt alle zuvor durchgeführten Untersuchungen wiederholte (Rettig & Schiefele, 2023). Während dieser Studie wurden sowohl die intrinsische sowie extrinsische Motivation als auch die Lesemenge der Kinder untersucht. In der Studie konnten vier relevante Ergebnisse festgestellt werden. Ein positiver Zusammenhang stellt die intrinsische Motivation mit der Leseeffizienz dar (Rettig & Schiefele, 2023). Schülerinnen und Schüler mit hoher intrinsischer Motivation weisen sowohl kurz- als auch langfristig eine kürzere Fixationsdauer, eine geringe Gesamtlesezeit sowie weniger Regressionen im Text auf (Rettig & Schiefele, 2023). Zudem zeigt sich im zweiten Untersuchungsjahr, dass es einen wechselseitigen Effekt zwischen der Motivation und der Leseeffizienz gibt (Rettig & Schiefele, 2023). So steigt durch effizientes Lesen im zweiten Jahr die intrinsische Motivation. Bei der extrinsischen Motivation hingegen zeigt sich kein oder gar ein negativer Effekt auf die Leseeffizienz (Rettig & Schiefele, 2023). Die extrinsische Motivation hingegen leistet somit keinen Beitrag zur Entwicklung von Lesekompetenz, da Belohnungen eher ablenkend wirken (Rettig & Schiefele, 2023). Zudem weisen einige extrinsisch motivierte Kinder eine längere Fixationsdauer auf (Rettig & Schiefele, 2023). Im Hinblick auf die Lesemenge der Lernenden zeigt sich keine Wirkung auf die Leseeffizienz, was jedoch laut der Autorinnen und Autoren daran liegen könnte, dass sich die Lesepraxis zwischen den einzelnen Kindern nicht signifikant unterschied (Rettig & Schiefele, 2023).

Hautala et al. (2024) untersuchten, wie sich die Prozesse der Worterkennung bei 81 Grundschulkindern (dritte und vierte Klasse) im Verlauf eines halben Jahres entwickelten. Besonders wurde die zunehmende Leseflüssigkeit betrachtet. Mithilfe von Eye-Tracking-Daten beim Lesen von Texten analysierten sie, wie sich die Erstfixationen, Refixationsdauern und Fixationsanzahlen verändern (Hautala et al., 2024). Darüber hinaus betrachteten sie die Effekte von Wortfrequenz und Wortlänge, um Rückschlüsse auf die Entwicklung orthografischer Verarbeitung und Dekodierfähigkeit zu ziehen (Hautala et al., 2024). Ziel war es zu erkennen, welche spezifischen Veränderungen in den Augenbewegungen die Entwicklung der Leseflüssigkeit vorhersagen und wie effizient Kinder Buchstabenfolgen im Arbeitsgedächtnis verarbeiten (Hautala et al., 2024). Die Ergebnisse stützen sich dabei auf

theoretische Ansätze wie das Dual-Route-[5] und das Dual-Stage-Modell[6] der Worterkennung. Die Forscherinnen und Forscher zeigten, dass die Verkürzung der Erstfixations- und Refixationsdauern eng miteinander verknüpft sind und den größten Einfluss auf den Lesefortschritt haben (Hautala et al., 2024). Überraschenderweise hemmte die Reduktion der Fixationsdauer sogar die gleichzeitige Abnahme der Fixationsanzahl. Nur Kinder, die beide Parameter reduzieren konnten, entwickelten sich deutlich weiter in ihrer Leseflüssigkeit. Darüber hinaus fand die Studie Hinweise darauf, dass diese Entwicklung durch eine zunehmende Effizienz in der Verarbeitung von Buchstabenfolgen im Arbeitsgedächtnis ermöglicht wird. Einzelne Schülerinnen und Schüler zeigten eine besonders fortgeschrittene Fähigkeit, Wörter mit weniger und kürzeren Fixationen zu erfassen, was auf einen höheren Grad an orthografischer Verarbeitungskompetenz hindeutet (Hautala et al., 2024).

Auch im Kontext von Legasthenie (Dyslexie) und damit verbundenen Leseschwierigkeiten hat sich die Methodik des Eye-Trackings als vielversprechend erwiesen. Smynarski et al. (2017) untersuchten in einer Stichprobe von 69 Kindern im Alter von 8;6 bis 12;6 Jahren (dritte bis sechste Schulstufe) die Eignung charakteristischer Augenbewegungsmuster für ein Screeningverfahren zur Identifikation von Legasthenie. Dabei wurden 32 klinisch diagnostizierte Kinder mit Legasthenie einer Kontrollgruppe von 37 Kindern ohne Leseschwierigkeiten gegenübergestellt, deren unauffälliges Leseverhalten durch eine sonderpädagogische Fachkraft bestätigt worden war (Smyrnakis et al., 2017). Als methodische Grundlage diente die RADAR-Methode (Rapid Assessment of Difficulties and Abnormalities in Reading), die während des stillen Lesens eines standardisierten, altersgerechten Textes Eye-Tracking-Daten erhebt und auswertet. Die RADAR-Methode bewertet zweimal differenzierte Parametergruppen. Die erste Gruppe stellt die nicht wortspezifischen Kennwerte wie die Fixationsdauer, Sakkadenlänge, kurze Refixationen und Gesamtzahl der Fixationen dar. Als weitere Gruppe werden die wortspezifischen Kennwerte betrachtet, die sich aus der Blickdauer beim Erstbesuch eines Wortes und der Anzahl der Wiederholungsbesuche auf dasselbe Wort zusammensetzen (Smyrnakis et al., 2017). Aus diesen Messgrößen wird eine multivariate Punktzahl berechnet, anhand derer Probandinnen und Probanden als wahrscheinlich dyslektisch oder vermutlich nicht dyslektisch klassifiziert

---

[5] Nach dem Dual-Route-Modell läuft die Worterkennung parallel über zwei Wege ab. Der indirekte Weg wandelt Buchstaben systematisch in Laute um und erschließt so auch unbekannte Wortfolgen. Der direkte Weg aktiviert anhand vertrauter Buchstabenmuster sofort orthografische Einträge im mentalen Lexikon, ohne Lautumwandlung (Coltheart et al., 2001).

[6] Das Dual-Stage-Modell geht von zwei aufeinanderfolgenden Phasen aus: einer frühen orthografischen Aktivierung und einer anschließenden Graphem-Phonem-Konversion. Diese zeitliche Abfolge stützt ein gestuftes Zusammenspiel lexikalischer und nicht-lexikalischer Prozesse gegenüber dem klassischen parallelen Dual-Route-Ansatz (Hautala et al., 2024).

werden (Smyrnakis et al., 2017). Im Versuchsaufbau lasen die Kinder den Text von einem Bildschirm und beantworteten im Anschluss fünf Fragen zum Textverständnis, die jedoch ausschließlich der Kontrolle dienten, ob der Text wirklich erfasst wurde (Smyrnakis et al., 2017). Die Auswertung ergab, dass Kinder mit Legasthenie im Vergleich zur Kontrollgruppe deutlich längere Fixationen, kürzere Sakkaden, eine höhere Anzahl von Rücksprungfixationen und insgesamt mehr Fixationen aufweisen (Smyrnakis et al., 2017). Diese Befunde zeigten sich sowohl bei griechischsprachigen als auch bei englischsprachigen Probandinnen und Probanden und bestätigen die sprachübergreifende Validität der RADAR-Methode (Smyrnakis et al., 2017). Zugleich wird deutlich, dass RADAR als automatisiertes Erstscreening konzipiert ist und nicht alle relevanten kognitiven Aspekte einer vollständigen Dyslexie-Diagnostik erfasst (Smyrnakis et al., 2017). Wichtige Parameter wie Arbeitsgedächtnis, Genauigkeit der Wortdekodierung, Rechtschreibung und weitere sprachbezogene kognitive Fähigkeiten bleiben unberücksichtigt. Deshalb dient RADAR primär der effizienten Vorauswahl von Kindern, die für eine umfassendere, zeit- und fachintensive Diagnostik infrage kommen (Smyrnakis et al., 2017).

Die bisherige Forschung zeigt, dass die Eye-Tracking-Methode bereits in zahlreichen Studien zur Untersuchung von Leseprozessen erfolgreich eingesetzt wurde. Auffällig ist jedoch, dass entsprechende Untersuchungen bislang kaum in den ersten beiden Grundschulklassen durchgeführt werden, obwohl gerade diese frühen Phasen entscheidend für die Entwicklung zentraler Leseprozesse sind. Ein weiterer kritischer Punkt besteht darin, dass Eye-Tracking-Studien häufig primär die visuellen Parameter des Lesens analysieren, während das Textverstehen als kognitive Zielgröße oft unberücksichtigt bleibt. Dennoch zeigt sich, dass Eye-Tracking ein hochwirksames Verfahren darstellt, um text- und aufgabenunabhängige Einblicke in die zugrundeliegenden kognitiven Mechanismen des Lesens zu gewinnen. So kann das Eye-Tracking dazu genutzt werden, Aussagen über die Lesefähigkeit zu treffen und Leseschwierigkeiten oder auch Anzeichen für Legasthenie zu ermitteln.

Der aktuelle Forschungsstand zeigt die Wirksamkeit von Eye Tracking als Methode zur Identifikation und Quantifizierung zentraler Leseprozesse, indem diese Einblicke in Worterkennung und Leseflüssigkeit gewährt. In den meisten Untersuchungen stehen visuelle Kennwerte wie Fixationsdauer, Sakkadenlänge und Regressionshäufigkeit im Vordergrund. Studien wie Smynarski et al. (2017) zeigen jedoch, dass sich Lesende mit Wortverarbeitungsproblemen bereits in der Grundschule anhand dieser Parameter zuverlässig von typischen Lesenden abgrenzen lassen. Screeningverfahren wie die RADAR-Methode erreichen dabei Sensitivitäten und Spezifitäten von über 90 % und ermöglichen ein automatisiertes Erstscreening ohne verbale oder schriftliche Rückmeldungen. Neuere Befunde differenzieren zudem die Mechanismen, die der Leseflüssigkeit zugrunde liegen.

Hautala et al. (2024) identifizierten die simultane Verkürzung von Erst- und Refixationsdauer als entscheidenden Prädiktor für Lesefortschritt, vermittelt durch effizientere arbeitsgedächtnisbasierte Verarbeitung von Buchstabenketten und gestärkte orthografische Kompetenz. Rettig & Schiefele (2023) zeigen, dass intrinsische Motivation mit geringeren Fixationsdauern, geringerem Regressionsaufkommen und verkürzter Gesamtlesezeit korreliert, während extrinsische Anreize diese Effekte nicht fördern oder sogar hemmen. Entwicklungs- und Sprachhintergrund modifizieren das Leseprofil zusätzlich: Jüngere und nicht-muttersprachliche Kinder weisen höhere Fixations- und Regressionsraten auf, ein Unterschied, der bis zur fünften Klasse sinkt (Rayner, 1998; Wertli et al., 2023). Fortgeschrittene Leserinnen und Leser nutzen ihr peripheres Sichtfeld effektiv (Häikiö et al., 2009) und deutsche Muttersprachler fokussieren stärker auf Silbenstruktur als auf Morpheme (De Simone et al., 2023). Mézière et al. (2023) und de la Peña (2024) unterstreichen schließlich, dass Eye Tracking Leseprozesse abbildet, die standardisierte Lesetests nur unzureichend erfassen.

Trotz dieser Fortschritte bleiben relevante Einflussgrößen wie der rezeptive Wortschatz, der sozioökonomische Status, die Quantität des sprachlichen Inputs und insbesondere das tatsächliche Textverstehen in aktuellen Eye-Tracking-Studien unberücksichtigt. Darüber hinaus ist die Evidenzlage für die schulische Anfangsphase (Klassen eins und zwei) bislang unzureichend, obwohl gerade hier die Basis für alle nachfolgenden Leseleistungen gelegt wird. Für eine aussagekräftige Diagnostik und individualisierte Förderplanung sollte Eye Tracking daher als integraler Bestandteil eines multimodalen Ansatzes verstanden werden. Künftige Forschungen sollten die methodische Bandbreite erweitern und neben Parametern der Leseflüssigkeit systematisch semantische Verstehensleistungen und sprachliche Voraussetzungen in den Blick nehmen sowie spezifisch jene Kinder untersuchen, die sich im frühen Leseerwerb befinden.

3 Empirischer Teil

Experimentelle Methoden finden zunehmend Anwendung in der psycholinguistischen Forschung (Mertins, 2016). Auch bei der vorliegenden Masterarbeit handelt es sich um eine empirische Untersuchung, da sie auf der systematischen Sammlung und Auswertung von Daten basiert, um eine wissenschaftliche Fragestellung zu beantworten (Albert & Marx, 2016). Ziel ist es, auf Grundlage dieser experimentellen Untersuchung wissenschaftlich fundierte Aussagen über den Einfluss unterschiedlicher Faktoren auf die Lesekompetenz zu gewinnen.

Die Untersuchung folgt einem quantitativen Forschungsansatz, indem die aufgenommenen Daten statistisch ausgewertet und interpretiert werden (Rothstein et al., 2022). Um vergleichbare und objektive Ergebnisse zu gewinnen, werden standardisierte Messinstrumente wie ein standardisierter Wortschatztest sowie das Eye-Tracking verwendet (Rothstein et al., 2022). Die statistische Auswertung soll dazu beitragen, Zusammenhänge zwischen den untersuchten Variablen zu identifizieren und die aufgestellten Hypothesen zu überprüfen (Rothstein et al., 2022). Zudem weist die Studie experimentelle Elemente auf, da mit Hilfe des Eye-Trackings die Augenbewegungen unter kontrollierten Bedingungen gezielt erfasst werden (Albert & Marx, 2016). Die identische Testsituation ermöglicht eine gezielte Untersuchung der Lesekompetenz, um im Rahmen der experimentellen Untersuchung kausale Rückschlüsse auf die zentralen Einflussfaktoren der Lesekompetenz zu ziehen (Albert & Marx, 2016; Rothstein et al., 2022).

3.1 Fragestellung und Hypothesen

Die Lesekompetenz ist ein vielschichtiges Konstrukt, das über das reine Dekodieren von Wörtern hinausgeht. Sie umfasst sowohl die Fähigkeit zur schnellen und fehlerfreien Worterkennung als auch das tiefgehende Verstehen und Interpretieren von Textinhalten (Rosebrock & Nix, 2025). Lenhard (2024) betont, dass automatisierte Dekodierprozesse das Arbeitsgedächtnis entlasten und so den Weg für komplexe Verständnisleistungen ebnen. Rosebrock und Nix (2025) beschreiben Lesekompetenz als hierarchisches Zusammenspiel von Leseflüssigkeit, lokaler sowie globaler Kohärenzbildung, dabei bilden automatisierte Worterkennung und prosodisches Sequenzieren die Basis für höhere inferenzielle Prozesse. McElvany und Schneider (2009) heben die wechselseitige Beziehung zwischen Wortschatzentwicklung und Textverstehen hervor: Ein sicherer Zugriff auf das mentale Lexikon ermöglicht nicht nur schnellere Leseprozesse, sondern fördert auch das Erschließen unbekannter Wörter aus dem Kontext. Lenhard und Artelt (2009) zeigen, dass durch diese Kontexteinbindung der Wortschatz weiter wächst und langfristig zur Verbesserung der Lesekompetenz beiträgt. Familiäre Einflussgrößen wirken sich indirekt über das kulturelle und soziale Kapital des Elternhauses aus, indem das phonologische Bewusstsein, der rezeptive

Wortschatz und das Vorwissen gefördert wird (Jansen & Marx, 1999, zitiert nach Schaffner, 2009). Hurrelmann (2004, zitiert nach Schaffner, 2009) zeigt, dass Kinder aus sozial schwächeren Familien seltener frühkindlich sprachlich-kognitiv gefördert werden.

Vor diesem Hintergrund hat die Forschung in den letzten Jahren verstärkt Eye-Tracking-Methoden eingesetzt, um kognitive Prozesse beim Lesen zu untersuchen (u.a. De Simone et al., 2023; de-la-Peña, 2024; Häikiö et al., 2009; Hautala et al., 2024; Krause & Ritter, 2022; Mézière et al., 2023; Rettig & Schiefele, 2023; Wertli et al., 2023). Diese Studien liefern wertvolle Erkenntnisse über okulomotorische Muster wie Fixationsdauer und -anzahl. Allerdings konzentrieren sich viele dieser Untersuchungen auf ältere Altersgruppen und vernachlässigen die Schuleingangsphase. Lediglich Wertli et al. (2023) beziehen eine kleine Stichprobe von Zweitklässlerinnen und Zweitklässlern ein. Weitere Einflussfaktoren wie Motivation, Alter und Buchstabenidentitätsspanne wurden zwar untersucht (Krause & Ritter, 2022; Rettig & Schiefele, 2023; Häikiö et al., 2009), doch bleibt die Verbindung dieser Variablen zur Lesekompetenz bei jüngeren Kindern unzureichend erforscht.

Die bestehenden Eye-Tracking-Studien liefern wertvolle methodische Erkenntnisse, decken jedoch nicht das gesamte Spektrum theoretisch wichtiger Einflussfaktoren ab. Vor allem fehlt eine ganzheitliche Untersuchung der Lesekompetenz von Zweitklässlerinnen und Zweitklässlern, die automatisierte Leseprozesse, Wortschatzeffekte, Textverstehen und soziale Rahmenbedingungen gleichzeitig berücksichtigt.

Ausgehend von den hier skizzierten theoretischen Modellen (Lenhard, 2024; McElvany & Schneider, 2009; Rosebrock & Nix, 2025; Schaffner, 2009) und der identifizierten Forschungslücke ergibt sich die Notwendigkeit, die Lesekompetenz von Kindern am Ende der zweiten Klasse mithilfe von Eye-Tracking-Daten zu erfassen und deren Zusammenhang mit Verstehensleistungen zu analysieren. Mit Blick auf die vorgestellten theoretischen Konzepte und den identifizierten Forschungsbedarf richtet sich die vorliegende Studie gezielt auf die bislang wenig untersuchte Altersgruppe. Anhand der Eye-Tracking-Methode sollen feingliedrige okulomotorische Kennwerte wie Lesedauer, Fixationsanzahl und Fixationsdauer systematisch mit den erzielten Verstehensleistungen verknüpft werden. Darüber hinaus werden auch kognitive und soziale Einflussgrößen wie der rezeptive Wortschatz sowie die Quantität des sprachlichen Inputs in die Analyse einbezogen. Auf dieser Grundlage werden im Anschluss differenzierte Forschungsfragen und Hypothesen entwickelt, die sowohl die beobachtbaren Augenbewegungsmuster als auch die inhaltliche Textverarbeitung gleichermaßen in den Blick nehmen. Die übergeordnete Forschungsfrage lautet: Welche Faktoren nehmen Einfluss auf die Lesekompetenz von Schülerinnen und Schülern am Ende

der zweiten Klasse? Im Einzelnen werden folgende detaillierte Fragestellungen betrachtet, um die aufgestellten Hypothesen zu bestätigen oder zu falsifizieren:

Forschungsfrage 1: Welchen Einfluss hat die Gesamtlesedauer eines Textes auf das Textverstehen der Schülerinnen und Schüler am Ende des zweiten Schuljahres?

> *Hypothese 1:* Je kürzer die Gesamtlesedauer eines Textes ist, desto höher ist die Wahrscheinlichkeit für mehr korrekte Antworten zum Textinhalt.

Forschungsfrage 2: Welchen Einfluss hat die Quantität des sprachlichen Inputs auf die Leseflüssigkeit und das Textverstehen der Schülerinnen und Schüler am Ende des zweiten Schuljahres?

> *Hypothese 2a:* Je höher die Quantität des sprachlichen Inputs ist, desto höher ist die Wahrscheinlichkeit für mehr korrekte Antworten zum Textinhalt.

> *Hypothese 2b:* Je höher die Quantität des sprachlichen Inputs ist, desto geringer ist die Lesedauer des ersten Lesedurchganges.

Forschungsfrage 3: Welchen Einfluss hat der rezeptive Wortschatz auf die Leseflüssigkeit und das Textverstehen der Schülerinnen und Schüler am Ende des zweiten Schuljahres?

> *Hypothese 3a:* Je größer der rezeptive Wortschatz ist, desto höher ist die Wahrscheinlichkeit für mehr korrekte Antworten zum Textinhalt.

> *Hypothese 3b:* Je größer der rezeptive Wortschatz ist, desto geringer ist die Lesedauer des ersten Lesedurchganges.

Forschungsfrage 4: Welchen Einfluss nimmt die Schwierigkeit des Textes auf die Leseflüssigkeit und das Textverstehen der Schülerinnen und Schüler am Ende des zweiten Schuljahres?

> *Hypothese 4a:* Je einfacher der Text ist, desto höher ist die Wahrscheinlichkeit für mehr korrekte Antworten zum Textinhalt.

> *Hypothese 4b:* Je einfacher der Text ist, desto geringer ist die Lesedauer des ersten Lesedurchganges.

Forschungsfrage 5: Welchen Einfluss nimmt die durchschnittliche Anzahl an Fixationen auf die Leseflüssigkeit und das Textverstehen der Schülerinnen und Schüler am Ende des zweiten Schuljahres?

> *Hypothese 5a:* Je geringer der Average Fixation Count pro Wort ist, desto höher ist die Wahrscheinlichkeit für mehr korrekte Antworten zum Textinhalt.

Hypothese 5b: Je geringer der Average Fixation Count pro Wort ist, desto geringer ist die Lesedauer des ersten Lesedurchganges.

Forschungsfrage 6: Welchen Einfluss hat die durchschnittliche Fixationsdauer auf das Textverstehen der Schülerinnen und Schüler am Ende des zweiten Schuljahres?

Hypothese 6: Je geringer die Average Fixation Duration ist, desto höher ist die Wahrscheinlichkeit für mehr korrekte Antworten zum Textinhalt.

Forschungsfrage 7: Welchen Einfluss hat die Anzahl an Lesedurchgängen auf das Textverstehen der Schülerinnen und Schüler am Ende des zweiten Schuljahres?

Hypothese 7: Je häufiger die Kinder den Text lesen, desto höher ist die Wahrscheinlichkeit für mehr korrekte Antworten zum Textinhalt.

3.2 Operationalisierung

Um die im theoretischen Teil dargestellten Variablen im Rahmen der empirischen Untersuchung messbar zu machen, ist eine präzise Operationalisierung erforderlich. Im Folgenden wird daher erläutert, wie die zentralen Variablen definiert und in erfassbare Indikatoren überführt wurden.

Die Lesekompetenz besteht aus einer Vielzahl von Teilkompetenzen (siehe Kapitel 2.1). Im Rahmen dieser Studie werden zwei Bereiche der Lesekompetenz betrachtet, bestehend aus der Leseflüssigkeit und dem Textverstehen. Diese Teilbereiche wurden ausgewählt, da sie quantifizierbar sind und eine differenzierte Messung der Lesekompetenz ermöglichen. Sie werden unabhängig voneinander operationalisiert und ausgewertet. Die getrennte Betrachtung ist notwendig, da nicht zwangsläufig davon ausgegangen werden kann, dass eine hohe Leseflüssigkeit automatisch mit einem guten Textverstehen einhergeht. Eine hohe Leseflüssigkeit muss nicht zwingend auf eine korrekte Informationsverarbeitung und das Verständnis dieser hindeuten. Da Sprache stets kontextbezogen verarbeitet wird, erfolgt die Betrachtung der Leseflüssigkeit und des Textverstehens nicht wortbasiert, sondern bezogen auf zusammenhängende Texte. Das Textverstehen wird anhand der korrekten Beantwortung von Verständnisfragen operationalisiert. Je höher die Anzahl der korrekt beantworteten Fragen zu einem Text ist, desto besser wird das Textverstehen eingeschätzt. Diese Methode erlaubt eine direkte Messung des inhaltlichen Erfassens des Gelesenen und dient somit als Indikator für das Textverstehen. Die Leseflüssigkeit wird über den Parameter der Lesezeit beim erstmaligen Lesen des Textes ermittelt. Je kürzer diese Zeit ist, desto höher wird die Leseflüssigkeit eingeschätzt. Grundlage bilden hierzu die Studien von Rayner (1998), Rettig und Schiefele (2023) sowie Hautala et al. (2024). Hautala et al. (2024) untersuchten die Leseflüssigkeit, wobei deren Indikator zur Messung der Leseflüssigkeit die Gesamtlesedauer

betrachtet (Hautala et al., 2024). Da die Kinder in dieser Studie mehrfach lesen durften, wird hier ausschließlich die Lesezeit des ersten Durchganges beachtet. Die Leseflüssigkeit lässt sich nicht in klar abgrenzbare Kompetenzstufen einteilen. Die Übergänge zwischen niedrigem und hohem Niveau sind fließend, sodass eine Einteilung in definierte Bereiche nicht sinnvoll erscheint.

Im Gegensatz dazu lässt sich der rezeptive Wortschatz von Schülerinnen und Schüleren in konkrete Kompetenzbereiche einteilen. Dieser wird mit Hilfe des PPVT erhoben, einem standardisierten und normierten Verfahren zur Erfassung des rezeptiven Wortschatzes (Lenhard et al., 2015). Zur altersunabhängigen Vergleichbarkeit wird der sogenannte T-Wert[7] anstelle des Rohwertes verwendet. Die Kinder werden gemäß den folgenden Normbereichen klassifiziert: *stark unterdurchschnittlich, unterdurchschnittlich, unterer Normalbereich, oberer Normalbereich, überdurchschnittlich, stark überdurchschnittlich* (Lenhard et al., 2015).

Die Grundlage zur Erhebung der Mehrsprachigkeit, des sprachlichen Inputs sowie des sozioökonomischen Status stellt ein Fragebogen dar, der in Kapitel 3.4.1 erläutert wird. Anhand einer siebenstufigen Lickert-Skala können dem sprachlichen Input sowie dem Grad der Mehrsprachigkeit eindeutige Werte zugeordnet werden. Die Operationalisierung der Eye-Tracking-Variablen wird in Kapitel 3.4.3 erläutert. Anhand welcher Daten die Eye-Tracking-Variablen gemessen werden, ist der Tabelle 1 zu entnehmen.

Konstrukt	**Operationalisierung**
Sozioökonomischer Status	Die Messung erfolgt über einen Elternfragebogen anhand dessen eine Einordnung mit Punkten (3-21) gemäß Lampert et al. (2018) ermöglicht wird.
Grad des Bilingualismus	Die Messung erfolgt über einen Elternfragebogen in Hinblick auf den sprachlichen Kontakt sowie die sprachlichen Fähigkeiten.
Sprachlicher Input	Die Messung erfolgt über einen Elternfragebogen bezüglich der wöchentlichen Quantität des sprachlichen Inputs im Kontext verschiedener Aktivitäten.
Rezeptiver Wortschatz	Anhand des T-Wertes gemäß des PPVTs kann eine Einteilung in sechs Leistungsbereiche erfolgen.
Schwierigkeitsgrad	Die Texte unterscheiden sich in zwei Schwierigkeitsniveaus: leicht und schwer.

[7] T-Werte stellen standardisierte Normwerte dar, die aus Rohwerten abgeleitet werden, um individuelle Leistungen im Verhältnis zu einer Vergleichsgruppe einordnen zu können. Grundlage bildet die Annahme einer Normalverteilung der Testergebnisse (Lenhard et al., 2015).

Leseflüssigkeit	Die Messung erfolgt anhand der Lesegeschwindigkeit (in Sekunden) beim erstmaligen Lesen des Textes (Reading Time First Trial).
Textverstehen	Die Messung erfolgt anhand der Anzahl korrekt beantworteter Fragen zu jedem Text.
Eye-Tracking-Daten	
Gesamtlesedauer	Die Messung erfolgt anhand der Lesedauer (in Sekunden), die die Kinder für einen Text benötigt haben.
Average Fixation Count pro Word	Die Messung erfolgt anhand der Gesamtzahl der Fixationen im Text geteilt durch Anzahl der Wörter im Text.
Average Fixation Duration	Die Messung erfolgt über die Summe der Dauer aller Fixationen in einem Text (in Millisekunden) geteilt durch Anzahl der Fixationen im Text.

Tabelle 1: Operationalisierung der Zielvariablen und Einflussfaktoren

3.3 Stichprobe

Die Stichprobe dieser Studie setzte sich aus 42 Schülerinnen und Schülern der zweiten Klasse zusammen. Alle Probandinnen und Probanden besuchten die katholische Grundschule an der Bergstraße in Herne, bei der es sich um einen lokalen Kooperationspartner des *BeeMEHR-Projektes*[8] handelt. Die Stichprobe teilte sich gleichmäßig auf beide zweiten Klassen der Schule auf, sodass 21 Kinder der Klasse 2a und weitere 21 Kinder der Klasse 2b an der Studie teilnahmen. Trotz Teilnahme an der Studie konnten für die Auswertung nur 29 Kinder herangezogen werden, da zehn Elternfragebögen nicht eingereicht wurden und drei Kinder krankheitsbedingt an den Tagen der Testung abwesend waren. Da der Fragebogen Informationen über die Mehrsprachigkeit, den sozioökonomischen Status sowie den sprachlichen Input enthält, konnten Kinder mit fehlenden Elternfragebögen nicht in die Auswertung dieser Studie einbezogen werden.

Die für die Stichprobe relevanten 29 Kinder wiesen ein Alter zwischen 7;6 Jahren und 8;7 Jahren auf. Darüber hinaus zeigt sich eine klare Verteilung der Geschlechter in der Stichprobe.

[8] Das Projekt *BeeMEHR* setzt sich für einen differenzierten Umgang mit Mehrsprachigkeit ein und basiert auf aktuellen (psycho-)linguistischen Erkenntnissen, wonach nicht die Mehrsprachigkeit selbst, sondern die Qualität und Quantität des sprachlichen Inputs für die Entwicklung entscheidend sind. In Kooperation mit lokalen Partnern im Ruhrgebiet werden pädagogische Fachkräfte durch Seminare sensibilisiert und eine digitale Lernplattform entwickelt, die über eine App das metalinguistische Bewusstsein und die kognitive Aufmerksamkeitskontrolle von Kindern fördert. Ziel ist es bestehende Vorurteile über Mehrsprachigkeit aufzubrechen und in inklusive Handlungspraxen zu überführen. (BeeMehr, 2025).

Neun männliche Probanden und 20 weibliche Probandinnen nahmen an der Studie teil (siehe Tabelle 2).

	7 Jahre	**8 Jahre**	**Gesamt**
männlich	n=4	n=5	n=9
weiblich	n=12	n=8	n=20
Gesamt	**n=16**	**n=13**	**n=29**

Tabelle 2: Zusammensetzung der Stichprobe

Die katholische Grundschule an der Bergstraße wurde dem Schulsozialindex eins zugeordnet (MSB NRW, 2024). Der Sozialindex dient dazu, „die unterschiedlichen Arbeits- und Rahmenbedingungen der Schulen […] [abzubilden], die sich durch eine ungleiche soziale Zusammensetzung der jeweiligen Schülerschaft ergeben“ (Schräpler & Jeworutzki, 2021, S. 2). Dieser setzt sich aus vier Indikatoren zusammen: Kinder- und Jugendarmut, Kinder mit überwiegend nicht-deutscher Familiensprache, Kinder mit eigenem Zuzug aus dem Ausland und Kinder mit den sonderpädagogischen Förderschwerpunkten Lernen, Sprache und emotionale und soziale Entwicklung (Schräpler & Jeworutzki, 2021). Der Sozialindex setzt sich aus neun Stufen zusammen, die die Anforderungen für Lehrkräfte abbilden soll. Dabei stellt die Stufe eins die niedrigste Anforderung dar (MSB NRW, 2025). Da die Schule der Probandinnen und Probanden mit einem Sozialindex von eins bewertet wurde, kann davon ausgegangen werden, dass die soziale Anforderung an dieser Einrichtung relativ niedrig ist.

Alle Probandinnen und Probanden nahmen freiwillig an dieser Studie teil. Erst nach einem Informationsbrief an die Eltern und der Einreichung der elterlichen Einverständniserklärung wurde die Testung mit den Kinder durchgeführt. Darüber hinaus wurden alle Daten anonymisiert, sodass keine Rückschlüsse auf die Kinder und Familien gezogen werden können.

3.4 Methoden

Das folgende Kapitel beschreibt die methodischen Vorgehensweisen zur Untersuchung der Einflussfaktoren auf die Leseflüssigkeit und das Textverstehen von Grundschulkindern am Ende der zweiten Klasse. Mittels eines Fragebogens wurden zentrale soziodemografische Merkmale erhoben, darunter der sozioökonomische Status, die Mehrsprachigkeit sowie der sprachliche Input. Zur Erhebung des rezeptiven Wortschatzes der Kinder kam der PPVT zum Einsatz. Darüber hinaus wurden mithilfe eines Eye-Tracking-Verfahrens Blickbewegungsdaten aufgezeichnet, die für die Analyse der Leseflüssigkeit eine zentrale Rolle spielen. Die Variable des Textverstehens wurde anhand inhaltlicher Fragen

erhoben, deren Konzeption und Anwendung im Kapitel 3.5 im Zusammenhang mit dem eingesetzten Untersuchungsmaterial näher erläutert werden.

3.4.1 Fragebogen

Mit Hilfe eines digitalen Fragebogens sollen die Variablen des sozioökonomischen Status, der Mehrsprachigkeit und des sprachlichen Inputs erhoben werden. Hierzu wurden zwei standardisierte Fragebögen genutzt, die von den *psycholinguistics laboratories* des Teams Prof. Dr. Barbara Mertins entwickelt wurden (siehe Anhang B1). Den Fragebogen gibt es in zwei Ausführungen, sodass sich ein Fragebogen an die Erziehungsberechtigten einsprachig aufgewachsener Kinder und der zweite Fragebogen an die Erziehungsberechtigten mehrsprachig aufgewachsener Kinder richtet. Die Teilnahme an der Befragung erfolgt über die Online-Plattform LimeSurvey, die durch die Technische Universität Dortmund zur Verfügung gestellt wird. Den Zugang erhalten die Erziehungsberechtigten über QR-Codes, die auf einem Elternbrief bereitgestellt sind. Die Fragebögen enthalten sowohl offene Fragen als auch geschlossene Fragen auf einer 7-Punkte-Likert-Skala (Rothstein et al., 2022). Die Wahl der 7-Punkte-Skala gegenüber der herkömmlichen 5-Punkte-Variante erfolgte, um durch eine erweiterte Skalenauflösung eine höhere Präzision und Genauigkeit der elterlichen Rückmeldungen zu erzielen. Die Zuordnung der Fragebögen zur Datenauswertung wird anhand eines anonymisierten Namencodes ermöglicht.

Der Fragebogen für die einsprachig aufgewachsenen Kinder umfasst insgesamt 27 Fragen, wohingegen der Fragebogen für mehrsprachig aufgewachsene Kinder insgesamt 61 Fragen umfasst. In Tabelle 3 sind die Bereiche zu sehen, die in den jeweiligen Fragebögen abgefragt werden. Der Unterschied zwischen den beiden Fragebögen besteht darin, dass der sprachliche Input und die sprachlichen Kompetenzen in beiden Sprachen abgefragt werden. Zudem Erhalten die Elternteile mehrsprachiger Kinder Fragen zum Sprachkontakt der Kinder, um die Balance der Sprachen zu ermitteln.

Fragebogen für einsprachige Kinder		Fragebogen für mehrsprachige Kinder	
Fragenbereich	Fragenanzahl	Fragenbereich	Fragenanzahl
sozioökonomischer Status	9	sozioökonomischer Status	9
Teilnehmer*inneninformationen	4	Teilnehmer*inneninformationen	4
sprachlicher Input	7	sprachlicher Input	14
sprachliche Kompetenzen	4	sprachliche Kompetenzen	8
		Sprachkontakt	26

Tabelle 3: Zusammenfassung des Online-Elternfragebogens

Die Bereiche des sozioökonomischen Status und der Teilnehmer*inneninformationen gestalten sich in beiden Fragebögen gleich. Um den sozioökonomischen Status zu ermitteln entwickelten Lampert et al. (2018) einen umfassenden Fragebogen, der die Bereiche Bildung, Beruf und Einkommen umfasst (Lampert et al., 2018). Jede Dimension erhält eine Punktzahl zwischen eins und sieben, die im Anschluss zu einer Gesamtpunkteanzahl zwischen drei und 21 Punkten zusammengerechnet werden (Lampert et al., 2018). Für die Berechnung des sozioökonomischen Status wird der Elternteil mit der höheren Punktzahl in dem jeweiligen Bereich genutzt. Die Punktzahlen lassen sich in eine dreistufige Skala einteilen: niedriger sozioökonomischer Status, mittlerer sozioökonomischer Status und hoher sozioökonomischer Status (siehe Tabelle 4).

Bezeichnung der Kategorie	**Quintil des SES**	**Niedrigster Punktwert**	**Höchster Punktwert**	**Anteil gewichtet**
Niedrig	1. Quintil	3,2	8,7	20,1%
Mittel	2. Quintil	8,8	11,3	20,1%
	3. Quintil	11,4	13,7	20,5%
	4. Quintil	13,8	16,9	19,4%
Hoch	5. Quintil	17,0	21,0	20,0%

Tabelle 4: Skala des sozioökonomischen Status nach Lampert et al. (2018)

Die einzelnen Bereiche des Fragebogens wurden auf Grundlage verschiedener wissenschaftlicher Daten erstellt. Der Bereich der Bildungsabschlüsse der Eltern wurde auf Grundlage „der internationalen Klassifikation Comparative Analysis of Social Mobility in Industrial Nations entwickelt“ (Lampert et al., 2018, S. 117). Die Punktzahlen spiegeln hier das durchschnittliche Gehalt wider, das mit dem Berufs- und Bildungsabschluss erlangt werden kann. Die Grundlage des zweiten Bereiches der beruflichen Stellung bildet „der International Socio-Economic-Index of Occupational Status (ISEI) nach Ganzeboom und Treimann“ (Lampert et al., 2018, S. 117). Die letzte Dimension des Einkommens wurde anhand der „Vorgaben der Armuts- und Reichtumsberichterstattung der Bundesregierung“ (Lampert et al., 2018, S. 117) ermittelt. Aus ethischen Gründen wurde der Bereich des Einkommens nicht erfragt, sondern durch die Erwerbstätigkeit (Vollzeit, Teilzeit, geringfügig beschäftigt, keine Beschäftigung) ersetzt. Auch der subjektive Sozialstatus, wie die Familie ihre Stellung in der Gesellschaft wahrnimmt, wurde erfasst, jedoch nicht zur Auswertung herangezogen.

Die Fragegruppe der Teilnehmer*inneninformationen umfasst vier Fragen. Nachdem der anonymisierte Namencode erstellt wurde, wird das Geschlecht des Kindes abgefragt. Zudem werden Angaben über Schwierigkeiten während der Schwangerschaft, der Geburt und in den

ersten Lebensmonaten abgefragt. Falls Schwierigkeiten auftraten, werden diese in einer offenen Frage abgefragt.

Der sprachliche Input und die sprachlichen Kompetenzen werden in einem gemeinsamen Bereich des Fragebogens anhand von insgesamt elf Fragen (sieben Fragen zum sprachlichen Input, vier Fragen zu den sprachlichen Kompetenzen) behandelt. Anhand einer 7-Punkte-Likert-Skala schätzen die Eltern den sprachlichen Input und die sprachlichen Kompetenzen des Kindes ein. Der sprachliche Input wird anhand von wöchentlichen Tätigkeiten abgefragt. Bei diesen Tätigkeiten handelt es sich um das Vorlesen, das eigenständige Lesen, das Spielen von Gesellschaftsspielen, das Hören von Hörbüchern, das Schauen von Serien oder Filmen, die Durchführung von Sprachübungen und das eigenständige Schreiben. Die Angaben zu den Tätigkeiten werden anhand wöchentlicher Stundenzeiten angegeben. Die sprachlichen Kompetenzen hingegen werden von *gar nicht* bis hin zu *ausgezeichnet* eingeordnet. Abgefragt werden die Kompetenzen in den Bereichen des Lesens, Schreibens, Verstehens und Sprechens. Der Fragebogen für mehrsprachig aufgewachsene Kinder umfasst die gleichen Fragen, jedoch verdoppelt sich die Anzahl, da beide gesprochenen Sprachen abgefragt werden.

Im Rahmen des Fragebogens für mehrsprachige Kinder wird der Themenbereich Sprachkontakt durch einen zusätzlichen Abschnitt mit insgesamt 26 Items abgebildet. Zu Beginn erfolgt eine Erhebung der verwendeten Sprachen sowie des Erstkontakts mit beiden Sprachen. Darüber hinaus wird die Intensität des Sprachgebrauchs in der Zeit vor der Einschulung und seit Beginn der Schulzeit systematisch erfasst. Um ein differenziertes Bild des sprachlichen Umfelds zu erhalten, werden den Eltern Aussagen zu sprachlichen Interaktionen mit Familienmitgliedern wie Eltern, Geschwistern, Großeltern, oder Cousins und Freunden vorgelegt. Die Bewertung dieser Aussagen erfolgt ebenfalls anhand einer 7-Punkte-Likert-Skala, die von *Ich stimme gar nicht zu* bis *ich stimme voll und ganz zu* reicht.

3.4.2 Peabody Picture Vocabulary Test

Um den rezeptiven Wortschatz zu ermitteln, wurde die deutschsprachigen Adaption des PPVTs verwendet (Lenhard at al., 2015, zitiert nach Lenhard & Lenhard, 2022). Den Test entwickelten in den 1960er Jahren das Ehepaar Llyod und Leota Dunn und er wurde in den folgenden Jahren ausgebaut und überarbeitet (Lenhard & Lenhard, 2022). Zur Ermittlung des Wortschatzes in dieser Arbeit wurde die vierte Auflage der deutschsprachigen Adaption des PPVT genutzt (Lenhard & Lenhard, 2022). Bei dem PPVT handelt es sich um ein standardisiertes Messinstrument, dass international verwendet wird und auf verschiedene Sprachen angepasst wurde. Der PPVT dient der Erfassung des rezeptiven Wortschatzes und

liefert standardisierte Werte auf Basis einer repräsentativen Stichprobe von n = 3555 Kindern und Jugendlichen im Alter von 3;0 bis 16;11 Jahren. Das Verfahren umfasst 228 Items, die der Schwierigkeit nach in 19 Blöcke zu je zwölf Aufgaben gruppiert sind. Jeweils wird eine Bildkarte mit vier Motiven präsentiert und die Testperson soll das Bild markieren, das zum vorgelesenen Wort passt (Lenhard & Lenhard, 2022). Der Beginn des Tests ist altersabhängig, sodass beispielsweise Dreijährige in Block eins, Fünfjährige in Block vier und Jugendliche ab 14 Jahren in Block elf starten. Erreicht die Testperson im Einstiegsblock mindestens elf von zwölf Items korrekt, gelten alle niedrigeren Blöcke als bestanden. Bei weniger als elf richtigen Antworten wird schrittweise auf vorherige Blöcke zurückgegangen, bis das Kriterium erfüllt ist (Lenhard & Lenhard, 2022). Anschließend wird in aufsteigender Reihenfolge weitergetestet, bis in einem Block acht oder mehr Fehler auftreten, womit das Verfahren endet (Lenhard & Lenhard, 2022). Für jede Aufgabe stehen etwa zehn Sekunden zur Verfügung, falls die Testperson nicht reagiert, erfolgt eine einmalige Ermutigung, danach wird die Antwort als Fehler gewertet und zum nächsten Item übergegangen (Lenhard & Lenhard, 2022).

Die durchschnittliche Dauer der Durchführung des PPVT hängt laut Lenhard und Lenhard (2022) stark von der Testperson ab und kann zwischen etwa zehn und 20 Minuten variieren. Das Ergebnis, das das Kind erzielt, kann mit einer Normtabelle verglichen werden, um Aufschluss über den rezeptiven Wortschatz im Vergleich zu der selben Altersklasse zu erhalten.

3.4.3 Eye-Tracking

Um die Variable der Lesekompetenz zu erfassen, wurde für die vorliegende Studie die Methodik des Eye-Trackings gewählt. Bei dem für die vorliegende Studie eingesetzte Eye-Tracking-System handelt es sich um das System SMI RED250 mobile der Marke SensoMotoric Instruments. Hierbei handelt es sich um ein Remote-System (siehe Kapitel 2.4.2), das mit einem Bildschirm sowie einem Laptop verbunden wird. Der Eye-Tracker wird während der Untersuchung mit Hilfe von Magneten fest an dem Bildschirm montiert, auf dem den Schülerinnen und Schülern der Text präsentiert wird. Dadurch kann die Augenbewegung während des Lesens präzise erfasst werden. Für eine valide Erhebung ist es erforderlich, dass die Probandinnen und Probanden in einem Abstand von etwa 60 Zentimetern zentral im Erfassungsbereich des Gerätes positioniert sind. Die Ausrichtung erfolgt mit Hilfe der zughörigen Software der Firma SensoMotoric Instruments, die eine genaue Kalibrierung der Augen- und Sitzposition jedes Kindes ermöglicht. Da es sich um ein flexibel einsetzbares System handelt, kann der Eye-Tracker mobil in schulischen Kontexten und Klassenräumen genutzt werden. Aufgrund seiner nicht-intrusiven Bauweise ist es möglich, die Sitzposition der Schülerinnen und Schüler mit Hilfe eines verstellbaren Stuhls oder auch den Bildschirm durch

eine Höheneinstellung so anzupassen, dass die Probandinnen und Probanden sich optimal im Erfassungsbereich des Eye-Trackers befinden.

Die Analyse der Blickdaten erfolgt unter Verwendung von Gaze Plots und Gaze Replays (siehe Kapitel 2.4.3). Beide dynamischen Visualisierungen lassen sich in der Software nutzen und ermöglichen eine differenzierte Darstellung von Fixationen und Sakkaden. Darüber hinaus wird die Möglichkeit geboten, mehrere Gaze Plots übereinanderzulegen, um Unterschiede oder Gemeinsamkeiten hinsichtlich der Blickverläufe und Fixationspunkte zwischen den Lernenden sichtbar zu machen.

3.5 Untersuchungsmaterial

Das vorliegende Kapitel erläutert die Konzeption des Untersuchungsmaterials, das im Rahmen des Eye-Tracking-Verfahrens den teilnehmenden Kindern präsentiert wurde. Im Zentrum stehen die eigens entwickelten Lesetexte, die inhaltlichen Verständnisfragen zur Erfassung des Textverstehens sowie die eingesetzten Filleritems, die der kognitiven Entlastung während der Datenerhebung dienen. Zur Überprüfung der Verständlichkeit und Angemessenheit des Materials wurde im Vorfeld eine Pilotstudie mit zwölf Erwachsenen sowie fünf Kindern der dritten Jahrgangsstufe durchgeführt. Die Ergebnisse der Pilotierung zeigten, dass nur geringe inhaltliche noch strukturelle Anpassungen der Fragen erforderlich waren, sodass das Material in seiner ursprünglichen Form in die Hauptstudie übernommen werden konnte.

Für die Untersuchung wurden vier Texte erstellt, die sich in Thema und Schwierigkeitsgrad unterscheiden. Da, nach Schaffner (2009), die Verstehensleistung positiv mit dem thematischen Interesse korreliert, wurden zwei verschiedene inhaltliche Bereiche für die Texte gewählt (siehe Kapitel 2.1.3). Zwei Texte behandeln das Thema *Leben auf dem Bauernhof* (siehe Anhang A6.1) und die anderen beiden Texte das Thema Ein Sommertag (siehe Anhang A6.2). Es wurden Themen gewählt, die den Kindern zumeist durch den Alltag oder auch durch die Schule bekannt sind. Des Weiteren wurden beide Texte des jeweiligen Themas in verschiedenen Schwierigkeitsgraden verfasst. Dabei wurde darauf geachtet, dass sich beide Texte mit geringem sowie höherem Schwierigkeitsgrad zwischen den Themenbereichen ähneln, um eine Vergleichbarkeit herzustellen. Die Texte des niedrigeren Schwierigkeitsgrades weisen eine durchschnittliche Wortlänge von 4,60 und 4,57 Zeichen auf mit einer minimalen Wortfrequenzklasse von drei, was bedeutet, dass diese Wörter eher häufig genutzt werden (siehe Tabelle 5). Der höhere Schwierigkeitsgrad weist eine durchschnittliche Wortlänge von 4,97 und 4,94 Zeichen mit einer minimalen Frequenzklasse von eins beziehungsweise zwei auf. Die Gesamtlänge der Texte unterscheidet sich um 50 beziehungsweise 118 Zeichen (siehe Tabelle 5).

Texte	Wortanzahl gesamt	Durchschnittliche Wortlänge	Frequenzklassen der Wörter (Worthäufigkeit)	Wortanzahl in den Frequenzklassen					
				F1	F2	F3	F4	F5	F6
Niedriger Schwierigkeitsgrad									
Sommertag 1	51	4,57	3-6	0	0	8	10	4	2
Leben auf dem Bauernhof 1	55	4,60	3-6	0	0	11	14	3	1
Höherer Schwierigkeitsgrad									
Sommertag 2	71	4,94	1-6	1	2	9	12	8	3
Leben auf dem Bauernhof 2	61	4,97	2-6	0	4	8	10	5	1

Tabelle 5: Wort-, Zeichen- und Frequenzwerte der Texte

Die Komplexität wurde mit Hilfe der Textlänge, der Wortfrequenz sowie der Nutzung von Anaphern und Nebensätzen angepasst. Abbildung 6 stellt den Bauernhoftext des niedrigeren Schwierigkeitsgrades dar. In Abbildung 7 wird der höhere Schwierigkeitsgrad präsentiert, in dem die Begriffe mit höheren Schwierigkeitsanforderungen gelb markiert sind. Diese Hervorhebungen kennzeichnen auch syntaktische Mittel wie Nebensätze und Anaphern, die für die textuelle Kohärenz von zentraler Bedeutung sind. Nach der Profilanalyse von Wilhelm Grießhaber wird die syntaktische Komplexität mittels sechs ausgewählter Wortstellungsmuster erfasst, die in Profilstufen null bis sechs untergliedert sind (Grießhaber, 2024). Der wiederholte Einsatz kausaler Konjunktionen wie *weil* und konditionaler Konjunktionen wie *wenn* korrespondiert mit dem Nebensatzmuster der Profilstufe vier und deutet auf eine erhöhte Komplexität hin (Grießhaber, 2024). Die mehrfachen Pronomen wie *sie* und *seine* verdeutlichen den Einsatz von Anaphern, die laut Grießhaber als Kohärenzmittel gelten und die textuelle Vernetzung stärken (Grießhaber 2019b). Grießhaber legt fest, dass die Erwerbsstufe eines Textes durch die höchste mindestens dreimal verwendete Profilstufe bestimmt wird (Grießhaber 2014). So lässt sich der Text des höheren Schwierigkeitsgrades der Erwerbsstufe vier einordnen (Grießhaber, 2024). Im Vergleich dazu stellt der Text der niedrigeren Schwierigkeitsstufe (siehe Abbildung 6) eine Erwerbsstufe von drei dar, indem keine verschachtelte Sprachverwendung oder Nebensatzmuster auftreten (Grießhaber, 2024).

Auf dem Bauernhof leben viele Tiere. Jedes Tier hat eine Aufgabe.

Die Katzen fangen Mäuse und die Hunde passen auf. Darum fühlen sich die Tiere sicher.

Der Bauer kümmert sich um die Tiere. Er füttert die Hühner und sammelt die Eier ein.

Die Kühe fressen Gras und geben Milch. Auf dem Bauernhof ist viel los.

Abbildung 6: Lesetext - Bauernhof geringerer Schwierigkeitsgrad

Auf dem Bauernhof gibt es auch andere Tiere. Die Schweine wühlen im Schlamm, weil sie schwitzen.

Die Gänse schnattern laut, wenn ein Besucher kommt. Sie haben immer viel zu erzählen.

Die Pferde laufen auf der Weide. Manchmal reitet der Bauer auf einem Pferd.

Seine Tochter zeigt den Kindern, wie sie das Heu erntet. Sie helfen ihr, weil sie gerne mitarbeiten.

Abbildung 7: Lesetext - Bauernhof höherer Schwierigkeitsgrad

Die Texte wurden den Kinder auf einem 24-Zoll Monitor präsentiert. Für die Darstellung der Texte wurde sich an verschiedenen Studien orientiert. Es wurde eine schwarze Schrift mit einem Zeilenabstand von 2,5 auf einem grauen Hintergrund gewählt (Krause & Ritter, 2022; Smyrnakis et al., 2017). Vorläufig wurde wie bei Krause & Ritter (2022) die Schriftgröße 20 gewählt, jedoch stellte sich während der Pilotierung in der dritten Klasse heraus, dass eine Schriftgröße von 30 für Kinder der zweiten Klasse geeigneter ist (Smyrnakis et al., 2017). Da sich die Kinder auch in einem Abstand von 60 Zentimetern vor dem Bildschirm befinden, konnte so sichergestellt werden, dass eine zu geringe Schriftgröße die Ergebnisse der Studie nicht beeinflusst. Im Hinblick auf die Schriftart stellt das System des Eye-Trackers nur eine geringe Auswahl von Schriftarten, insbesondere für die Grundschule, zu Verfügung. Aufgrund dessen, dass eine Vielzahl an Schriftarten eine andere Schreibweise der Kleinschreibung der Buchstaben <a>, <l> und <t> aufweist, wurde sich für die Schrift Comic Sans entschieden. Diese Schriftart ermöglicht die Schreibweise des *As* (a / ɑ), die auch in der Grundschule verwendet wird (siehe Abbildung 6 & 7). In der Pilotierung zeigte sich, dass die Kinder keine Probleme beim Erkennen des *Ts* (t / ŧ) und des *Ls* (l / ł) hatten, sodass diese Schriftart genutzt werden kann. Darüber hinaus wurde der Text den Lesenden nicht Satz für Satz präsentiert, damit ein zusammenhängendes Textbild entsteht. Zudem hatte dies auch Platzgründe, sodass sich der Text nicht über mehrere Folien erstreckt. Viele Studien arbeiten mit dem selbstständigen Weiterklicken der Probandinnen und Probanden nach dem Lesen, jedoch folgt in dieser Studie eine leichte Abwandlung (Bosma & Nota, 2020; Häikiö et al., 2009; Krause & Ritter, 2022). Nach dem Beenden der jeweiligen Lese-, Frage- oder Filler-Aufgabe teilen die Probandinnen und Probanden dies mit und die Versuchsleitung setzte die Studie mit einem

Klick fort. Diese Methode wurde ausgewählt, damit die Kinder ihre weitgehend unbewegte Sitzposition beibehalten können, um eine erneute Kalibrierung des Eyetrackers zu vermeiden.

Zu den jeweiligen Texten wurden passende Fragestellungen entwickelt. Um eine Vergleichbarkeit zwischen den Fragen zu erhalten, wurden stets drei Fragen nach einem festen Schema erstellt. Hierzu wurden drei W-Fragen gewählt: *Wer, Was und Warum* (siehe Tabelle 6). Die Reihenfolge der Fragen bleibt bei jedem Text gleich und umfasst stets eine ähnliche Wortanzahl. So kann sichergestellt werden, dass das Ergebnis der Untersuchung keine Abhängigkeit von den Fragen aufweist. Darüber hinaus wurden die Fragen so gewählt, dass aus dem Text eine oder mehrere klare Antworten zu erkennen sind. Die Tabelle stellt die Fragen mitsamt der Antworten dar (siehe Tabelle 6). Eine Ausnahme bildet die Frage sechs, da sich in der Pilotierung gezeigt hat, dass der Textinhalt interpretiert wurde. Der Text gibt die in der Abbildung genannten Antworten wieder, jedoch führte die Interpretation dazu, dass die Antwort gegeben wurde: „Weil es schon spät ist“. Diese Antwort zählen wir als korrekt, da es sich um eine Interpretation der Situationen, der untergehenden Sonne, der Kälte und der Müdigkeit handelt. Die Beantwortung der Fragen muss nicht in ganzen Sätzen erfolgen, als korrekt werden auch einzelne Aussagen, wie „die Katze“ gewertet. Die Fragen wurden den Kindern mündlich gestellt. Die Antwort wurde direkt händisch in eine Tabelle eingetragen. Der Text wurde während der Fragesequenz nicht eingeblendet, da ein wiederholtes Lesen des Textes nicht darauf hinweisen muss, dass dieser nicht verstanden wurde, sondern auch eine Rückversicherung für die Kinder darstellen kann. Zudem ist das Arbeitsgedächtnis ein wichtiger Faktor für die Lesekompetenz, sodass diese nur ausgeprägt ist, wenn die Inhalte im Gedächtnis verweilen und im Nachhinein wiedergegeben werden können (Tischler et al., 2017). Damit sich die Kinder während der Fragensequenz nicht bewegen, wurde ihnen ein Fixationskreuz gezeigt, auf das der Blick gerichtet bleibt (Holmqvist & Andersson, 2017).

Fragen	**Antworten**
Ein Sommertag – 1	
Wer fährt zum See?	Tim und Paula fahren zum See.
Was essen Tim und Paula?	Sie essen ein Eis.
Warum ist Tim traurig?	Sein Eis fällt auf den Boden.
Ein Sommertag – 2	
Wer baut einen Turm?	Tim baut einen Turm.
Was sammelt Paula?	Paula sammelt Muscheln.

Warum fahren sie nach Hause?	Weil die Sonne untergeht, es kalt wird und Paula müde ist.
Leben auf dem Bauernhof – 1	
Wer fängt die Mäuse?	Die Katze fängt die Mäuse.
Was sammelt der Bauer?	Er sammelt die Eier ein.
Warum fühlen sich die Tiere sicher?	Weil der Hund aufpasst.
Leben auf dem Bauernhof – 2	
Wer schnattert laut?	Die Gänse schnattern laut.
Was macht der Bauer?	Er reitet auf dem Pferd.
Warum wühlen die Schweine im Schlamm?	Weil die Schweine schwitzen.

Tabelle 6: Fragen mit den korrekten Antworten für jeden Text

Nach dem Lesen eines Textes und der Beantwortung der zugehörigen Fragen, werden den Probandinnen und Probanden sogenannte Filler präsentiert. In dieser Studie handelt es sich dabei um Wimmelbilder, die den Kindern eine Erholung beziehungsweise einen neutralen Reiz zwischen dem Lesen bieten sollen, um die kognitiven Ressourcen zu entspannen (Holmqvist & Andersson, 2017). Bei der Erstellung der Wimmelbilder wurde darauf geachtet, dass diese keine inhaltlichen Anknüpfungspunkte mit den Texten haben. Hierbei wurden die Themen Eislaufen, Ostern, Schulhof und Klassenzimmer behandelt. Da bei der Pilotierung nicht allen Kindern ein Wimmelbild bekannt war, wurde ein Probeitem eingefügt, mit dem den Kindern die Vorgehensweise bei einem Wimmelbild erläutert wurde.

Alle Teile wurden in einer bestimmten Abfolge präsentiert. Ein Block beginnt jeweils mit dem Text, gefolgt von drei Fragen und einem Wimmelbild, dann beginnt der nächste Block. Während der Fragen wurde den Kinder ein Fixationskreuz präsentiert, sodass der Blickfokus auf dem Bildschirm blieb. Durch die vier Texte ergaben sich vier einzelne Blöcke, die hintereinander ohne Pause durchgeführt wurden. Die Abfolge der Texte wurde randomisiert, sodass ausgeschlossen werden kann, dass ein unterschiedliches Abschneiden zwischen den Themen durch einen Reihenfolgeeffekt oder auch die Konzentrationsfähigkeit beeinflusst wird (Holmqvist & Andersson, 2017). Trotz der Zufälligkeit der Themen wurde darauf geachtet, dass stets mit dem geringeren Schwierigkeitsgrad begonnen wurde, gefolgt von dem selben Thema mit dem höheren Schwierigkeitsgrad, sodass ein inhaltliches Thema immer abgeschlossen wurde.

3.6 Durchführung

Zu Beginn der Studie erhielten die Eltern der teilnehmenden Kinder eine schriftliche Information über Zielsetzung und Ablauf der Untersuchung sowie eine Einverständniserklärung zur Teilnahme an der Studie. Nach Rückgabe der unterzeichneten Einverständniserklärung wurde den Eltern ein QR-Code ausgehändigt, der Zugang zu einem digitalen Fragebogen ermöglichte. Dieser Fragebogen diente der Erhebung der unabhängigen Variablen: Mehrsprachigkeit, sprachlicher Input sowie sozioökonomischer Status. Das Ausfüllen erfolgte in der Regel etwa zwei Wochen vor Beginn der ersten Testphase (PPVT). In Fällen, in denen der Fragebogen nicht fristgerecht eingereicht wurde, wurde den Eltern eine Nachreichfrist eingeräumt, während die Kinder dennoch am Testverfahren teilnehmen konnten.

Die erste Testphase bestand aus der Durchführung des PPVT, welcher etwa zwei Wochen vor der Erhebung der Eye-Tracking-Daten stattfand. Die Schule stellte hierfür einen ruhigen, ungenutzten Klassenraum zur Verfügung. Der Raum wurde so vorbereitet, dass Ablenkungen minimiert und standardisierte Bedingungen gewährleistet wurden: Ein Tisch wurde so positioniert, dass die Kinder keinen Einblick in den Bewertungsbogen hatten und zwei Stühle wurden gegenüberliegend platziert, um eine direkte Sicht auf die Bildkarten des PPVT zu ermöglichen. Nach Absprache mit den jeweiligen Klassenlehrkräften stellte sich die Versuchsleitung in der Klassengemeinschaft vor. Anschließend wurde das erste Kind in den Testraum begleitet. Dort wurde zunächst überprüft, ob die Sitzposition eine uneingeschränkte Sicht auf die Bildkarten zuließ. Gegebenenfalls wurden die Materialien oder die Sitzgelegenheit angepasst. Die Versuchsleitung erklärte dem Kind, dass gemeinsam herausgefunden werden soll, wie viele Wörter es bereits kennt. Dabei hört das Kind ein gesprochenes Wort und wählt aus vier Bildern dasjenige aus, das am besten zum gehörten Begriff passt. Nach einem Probeitem begann die eigentliche Testung. Vor dem Wechsel zum nächsten Bildquartett dokumentierte die Versuchsleitung die Antwort auf dem Bewertungsbogen. Die Testung wurde beendet, sobald ein Kind innerhalb eines Sets acht oder mehr Items falsch zugeordnet hatte. Abschließend bedankte sich die Versuchsleitung bei dem Kind, das daraufhin in die Klasse zurückkehrte und das nächste Kind in den Testraum schickte. Die Dauer der Testung pro Kind betrug durchschnittlich etwa 20 Minuten. Es wurde darauf geachtet, dass die Durchführung nicht mit Pausenzeiten kollidierte, um eine ungestörte Testatmosphäre zu gewährleisten.

Nach der Erhebung des rezeptiven Wortschatzes wurde in einer zweiten Testphase die Lesekompetenz mit der Eye-Tracking-Methode erfasst. Hierzu wurde ein ruhiger Raum im Dachgeschoss der Grundschule gewählt und die Fenster sowie die Tür geschlossen, sodass

keine Störungen auftreten können. Zudem wurde darauf geachtet, dass die Dauer der Studie nicht in Pausenzeiten fällt, damit weder die Klingel noch spielende Kinder Unruhe und Ablenkung in die Untersuchung bringen konnten. Um visuelle Ablenkungen in der Blickrichtung der Kinder zu vermeiden, wurde der Eye-Tracker auf einem Tisch in Richtung der Wand platziert. Im Raum wurden zwei Stühle in unterschiedlichen Größen bereitgestellt, sodass je nach Körpergröße der Kinder die Sitzhöhe angepasst werden konnte. Die gesamten Teilnehmerinnen und Teilnehmer wurden zu Beginn in der Klassengemeinschaft begrüßt. Eine erneute Vorstellung der Versuchsleitung war nicht notwendig, da diese den Kindern bereits durch den Wortschatztest bekannt war. Das erste Kind ging gemeinsam mit der Versuchsleitung in den Untersuchungsraum. Im Gegensatz dazu wurden die nachfolgenden Kinder jeweils vom vorherigen Kind in den richtigen Raum geführt. Nach dem Eintreffen in dem Raum der Studiendurchführung wurden die Kinder erneut begrüßt. Bevor die Instruktionen begannen, probierten die Kinder die verschiedenen Stühle aus, um den passenden auszuwählen. Anschließend las die Versuchsleitung die Instruktionen vor. Das Vorlesen dieser wurde gewählt, damit jedes Kind die selben Informationen erhält und so die Daten vergleichbar sind (siehe Anhang B 4). Nachdem die Instruktionen durchgesprochen wurden, setze sich die Versuchsleitung gegenüber der Kinder hinter den Bildschirm und kalibrierte den Eye-Tracker. Überprüft wurde, ob die Höhe des Eye-Trackers und die Zentrierung sowie Entfernung der Probandinnen und Probanden korrekt eingestellt sind. Zur Kalibrierung wurden Punkte an verschiedenen Orten des Bildschirmes anzeigt, die durch die Kinder laut Instruktion der Leitung mit den Augen gefangen werden sollen, ohne den Kopf zu bewegen. Bei einer unzureichenden Kalibrierung wurde diese wiederholt, um die Augenbewegung der Kinder möglichst genau erfassen zu können. Nach der Kalibrierung startete das Experiment. Die Texte, Fragen beziehungsweise Fixationskreuze und Filler wurden in einer bestimmten Abfolge präsentiert. Die Texte wurden in zwei Themenblöcke geteilt, sodass jeweils ein Thema vollends durchgeführt wurde. Wie bereits in Kapitel 3.5 erläutert, wurde die Reihenfolge der Themenblöcke randomisiert, um eine Abhängigkeit der Ergebnisse an der Reihenfolge auszuschließen. Nachdem ein Text gelesen wurde, beantworteten die Kinder die gestellten Fragen. Im Anschluss daran lösten sie ein Wimmelbild. Erst nach der Lösung des Wimmelbildes begann der nächste Text. Alle Texte beziehungsweise Themenblöcke wurden hintereinander ohne Pause durchgeführt. In der Pilotierung stellte sich heraus, dass die Kinder keine Pause benötigen und die Konzentration für die circa 15 Minuten der Testdurchführung ausreicht. Nach jeder Eye-Tracking-Erhebung wurden die aufgezeichneten Blickbewegungsmuster automatisiert verarbeitet (siehe Abbildung 8). Abbildung 8 veranschaulicht exemplarisch, wie das Eye-Tracking-System Sakkaden durch Linien und die variierenden Fixationsdauern durch unterschiedlich große Kreise darstellt. Die Speicherung erfolgte extern und unmittelbar nach Abschluss der jeweiligen Sitzung, sodass ein

Datenverlust ausgeschlossen werden konnte. Die Datenstruktur ermöglichte eine eindeutige Zuordnung zum jeweiligen Kind und gewährleistete die Nachvollziehbarkeit der individuellen Blickverläufe.

Abbildung 8: Aufnahme der Blickbewegung mit Gaze Plots

3.7 Datenauswertung

Die Datenauswertung erfolgte schrittweise unter Anwendung standardisierter und nachvollziehbarer Verfahren. Zunächst wurde der eingesetzte Fragebogen hinsichtlich relevanter Merkmale ausgewertet, darunter der sozioökonomische Status, die Mehrsprachigkeit sowie der sprachliche Input im häuslichen Umfeld.

Im Anschluss erfolgte die Auswertung des rezeptiven Wortschatztests basierend auf den Normwerten nach Lampert et al. (2015). Für jedes Kind wurden sowohl der Rohwert als auch der entsprechende T-Wert ermittelt und systematisch in einer Excel-Tabelle[9] dokumentiert. Die weitere Analyse basierte ausschließlich auf dem T-Wert, da dieser bereits altersnormiert ist und somit eine vergleichbare Bewertung der Wortschatzleistungen über unterschiedliche Altersstufen hinweg ermöglicht.

Im Anschluss an die Präsentation jedes Lesetextes wurden den Kindern Verständnisfragen gestellt, um das Textverstehen zu erfassen. Die Antworten wurden unmittelbar durch die Versuchsleitung dokumentiert. Dabei wurden korrekte Antworten mit einem Häkchen (✓) und inkorrekte Antworten mit einem Kreuz (X) gekennzeichnet (siehe Anhang B 5). Für die Auswertung wurde zwischen den vier Lesetexten differenziert, da diese unterschiedliche Schwierigkeitsgrade aufwiesen. Zur Quantifizierung der Leistungen wurde für jeden Text der

[9] Eine Übersicht sämtlicher erhobener Daten zum Wortschatz, zu den Antwortverhalten sowie zu den Eye-Tracking-Parametern ist im externen Anhang B6 dokumentiert.

prozentuale Anteil korrekter Antworten berechnet. Diese Werte wurden ebenfalls in der zentralen Excel-Tabelle erfasst und dienten als Grundlage für die weitere Analyse.

Auch die Eye-Tracking-Daten wurden gesondert für jeden Text systematisch dokumentiert. In die Auswertung gingen folgende Parameter ein: die Dauer der Fixationen in Millisekunden, die gesamte Lesedauer in Sekunden sowie die Dauer des ersten Lesedurchgangs in Sekunden und die durchschnittliche Anzahl an Fixationen pro Wort. Die durchschnittliche Anzahl von Fixationen pro Wort wurde durch die Division der Gesamtanzahl der Fixationen pro Text durch die Anzahl an Wörtern des jeweiligen Textes ermittelt. Zur Sicherstellung der Datenqualität wurde für jede Probandin bzw. jeden Probanden die individuelle Tracking-Ratio[10] pro Text geprüft. Da keines der getesteten Kinder eine Tracking-Ratio von unter 80 % aufwies, konnten alle Kinder (n = 29) in die Auswertung einbezogen werden.

Auf Grundlage dieser Datenerhebung konnte anschließend die Analyse der Einflussfaktoren auf die beiden Zielvariablen der Leseflüssigkeit und des Textverstehens vorgenommen werden. Hierfür wurden zwei Mixed-Model-Analysen durchgeführt, die eine differenzierte Betrachtung der Einflussfaktoren ermöglichen[11]. Eine detaillierte Darstellung der Modellstruktur und der Ergebnisse erfolgt in Kapitel 3.8.

3.8 Darstellung der Ergebnisse

Eingangs erfolgt eine deskriptive Analyse ausgewählter Variablen. Das Ziel besteht darin, einen fundierten Überblick über die Zusammensetzung der Stichprobe sowie über zentrale Charakteristika der erhobenen Daten zu ermöglichen. Darüber hinaus erlaubt die Betrachtung dieser Verteilungen eine methodisch nachvollziehbare Begründung für den Ausschluss bestimmter Variablen aus der weiteren Analyse. Anschließend werden die Resultate zweier gemischter Modelle präsentiert, die im Rahmen der statistischen Auswertung zur Beantwortung der Forschungsfragen eingesetzt wurden. Der Fokus liegt dabei auf der Identifikation signifikanter Einflussgrößen auf die Variablen der Leseflüssigkeit sowie des Textverstehens.

---

[10] Die Tracking-Ratio dient als Qualitätsmaß für Eye-Tracking-Daten und beschreibt das Verhältnis gültiger zu erwarteten Blickdaten (vgl. Müller 2013). Holmqvist und Andersson (2017) betonen zusätzlich, dass eine hohe Tracking-Ratio essentiell für die Validität von Blickbewegungsanalysen ist, da sie die Zuverlässigkeit der Datenbasis widerspiegelt.

[11] Die Auswertung der Daten im Rahmen der statistischen Analyse wurde von Vincent Schröder vorgenommen.

3.8.1 Deskriptive Auswertung der Ergebnisse

Im Bereich der deskriptiven Analyse werden der rezeptive Wortschatz, der sprachliche Input und der sozioökonomische Status der Kinder betrachtet. Um die Verteilung dieser Variablen in der vorliegenden Stichprobe (n=29) sichtbar zu machen, werden Boxplots genutzt (siehe Abbildung 9 und Abbildung 10).

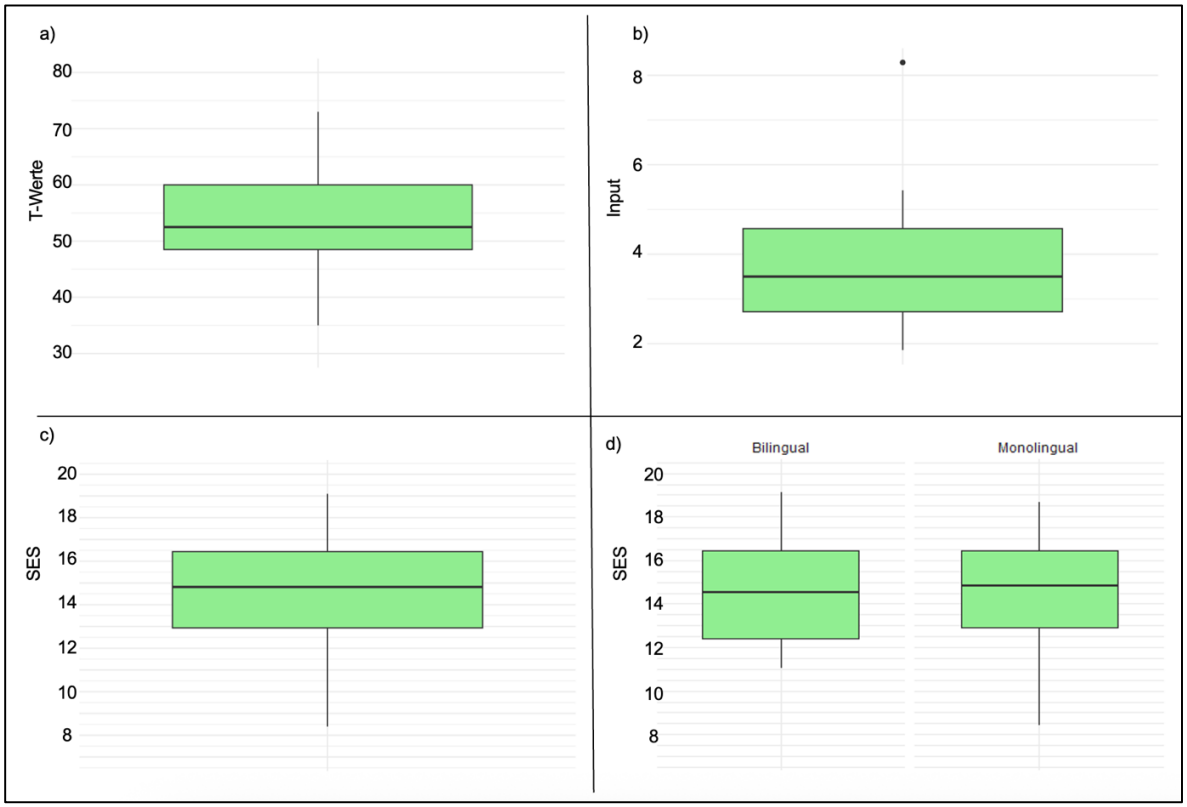

Abbildung 9: Boxplots – a) Verteilung des T-Wertes (rezeptiver Wortschatz); b) Verteilung des sprachlichen Inputs in Stunden pro Tag; c) Verteilung des sozioökonomischen Status nach Lampert et al. (2018); d) Verteilung des sozioökonomischen Status zwischen bilingualen und monolingualen Kindern

Beginnend mit dem Wortschatz wird in dem vorliegenden Boxplot der erreichte T-Wert dargestellt (siehe Abbildung 9a)). Dieser T-Wert wurde mit Hilfe des PPVT ermittelt (siehe Kapitel 3.4.2). Der Boxplot zeigt eine Verteilung der T-Werte von 35 bis 73, wobei der Median bei 52,50 und der Mittelwert bei 53,07 liegt. 50% der Probandinnen und Probanden befinden sich in der Interquartilsspanne zwischen circa 47,5 und 59. Die Whisker reichen vom Minimum (35) bis hin zum Maximum (73) des T-Wertes, was auf eine homogene Stichprobe ohne Ausreißer hindeutet (siehe Abbildung 9a)). Die Standardabweichung liegt bei 8,36. Da der rezeptive Wortschatz der Kinder, wie in Kapitel 3.2 erläutert, in Normbereiche eingeteilt werden kann, wurde zusätzlich zum Boxplot ein Balkendiagramm herangezogen (siehe Abbildung 10). In diesem Diagramm ist die Verteilung der Kinder auf die verschiedenen Normbereiche des PPVTs zu sehen. Deutlich zu erkennen ist, dass circa 69% der Kinder im oberen Normalbereich und darüber verortet sind, wohingegen ausschließlich 31% der Kinder im unteren Normalbereich beziehungsweise unterdurchschnittlich einzuordnen sind (siehe

Abbildung 10). Mit über zwei Drittel der Kinder, die sich im oberen Leistungsbereich des rezeptiven Wortschatzes bewegen, zeigt sich die Gruppe als sprachlich stark.

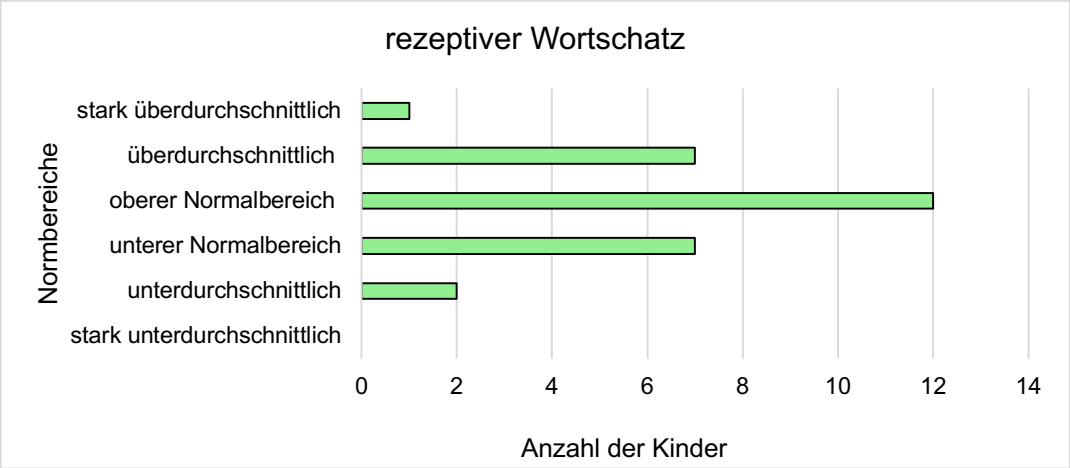

Abbildung 10: Einteilung der Kinder in die Normbereiche des PPVTs auf Basis der T-Werte

Der Boxplot in Abbildung 9b) veranschaulicht die Verteilung des täglichen sprachlichen Inputs in Stunden, dem die Kinder außerhalb der schulischen Umgebung ausgesetzt sind. Die Daten wurden aus dem Elternfragebogen entnommen (siehe Kapitel 3.4.1). Die deskriptive Statistik der wöchentlichen sprachlichen Inputzeit zeigt, dass die Interquartilsspanne der Stichprobe zwischen 2,9 und 4,8 Stunden liegt, wobei der Median bei 3,5 Stunden pro Woche verortet ist. Der minimale Wert beläuft sich auf 1,86 Stunden, der maximale auf 8,29 Stunden und das arithmetische Mittel beträgt 3,8 Stunden. Die Standardabweichung von 1,38 Stunden signalisiert eine moderate Streuung um den Mittelwert. Ein einzelner Ausreißer (8,29 Stunden) am oberen Ende der Verteilung deutet auf eine atypisch hohe Inputmenge hin. Insgesamt lässt sich sagen, dass die Verteilung relativ ausgewogen erscheint, mit einem klaren Zentrum zwischen circa drei und fünf Stunden, abgesehen des einzelnen Ausreißers nach oben.

Die Verteilung des sozioökonomischen Status wird anhand des Punkte-Scores nach Lampert et al. (2018) dargestellt (für die Erläuterung des Punkte-Scores siehe Tabelle 4). Auch diese Daten wurden dem Elternfragebogen entnommen (siehe Kapitel 3.4.1). Die deskriptive Analyse der Stichprobe zeigt, dass die Werte zwischen 9,4 und 20,10 Punkten variieren (siehe Abbildung 9c)). Die Interquartilsspanne erstreckt sich von 14,00 bis 17,40 Punkten, wobei der Median bei 15,82 Punkten liegt. Das arithmetische Mittel beträgt 15,61 Punkte, die Standardabweichung beläuft sich auf 2,62 Punkte. Der Boxplot zeigt eine ausgewogene Verteilung der Werte des sozioökonomischen Status ohne erkennbare Ausreißer. Die analysierten Daten lassen darauf schließen, dass die untersuchte Stichprobe überwiegend aus Kindern mit einem hohen sozioökonomischen Hintergrund besteht. Dieses Ergebnis korrespondiert mit der Einordnung der betreffenden Grundschule in den Schulsozialindex 1 (siehe Kapitel 3.3).

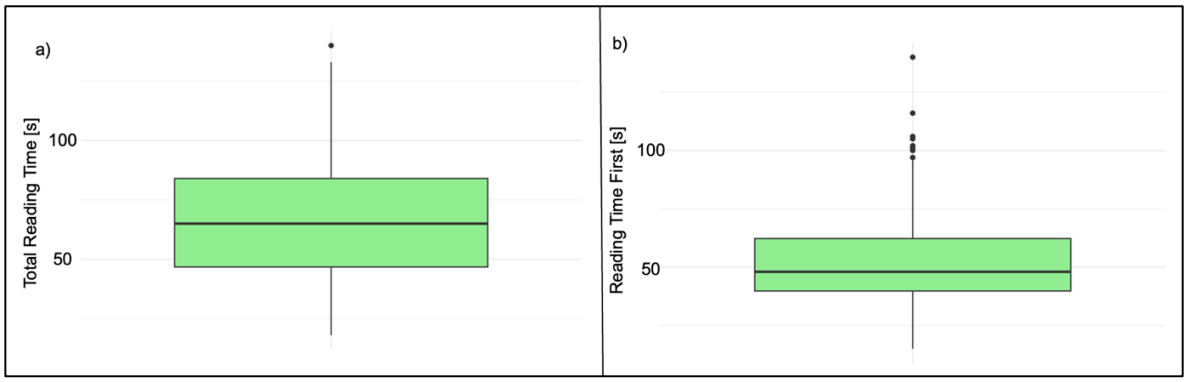

Abbildung 11a): Boxplot - Verteilung der Gesamtlesezeit

Abbildung 11b): Boxplot – Verteilung der Lesezeit des ersten Durchganges

Im Weiteren werden die Ergebnisse der Eye-Tracking-Daten betrachtet. Die Boxplots in Abbildung 11a) und 11b) zeigen die Verteilung der Gesamtlesezeit (Total Reading Time in Sekunden) sowie der Lesezeit des ersten Durchganges an (Reading Time First Trial in Sekunden). Hierzu wurden pro Kind jeweils alle vier Texte beachtet und flossen in den Boxplot mit ein. Die Gesamtlesezeit, die die Kinder für die Bearbeitung der vier Texte benötigten, liegt zwischen 18 und 133 Sekunden. Innerhalb der Interquartilsspanne, die sich ungefähr zwischen 48 und 85 Sekunden erstreckt, befinden sich 50% der Probandinnen und Probanden mit einem Medianwert von 65 Sekunden (siehe Abbildung 11a)). Ein Ausreißer mit einer Lesezeit von etwa 140 Sekunden überschreitet deutlich den oberen Whisker und sticht dadurch als besonders abweichender Wert hervor.

Die durchschnittliche Lesezeit beim ersten Durchgang zeigt eine deutlich erhöhte Anzahl an Ausreißern (siehe Abbildung 11b)). Die Lesedauer der Probandinnen und Probanden liegt zwischen 15 und 96 Sekunden, wobei der Medianwert bei 48 Sekunden liegt. Die zentralen 50% der Kinder bewegen sich innerhalb einer Interquartilsspanne von etwa 40 bis 65 Sekunden (siehe Abbildung 11b)). Auffällig ist, dass der höchste Ausreißer in beiden Boxplots auf dasselbe Kind zurückzuführen ist. Beim ersten Lesen sind die Unterschiede zwischen den Schülerinnen und Schülern größer, was sich in mehreren Ausreißern zeigt.

	Arithmetisches Mittel	Median	SD	Minimum	Maximum
T-Wert (PPVT)	53,07	52,50	8,36	35,00	73,00
Input (in h/Tag)	3,80	3,50	1,38	1,86	8,29
SES	15,61	15,82	2,62	9,40	20,10
Total Reading Time (in s)	67,06	65,00	26,93	18,00	140,00
Reading Time First Trial (in s)	53,12	48,00	21,79	15,00	140,00

Tabelle 7: Übersicht der deskriptiven Werte; SD = standard derivation

Innerhalb der Stichprobe ließen sich keine signifikanten Unterschiede im sozioökonomischen Status feststellen. Auch zwischen monolingualen und bilingualen Kindern zeigen sich keine bedeutsamen Differenzen hinsichtlich des sozioökonomischen Status, des sprachlichen Inputs sowie des rezeptiven Wortschatzes. Da die beobachteten Unterschiede in den Bereichen sozioökonomischer Status und Mehrsprachigkeit nicht signifikant sind und somit keine aussagekräftigen Aufschlüsse auf einen Zusammenhang mit der Lesekompetenz zulassen, werden diese Variablen im weiteren Verlauf nicht in das gemischte Modell einbezogen (siehe Abbildung 9d)). Die deskriptive Analyse zeigt, dass im Rahmen der gemischten Modelle eine Differenzierung der betrachteten Lesezeiten erforderlich ist. Im ersten Modell dient die Gesamtlesezeit als Einflussvariable. Diese kann jedoch nur sinnvoll berücksichtigt werden, wenn die Anzahl der Lese-Durchgänge mit in die Modellierung einbezogen wird, da sie einen direkten Einfluss auf die Gesamtzeit nimmt. Im zweiten Modell hingegen wird ausschließlich die Lesezeit des ersten Durchgangs analysiert. Da hierbei wirklich nur die Zeit des ersten Lesens aus den Eye-Tracking-Daten entnommen wird, entfällt die Notwendigkeit, die Anzahl der Durchgänge zu berücksichtigen.

3.8.2 Auswertung der gemischten Modelle

Im Rahmen der inferenzstatistischen Auswertung wurden zwei gemischte Modelle berechnet: ein lineares gemischtes Modell (LMM) für metrische Zielvariablen sowie ein verallgemeinertes lineares gemischtes Modell (GLMM) für kategoriale bzw. nicht-normalverteilte Zielvariablen. Zum Verständnis der beiden gemischten Modelle sind einige Definitionen wichtig. So handelt es sich bei dem Schätzwert (Estimate), der in beiden Tabellen zu finden ist, um den geschätzten Einfluss, die der jeweilige feste Effekt auf die Zielvariable hat (Bates et al., 2015). Die Standardabweichung (SD) stellt die Unsicherheit des Schätzwertes dar. Der z-Wert bildet der Quotient aus dem Schätzwert sowie dem Standardfehler und wird verwendet, um die Signifikanz der Effekte zu prüfen. Der p-Wert lässt

sich aus dem z-Wert berechnen und stellt die statistische Signifikanz dar. Dieser gibt an, wie hoch die Wahrscheinlichkeit ist, dass der auftretende Effekt zufällig ist. Es wird zwischen fünf Signifikanzniveaus unterschieden (siehe Tabelle 8).

p-Wert-Bereich	**Code**	**Signifikanzniveau**
p ≤ 0,001	***	hoch signifikant
p ≤ 0,01	**	sehr signifikant
p ≤ 0,05	*	signifikant
p ≤ 0,1	.	tendenziell signifikant
p > 0,1		nicht signifikant

Tabelle 8: Zuordnung der p-Werte zu den Signifikanzniveaus

Das GLMM wurde zur Identifikation potentieller Einflussfaktoren auf das Textverstehen berechnet, operationalisiert als die Anzahl an korrekten Antworten (siehe Tabelle 9). Die folgenden Ergebnisse beziehen sich auf die festen Effekte (Fixed-Effects[12]) des Modells, wobei alle Faktoren bis auf die Anzahl an Lesedurchgängen (TrialCount) und die Schwierigkeit des Textes (difficult) z-standardisiert wurden. Dies erlaubt eine direkte Vergleichbarkeit der Effektstärken.

Die Variable *PPVT_z*, die den T-Wert des rezeptiven Wortschatzes darstellt, weist einen hochsignifikanten positiven Effekt mit p = 0,000625 auf das Textverstehen auf. Auch die Variable *difficult* weist einen stark signifikanten Effekt (p = 0,003649) auf. Hierbei handelt es sich jedoch um einen negativen Effekt mit einem Schätzwert von -1,11524 auf die korrekte Antwortzahl.

Die übrigen vier festen Effekte zeigen keine statistisch signifikanten Effekte (p > 0,1). Der sprachliche Input (Input_z, p = 0,738829), die durchschnittliche Fixationsdauer (Avg_FixDur_z, p = 0,593883), die durchschnittliche Anzahl an Fixation (Avg_FixCount_z, p = 0,815233) und die Anzahl an Lesedurchgängen (TrialCount, p = 0,284526) zeigen keine statistisch signifikanten Effekte auf das Textverstehen der Schülerinnen und Schüler. Tendenziell weist die Gesamtlesezeit (ReadingTime_z) mit einem positiven Koeffizienten von 0,85 auf eine bessere Leistung bei längerer Lesedauer hin, jedoch reicht der p-Wert (p = 0,105) nicht aus, um diesen Zusammenhang zu bestätigen.

[12] Fixed-Effects sind die durchschnittlichen Werte des Intercept und der Slope, mit denen die Prädiktorvariablen im Mittel die Zielgröße beeinflussen. Sie beschreiben damit den generellen Effekt über alle statistischen Einheiten hinweg (Bates et al., 2015).

Einflussfaktor	Estimate (E)	SD	z-Wert	p-Wert
(Intercept)	3,06795	1,13687	2,699	0,006963 **
Input_z	- 0,06788	0,20358	- 0,333	0,738829
ReadingTime_z	0,85286	0,52685	1,619	0,105492
PPVT_z	0,79402	0,23213	3,421	0,000625 ***
Avg_FixDur_z	- 0,21112	0,39593	- 0,533	0,593883
Avg_FixCount_z	0,09424	0,40328	0,234	0,815233
TrialCount	- 0,74909	0,69995	- 1,070	0,284526
difficult	- 1,11524	0,38364	- 2,907	0,003649 **

Tabelle 9: Generalized linear mixed model: Textverstehen – fixed effects; SD = standard derivation

Zur Analyse der Einflussfaktoren auf die Leseflüssigkeit wurde ein LMM berechnet. Als Zielvariable diente die Zeit, die die Kinder für das Lesen des jeweiligen Textes im ersten Durchgang benötigten (siehe Kapitel 3.2). Ziel des LLMs ist es Einflussfaktoren zu identifizieren, die signifikant mit der Leseflüssigkeit zusammenhängen (siehe Tabelle 10). Als feste Effekte gingen die Faktoren des sprachlichen Inputs (Input_z), des rezeptiven Wortschatzes (PPVT_z), des Schwierigkeitsgrads und der durchschnittlichen Anzahl an Fixationen pro Wort (Avg_FixCount_z) ein. Die Einflussfaktoren wurden z-standardisiert. Die Variable *difficult* wurde nicht standardisiert, sondern als binäre Variable kodiert (1 = leicht; 2 = schwer).

Das Modell ergab, dass ausschließlich der Schwierigkeitsgrad einen statistisch hochsignifikanten Einfluss auf die Leseflüssigkeit aufweist (Estimate = 14,1054, t = 9,939, p = $1{,}11e^{-15}$). Ein höherer Wert in *difficult* geht mit einer höheren Lesezeit einher, was sich in einem positiven Schätzwert widerspiegelt. Die weiteren Einflussfaktoren *Input_z* (p = 0,677), *PPVT_z* (p = 0,733) sowie *Avg_FixCount_z* (p = 0,532) zeigten keinen statistisch signifikanten Zusammenhang mit der Leseflüssigkeit in der untersuchten Stichprobe.

Einflussfaktor	Estimate (E)	SD	df	t-Wert	p-Wert
(Intercept)	46,0634	4,0054	25,3725	11,500	$1{,}47e^{-11}$ ***
Input_z	- 1,6547	3,9216	25,0054	- 0,422	0,677
PPVT_z	- 1,3538	3,9231	25,0424	- 0,345	0,733
difficult	14,1054	1,4192	81,0997	9,939	$1{,}11e^{-15}$ ***
Avg_FixCount_z	- 0,6195	0,9861	81,3930	- 0,628	0,532

Tabelle 10: Linear mixed model - Leseflüssigkeit; df = degress of freedom

Die Analyse der beiden Modellvarianten verdeutlicht, dass der Schwierigkeitsgrad der Texte einen signifikanten Einfluss sowohl auf die Leseflüssigkeit als auch auf das Textverstehen der Kinder ausübt. Darüber hinaus zeigt sich, dass der rezeptive Wortschatz der Lernenden einen eigenständigen und bedeutsamen Beitrag zum Textverstehen leistet, während er die Leseflüssigkeit nicht signifikant beeinflusst. Im Gegensatz dazu erweisen sich die weiteren untersuchten Variablen in beiden Modellen als statistisch nicht signifikant.

3.9 Diskussion

Der vorliegenden Untersuchung lagen mehrere untersuchungsleitende Fragestellungen zugrunde, die den Einfluss verschiedener Variablen auf die Lesekompetenz von Lernenden am Ende der zweiten Klasse in den Fokus nahmen. Zur differenzierten Betrachtung wurde die Lesekompetenz in die Komponenten der Leseflüssigkeit und des Textverstehens untergliedert. Die deskriptive Auswertung wurde der Analyse mittels gemischter Modelle bewusst vorgelagert, um die Eignung der Verteilung aller erhobenen Einflussfaktoren innerhalb der Stichprobe zu überprüfen.

3.9.1 Interpretation der deskriptiven Ergebnisse

Deskriptiv untersucht wurden der rezeptive Wortschatz, der sprachliche Input, der sozioökonomische Status, die Gesamtlesedauer der Kinder und die Lesedauer des ersten Durchganges. Die Analyse des Wortschatzes zeigt, dass sich mehr als zwei Drittel der Stichprobe im oberen Normalbereich bis hin zum stark überdurchschnittlichen Bereich befinden. Dieses Ergebnis weist auf eine insgesamt gut entwickelte rezeptive sprachliche Kompetenz innerhalb der Stichprobe hin. Es könnte darauf schließen lassen, dass die Lernumgebung den Erwerb rezeptiver Wortschatzkenntnisse besonders begünstigt.

Im Bereich des sprachlichen Inputs zeigt sich eine insgesamt homogene Verteilung. Die Kinder erhalten durchschnittlich zwischen circa drei bis fünf Stunden sprachlichen Input pro Tag. Dabei bezieht sich die Erhebung ausschließlich auf die Quantität des Inputs. Qualitative

Unterschiede etwa zwischen sprachlich anspruchsvollen Aktivitäten wie dem Vorlesen und alltagsbezogenen Interaktionen wie dem Spielen mit Spielzeug wurden nicht berücksichtigt. Zwar zeigen Studien wie die von Hoff et al. (2024), dass die Qualität des sprachlichen Inputs sich in den sprachlichen Aktivitäten unterscheidet. Jedoch ist die Erhebung qualitativer Merkmale über Elternfragebögen methodisch herausfordernd, da zentrale Dimensionen wie Interaktivität, linguistische Komplexität und konzeptionelle Tiefe (Rowe & Snow, 2020) schwer objektiv zu erfassen sind und stark von subjektiven Einschätzungen der Eltern abhängen. Eine differenzierte Bewertung dieser qualitativen Aspekte erfordert kontextgebundene Beobachtungen, die über standardisierte Selbstauskünfte hinausgehen. Aus diesem Grund fokussiert die vorliegende Untersuchung die Quantität des Inputs. Die Analyse zeigt, dass es sich um eine weitgehend homogene Gruppe handelt, in der die Kinder vergleichbare tägliche Inputzeiten erhalten, was sich auch in der geringen Streuung innerhalb des Boxplots widerspiegelt.

Die deskriptiven Ergebnisse des sozioökonomischen Status zeigt, dass es sich bei der untersuchten Stichprobe um Kinder mit einem insgesamt hohen sozioökonomischen Status handelt. Der Durchschnitt befindet sich in einem Bereich zwischen 14 und 17 Punkten auf einer Skala von drei bis 21 Punkten. Der niedrigste Wert liegt bei 9,4, was ebenfalls auf eine eingeschränkte Streuung hinweist. Diese geringe Varianz im sozioökonomischen Status erschwert es, belastbare Aussagen über den Einfluss des sozioökonomischen Status auf die Lesekompetenz zu treffen, da mögliche Unterschiede innerhalb der Stichprobe statistisch kaum abbildbar sind. Aus diesem Grund wird der sozioökonomische Status in der weiteren Untersuchung mittels gemischter Modelle nicht beachtet. Unabhängig von der eingeschränkten Varianz zeigt sich jedoch, dass ein hoher sozioökonomischer Status mit einem erweiterten Wortschatz einhergeht. Dies spiegelt die bisherigen empirischen Befunde wider (Aktas, 2020; Calvo & Bialystok, 2014; Ehl & Grosche, 2020; McElvany et al., 2009). In den Studien wurde deutlich, dass familiäre Strukturmerkmale wie Bildung, Einkommen aber auch sprachlicher Input der Eltern signifikant mit der sprachlichen Entwicklung von Kindern korrelieren, insbesondere mit dem Wortschatz.

Auch im Hinblick auf die Variable der Mehrsprachigkeit zeigten sich in den Bereichen des Wortschatzes, des sozioökonomischen Status sowie der Quantität des sprachlichen Inputs keine signifikanten Unterschiede zwischen ein- und mehrsprachigen Kindern. Es ist zudem zu berücksichtigen, dass lediglich sechs Kinder der Stichprobe mehrsprachig aufwachsen, wodurch keine statistisch belastbare oder repräsentative Aussage über diese Gruppe getroffen werden kann.

3.9.2 Interpretation der GLMM-Ergebnisse – Textverstehen

Nach der deskriptiven Analyse einzelner Variablen erfolgte die Auswertung der gemischten Modelle, wie in Kapitel 3.8 erläutert. Ziel war es, die Einflussfaktoren auf die Lesekompetenz differenziert zu untersuchen. Dabei wurde die Lesekompetenz in die Dimensionen Textverstehen und Leseflüssigkeit unterteilt. Die folgenden Abschnitte widmen sich der Interpretation der Ergebnisse im Hinblick auf die aufgestellten Hypothesen. Hierzu werden die beiden Modelle zuerst getrennt voneinander interpretiert.

Das GLMM zur Untersuchung des Textverstehens analysiert eine Reihe potenzieller Einflussfaktoren auf die Anzahl korrekter Antworten, die Kinder im Anschluss an einen Lesetext gegeben haben. Zur Erinnerung an die Operationalisierung: Das Textverstehen wurde über die quantifizierte Anzahl richtiger Antworten erfasst (siehe Kapitel 3.2). Im Modell berücksichtigt wurden folgende Einflussfaktoren: die Gesamtlesedauer, die Quantität des sprachlichen Inputs, der rezeptive Wortschatz, die Schwierigkeit des Textes, die Anzahl der Lesedurchgänge sowie zwei Eye-Tracking-Metriken: die durchschnittliche Anzahl an Fixationen pro Wort und die durchschnittliche Fixationsdauer.

Die erste Hypothese stellt einen positiven Zusammenhang zwischen der Gesamtlesedauer und dem Textverstehen dar (H1). Erwartet wurde, dass eine kürzere Lesedauer mit einer höheren Anzahl an korrekten Antworten korreliert. Die Ergebnisse des GLMM zeigen jedoch keinen signifikanten Einfluss der Gesamtlesedauer auf das Textverstehen der Kinder. Obwohl der positive Regressionskoeffizient (E = 0,85) der z-standardisierten Gesamtlesezeit eine Tendenz zu verbesserten Leseleistungen bei längerer Lesedauer anzeigt (vgl. Kapitel 3.8.2), erreicht dieser mit p = 0,105492 nicht die konventionelle Signifikanzgrenze. Dieser Befund steht im Widerspruch zu früheren Untersuchungen, nach denen geübte Leserinnen und Leser durchschnittlich kürzere Gesamtlesezeiten aufweisen als weniger erfahrene Lesende (Rayner 1998; Wertli et al., 2023). Eine plausible Erklärung für den beobachteten positiven Zusammenhang ist die sogenannte Wiederholungsstrategie, bei der Lernende einen Text mehrfach rekapitulieren, um Verständnislücken zu schließen oder Lesegenauigkeit zu überprüfen. Diese Strategie kann die Wahrscheinlichkeit korrekter Antworten erhöhen, ohne dass daraus ein kausaler Effekt der Lesedauer auf das Textverstehen abgeleitet werden kann (Jesch, 2020).

Auch die Hypothese (H2a), dass eine höhere Quantität des sprachlichen Inputs mit einer höheren Wahrscheinlichkeit für mehr korrekte Antworten einhergeht, kann statistisch nicht bestätigt werden (p= 0,738829). Hierbei ist zu beachten, dass ausschließlich die Quantität des Inputs als Einflussfaktor berücksichtigt wurde, ohne die Qualität des Inputs heranzuziehen, welcher durch den Fragebogen nicht objektiv zu erfassen war. Dieser Befund legt nahe, dass

qualitative Merkmale des sprachlichen Inputs möglicherweise eine größere Rolle für das Textverstehen spielen. So legt die aktuelle Forschung nahe, dass die Qualität des sprachlichen Inputs entscheidender ist als dessen Häufigkeit (Masek et al., 2021; Richardson, 2023). Hoff et al. (2024) zeigen in einer Untersuchung mit 155 Eltern-Kind-Paaren, dass die sprachliche Elaboriertheit beim gemeinsamen Lesen und autobiografischen Erinnern signifikant höher ist als im freien Spiel. Vor diesem Hintergrund erscheint die Einbeziehung qualitativer Inputkomponenten, etwa spezifischer Interaktionskontexte oder der syntaktischen Tiefe, für das Textverstehen von zentraler Bedeutung, lässt sich im Rahmen der vorliegenden Studie jedoch nicht vertiefend analysieren.

Im Gegensatz zum sprachlichen Input zeigt sich beim rezeptiven Wortschatz ein signifikanter Einflussfaktor auf das Textverstehen der Kinder. Die Ergebnisse zeigen, dass Kinder mit einem höheren Wortschatz eine höhere Wahrscheinlichkeit aufweisen, mehr korrekte Antworten zu geben (H3a). Mit einem hohen Schätzwert (E = 0,794) weist der rezeptive Wortschatz auf einen substanziellen positiven Effekt hin. Die Hochsignifikanz des Zusammenhangs (p = 0,000625) unterstreicht die Relevanz dieses Befundes. Eine mögliche Erklärung hierfür ist, dass Kinder mit einem größeren Wortschatz über mehr semantisches Vorwissen verfügen und somit Texte besser ganzheitlich sowie satzübergreifend erfassen können. Darüber hinaus könnte ein umfassender Wortschatz zu einer geringeren Belastung des Arbeitsgedächtnisses führen, wodurch kognitive Ressourcen für das Textverstehen statt für das Dekodieren frei werden (Lenhard, 2024).

Auch die Hypothese, dass ein einfacher Text dazu führt, dass mehr korrekte Antworten gegeben werden, kann anhand der Ergebnisse des GLMMs statistisch bestätigt werden (H4a). Die Ergebnisse zeigen, dass die Schwierigkeit eines Textes einen negativen Einfluss auf das Textverstehen hat (E = -1,11524). Das heißt, dass mit zunehmender Schwierigkeit auch die Wahrscheinlichkeit an korrekten Antworten substantiell abnimmt. Hierbei handelt es sich um einen sehr signifikanten Einfluss (p = 0,003649). Als zentrale Einflussfaktoren für die Textschwierigkeit wurden Anaphern, Nebensätze und die Wortfrequenz genutzt, um höhere Profilstufen zu erreichen (Grießhaber, 2024). Diese Merkmale erschweren nicht nur die semantische Verarbeitung, sondern führen zu einer erhöhten kognitiven Belastung: Kinder müssen mehr Begriffe dekodieren, satzübergreifende Bezüge herstellen und unbekannte Wörter verknüpfen (Lenhard, 2024; Rosebrock & Nix, 2025; Schaffner, 2009). Dadurch wird das Arbeitsgedächtnis stärker beansprucht, was die Verarbeitung zeitlich verlängert und die inhaltliche Erschließung erschwert.

Im Bereich des Eye-Tracking wurden die Einflussfaktoren der durchschnittlichen Anzahl an Fixationen pro Wort und die durchschnittliche Fixationsdauer untersucht. Im Vorhinein wurde

angenommen, dass eine geringere Anzahl an Fixationen mit einer höheren Wahrscheinlichkeit für mehr korrekte Antworten einhergeht (H5a). Entgegen der formulierten Hypothese zeigt der positive Schätzwert (E = 0,09), dass eine höhere durchschnittliche Fixationsanzahl pro Wort tendenziell mit einer besseren Beantwortung der Verständnisfragen einhergeht. Dieser Effekt ist jedoch statistisch nicht signifikant (p = 0,815), sodass kein Zusammenhang abgeleitet werden kann. Möglicherweise reflektiert eine erhöhte Anzahl an Fixationen intensivere kognitive Verarbeitung oder eine stärkere Auseinandersetzung mit anspruchsvollen Textstellen. Somit könnten hohe Fixationszahlen nicht zwangsläufig auf Schwierigkeiten hinweisen, sondern auch auf ein tieferes Verstehen der Inhalte. Allerdings stellen bisherige Studien heraus, dass eine geringere Fixationsanzahl mit einem effizienteren Lesen einhergeht (Häikiö et al., 2009; Hautala et al., 2024; Krause & Ritter, 2022; Rettig & Schiefele, 2023; Wertli et al., 2023).

Hypothese H6 besagt, dass eine geringere durchschnittliche Fixationsdauer mit besseren Verständnisleistungen einhergeht. Zwar weist der negative Schätzwert (E = -0,21112) auf einen solchen Zusammenhang hin, konnte jedoch ebenfalls nicht signifikant bestätigt werden (p = 0,594). Dieses Ergebnis stützt die aktuellen Eye-Tracking-Befunde, wonach verkürzte Erst- und Refixationsdauern auf automatisierte, arbeitsgedächtnisgestützte Wortverarbeitung und eine gestärkte orthografische Kompetenz hinweisen. Dadurch werden kognitive Ressourcen frei, die für tiefergehende semantische Verarbeitungsprozesse notwendig sind. Erfahrene Leserinnen und Leser zeigen kürzere Fixationen und nutzen ihr peripheres Sichtfeld effektiver, was flüssigeres Lesen und eine bessere Integration von Textinformationen ermöglicht (Häikiö et al., 2009; Hautala et al., 2024; Krause & Ritter, 2022; Rettig & Schiefele, 2023; Wertli et al., 2023). Die fehlende Signifikanz lässt sich durch methodische Einschränkungen wie der geringen Stichprobengröße erklären.

Die letzte Hypothese im Bereich des Textverstehens (H7) bezog sich auf die Anzahl der Lesedurchgänge. Da den Kindern freigestellt war, ob sie den Text einmal oder zweimal lesen, konnte untersucht werden, ob mehrfaches Lesen mit mehr korrekten Antworten einhergeht. Entgegen der Erwartung zeigten Kinder mit nur einem Lesedurchgang mehr korrekte Antworten (E = -0,74909). Dieser Effekt ist jedoch statistisch nicht signifikant, sodass kein belastbarer Einfluss auf das Textverstehen abgeleitet werden kann (p = 0,284526). Ein möglicher Erklärungsansatz für die ausbleibende Signifikanz liegt in der geringen Anzahl doppelt gelesener Texte (36 von 116), wodurch die Vergleichbarkeit und Aussagekraft dieser Variable eingeschränkt ist.

Zusammenfassend zeigen die Ergebnisse des GLMMs, dass bei der vorliegenden Stichprobe lediglich zwei Variablen einen signifikanten Einfluss auf das Textverstehen von Kindern am

Ende der zweiten Klasse ausüben: der rezeptive Wortschatz sowie das Schwierigkeitsniveau der Lesetexte.

3.9.3 Interpretation der LMM-Ergebnisse – Leseflüssigkeit

Den zweiten Teilaspekt der Lesekompetenz stellt in dieser Studie die Leseflüssigkeit dar, die anhand der Lesedauer des ersten Lesedurchganges gemessen wurde. Die gemessenen Einflussfaktoren stellen in diesem LMM die Quantität des sprachlichen Inputs, der rezeptive Wortschatz, die Schwierigkeit des Textes sowie die durchschnittliche Anzahl an Fixationen pro Wort dar.

Im Rahmen der Untersuchung wurde die Hypothese aufgestellt, dass eine höhere Quantität des sprachlichen Inputs mit einer geringeren Lesedauer und folglich einer höheren Leseflüssigkeit einhergeht (H2b). Diese Hypothese kann jedoch durch die Ergebnisse des LMM statistisch nicht bestätigt werden. Konkret zeigt das LMM, dass der sprachliche Input keinen signifikanten Einfluss auf die Leseflüssigkeit der Schülerinnen und Schüler hat ($p = 0{,}677$). Zudem weist der negative Schätzwert ($E = -1{,}6547$) auf einen negativen Zusammenhang hin. Dieser lässt darauf schließen, dass die Qualität des sprachlichen Inputs eine wichtigere Rolle einnimmt als die Quantität des Inputs. Im Kontext bestehender Forschungen zum sprachlichen Input wird vermehrt darauf hingewiesen, dass insbesondere die Qualität des Inputs einen bedeutsamen Einfluss auf die Sprach- und Lesekompetenz von Kindern haben könnte als die reine Quantität (Calvo & Bialystok, 2014; Hoff et al., 2024; Masek et al., 2021; Richardson, 2023; Rowe & Snow, 2020). In der vorliegenden Studie wurde jedoch keine Differenzierung hinsichtlich qualitativer Merkmale sprachbezogener Aktivitäten vorgenommen. Beispielsweise kann ein Kind, das wöchentlich vier Stunden mit Vorlesen oder Lesen verbringt, formal einen geringeren Input aufweisen als ein Kind, das acht Stunden wöchentlich Serien konsumiert. In diesem Fall liegt zwar eine höhere Quantität vor, jedoch handelt es sich laut Literatur um sprachlich weniger hochwertigen Input, der gegebenenfalls kaum zur Förderung der Leseflüssigkeit beiträgt (Hoff et al., 2024).

In Bezug auf den rezeptiven Wortschatz zeigt die Analyse der Daten, dass die Hypothese, dass ein größerer rezeptiver Wortschatz mit einer geringeren Lesedauer einhergeht, nicht statistisch bestätigt werden kann (H3b). Der entsprechende p-Wert liegt deutlich über der Signifikanzgrenze ($p = 0{,}733$), was auf einen fehlenden signifikanten Einfluss hindeutet. Darüber hinaus zeigt sich auch in diesen Daten, dass es sich um einen negativen Schätzwert handelt ($E = -1{,}3538$). Dieses Ergebnis legt nahe, dass der Dekodierprozess auch ohne umfangreiches Wortschatzwissen effizient ablaufen kann und stattdessen vermehrt auf phonologische und orthografische Strategien zurückgegriffen wird (Rosebrock & Nix, 2025). Dieser Befund deutet darauf hin, dass sprachliche Einheiten zwar zügig identifiziert,

jedoch unter Umständen nicht semantisch verarbeitet werden. Daraus ergibt sich, dass die Leseflüssigkeit allein kein valider Indikator für das Textverstehen ist, da auch Kinder mit einem umfangreichen Wortschatz mitunter längere Lesezeiten aufweisen. Aufgrund der fehlenden statistischen Signifikanz ist dieser Zusammenhang jedoch mit Vorsicht zu interpretieren.

Als weiterer Einflussfaktor auf die Leseflüssigkeit wurde der Schwierigkeitsgrad der Texte betrachtet. Im Vorhinein wurde die Hypothese aufgestellt, dass ein geringerer Schwierigkeitsgrad dazu führt, dass die Lesedauer des ersten Lesedurchganges geringer ist (H4b). Das Ergebnis des LMM zeigt einen hohen positiven Schätzwert (E = 14,1054). Dieser bestätigt die Richtigkeit der Hypothese, dass eine geringere Schwierigkeit mit einer höheren Leseflüssigkeit einhergeht. Zudem bestätigt der p-Wert ($p = 1{,}11e^{-15}$ ***), dass es sich um einen hochsignifikanten Einflussfaktor auf die Leseflüssigkeit handelt, sodass die Hypothese statistisch verifiziert werden kann. Ein möglicher Erklärungsansatz für den hochsignifikanten Einfluss des Schwierigkeitsgrads auf die Leseflüssigkeit liegt nicht allein in der geringeren Wortanzahl der einfachen Texte, sondern auch in deren sprachlicher Gestaltung. Texte mit höherem Schwierigkeitsgrad enthalten häufiger niedrigfrequente Wörter, komplexe Satzstrukturen mit Nebensätzen sowie Anaphern (Grießhaber, 2024). Diese führen zu einem höheren Maß an kognitiver Verarbeitung, um den Text kohärent erfassen zu können (Rosebrock, 2022). Im Gegensatz dazu zeichnen sich die einfacheren Texte durch eine klare Syntax, häufige Wortformen und weniger Satzverschachtelung aus, was ein flüssiges Lesen begünstigt (Grießhaber, 2024).

Die abschließende Hypothese befasst sich mit der durchschnittlichen Anzahl an Fixationen pro Wort und sagt aus, dass eine geringere Anzahl an Fixationen mit einer geringeren Lesedauer einhergeht (H5b). Diese Annahme konnte durch die vorliegenden Daten nicht bestätigt werden, da kein signifikanter Zusammenhang festgestellt wurde (p = 0,532). Der negative Schätzwert (E = -0,6195) deutet jedoch auf einen inversen Zusammenhang hin. Ein möglicher Grund liegt in der unterschiedlichen zeitlichen Ausprägung der Fixationen: Probandinnen und Probanden mit einer geringeren Fixationsanzahl zeigen tendenziell längere durchschnittliche Fixationsdauern, während Probandinnen und Probanden mit einer höheren Fixationsfrequenz ihre Blickbewegungen innerhalb eines vergleichsweise kürzeren Zeitraums realisieren. Dadurch können erstere zwar verstärkt auf periphere Informationen zurückgreifen, benötigen jedoch mehr Zeit für die kognitive Verarbeitung der gleichzeitig aufgenommenen Reize (Häikiö et al., 2009; Rosebrock, 2022). Im Gegensatz dazu weisen Probandinnen und Probanden mit hoher Fixationsanzahl eine geringere Gesamtlesedauer auf, sodass sie innerhalb der einzelnen Fixationen kleinere Informationseinheiten erfassen und diese automatisiert und schneller verarbeiten können als größere, zusammenhängende Einheiten (Lenhard, 2024; Rosebrock & Nix, 2025). Da dieser Effekt jedoch statistisch nicht signifikant ist, kann kein

valider Rückschluss auf die Grundgesamtheit der Zweitklässlerinnen und Zweitklässler gezogen werden.

Für das LMM in dieser Stichprobe zeigt sich, dass ausschließlich die Schwierigkeit des Textes einen signifikanten Einfluss auf die Leseflüssigkeit der Schülerinnen und Schüler am Ende des zweiten Schuljahres hat. Dieser Einfluss stellt sich als hochsignifikant heraus.

3.9.4 Gesamtdiskussion

Da das Textverstehen und die Leseflüssigkeit zentrale Teilaspekte der Lesekompetenz darstellen, ist eine zusammenführende Betrachtung der Modellbefunde von besonderer Relevanz. Im Folgenden werden die Ergebnisse der beiden Modelle gegenübergestellt und im Kontext bestehender Forschungsliteratur diskutiert.

Mit Blick auf die Ergebnisse beider gemischten Modelle zeigt sich, dass lediglich eine Variable sowohl auf das Textverstehen als auch auf die Leseflüssigkeit einen signifikanten Einfluss ausübt. Dabei handelt es sich um den Schwierigkeitsgrad der Texte. Ein geringerer Schwierigkeitsgrad korreliert dabei sowohl mit einer höheren Anzahl korrekter Antworten als auch mit einer geringeren Lesedauer. Als Ursache wurde bereits bei der Interpretation der Modelle die Struktur der Texte aufgegriffen. Texte mit höherem Schwierigkeitsgrad zeichnen sich durch komplexere Satzstrukturen, niedrigfrequente Wortformen sowie Anaphern aus. Diese Merkmale erhöhen die kognitive Belastung beim Lesen, da Kinder mehr Zeit benötigen, um Wörter und syntaktische Konstruktionen zu dekodieren, was sich in einer verlängerten Lesedauer niederschlägt. Dieses Ergebnis ist konsistent mit bisherigen Befunden aus der Leseforschung. Laut Jesch (2020) sowie Lenhard (2024) werden vertraute und häufig vorkommende Wörter am Ende der Schuleingangsphase zunehmend automatisiert dekodiert. Schwierige Texte hingegen enthalten häufig neue oder niedrigfrequente Wörter, deren Dekodierung noch nicht automatisiert erfolgt. Im Hinblick auf das Textverstehen bedeutet die verlängerte Dekodierungszeit, dass Kinder während des Lesens bereits erste Textinformationen aus dem Arbeitsgedächtnis verlieren könnten. Dies beeinträchtigt die Fähigkeit, Inhalte zu verknüpfen und sinnvoll zu interpretieren. Auch dies wird durch die Literatur gestützt: Die fehlende Automatisierung der Worterkennung führt dazu, dass kognitive Ressourcen verstärkt für das Entschlüsseln komplexer Sätze und neuer Wörter aufgewendet werden müssen, wodurch weniger Kapazitäten für die inhaltliche Verarbeitung verfügbar sind (Jesch, 2020; Lenhard, 2024; Schaffner, 2009).

Keine weitere Variable ist sowohl für das Textverstehen als auch für die Leseflüssigkeit signifikant. Der rezeptive Wortschatz ist der einzige Faktor, der ebenfalls eine hohe Signifikanz zeigt, jedoch nur beim Textverstehen. Im Bereich der Leseflüssigkeit stellt dieser sich als nicht

signifikant heraus. Wie bereits in Kapitel 3.9.2 erläutert, kann das Vorwissen möglicherweise eine Erklärung dafür sein. Ein größerer rezeptiver Wortschatz führt dazu, dass das Arbeitsgedächtnis entlastet werden kann, da weniger neue Bedeutungen kennengelernt, verarbeitet sowie in Beziehung zu den schon bestehenden gesetzt werden müssen (Lenhard, 2024). Dieses Ergebnis spiegelt vorherige Forschungsresultate wider, in denen festgestellt wurde, dass ein hoher Wortschatz zu einer Entlastung des Arbeitsgedächtnis führen kann und höhere kognitive Verstehensprozesse stattfinden können (Lenhard, 2024; McElvany & Schneider, 2009; Rosebrock & Nix, 2025). Da der rezeptive Wortschatz jedoch keinen signifikanten Einfluss auf die Leseflüssigkeit zeigt, lässt sich ableiten, dass dieser für den reinen Lesevorgang, im Sinne der technischen Dekodierung auf Wort- und Satzebene, von untergeordneter Bedeutung ist. Als möglicher Erklärungsansatz kommt in Betracht, dass die phonologische und orthografische Wortbildung bei den untersuchten Kindern bereits weitgehend automatisiert ist. In diesem Fall würde ein umfangreicher rezeptiver Wortschatz nicht zwingend zu schnelleren Decodierungsprozessen führen, da die notwendigen Worterkennungsmuster bereits etabliert sind (Rosebrock & Nix, 2025). Eine alternative Deutung wäre, dass die Schülerinnen und Schüler den vorgelegten Text vorwiegend oberflächlich lesen, ohne ein tiefgehendes Textverstehen anzustreben (Rayner et al., 2016). Die vorliegenden Ergebnisse legen nahe, dass die Leseflüssigkeit keinen zuverlässigen Prädiktor für eine umfassende Lesekompetenz darstellt. Besonders bei Kindern mit hohem rezeptivem Wortschatz und ausgeprägtem Textverstehen zeigt sich, dass eine erhöhte Lesegeschwindigkeit nicht zwangsläufig gegeben ist. Dieses Befundmuster verweist auf eine notwendige Unterscheidung zwischen basaler Dekodierfähigkeit und tiefergehender semantischer Verarbeitung. Auffällig ist, dass dieser Aspekt in bisherigen Eye-Tracking-Studien weitgehend unberücksichtigt blieb.

Auch die Ergebnisse im Bereich der Gesamtlesezeit zeigen, dass diese keinen signifikanten Einfluss auf das Textverstehen hat. Daraus lässt sich schließen, dass ein höheres Lesetempo nicht zwangsläufig mit einem besseren Textverstehen einhergeht. Dieses Resultat steht im Widerspruch zu gängigen standardisierten Verfahren im Grundschulbereich, wie etwa dem Stolperlesetest, der simultan Lesetempo, Lesegenauigkeit und Leseverständnis zu erfassen versucht (Metze, 2024). Die Befunde der aktuellen Studie stellen jedoch die Validität einer solchen kombinierten Messung infrage, da das Lesetempo offenbar keinen entscheidenden Beitrag zum Textverstehen leistet. Es ist somit zu hinterfragen, inwieweit ein hoher Zeitdruck in der Testdurchführung Kinder zu oberflächlichem Lesen verleitet und dabei das tatsächliche Verständnis des Textinhalts beeinträchtigt (Mézière et al., 2025).

In beiden gemischten Modellen erweist sich die Quantität des sprachlichen Inputs als kein signifikanter Faktor. Es zeigt sich, dass dessen reine Quantität keinen bedeutsamen Einfluss

auf die Leseflüssigkeit oder das Textverstehen nimmt. Dieser Befund steht im Einklang mit der aktuellen Forschungslage, die die Qualität des Inputs gegenüber der bloßen Häufigkeit priorisiert (Hoff et al., 2024; Masek et al., 2021; Rowe & Snow, 2020). Zwar stellt häufiger Sprachkontakt eine grundlegende Voraussetzung für den Erwerb sprachlicher und schriftsprachlicher Kompetenzen dar, jedoch scheint die Qualität dieses Kontakts der entscheidende Faktor zu sein. Kinder profitieren demnach in besonderem Maße von sprachlich hochwertigen Aktivitäten, beispielsweise dem Lesen oder elaborierten altersgerechten Konversation, was sich in einem höheren Kompetenzniveau zeigt (Hoff et al., 2024; Masek et al., 2021; Rowe & Snow, 2020). Die Ergebnisse der vorliegenden Studie bestätigen diese Einschätzung, indem sie zeigt, dass die Quantität des Inputs weder auf die Leseflüssigkeit noch auf das Textverstehen einen signifikanten Einfluss ausübt.

Im Hinblick auf die Häufigkeit der Lesedurchgänge konnten ebenfalls keine statistisch signifikanten Effekte auf das Textverstehen festgestellt werden. Dieses Ergebnis legt nahe, dass ein freiwilliges Mehrfachlesen nicht zwingend zu einer tieferen inhaltlichen Verarbeitung führt, sondern vielmehr als Wiederholungsstrategie zu verstehen sein könnte, um die bereits im Arbeitsgedächtnis gespeicherten Textinformationen zu verifizieren (Jesch, 2020). Ebenso zeigt sich, dass die durchschnittliche Fixationsdauer während des Lesens keinen nennenswerten Einfluss auf das Textverstehen ausübt. Während frühere Studien übereinstimmend berichten, dass höhere Lesekompetenz mit kürzeren Fixationszeiten korreliert (Rayner, 1998; Wertli et al., 2023), bleiben diese Befunde in der vorliegenden Untersuchung aus. Ein wesentlicher Unterschied besteht darin, dass hier nicht nur dekodierende Fähigkeiten, sondern explizit das Textverstehen als abhängige Variable herangezogen wurde. Demnach erweist sich die Fixationsdauer hingegen bisheriger Studien allein nicht als hinreichender Faktor für die Lesekompetenz.

In Anlehnung an frühere Lesestudien wurde in der vorliegenden Untersuchung geprüft, ob die durchschnittliche Anzahl an Fixationen pro Wort einen signifikanten Einfluss auf das Textverstehen und die Leseflüssigkeit ausübt. Die Ergebnisse zeigen jedoch, dass weder auf das Textverstehen noch auf die Leseflüssigkeit statistisch signifikante Effekte nachzuweisen sind. Dieses Ergebnis widerspricht zahlreichen vorherigen Studien, in denen eine geringere Fixationsanzahl als Indikator für höhere Lesekompetenz und ein effizientes Lesen identifiziert wurde (Hautala et al., 2024; Krause & Ritter, 2022; Rayner, 1998; Wertli et al., 2023). Vor diesem Hintergrund erscheint die Annahme, dass Blickbewegungsdaten allein ausreichen, um valide Rückschlüsse auf die Lesekompetenz zu ziehen, zunehmend fragwürdig. Insbesondere das Textverstehen als zentrale Dimension der Lesekompetenz wurde in vielen bisherigen Studien nur unzureichend operationalisiert oder gar vernachlässigt. Die vorliegenden Befunde legen nahe, dass eine effizientere Blickführung, gemessen an der Fixationsanzahl, nicht

zwangsläufig mit einem besseren Textverstehen oder einer höheren Leseflüssigkeit einhergeht. Vielmehr bedarf es einer differenzierten Betrachtung kognitiver Verarbeitungsprozesse, die über rein visuelle Parameter hinausgehen und das Textverstehen stärker berücksichtigen.

Insgesamt zeigt die vorliegende Studie, dass hauptsächlich der rezeptive Wortschatz und der Textschwierigkeitsgrad in dieser Stichprobe als signifikante Einflussfaktoren auf die Lesekompetenz von Schülerinnen und Schülern am Ende des zweiten Schuljahres statistisch bestätigt werden können. Darüber hinaus zeigt die Studie wichtige Aspekte für weitere Studien in Bezug auf die Messung der Lesekompetenz im Bereich des Eye-Trackings auf. So stellt die Studie heraus, dass im Gegensatz zum früheren Verständnis in Eye-Tracking-Studien die Lesekompetenz nicht nur als effizientes Dekodieren von Wörtern, Sätzen und Texten verstanden werden kann. Gute Leserinnen und Leser zeichnen sich demnach nicht nur durch schnelles, ökonomisches und flüssiges Lesen aus. Die vorliegende Untersuchung zeigt, dass Leseflüssigkeit allein nicht als hinreichender Indikator für hohe Lesekompetenz gelten kann. Vielmehr müssen flüssiges Lesen und sinnentnehmendes Textverstehen gemeinsam betrachtet werden. Vorangegangene Arbeiten haben meist nur die Leseflüssigkeit als Indikator herangezogen und auf umfassende Verständnistests verzichtet. Abschließend unterstreichen die vorliegenden Ergebnisse, dass die Lesekompetenz mehrdimensional ist und nicht allein über Leseflüssigkeit erhoben werden kann, sondern eine gemeinsame Erfassung von flüssigem Dekodieren und sinnentnehmendem Textverständnis erfordert.

3.9.5 Limitationen und Ausblick auf zukünftige Studien

Eine transparente Darstellung dieser Limitationen fördert die Nachvollziehbarkeit und zeigt auf, inwieweit die Befunde generalisierbar sind. Ziel dieses Kapitels ist es, die wesentlichen Beschränkungen der vorliegenden Studie systematisch aufzuzeigen, insbesondere in den Bereichen der Stichproben-Limitation, methodischen Limitation und der Störvariablen.

Eine zentrale Einschränkung der vorliegenden Studie betrifft die Stichprobengröße. Obwohl zunächst 42 Zweitklässlerinnen und Zweitklässler an der Studie teilgenommen haben und mit einer Tracking-Ratio von mindestens 80% erfolgreich erfasst wurden, konnten für die endgültige Auswertung lediglich 29 Datensätze herangezogen werden. Die Reduktion ist nicht auf die Qualität der Eye-Tracking-Daten zurückzuführen, sondern auf fehlende Elternfragebögen, ohne deren Angaben wichtige Variablen nicht berechnet werden konnten. Eine Stichprobe von 29 Kindern ist im Vergleich zu Studien, die nur ältere Altersgruppen oder lediglich sechs Zweitklässlerinnen und Zweitklässler einschließen, zwar verhältnismäßig groß, doch limitiert diese Fallzahl die statistische Aussagekraft und erschwert die Übertragbarkeit

der Befunde auf die Gesamtheit der Zweitklässlerinnen und Zweitklässler (De Simone et al., 2023; Häikiö et al., 2009; Krause & Ritter, 2022; Wertli et al., 2023). Darüber hinaus wurde die gesamte Stichprobe ausschließlich an einer Grundschule mit niedrigem Sozialindex rekrutiert, was sich in der Verteilung des sozioökonomischen Status der Teilnehmenden widerspiegelt (siehe Kapitel 3.8.1). Zudem weisen die Kinder im Mittel einen überaus hohen rezeptiven Wortschatz auf, da über die Hälfte der Schülerinnen und Schüler Werte im oberen Normalbereich oder darüber erzielten. Folglich ist die externe Validität der Ergebnisse für Schulen mit niedrigerem Sozialindex eingeschränkt. Darüber hinaus umfasste die Stichprobe lediglich sechs mehrsprachig aufwachsende Kinder, weshalb diese Variable aufgrund der unzureichenden Fallzahl nicht in die statistischen Analysen einbezogen werden konnte.

Die methodischen Limitationen dieser Studie zeigen sich in der Erhebung der Augenbewegungen mittels Eye-Tracking. Die teilnehmenden Schülerinnen und Schüler des zweiten Schuljahres sind weder an das Lesen längerer Texte auf Bildschirmen gewöhnt noch daran, ohne Orientierungshilfen zu lesen. Viele Kinder lesen in dieser Altersgruppe mithilfe des Zeigefingers, um ein Verrutschen in der Textzeile zu vermeiden (Holstein, 2024). Um diesem Leseverhalten möglichst gerecht zu werden, wurde der Zeilenabstand im Experiment deutlich vergrößert (siehe Kapitel 3.5). Eine weitere methodische Schwierigkeit ergibt sich aus den Anforderungen an die Körperhaltung bei der Nutzung eines nicht-invasiven Eye-Trackers. Um valide Messdaten zu gewährleisten, müssen die Probandinnen und Probanden während der Aufzeichnung möglichst reglos sitzen, was insbesondere für jüngere Kinder eine Herausforderung darstellte (siehe Kapitel 2.3.2). Als Ausgleich kamen Filler-Aufgaben mit Wimmelbildern zum Einsatz, die den Kindern erlaubten, sich zwischen den Lesephasen freier zu bewegen. Trotz dieser Limitationen konnten alle Datensätze der Studie verwendet werden. Eine weitere methodische Einschränkung ergibt sich aus der möglicherweise unterschiedlichen inhaltlichen Relevanz der gewählten Textthemen (Sommer und Bauernhof) für die teilnehmenden Schülerinnen und Schüler. Weil die wahrgenommene Relevanz von Lesetexten einen zentralen Einfluss auf die intrinsische Lesemotivation ausübt, kann ein mangelndes Interesse zu einer Abnahme der Lesebereitschaft führen. Folglich besteht die Gefahr, dass eine reduzierte Motivation die Leseleistung beeinträchtigt und somit die gemessenen Ergebnisse zwischen Teilnehmenden mit unterschiedlicher thematischer Affinität verzerrt. Schließlich erschwerte die freie Wahl der Anzahl an Lesedurchgängen die Vergleichbarkeit der Eye-Tracking-Daten. Diese Entscheidung wurde getroffen, um den Lernenden den Druck zu nehmen und ihnen die Selbstbestimmung zu ermöglichen, bei Bedarf einen Text erneut zu lesen. Für Folgeuntersuchungen empfiehlt es sich jedoch, die Anzahl der Lesevorgänge zu standardisieren, um eine einheitliche Versuchsdurchführung und eine konsistentere Datengrundlage sicherzustellen.

In Bezug auf potenzielle Störvariablen sind insbesondere die Umgebungsbedingungen und die Elternbefragung zu nennen, da beide nicht durch die Versuchsleitung kontrolliert werden konnten. Die Erhebung fand in einem Dachgeschossraum statt, in dem die Temperatur an heißen Tagen bis zu 30 °C anstieg und dadurch die Konzentrationsfähigkeit der Kinder negativ beeinflusst worden sein könnte. Zudem setzten sich die Teilnehmenden aus zwei verschiedenen Klassen mit unterschiedlichen Deutschlehrkräften zusammen, was zusätzliche Varianz in den Lernbedingungen erzeugt. Schließlich beruht die Datenerhebung zum sozioökonomischen Status und zum sprachlichen Input auf selbstständig ausgefüllten Elternfragebögen. Hierbei kann nicht ausgeschlossen werden, dass Angaben ungenau sind oder durch soziale Erwünschtheit verzerrt wurden.

Auf der Basis der vorliegenden Untersuchung und ihrer identifizierten Limitationen lassen sich für zukünftige Studien mehrere zentrale Handlungsfelder formulieren. Erstens ist eine substanzielle Erweiterung der Stichprobe erforderlich, um belastbare Aussagen zur Lesekompetenz von Grundschülerinnen und Grundschülern am Ende der zweiten Klasse treffen zu können. Zweitens empfiehlt sich ein Längsschnittdesign, in dem Entwicklungsverläufe von der ersten bis zum Ende der zweiten Klasse systematisch erhoben werden. Drittens sollte die Rekrutierung über mehrere Schulen mit unterschiedlichen Schulsozialindizes erfolgen, um eine größere Bandbreite sozioökonomischer Merkmale abzubilden und die Generalisierbarkeit der Befunde zu erhöhen. Darüber hinaus ist eine methodische Standardisierung der Lesedurchgänge anzustreben. Die Vorgabe eines einmaligen Lesedurchgangs kann das Verzerren der Eye-Tracking-Daten vermindern. In der vorliegenden Studie lasen etwa zwei Drittel der Kinder den Text nur einmal, was die Relevanz dieser Vorgabe unterstreicht. Ein weiterer Forschungsschwerpunkt sollte auf der differenzierten Erfassung des sprachlichen Inputs liegen. Neben quantitativen Erhebungen zum elterlichen Sprachangebot ist eine Analyse der Qualität des sprachlichen Inputs in Bezug auf die sprachliche Komplexität oder die sprachlichen Aktivitäten anzustreben. Ferner empfiehlt sich für alle weiteren Studien die Operationalisierung der Lesekompetenz mittels Leseflüssigkeit sowie Textverstehen. Zur Ermittlung des Textverstehens eignen sich standardisierte Verständnisaufgaben, um eine Vergleichbarkeit zu schaffen.

4 Fazit und Ausblick

Im Rahmen dieser Masterarbeit sollte untersucht werden, welche Faktoren die Entwicklung der Lesekompetenz im frühen Grundschulalter beeinflussen. Hierzu wurde zunächst die Lesekompetenz in die Teilbereiche des Textverstehens und der Leseflüssigkeit differenziert, um zu vermeiden, dass die Lesekompetenz ausschließlich über das technische Dekodieren definiert wird. Da die Lesekompetenz mehrdimensional ist, bedarf es einer kombinierten Operationalisierung aus Leseflüssigkeit und Textverstehen. Ein effizienter Leser oder eine effiziente Leserin liest nicht nur flüssig, sondern vor allem sinnentnehmend. In der vorliegenden Arbeit wurde insbesondere die bislang wenig berücksichtigte Altersgruppe der Zweitklässlerinnen und Zweitklässler untersucht. Durch die Kombination aus Eye-Tracking-Daten und der Analyse des Textverstehens der Probandinnen und Probanden konnte ein Beitrag zur differenzierten Betrachtung kognitiver und inhaltlicher Leseprozesse geleistet werden. Im Detail wurden in dieser Masterarbeit sieben Forschungsfragen behandelt, um die unterschiedlichen Einflussfaktoren auf die Lesekompetenz von Schülerinnen und Schülern am Ende des zweiten Schuljahres zu untersuchen. Das Textverstehen wurde anhand inhaltlicher Fragen ermittelt, wohingegen die Leseflüssigkeit mittels des Eye-Trackers über die Lesedauer des ersten Lesedurchganges erfasst wurde. Zur Untersuchung der Forschungsfragen wurden Einflussgrößen aus verschiedenen Bereichen ermittelt: individuelle Faktoren wie die Quantität des sprachlichen Inputs und der rezeptive Wortschatz, textbezogene Merkmale wie Textschwierigkeit und Lesedauer sowie kognitive Augenbewegungsparameter wie Fixationsdauer, Fixationsanzahl und Lesewiederholungen. Die Datenerhebung erfolgte über einen digitalen Elternfragebogen, einen standardisierten Wortschatztest sowie mittels Eye-Tracking-Technologie. Die Stichprobe umfasste 29 Zweitklässlerinnen und Zweitklässler der katholischen Grundschule an der Bergstraße in Herne.

Die Auswertung der erhobenen Daten zeigt, dass nicht alle oben genannten Faktoren einen signifikanten Einfluss auf die Leseflüssigkeit und das Textverstehen aufweisen. Der Schwierigkeitsgrad ist die einzige Variable, die sowohl signifikanten Einfluss auf die Leseflüssigkeit als auch auf das Textverstehen zeigt. Somit lassen sich die Hypothesen, dass Texte mit geringerem Schwierigkeitsgrad sowohl eine höhere Leseflüssigkeit als auch ein besseres Textverstehen bedingen, statistisch bestätigen. Im Bereich des rezeptiven Wortschatzes konnte ausschließlich die Hypothese, dass ein größerer Wortschatz besseres Textverstehen ermöglicht, statistisch verifiziert werden. Aus den Daten konnte jedoch kein Einfluss in Bezug auf die Leseflüssigkeit abgeleitet werden. Ein größerer rezeptiver Wortschatz kann das Arbeitsgedächtnis entlasten, da weniger neue Wortbedeutungen erschlossen werden müssen. Dennoch zeigt sich kein signifikanter Einfluss auf die Leseflüssigkeit. Der Wortschatz ist demnach für die rein technische Dekodierung von

untergeordneter Bedeutung. Vielmehr beeinflusst ein umfangreicher rezeptiver Wortschatz das Textverstehen auf inhaltlicher Ebene, da semantische Prozesse maßgeblich durch die Verfügbarkeit lexikalischer Kenntnisse bestimmt werden. Dieses Befundmuster stellt sich als überraschend heraus, da allgemein angenommen wird, dass ein größerer Wortschatz das Arbeitsgedächtnis über automatisierte Prozesse entlastet. Es zeigt sich weiterhin, dass weder die reine Quantität des sprachlichen Inputs noch das Lesetempo, die Häufigkeit von Lesewiederholungen oder Blickbewegungsparameter wie Fixationsdauer und -anzahl signifikanten Einfluss auf das Textverstehen ausüben. Auch beeinflussen die Leseflüssigkeit weder die Quantität des sprachlichen Inputs noch die durchschnittliche Anzahl an Fixationen pro Wort signifikant.

Zugleich wurde deutlich, dass ein höheres Lesetempo nicht automatisch mit einem besseren Textverstehen korreliert, was die Aussagekraft zeitbasierter Messverfahren der Lesekompetenz infrage stellt. Zudem zeigt sich prägnant, dass die reine Quantität des sprachlichen Inputs keinen Einfluss auf die Lesekompetenz von Zweitklässlerinnen und Zweitklässlern aufweist. Dies könnte die These unterstützen, dass die Qualität des Inputs für die Lesekompetenz von entscheidender Relevanz ist. Insgesamt unterstreichen diese Ergebnisse jedoch auch, dass die Lesekompetenz nicht ausschließlich mit Eye-Tracking-Daten im Bereich des flüssigen Dekodierens betrachtet werden kann. Sowohl in der Forschung als auch in der Diagnostik im schulischen Kontext ist es von großer Bedeutung, dass sowohl die Leseflüssigkeit als auch das sinnentnehmende Lesen erhoben werden.

Im Einklang mit früheren Studien zeigt die vorliegende Arbeit, dass ein umfassender rezeptiver Wortschatz das Textverstehen erheblich stärkt (Rosebrock & Nix, 2025). Im Kontrast zeigt sich jedoch, dass ein umfangreicher rezeptiver Wortschatz keinen Einfluss auf die technische Leseflüssigkeit ausübt. Kinder verfügen demnach über semantische Repräsentationen, ohne dass die orthografische Dekodierung automatisiert abläuft (Rosebrock & Nix, 2025). Gleichzeitig bestätigt sich der Textschwierigkeitsgrad als zentraler Prädiktor: Einfache Satzstrukturen und häufige Wörter reduzieren die kognitive Last, beschleunigen das Lesen und verbessern gleichzeitig das Textverstehen, während komplexe Konstruktionen durch verlängerte Dekodierungszeiten Arbeitsgedächtnisressourcen binden und die semantische Verarbeitung erschweren (Jesch, 2020; Schaffner, 2009). Im Gegensatz zu Arbeiten, die allein auf Eye-Tracking-Kennwerte wie Fixationsdauer und -häufigkeit oder auf zeitbasierte Lesetests setzen, korrelieren diese Metriken sowie die gemessene Lesegeschwindigkeit weder mit Leseflüssigkeit noch mit Textverstehen (Metze, 2024; Hautala et al., 2024; Wertli et al., 2023). Damit erweisen sie sich als unzureichende Indikatoren für die Lesekompetenz. Ein weiteres zentrales Ergebnis dieser Studie bestätigt, in Übereinstimmung mit vorangegangenen Untersuchungen (Hoff et al., 2024; Rowe & Snow, 2020), dass nicht das

quantitative Ausmaß, sondern vermutlich die Qualität des sprachlichen Inputs den maßgeblichen Einfluss auf die Lesekompetenzentwicklung ausübt. Für die Diagnostik und den Unterricht empfiehlt sich daher ein mehrdimensionaler Ansatz, der die automatisierte Dekodierfähigkeit und die semantische Tiefenerschließung beachtet.

Die vorliegende Studie leistet einen wichtigen Beitrag zur bestehenden Forschung, indem sie die Lesekompetenz von Schülerinnen und Schülern am Ende des zweiten Schuljahres differenziert untersucht hat, eine Zielgruppe, die in vielen bisherigen Studien kaum systematisch berücksichtigt wurde. Im Gegensatz zu früheren Arbeiten, die Lesekompetenz häufig über die Leseflüssigkeit definieren, zeigt diese Untersuchung, dass flüssiges Lesen allein kein hinreichender Indikator für tiefes Textverstehen ist. Die Studie verschiebt den Fokus von rein technischen Lesemaßen hin zu semantischen Verstehensprozessen. Besonders hervorzuheben ist die empirische Bestätigung der Bedeutung des rezeptiven Wortschatzes für das Textverstehen allerdings ohne nachweisbaren Einfluss auf die Leseflüssigkeit. Dieses Spannungsfeld eröffnet neue Perspektiven für das Zusammenspiel von lexikalischem Wissen und kognitiver Entlastung beim Lesen. Es deutet darauf hin, dass Kinder zwar die Bedeutung eines Wortes kennen, dessen schriftliche Form jedoch noch nicht automatisiert erkennen können. Darüber hinaus hinterfragt die Studie die Aussagekraft etablierter Diagnostikinstrumente wie zeitlich limitierter Lesetests und Eye-Tracking-Metriken. Die Ergebnisse zeigen, dass Lesetempo, Fixationsdauer und Anzahl der Lesedurchgänge keine signifikanten Erklärungswerte für das Textverstehen liefern. Dies widerspricht früheren Studien und unterstreicht die Notwendigkeit einer umfassenderen diagnostischen Betrachtung von Lesekompetenz. Im Gegensatz zur Einschätzung von Rupp (2022), der eine synonyme Verwendung der Begriffe Leseverstehen, Lesefähigkeit und Lesekompetenz befürwortet, zeigt die vorliegende Untersuchung, dass gerade die Differenzierung dieser Konzepte wesentlich ist. Insgesamt erweitert die Studie den Forschungsstand, indem sie technische und semantische Dimensionen der Lesekompetenz miteinander in Beziehung setzt und empirisch belegt.

Wie bereits detailliert erläutert, ist die Untersuchung durch eine kleine Stichprobe (n = 29) an einer Einzelschule mit hohem Sozialindex und überdurchschnittlichem Wortschatz der Teilnehmenden nur eingeschränkt generalisierbar. Methodische Hürden des Eye-Trackings, etwa die ungewohnte Bildschirmlesesituation, die erforderliche Körperhaltung und die freie Wahl der Lesedurchgänge vermindern zudem die Vergleichbarkeit der Daten. Hinzu kommen potenzielle Störfaktoren wie hohe Raumtemperaturen, unterschiedliche Deutschlehrkräfte und mögliche Verzerrungen in den Elternbefragungen. Die methodischen Hürden im Bereich des Eye-Trackings hatten jedoch in dieser Studie keine Auswirkungen auf die Validität der Daten.

Auf Basis der vorliegenden Erkenntnisse erweist sich zunächst eine differenzierte Untersuchung der Qualität des sprachlichen Inputs als notwendig, bei der quantitative Erhebungen zur Häufigkeit elterlicher Sprachangebote mit qualitativen Analysen zu Komplexität, Vokabularreichtum und dialogischer Interaktion kombiniert werden, um präzise Rückschlüsse auf deren Einfluss auf frühe Lesefähigkeiten zu ziehen. Parallel dazu sollte die Stichprobe deutlich erweitert und stärker heterogenisiert werden, indem neben einsprachig deutschsprachigen Zweitklässlerinnen und Zweitklässlern aus Schulen mit unterschiedlichem Sozialindex gezielt mehrsprachig aufwachsende Kinder sowie Schülerinnen und Schüler verschiedener Regionen und sozioökonomischer Hintergründe einbezogen werden, um differenzierte Aussagen über soziokulturelle Effekte zu ermöglichen. Ein weiterer Forschungsstrang muss die Validität gängiger standardisierter Lesetests kritisch hinterfragen und prüfen, ob reine Dekodieraufgaben die natürliche Lesekompetenz angemessen abbilden oder ob adaptive Testverfahren und dynamische Assessments realitätsnähere Einblicke gewähren. Dabei könnten Eye-Tracking-Daten ergänzend dazu dienen, Lesestrategien in Echtzeit zu analysieren. Schließlich verspricht eine vierjährige Längsschnittstudie von Klasse eins bis vier Erkenntnisse über die Entwicklung von Lesekompetenz über die frühen Schuljahre hinweg und wie sich die Einflussfaktoren auf die Lesekompetenz im Laufe der vier Jahrgänge verändern, sofern Leseflüssigkeit und Textverstehen gemeinsam operationalisiert werden.

5 Literaturverzeichnis

Aitchison, J. (1997). *Wörter im Kopf: Eine Einführung in das mentale Lexikon* (M. Wiese, Übers.). Niemeyer.

Aktas, M. (2020). Voraussetzungen und Bedingungen eines erfolgreichen Spracherwerbs. In S. Sachse, A.-K. Bockmann, & A. Buschmann (Hrsg.), *Sprachentwicklung: Entwicklung – Diagnostik – Förderung im Kleinkind- und Vorschulalter* (S. 45–64). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-60498-4_2

Albert, R., & Marx, N. (2016). *Empirisches Arbeiten in Linguistik und Sprachlehrforschung: Anleitung zu quantitativen Studien von der Planungsphase bis zum Forschungsbericht* (3., überarb. u. aktual. Aufl.). Narr Francke Attempto Verlag.

Backhaus, A. (2005). Der Stolperwörter-Lesetest und der Siegener Lesetest. In M. Götz & K. Müller (Hrsg.), *Grundschule zwischen den Ansprüchen der Individualisierung und Standardisierung* (S. 137–144). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-322-80804-2

Baddeley, A. (2000). The episodic buffer: A new component of working memory? Trends in Cognitive Sciences, 4(11), 417–423. https://doi.org/10.1016/S1364-6613(00)01538-2

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1). 1–48. https://doi.org/10.18637/jss.v067.i01

Becker, B. (2017). 1 Elternmerkmale: Sozioökonomischer Status und Migrationshintergrund. In U. Hartmann, M. Hasselhorn, & A. Gold (Hrsg.), *Entwicklungsverläufe verstehen – Kinder mit Bildungsrisiken wirksam fördern: Forschungsergebnisse des Frankfurter IDeA-Zentrums* (S. 32–50). Kohlhammer Verlag.

BeeMehr. (2025). *BeeMEHR: Warum Mehrsprachigkeit einfach mehr ist*. Abgerufen am 20. Juli 2025 von https://beemehr.kuwi.tu-dortmund.de/projekt/was-ist-bee-mehr/

Bosma, E., & Nota, N. (2020). Cognate facilitation in Frisian–Dutch bilingual children's sentence reading: An eye-tracking study. *Journal of Experimental Child Psychology*, 189, Artikel 104699. https://doi.org/10.1016/j.jecp.2019.104699

Calvo, A., & Bialystok, E. (2014). Independent effects of bilingualism and socioeconomic status on language ability and executive functioning. *Cognition*, 130(3), 278–288. https://doi.org/10.1016/j.cognition.2013.11.015

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, 108(1), 204–256. https://doi.org/10.1037/0033-295X.108.1.204

Cromley, J. G., & Azevedo, R. (2007). Testing and refining the direct and inferential mediation model of reading comprehension. *Journal of Educational Psychology*, 99(2), 311–325. https://doi.org/10.1037/0022-0663.99.2.311

Deilen, S. (2022). *Optische Gliederung von Komposita in Leichter Sprache: Blickbewegungsstudien zum Einfluss visueller, morphologischer und semantischer Faktoren auf die Verarbeitung deutscher Substantivkomposita.* Frank & Timme.

de-la-Peña, C. (2024). Eye-tracking contribution on processing of (implicit) reading comprehension. *Journal of New Approaches in Educational Research*, 13(1). Artikel 13. https://doi.org/10.1007/s44322-024-00013-w

De Simone, E., Moll, K., Feldmann, L., Schmalz, X., & Beyersmann, E. (2023). The role of syllables and morphemes in silent reading: An eye-tracking study. *Quarterly Journal of Experimental Psychology*, 76(11), 2493–2513. https://doi.org/10.1177/17470218231160638

Ditton, H., & Maaz, K. (2015). VIII-4 Sozioökonomischer Status und soziale Ungleichheit. In H. Reinders, H. Ditton, C. Gräsel, & B. Gniewosz (Hrsg.), *Empirische Bildungsforschung* (S. 229–244). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-19994-8_17

Ehl, B., & Grosche, M. (2020). Einbezug von Alter und Sprachkontaktdauer in die Wortschatzdiagnostik bei Mehrsprachigkeit. *Zeitschrift Für Entwicklungspsychologie Und Pädagogische Psychologie*, 52(1–2), 41–53. https://doi.org/10.1026/0049-8637/a000228

Grießhaber, W. (2024). Profilanalyse. In M. Szurawitzki & P. Wolf-Farré (Hrsg.), *Handbuch Deutsch als Fach- und Fremdsprache* (S. 217–232). De Gruyter. https://doi.org/10.1515/9783110690279-015

Groß Ophoff, J., Isaac, K., Hosenfeld, I., & Eichler, W. (2008*). Erfassung von Leseverständnis im Projekt VERA.* Checkpoint literacy. Tagungsband, 2, 36–51.

Häikiö, T., Bertram, R., Hyönä, J., & Niemi, P. (2009). Development of the letter identity span in reading: Evidence from the eye movement moving window paradigm. *Journal of Experimental Child Psychology*, 102(2), 167–181. https://doi.org/10.1016/j.jecp.2008.04.002

Hautala, J., Hawelka, S., & Ronimus, M. (2024). An eye movement study on the mechanisms of reading fluency development. *Cognitive Development*, 69, Artikel 101395. https://doi.org/10.1016/j.cogdev.2023.101395

Heine, J.-H., Heinle, M., Hahnel, C., Lewalter, D., & Becker-Mrotzek, M. (2023). Lesekompetenz in PISA 2022 Ergebnisse, Veränderungen und Perspektiven. In D. Lewalter, J. Diedrich, F. Goldhammer, O. Köller, & K. Reiss (Hrsg.), *PISA 2022: Analyse der Bildungsergebnisse in Deutschland* (S. 139–162). Waxmann.

Hessels, R. S., Nuthmann, A., Nyström, M., Andersson, R., Niehorster, D. C., & Hooge, I. T. C. (2024). The fundamentals of eye tracking part 1: The link between theory and research question. *Behavior Research Methods*, 57(1), 1–16. https://doi.org/10.3758/s13428-024-02544-8

Hochstadt, C., Krafft, A., & Olsen, R. (2022). *Deutschdidaktik: Konzeptionen für die Praxis* (3. überarb. u. erw. Aufl.). utb GmbH. https://doi.org/10.36198/9783838559414

Hoff, E., & Naigles, L. (2002). How Children Use Input to Acquire a Lexicon. *Child Development*, 73(2), 418–433. https://doi.org/10.1111/1467-8624.00415

Hoff, E., Trecca, F., Højen, A., Laursen, B., & Bleses, D. (2024). Context and education affect the quality of parents' speech to children. *Journal of Applied Developmental Psychology*, 91, Artikel 101632. https://doi.org/10.1016/j.appdev.2024.101632

Hoffmann-Erz, R. (2023). *Deutsch in der Grundschule: Eine Einführung*. J.B. Metzler. https://doi.org/10.1007/978-3-662-66653-1

Holmqvist, K., & Andersson, R. (2017). *Eye tracking: A comprehensive guide to methods, paradigms, and measures* (2. Aufl.). Lund Eye-Tracking Research Institute.

Holstein, M. (2024). *Zusammen ins Abenteuer Lesenlernen starten Außerunterrichtliche Leseförderung im Ganztag – Klassenstufe 2*. Stiftung Lesen.

Ilg, U., & Thier, P. (2006). 28 Zielgerichtete Augenbewegungen. In H.-O. Karnath & P. Thier (Hrsg.), *Neuropsychologie* (2., aktual. u. erw. Aufl., S. 296–307). Springer Medizin Verlag Heidelberg. https://doi.org/10.1007/3-540-28449-4

Jenkins, J. (1979). Four points to remember: A tetrahedral model of memory experiments. In L. S. Cermak & F. I. M. Craik (Hrsg.), *Levels of processing in human memory* (S. 429–446). Erlbaum.

Jesch, T. (2020). *Fachdidaktik Deutsch: Eine Einführung*. Narr Francke Attempto.

Jude, N., Hartig, J., Schipolowski, S., Böhme, K., & Stanat, P. (2013). Definition und Messung von Lesekompetenz. PISA und die Bildungsstandards. In N. Jude & E. Klieme (Hrsg.), *PISA 2009 – Impulse für die Schul- und Unterrichtsforschung* (S. 200–228). Beltz. https://doi.org/10.25656/01:7828

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge University Press.

Knickenberg, M. (2018). *Feedback und Attributionen im Grundschulunterricht: Bedeutung für Motivationen und Lesekompetenzen*. Vieweg.

Krause, M., & Ritter, N. (2022). Bilingual vs. Monolingual Readers: Insights from Eye-Tracking Data. *Journal of Home Language Research*, 5(1), 1–16. https://doi.org/10.16993/jhlr.36

Lampert, T., Hoebel, J., Kuntz, B., & Müters, S. (2018). Messung des sozioökonomischen Status und des subjektiven sozialen Status in KiGGS Welle 2. *Journal of Health Monitoring*, 3(1), 114–133. https://doi.org/10.17886/RKI-GBE-2018-016

Leimbrink, K. (2015). 2. Wortschatzerwerb. In U. Haß & P. Storjohann (Hrsg.), *Handbuch Wort und Wortschatz* (S. 27–52). De Gruyter. https://doi.org/doi:10.1515/9783110296013-002

Lenhard, W. (2024). *Leseverständnis und Lesekompetenz: Grundlagen - Diagnostik - Förderung* (3., erw. u. überarb. Aufl.). Verlag W. Kohlhammer. https://doi.org/10.17433/978-3-17-042763-1

Lenhard, W., & Artelt, C. (2009). Kapitel 1 Komponenten des Leseverständnisses. In W. Lenhard & W. Schneider (Hrsg.), *Diagnostik und Förderung des Leseverständnisses* (1. Aufl., S. 1–18). Hogrefe Verlag.

Lenhard, W., & Lenhard, A. (2022). Bedeutung und Diagnostik des Wortschatzes am Beispiel des Peabody Picture Vocabulary Test (PPVT-IV). *Bulletin Suisse de Linguistique Appliquee*, 2021(113), 9–27.

Lenhard, A., Lenhard, W., Segerer, R., & Suggate, S. (2015). *Peabody Picture Vocabulary Test – 4. Ausgabe: Deutsche Fassung (Manual)*. Pearson Assessment & Information GmbH.

Levelt, W. J. M. (2001). Spoken word production: A theory of lexical access. *Proceedings of the National Academy of Sciences*, 98(23), Artikel 13464–13471. https://doi.org/10.1073/pnas.231459498

Logothetis, N. K. (2006). 11 Neuronale Implementierung der Objekt- und Gesichtserkennung. In H.-O. Karnath & P. Thier (Hrsg.*), Neuropsychologie* (2., aktual. u. erw. Aufl., S. 117–127). Springer Medizin Verlag Heidelberg. https://doi.org/10.1007/3-540-28449-4

Lorenz, R., Frey, A., Trendtel, M., Ludewig, U., Schilcher, A., & McElvany, N. (2023). Kapitel 2 Ziele, Design, Instrumente und Durchführung der Internationalen Grundschul-Lese-Untersuchung (IGLU 2021). In DIPF, Leibniz-Institut für Bildungsforschung und Bildungsinformation, N. McElvany, R. Lorenz, A. Frey, F. Goldhammer, A. Schilcher, & T. C. Stubbe (Hrsg.), *IGLU 2021. Lesekompetenz von Grundschulkindern im internationalen Vergleich und im Trend über 20 Jahre* (S. 27–52). Waxmann. https://www.pedocs.de/frontdoor.php?source_opus=28075

MacLeod, A. A. N., & Demers, C. (2023). Transmitting white monolingual Anglo-American norms: A concept analysis of “quality of language” in parent-child interactions. *Applied Psycholinguistics*, 44(4), 431–459. https://doi.org/10.1017/s014271642300005x

Masek, L. R., Ramirez, A. G., McMillan, B. T. M., Hirsh-Pasek, K., & Golinkoff, R. M. (2021). Beyond counting words: A paradigm shift for the study of language acquisition. *Child Development Perspectives*, 15(4), 274–280. https://doi.org/10.1111/cdep.12425

McElvany, N., Becker, M., & Lüdtke, O. (2009). Die Bedeutung familiärer Merkmale für Lesekompetenz, Wortschatz, Lesemotivation und Leseverhalten. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 41(3), 121–131. https://doi.org/10.1026/0049-8637.41.3.121

McElvany, N., Lorenz, R., Frey, A., Goldhammer, F., Schilcher, A., & Stubbe, T. C. (2023). Kapitel 2 IGLU 2021: Zentrale Befunde im Überblick. In N. McElvany, R. Lorenz, A. Frey, F. Goldhammer, A. Schilcher, & T. C. Stubbe (Hrsg.), *IGLU 2021: Lesekompetenz von Grundschulkindern im internationalen Vergleich und im Trend über 20 Jahre* (S. 27–52). Waxmann.

McElvany, N., & Schneider, C. (2009). Kapitel 9 Förderung der Lesekompetenz. In W. Lenhard & W. Schneider (Hrsg.), *Diagnostik und Förderung des Leseverständnisses* (1. Aufl., S. 151–184). Hogrefe Verlag.

Mertins, B. (2016). The use of experimental methods in linguistic research: Advantages, problems and possible pitfalls. In T. Anstatt, A. Gattnar, & C. Clasmeier (Hrsg.), Slavic languages in psycholinguistics: Chances and challenges for empirical and experimental research (S. 15–33). Slavic Languages in the Black Box, Tübingen. Narr Francke Attempto.

Metze, W. (2024). *Stolperwörter-Lesetest*. Biss Transfer.

Mézière, D. C., Yu, L., Reichle, E. D., Von Der Malsburg, T., & McArthur, G. (2023). Using Eye-Tracking Measures to Predict Reading Comprehension. *Reading Research Quarterly*, 58(3), 425–449. https://doi.org/10.1002/rrq.498

Mézière, D. C., Yu, L., Von Der Malsburg, T., Reichle, E. D., & McArthur, G. (2025). Using Eye Movements From a “Read-Only” Task to Predict Text Comprehension. *Reading Research Quarterly*, 60(3), Artikel e70023. https://doi.org/10.1002/rrq.70023

Mihret, G., & Joshi, J. (2025). Understanding and Addressing Reading Comprehension Challenges: Magnitude, Causes and Solutions. *Journal of English as A Foreign Language Teaching and Research*, 5(1), 1–11. https://doi.org/10.31098/jefltr.v5i1.3220

Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen [MSB NRW] (Hrsg.). (2021). *Lehrpläne für die Primarstufe in Nordrhein-Westfalen*.

Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen [MSB NRW] (2024). *Übersicht über die Sozialindexstufen der Schulen nach Bezirksregierung, Kreis / kreisfreier Stadt und Schulform ab dem Schuljahr 2024/25*. Abgerufen am 25. Juli 2025 von https://www.schulministerium.nrw/system/files/media/document/file/schulsozialindex_schulliste_2024_2025.pdf

Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen [MSB NRW] (2025). *Schulsozialindex*. Abgerufen am 25. Juli 2025 von https://www.schulministerium.nrw/schulsozialindex

Müller, H. M. (2013). *Psycholinguistik - Neurolinguistik: Die Verarbeitung von Sprache im Gehirn*. Wilhelm Fink.

Müri, R., Cazzoli, D., & Nyffeler, T. (2019). Eye Movements in Neurology. In C. Klein & U. Ettinger (Hrsg.), *Eye Movement Research* (S. 749–774). Springer International Publishing. https://doi.org/10.1007/978-3-030-20085-5_17

Nation, P. (2017). How vocabulary is learned. Indonesian *JELT: Indonesian Journal of English Language Teaching*, 12(1), 1–14. https://doi.org/10.25170/ijelt.v12i1.1458

Negi, S., & Mitra, R. (2020). Fixation duration and the learning process: An eye tracking study with subtitled videos. *Journal of Eye Movement Research*, 13, 1–15. https://doi.org/10.16910/jemr.13.6.1

Pearson, B. Z., Fernandez, S. C., Lewedeg, V., & Oller, D. K. (1997). The relation of input factors to lexical learning by bilingual infants. *Applied Psycholinguistics*, 18(1), 41–58. https://doi.org/10.1017/s0142716400009863

Radach, R., Günther, T., & Huestegge, L. (2012). Blickbewegungen beim Lesen, Leseentwicklung und Legasthenie. *Lernen und Lernstörungen*, 1(3), 185–204. https://doi.org/10.1024/2235-0977/a000019

Rakoczi, G. (2012). Eye Tracking in Forschung und Lehre. Möglichkeiten und Grenzen eines vielversprechenden Erkenntnismittels. In G. S. Csanyi, F. Reichle, & A. Steiner (Hrsg.), *Digitale Medien—Werkzeuge für exzellente Forschung und Lehre* (S. 87–98). Waxmann.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124(3), 372–422. https://doi.org/10.1037/0033-2909.124.3.372

Rayner, K., Schotter, E. R., Masson, M. E. J., Potter, M. C., & Treiman, R. (2016). So Much to Read, So Little Time: How Do We Read, and Can Speed Reading Help? *Psychological Science in the Public Interest*, 17(1), 4–34. https://doi.org/10.1177/1529100615623267

Rettig, A., & Schiefele, U. (2023). Relations Between Reading Motivation and Reading Efficiency – Evidence From a Longitudinal Eye-Tracking Study. *Reading Research Quarterly*, 58(4), 685–709. https://doi.org/10.1002/rrq.502

Richardson, T. (2023). What constitutes a high-quality environment for speech and language development? *International Journal of Early Years Education*, 33(1), 45–61. https://doi.org/10.1080/09669760.2023.2284185

Römer, C., & Sturm, A. (2021). *Lexikologie*. Narr Francke Attempto Verlag.

Rosebrock, C. (2022). C1 Leseflüssigkeit. In J. Baurmann, C. Kammler, & A. Müller (Hrsg.), *Handbuch Deutschunterricht: Theorie und Praxis des Lehrens und Lernens* (3. Aufl., S. 108–111). Klett Kallmeyer.

Rosebrock, C., & Nix, D. (2025). *Grundlagen der Lesedidaktik und der systematischen schulischen Leseförderung* (10., erw. u. aktual. Aufl.). Schneider Verlag Hohengehren GmbH.

Rothstein, B., Stark, L., Betz, A., & Schuttkowski, C. (2022). *Wissenschaftliches Arbeiten in der Linguistik* (2., überarb. u. aktual. Aufl.). Narr Francke Attempto. https://doi.org/10.5555/9783823394839

Rowe, M. L., & Snow, C. E. (2020). Analyzing input quality along three dimensions: Interactive, linguistic, and conceptual. *Journal of Child Language*, 47(1), 5–21. https://doi.org/10.1017/s0305000919000655

Rupp, G. (2022). Lesekompetenz. In J. Baurmann, C. Kammler, & A. Müller (Hrsg.), *Handbuch Deutschunterricht: Theorie und Praxis des Lehrens und Lernens* (3. Aufl., S. 112–115). Klett Kallmeyer.

Schaffner, E. (2009). Kapitel 2 Determinanten des Leseverstehens. In W. Lenhard & W. Schneider (Hrsg.), *Diagnostik und Förderung des Leseverständnisses* (S. 19–44). Hogrefe Verlag.

Schräpler, J.-P., & Jeworutzki, S. (2021). *Konstruktion des Sozialindex für Schulen in Nordrhein-Westfalen.* https://doi.org/10.46586/rub.zefir.213.186

Schründer-Lenzen, A. (2004). *Schriftspracherwerb und Unterricht: Bausteine professionellen Handlungswissens*. VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-322-97604-8

Smyrnakis, I., Andreadakis, V., Selimis, V., Kalaitzakis, M., Bachourou, T., Kaloutsakis, G., Kymionis, G. D., Smirnakis, S., & Aslanides, I. M. (2017). RADAR: A novel fast-screening method for reading difficulties with special focus on dyslexia. *PLOS ONE*, 12(8), Artikel e0182597. https://doi.org/10.1371/journal.pone.0182597

Solheim, O., & Uppstad, P. (2011). Eye-tracking as a tool in process-oriented reading test validation. *International Electronic Journal of Elementary Education*, 4(1), 153–168.

Tamis-LeMonda, C. S., Bornstein, M. H., & Baumwell, L. (2001). Maternal Responsiveness and Children's Achievement of Language Milestones. *Child Development,* 72(3), 748–767. https://doi.org/10.1111/1467-8624.00313

Tarelli, I., & Bos, W. (2013). Kontextanalyse und empirische Bildungsforschung mit IGLU. In R. Becker & A. Schulze (Hrsg.), *Bildungskontexte: Strukturelle Voraussetzungen und Ursachen ungleicher Bildungschancen* (S. 117–144). Springer Fachmedien. https://doi.org/10.1007/978-3-531-18985-7

Tarkian, J., Maritzen, N., Eckert, M., & Thiel, F. (2019). Vergleichsarbeiten (VERA) – Konzeption und Implementation in den 16 Ländern. In F. Thiel, J. Tarkian, E.-M. Lankes, N. Maritzen, T. Riecke-Baulecke, & A. Kroupa (Hrsg.*), Datenbasierte Qualitätssicherung und -entwicklung in Schulen: Eine Bestandsaufnahme in den Ländern der Bundesrepublik Deutschland* (S. 41–103). Springer Fachmedien. https://doi.org/10.1007/978-3-658-23240-5_4

Tischler, T., Daseking, M., & Petermann, F. (2017). Kognitive Fähigkeiten und Lesekompetenz: Zur Bedeutung der Intelligenz für die Güte der Lesefertigkeit. *Kindheit und Entwicklung*, 26(1), 48–57. https://doi.org/10.1026/0942-5403/a000215

Wanzeck, C. (2010). *Lexikologie: Beschreibung von Wort und Wortschatz im Deutschen*. utb GmbH. https://doi.org/10.36198/9783838533162

Wertli, J., Schötzau, A., & Palmowski-Wolfe, A. (2023). The Influence of Age on Eye Movements during Reading in Early Elementary School Children. *Klinische Monatsblätter Für Augenheilkunde*, 240(04), 591–598. https://doi.org/10.1055/a-2045-7271

Zihl, J. (2006). 8 Zerebrale Blindheit und Gesichtsfeldausfälle. In H.-O. Karnath & P. Thier (Hrsg.), *Neuropsychologie* (2., aktual. u. erw. Aufl., S. 88–96). Springer Medizin Verlag Heidelberg. https://doi.org/10.1007/3-540-28449-4

6 Anhang A

6.1 Texte zum Bauernhof

6.1.1 Schwierigkeitsgrad 1

Auf dem Bauernhof leben viele Tiere. Jedes Tier hat eine Aufgabe.

Die Katzen fangen Mäuse und die Hunde passen auf. Darum fühlen sich die Tiere sicher.

Der Bauer kümmert sich um die Tiere. Er füttert die Hühner und sammelt die Eier ein.

Die Kühe fressen Gras und geben Milch. Auf dem Bauernhof ist viel los.

6.1.2 Schwierigkeitsgrad 2

Auf dem Bauernhof gibt es auch andere Tiere. Die Schweine wühlen im Schlamm, weil sie schwitzen.

Die Gänse schnattern laut, wenn ein Besucher kommt. Sie haben immer viel zu erzählen.

Die Pferde laufen auf der Weide. Manchmal reitet der Bauer auf einem Pferd.

Seine Tochter zeigt den Kindern, wie sie das Heu erntet. Sie helfen ihr, weil sie gerne mitarbeiten.

6.2 Texte zum Sommertag

6.2.1 Schwierigkeitsgrad 1

Die Sonne scheint. Tim und Paula wollen schwimmen gehen.

Sie fahren zu einem See. Beide haben Spaß. Sie bauen eine Burg.

Danach essen Tim und Paula ein Eis. Paula nimmt Vanille und Tim nimmt Schokolade.

Tims Eis fällt auf den Boden. Darum ist er traurig. Paula teilt ihr Eis mit Tim.

6.2.2 Schwierigkeitsgrad 2

Tim und Paula bleiben am See. Sie spielen im Sand, weil sie Sandburgen lieben.

Tim baut einen Turm. Paula sammelt Muscheln. Beide lachen, weil sie viel Spaß haben.

Die Sonne geht langsam unter. Es wird kälter. Tim schlägt vor, dass sie nach Hause fahren.

Paula nickt, weil sie müde ist. Sie steigen auf ihre Fahrräder und fahren los.

Beide finden den Sonnenuntergang schön. Paula erzählt ihrer Mutter von dem schönen Tag.