

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Unterschiede bei Leistungsbeurteilungen von Deutsch-
Schularbeiten - Ein Vergleich des Korrekturverhaltens
von Lehrer:innen aus dem AHS- und NMS-Bereich“

verfasst von / submitted by
Nina Straussberger BEd

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of
Master of Education (MEd)

Wien, 2022 / Vienna 2022

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

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

Betreut von / Supervisor:

UA 199 506 525 02

Masterstudium Lehramt Sek (AB)
Lehrverbund UF Deutsch
UF Psychologie und Philosophie
Ass.-Prof. Mag. Dr. Jutta Ransmayr

Danksagung

An dieser Stelle möchte ich mich bei allen Personen bedanken, die mich auf unterschiedliche Art und Weise beim Fertigstellen vorliegender Masterarbeit unterstützt haben.

Besonderer Dank gebührt an dieser Stelle Frau Mag. Dr. Jutta Ransmayr, die sich sofort dazu bereit erklärt hat, meine Masterarbeit zu betreuen. Als Expertin auf diesem Gebiet war sie mir mit ihren hilfreichen Anregungen während der Fertigstellung der Arbeit ein wahrer Wegweiser.

Ich bedanke mich bei all jenen, die diese Forschung durch ihre Teilnahme überhaupt erst möglich gemacht haben. Danke an alle Proband:innen und Schulen, die mir trotz Corona-Stress ihre Zeit geschenkt haben.

Ebenso möchte ich mich bei meinen Kommilitoninnen und Freundinnen Elly und Eli bedanken, die mich durch die Studienzeit begleitet und mir mit ihrer Hilfsbereitschaft und ihrem Interesse dabei geholfen haben, die Masterarbeit, wie sie in dieser Form vorliegt, fertigzustellen.

Auch allen anderen Herzensmenschen in meinem Leben möchte ich hiermit danken. Danke, dass ihr immer für mich da seid, ihr mich auch in meinen stressigsten Lebensphasen begleitet und dass ich euch meine Freund:innen nennen darf.

Darüber hinaus danke ich meinem Freund Alexander, der mich durch seinen eigenen Ehrgeiz immer angetrieben und motiviert hat. Ohne ihn würde diese Arbeit vermutlich noch immer in den Kinderschuhen stecken.

Abschließend möchte ich meiner Familie und besonders meinen Eltern danken, für ihre finanzielle Unterstützung all die Jahre, für ihre Geduld und ihr offenes Ohr. Ohne sie würde ich nicht da stehen, wo ich jetzt bin.

Inhaltsverzeichnis

1.	Einleitung	7
1.1	Relevanz vorliegender Forschung.....	9
1.2	Zielsetzung, Forschungsfrage und Hypothesen.....	10
1.3	Aufbau der vorliegenden Arbeit.....	11
2.	Theoretische Grundlagen	12
2.1	Sprachliche Normen.....	12
2.1.1	Sprachnorm	12
2.1.2	Sprachnormen im Deutschunterricht.....	14
2.1.3	Normsetzende Instanzen.....	15
2.1.4	Der sprachliche Fehler.....	18
2.2	Korrektur	21
2.2.1	Korrekturstrategien.....	22
2.2.2	Kriterienkataloge	23
2.3	Leistungsbeurteilung	26
2.3.1	Leistungsbegriff.....	26
2.3.2	Leistungsfeststellung.....	28
2.3.3	Leistungsbeurteilung	30
2.3.4	Leistungsbeurteilung an der Neuen Mittelschule.....	31
2.3.5	Notensystem	33
2.3.6	Bezugsnormen.....	37
2.4	Studienlage zu Korrekturverhalten und Beurteilungen von Deutschlehrer:innen....	40
3.	Methodik	45
3.1	Methodentriangulation	45
3.1.1	Quantitative Methode.....	45
3.1.2	Qualitative Methode.....	46
3.2	Datengrundlage	47
3.3	Datenerhebung.....	47
3.3.1	Pretest	47
3.3.2	Durchführung der Erhebung.....	48
3.4	Schularbeitsangabe und -text.....	49
3.5	Datenauswertung	50
3.6	Stichprobe.....	51
4.	Ergebnisse	54
4.1	Quantitative Aspekte: Fehlermarkierungen der Lehrpersonen	54
4.2	Qualitative Aspekte: Vergleich des Korrekturverhaltens und der Beurteilungen....	63

4.2.1	Darstellung der Korrekturen.....	63
4.2.2	Beurteilungen	86
5.	Interpretation und Diskussion der Ergebnisse.....	89
6.	Zusammenfassung und Ausblick.....	94
6.1	Zentrale Ergebnisse	94
6.2	Allgemeine Schlussfolgerungen für die Praxis	97
6.3	Ausblick	98
7.	Literaturverzeichnis.....	99
8.	Abbildungsverzeichnis	110
9.	Tabellenverzeichnis.....	110
10.	Anhang	111
10.1	Original-Schularbeitstext.....	111
10.2	Fragebogen	113
10.3	Abstract	116

1. Einleitung

Bis heute scheint das Sprichwort „Non vitae, sed scholae discimus“, auf Deutsch „Nicht für das Leben, sondern für die Schule lernen wir“ (Gunermann 1998), Gültigkeit zu besitzen. Trotz der Bildungsreform im Jahr 2008/09, welche die Einführung von Bildungsstandards verordnete beziehungsweise auf eine Kompetenzorientierung abzielte (vgl. BGBl. II Nr. 1/2009) und ungeachtet der Bemühungen seitens der Politik und Schulen, scheint das Denken um gute Noten noch immer fest in den Köpfen von Schüler:innen, Eltern und auch Lehrer:innen manifestiert zu sein. Die Relevanz von Leistungsbeurteilungen im Schulsystem ist daher in vielerlei Hinsicht, wie beispielsweise der pädagogischen, bildungspolitischen oder der rechtlichen Sichtweise, nicht abzustreiten.

Die Kritik (siehe etwa Jachmann 2003), die vor allem durch die Erkenntnisse von internationalen und nationalen Schulleistungsuntersuchungen wie PISA (ab 2000) und IGLU (ab 2001) laut geworden ist, betrifft unter anderem die schulische Leistungsbeurteilung und handelt von der Praxis der Notenvergabe. Mangelnde Objektivität, eine geringe Vergleichbarkeit und Ungerechtigkeiten werden oftmals den Lehrer:innen vorgeworfen (vgl. Tortik 2016: 19). Dies zieht eine nicht ausreichende Aussagekraft der gegebenen Noten für den weiteren Bildungsweg nach sich. Der Abstand zwischen tatsächlichen Kompetenzen und Noten scheint dabei immer wieder ins Kreuzfeuer der Kritik zu geraten. Warum dies so ist, hat viele Gründe. Einerseits kann dies durch das Drängen der Eltern aufgrund Möglichkeiten von weiterführenden Chancen und Schularten und den damit einhergehenden Konflikten mit den Schulen erklärt werden (vgl. Tortik 2016: 20). Andererseits müssen auch die Regierungsverantwortlichen betrachtet werden. Gute Bildungsabschlüsse und niedrige Drop-Out-Quoten führen nämlich dazu, dass Österreich im Vergleich zu anderen Staaten möglichst viele Schüler:innen an einen gewünschten Punkt bringt, ungeachtet den realen Leistungen (vgl. Tortik 2016: 20). Unter diesem Druck sehen sich die meisten Lehrer:innen gezwungen, die Notengebungen anzupassen und resignieren hinsichtlich einer gerechtfertigten Leistungsbeurteilung. Besonders in den Neuen Mittelschulen scheinen diese Beobachtungen großgeschrieben zu werden. So berichtete beispielsweise Bayrhammer (2014) in der Presse von einem „Fünferverbot“.¹ Verschiedenen Lehrer:innenaussagen nach zu urteilen wird dies

¹ Bayrhammer, Bernadette (2014): Migration und Schule: Von verschenkten Noten und Versagen, Die Presse, [online] <https://www.diepresse.com/3853938/migration-und-schule-von-verschenkten-noten-und-versagen> [abgerufen am 17.05.2022].

teilweise von Direktor:innen so vorgeschrieben, vor allem wenn es andere Lehrpersonen und Kolleg:innen genauso handhaben.

Aus rechtlicher Sicht kann gesagt werden, dass dieses Verhalten nicht im Sinne des Schulorganisationsgesetzes oder der Leistungsbeurteilungsverordnung ist. Dennoch dürfte dies in der Praxis häufig so gehandhabt werden, weil alle anderen Parteien einen Vorteil daraus ziehen. Auch für die Lehrperson selbst ist das „Verschenken“ von Noten oftmals ein erfolgreiches Umgehen von Aufwand und Ärger, sowohl mit der Schulleitung als auch den Eltern (vgl. Tortik 2016: 20). Werden die Aussagen Engelbrechts (2013: 15) betrachtet, lässt sich erkennen, dass solche Probleme nicht erst seit Kurzem so „gelöst“ werden. Die Schul- und Unterrichtsordnung wurde immer wieder als Antwort auf die hohen Repetentenzahlen angepasst. Die Schuldzuweisungen für diesen Zustand betrafen vor allem die Lehrpersonen. Wer viele „Nicht genügend“ verteilen muss, bekommt vielleicht den Vorwurf präsentiert, einen qualitativ schlechten Unterricht zu halten. Andere Komponenten, wie beispielsweise eine geringe Lernmotivation oder auch Überforderung seitens der Schüler:innen, weil sie sich wegen milderer Beurteilungen unberechtigt in einer bestimmten Schulkarriere befinden, werden dabei außer Acht gelassen (vgl. Neuweg 2014: 122).

Die gegensätzliche Perspektive zeigt ein Lehrer:innenverhalten auf, welches sich stark an den Gesetzen orientiert und auch dem Prestige eines gewissen Schultyps gerecht werden will. So kann es zu strengeren Urteilen kommen, vor allem in Gymnasien oder höheren Schulen, wo ein gewisses Maß an Leistung vorausgesetzt wird.

Bei Beachtung der Gesetze dürften die Ergebnisse der Beurteilungen nicht derart differieren, wie dies in der Schulpraxis teilweise zu passieren scheint. Aber auch wenn die Unsicherheiten und Unklarheiten bezüglich rechtlicher Punkte ausgeblendet werden, kann die Korrekturpraxis aufgrund pädagogischer Handlungsspielräume, gängiger Wahrnehmungsfehler und subjektiven Empfindens durchaus zu unterschiedlichen Notenvergaben führen. Vor allem in sprachlichen und nicht gänzlich standardisierten Unterrichtsfächern wie Deutsch werden weite Sprünge auf der Notenskala vermutet. In der Theorie gibt es Hinweise darauf, dass die Bewertung und Benotung von Texten im Deutschunterricht oftmals subjektiv getroffen wird (siehe etwa Birkel 2003; Baurmann 2006). Obwohl schriftliche Leistungen Vorteile, wie eine gewisse Objektivität in der Durchführung, vor allem im Vergleich zur mündlichen Leistungsüberprüfung, aufweisen, führen jedoch Urteilsfehler und mangelhafte Gütekriterien dazu, dass diese des Öfteren nicht sehr einheitlich ausfallen. So konnten vor allem im Bereich der Auswertungs- und Interpretationsobjektivität große Unterschiede bei den Beurteilenden ausgemacht werden (Ingenkamp & Lissmann 2008:

145). Auch weitere Studien, wie beispielsweise von Birkel (2003), konnten diese Annahmen bei der Analyse von Diktat-, Aufsatz-, und auch Mathematikleistungen weitgehend untermauern.

Grundsätzlich wird angenommen, dass hierbei durch die Festlegung von Kompetenzen und durch den vermehrten Einsatz von Beurteilungsrastern Abhilfe geschaffen wurde. Somit könnte gedacht werden, dass die Benotung von schriftlichen Arbeiten nun objektiver getroffen wird und die Korrektur weniger Probleme für die Lehrpersonen bereitet. Die Annahme über Unsicherheiten bezüglich Korrekturen geht aus bisheriger Forschung hervor, beispielsweise im Kontext von Rechtschreibung und Grammatik (siehe etwa Hennig 2012; Davies 2017). Die Literatur und Erfahrungen aus der Praxis zeigen aber, dass verschiedene Lehrer:innen eine Schularbeit unterschiedlich benoten würden (siehe Kapitel 2.4). Dahingehend stellte sich Ingenkamp (1995: 200) bereits die Frage:

Wann sagen wir ehrlich, dass kein Lehrherr aus den Zeugnissen eine vergleichbare Aussage über die Schulleistung von Schülern aus verschiedenen Klassen entnehmen kann? Wann geben unsere Universitäten zu, dass es unsinnig ist, unter zwei Bewerbern aus verschiedenen Schulen, Städten oder gar Bundesländern einen Studenten nach den Abiturnoten auszuwählen?

Um diese Aussage zu überprüfen und darüber hinaus Aufschlüsse über Unterschiede im AHS- und NMS-Bereich zu bekommen, ist es von Interesse, das Korrekturverhalten und die Leistungsbeurteilungen von Lehrer:innen aus beiden Schularten zu untersuchen.

1.1 Relevanz vorliegender Forschung

In der Literatur und im universitären Kontext gibt es bis dato keine vergleichbare Untersuchung, wie es sich vorliegende Masterarbeit zum Ziel gemacht hat. Während zu Leistungsbeurteilungen Studien existieren, die verschiedenste Gründe und Merkmale für bessere oder schlechtere Notengebungen aufdecken, wurde jedoch bisher kaum eine Untersuchung durchgeführt, wo es um reine Bewertungsunterschiede ging. Auch Forschungen zum Korrekturverhalten sind nur spärlich vorhanden. Da das Korrekturverhalten von Lehrenden und Leistungsbeurteilungen Hand in Hand gehen, soll vorliegende Masterarbeit erste Hinweise zu dieser Thematik liefern. Darüber hinaus soll dies als Anstoß für weitere Forschungen auf diesem Gebiet dienen.

Die Untersuchungen weisen auch eine praktische Relevanz auf, wenn bedacht wird, dass Noten nicht nur darüber entscheiden, welcher weiterer Bildungs- und Karriereweg verfolgt

werden kann (vgl. Baurmann 2006: 120, Hennig 2012: 130), sondern oft auch als Maß für Intelligenz, Eignung und Fähigkeiten herangezogen werden. Somit kann die Subjektivität bei der Benotung erhebliche Auswirkungen sowohl auf interne Faktoren, wie Motivation und Selbstbewusstsein (siehe etwa Poelchau 2004) als auch externe Faktoren, wie die weitere Bildungslaufbahn, haben. Insofern ist es wichtig, dass eine gerechte und objektive Beurteilung in den Schulen stattfindet.

Im Gegensatz zu anderen Studien wurde in vorliegender Masterarbeit darauf geachtet, notenbeeinflussende Faktoren, wie beispielsweise das Geschlecht, Erwartungshaltungen und den persönlichen Bezug zum Schüler oder zur Schülerin auszuschließen, indem ein anonymes Schüler:innentext an Deutsch-Lehrkräfte geschickt wurde. Diese Voraussetzung ermöglichte eine reine Bewertung des Textes ohne zusätzliche Einflussfaktoren. Lediglich eventuelle Auswirkungen der Handschrift auf die Notengebung konnten nicht ausgeschlossen werden.

1.2 Zielsetzung, Forschungsfrage und Hypothesen

Ziel vorliegender Masterarbeit ist es, das Korrekturverhalten und die Leistungsbeurteilungen verschiedener Lehrer:innen einerseits aus dem AHS und andererseits aus dem NMS-Bereich zu untersuchen. Von diesen Überlegungen ausgehend ergibt sich folgende Forschungsfrage:

Gibt es Unterschiede im Korrekturverhalten und in der Leistungsbeurteilung verschiedener Lehrer:innen für Schularbeiten im Fach Deutsch sowie bei Lehrenden zweier verschiedener Schultypen?

Darüber hinaus werden aufgrund der Literatur und Praxiserfahrungen aus dem Schulalltag folgende Hypothesen zu den Unterschieden im Korrekturverhalten aufgestellt:

1. Junglehrende markieren im Durchschnitt mehr Fehler als Lehrer:innen mit mehr Unterrichtserfahrung.
2. AHS-Lehrende beurteilen strenger als NMS-Lehrende.

Aus den in Kapitel 1.1 genannten Gründen ist neben der Beantwortung der Forschungsfrage und der Überprüfung der Hypothesen ein weiteres Ziel, ein Bewusstsein bei Lehrpersonen zu schaffen, das eigene Korrekturverhalten regelmäßig zu überprüfen und zu reflektieren.

Darüber hinaus sollte dadurch im besten Fall die stetige Revisionsbedürftigkeit der Beurteilungsvorgaben aufgezeigt werden.

1.3 Aufbau der vorliegenden Arbeit

Der erste Teil vorliegender Masterarbeit behandelt die theoretischen Ausführungen, sodass eine Grundlage für die empirische Forschung geschaffen wird. Dabei werden sowohl relevante Begrifflichkeiten, beispielweise „Sprache“, „Normen“ und „Fehler“ definiert, als auch Rechtliches zu Leistungsfeststellungen und -beurteilungen näher erläutert. Im Kapitel 2.4 wird die aktuelle Studienlage einerseits zum Korrekturverhalten von Lehrenden und andererseits zu den Notengebungen sowie deren Beeinflussungen dargestellt.

Im empirischen Teil dieser Arbeit wird zunächst im Kapitel 3.1 die methodische Vorgehensweise anhand des *mixed-method-designs* (vgl. Johnson, Onwuegbuzie & Turner 2007: 213) beschrieben. Dabei wird sowohl eine quantitative als auch qualitative Analyse durchgeführt. Darüber hinaus wird die Datengrundlage für vorliegende Untersuchung aufgezeigt. Diese beinhaltet den zu verbessernden Schularbeitstext und den Fragebogen für die Proband:innen.

Im darauffolgenden Kapitel 4 erfolgt die Veranschaulichung der Ergebnisse, welche mit Diagrammen visualisiert werden.

Im fünften Kapitel *Interpretation und Diskussion der Ergebnisse* werden die aus der Untersuchung gewonnenen Erkenntnisse interpretiert und mögliche Verbesserungsvorschläge diskutiert.

Abschließend wird im sechsten Kapitel *Zusammenfassung und Ausblick* nochmals auf die hier gestellten Fragen und Hypothesen eingegangen und die wichtigsten Erkenntnisse werden festgehalten. Darüber hinaus werden Schlussfolgerungen für die Praxis aufgezeigt und ein Ausblick auf weitere Forschungsmöglichkeiten in diesem Bereich gegeben.

2. Theoretische Grundlagen

2.1 Sprachliche Normen

Bevor der Begriff der Sprachnorm beleuchtet wird, sollte zunächst ein Blick auf die Sprache als dynamisches System geworfen werden. Sprachwissenschaftler:innen sind sich einig, dass sich Sprache stetig im Wandel befindet und dadurch auch immer Schwierigkeiten unterliegt, eine allgemeingültige und fortwährende Norm festzulegen. Es wurden bereits viele Modelle entwickelt, um das System Sprache erklären und definieren zu können. Hierbei soll vor allem das Schichtenmodell Wiesingers herangezogen werden, da es die für vorliegende Masterarbeit wichtigsten Begriffe enthält, ohne zu tief in diese Materie einzutauchen. Wiesinger (1988: 27) beschreibt Sprache als ein System, worin verschiedene Varietäten koexistieren und sich in ständiger Beziehung zueinander befinden, ineinander überlappen und überlaufen. Dabei stehen sich die Standardsprache und der Dialekt gegenüber und kennzeichnen somit die beiden äußeren Pole des dynamischen Systems Sprache (vgl. Ernst 2012: 13). Auf diese Begriffe soll unter anderem im folgenden Kapitel der Sprachnorm näher eingegangen werden.

2.1.1 Sprachnorm

Um sich an den umfassenden und schwer zu definierenden Begriff der Sprachnorm anzunähern, muss zunächst geklärt werden, was denn überhaupt eine Norm ist. Zu beidem gibt Kellermeier-Rehbein (2013: 6) eine Definition:

Eine Norm ist [...] gemeinhin eine (rechtlich) anerkannte, allgemeingültige und meist schriftlich fixierte Regel zur Handhabung eines Sachverhalts. Aus linguistischer Perspektive entspricht die Norm der Menge aller sprachlichen Einheiten und Regeln, die schriftlich im Kodex (Wörterbücher und Grammatiken) einer Sprache oder Varietät festgelegt sind. Normierung ist dementsprechend die Festlegung der Norm. (Kellermeier-Rehbein 2013: 6)

Diese sprachliche Norm bezieht sich dabei auf alle möglichen Bereiche der Linguistik, also beispielsweise auf die Aussprache, die Orthografie, die Grammatik und die Lexik (vgl. de Cillia & Ransmayr 2019: 54). Bei Betrachtung dieser Aspekte kann jedoch davon ausgegangen werden, dass maximal die kodifizierte Orthografie lückenlos durch die Vorgaben der deutschen Rechtschreibung als Norm angesehen werden kann (vgl. Rat für deutsche Rechtschreibung 2018). Schmidlin (2011) meint dahingehend, dass alle anderen Bereiche lediglich mit einer nicht kodifizierten Form angewendet werden können, was mit einem sogenannten Sprachgefühl gleichgesetzt wird. Dieses Gefühl entscheide dann schlussendlich, was „korrektes“ Deutsch sei (vgl. Schmidlin 2011: 52). Wenn bedacht wird,

dass jede Person subjektiv empfindet und etwas anderes für richtig hält beziehungsweise bei Lernenden kein Sprachgefühl vorausgesetzt werden kann, lässt sich bereits die Problematik der gesetzten Sprachnorm erkennen. Trotzdem scheint die sprachliche Norm einen festgesetzten Charakter in der Gesellschaft zu haben. Diese Annahme deckt sich auch mit der Meinung von Bartsch (1985: 149), dass Verständigung und Kommunikation ohne Sprachnormen nicht funktionieren könnten. Damit kann die verbindliche und notwendige Stellung von Sprachnormen in der Gesellschaft aufgezeigt werden. Eine reibungslose Kommunikation muss schließlich gewährleistet werden (vgl. Bartsch 1985: 149). Diese Verständigung sollte, vor allem im schulischen oder institutionellen Bereich, über die Standardsprache erfolgen. Diese macht innerhalb des dynamischen Systems von Sprache und ihren Varietäten nach Ernst (2012: 13) die gesellschaftliche Norm aus, wobei sie auch als Hoch- oder Schriftsprache bezeichnet werden kann (vgl. Ebner 2008: 7). Bereits John Earl Joseph (1987) postuliert, dass Standardsprache kodifiziert und deshalb auch stabil ist. Auch laut Höhle (2010: 319) ist sie

die in den Grammatiken und Wörterbüchern einer Sprache kodifizierte Ausprägung einer Einzelsprache, die für die betreffende Sprachgemeinschaft einer Nation verbindlich ist und in öffentlichen Situationen als sprachliche Norm gilt. Sie wird schriftlich und mündlich gebraucht und in formalen und Bildungskontexten als angemessen angesehen. (Höhle 2010: 319)

Diese Aussagen decken sich auch mit jener von Huesmann (1998: 10) und Ammon (2005: 32), welche von einer kodifizierten, bildungsnahen und überregional gebrauchten Form von Sprache ausgehen. Dennoch räumt Huesmann ein, dass sie nur zu 95 Prozent standardsprachlich sein könne, der Rest variiert (vgl. Huesmann 1998: 32). Darüber hinaus gelten in den verschiedenen Räumen der deutschen Sprache unterschiedliche linguistische Normvorstellungen. Deshalb wird beim Deutschen auch von einer plurizentrischen Sprache gesprochen, was bedeutet, dass es keine Einheitssprache ist (vgl. Wiesinger 2015: 91f.). Das österreichische Deutsch beispielsweise stellt die Standardsprache Österreichs dar, unterscheidet sich aber geringfügig durch Ansammlungen regionalspezifischer Varianten vom Bundesdeutschen oder beispielsweise dem Schweizerdeutschen (vgl. Wiesinger 2008: 6f.). Sogenannte Austriaismen, also österreichspezifische Varianten, leisten ihren übrigen Beitrag dazu, eine eigene Standardvarietät in Österreich vorzufinden. Durch die ständige Veränderung und Entwicklung der Sprache werden neu aufkommende Ausdrücke der sogenannten Umgangssprache zugeschrieben, welche im Schichtenmodell zwischen den beiden äußeren Polen, der Standardsprache und des Dialektes, lokalisiert wird. Im Variantenwörterbuch (2004) lässt sich eine Erklärung dazu finden:

Ähnlich wie in Süddeutschland ist das Sprachleben Österreichs geprägt vom fließenden Übergang zwischen rein standardsprachlichen und rein dialektalen Strukturen, [...]. Es hat sich als zweckmäßig erwiesen, die ganze Bandbreite dazwischen als „Umgangssprache“ zusammenzufassen, [...]. (Variantenwörterbuch 2004: 36)

Obwohl hinreichende Forschungen zur Umgangssprache bisher fehlen, kann im Allgemeinen gesagt werden, dass sie die im Alltag gebrauchte Sprache beziehungsweise gewisse Sprachformen (vgl. Bichel, zitiert nach Goldgruber 2011: 16) darstellt. Sie lässt sich zwar nicht mit der Standardsprache vergleichen, wird jedoch weitgehend akzeptiert und meistens überregional verwendet (vgl. Redaktion des Dudens 2022).

Wie bereits erwähnt gibt es neben der Standardsprache auch verschiedenste Varietäten, wie beispielsweise den Dialekt. Dieser stellt den Gegenpol zur Standardsprache und eine komplexe Schicht im Sprachenmodell dar (vgl. Wiesinger 2008: 26). Meistens wird er in gesprochenen und lokal beziehungsweise regional verschiedenen Formen einer Sprache vorgefunden. Gebraucht wird der Dialekt eher in kleinräumigen und privaten Rahmen und daher wird er auch als „Sprache der Nähe“ bezeichnet (Linke et al. 2004: 121). Somit ist das Charakteristikum des Dialekts im Gegensatz zur Standardsprache eine „nicht kodifizierte und flexible“ Sprachform, welche gerne im persönlichen Umgang miteinander verwendet wird. Laut Wiesinger (2008: 9) ist der Dialekt eine wenig geschätzte Gebrauchsform einer Sprache und deshalb auch in institutionellen Kontexten nicht gern gesehen. Deswegen wird nicht der Dialekt, sondern die Standardvarietät als Sprachnorm „förmlich gelehrt“ (Ammon 2005: 32), um den Schüler:innen zu einer angemessenen Ausdrucksweise zu verhelfen. Diese Realisierung trifft vor allem auf den Sprachunterricht zu, deshalb und aufgrund der Thematik vorliegender Masterarbeit sollen im nächsten Kapitel die Sprachnormen im Deutschunterricht näher betrachtet werden.

2.1.2 Sprachnormen im Deutschunterricht

Bei den sprachlichen Normen im Deutschunterricht geht es wie bei allen Normen um das Urteilen über „angemessen“ und „nicht angemessen“. Wenn das Korrigieren von Texten betrachtet wird, kommt den Deutschlehrenden eine besondere Rolle zu, denn sie müssen Entscheidungen bezüglich des Sprachgebrauchs treffen (vgl. Beuge 2021). Dabei kann zwischen dem normorientierten Ansatz und der gebrauchorientierten Auffassung des Standards unterschieden werden. Die normorientierte Standardauffassung geht dabei von lediglich im Wörterbuch festgehaltenen Sprachformen aus, während sich der gebrauchorientierte Ansatz stärker nach dem Zweck statt dem Stil richtet (vgl. Mroczynski

2013: 256). Er beschreibt somit einen eher demokratischen Zugang zur Sprache, bei welchem kein hohes Maß an Bildung nötig ist, sondern die Sprache im Alltag oder die Art, wie sie häufig von vielen gesprochen wird, betrachtet und gebraucht wird (vgl. Kellermeier-Rehbein 2013: 14f).

Sprachwissenschaftler:innen und Lehrpersonen, welche den gebrauchorientierten Zugang zur Sprache gutheißen, sehen vermehrt die Wirklichkeit und die Dynamik der Sprache und fordern aus diesem Grund auch eine Erweiterung des Standardbegriffs, sodass bestimmte sprachliche Phänomene in der Standardsprache mehr toleriert werden und darüber hinaus keine scharfen Abgrenzungen zu Nicht-Standardvarietäten mehr getroffen werden sollen (vgl. Kellermeier-Rehbein 2013: 15). Denn „wie immer man es dreht und wendet, es gibt sprachliche Problemfälle, die nicht eindeutig einer Normebene zugeordnet werden können und vielleicht auch nicht nach den Kategorien richtig oder falsch beurteilt werden müssen“ (Kellermeier-Rehbein 2013: 20). Dies kann geschehen, indem die Sprache mehr als Kontinuum gesehen wird, in welcher die hierarchischen Ebenen der Sprachschichten ineinander übergehen, wie es auch Peter Wiesinger über das System der Sprache postuliert (vgl. Wiesinger 1988: 27).

Durch die Herausforderungen, die eine Entscheidung zwischen dem gebrauchorientierten und dem normorientierten Ansatz mit sich bringen, gelten Lehrpersonen als Sprachnormautoritäten (vgl. Ammon 2005: 36). Darüber hinaus nennt Ammon (2005: 33) drei weitere Instanzen, welche Einfluss auf die Festlegung der Sprachnorm haben. Diese sollen im nächsten Kapitel näher erläutert werden.

2.1.3 Normsetzende Instanzen

Nach Ammon (1995: 80) lassen sich vier normsetzende Instanzen unterscheiden: Sprachnormautoritäten, Kodifizierer:innen, Modellsprecher:innen und -schreiber:innen und Sprachexpert:innen. Diese stehen zueinander in Beziehung und wirken in einem sozialen Kräftefeld aufeinander, welches im Folgenden zur Veranschaulichung aufgezeigt wird.

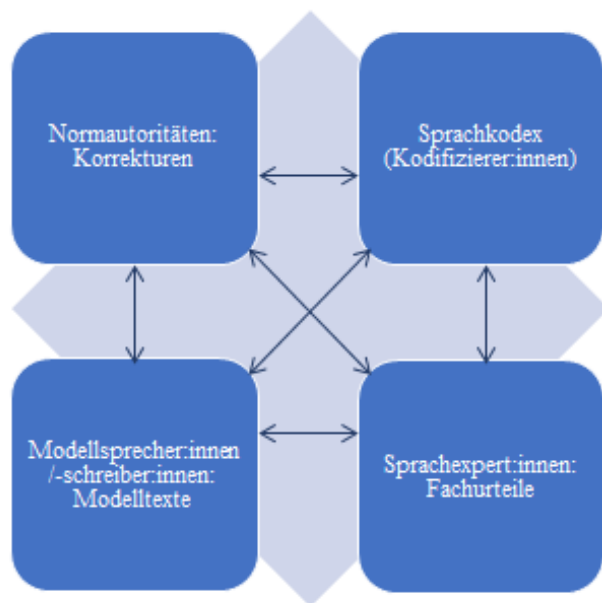


Abbildung 1: Soziales Kräftefeld einer Standardvarietät, eigene Darstellung nach Ammon (1995: 80)

Laut Ammon (2005) sind *Sprachnormautoritäten* alle Personen, welche

über ausreichende Macht verfügen oder dies glaubhaft machen können, um das Sprachhandeln anderer Personen (der Normsubjekte) zu korrigieren. [...] Für standardsprachliche Normen sind nun Sprachnormautoritäten charakteristisch, denen von Berufs wegen Sprachkorrekturen erlaubt oder sogar geboten sind. Die vielleicht wichtigsten sind die Lehrer, denen die Vermittlung der standardsprachlichen Norm in der Schule obliegt. (Ammon 2005: 36)

Wie aus diesem Zitat zu entnehmen ist, kommt den Lehrer:innen laut Ammon (2005: 36) als einflussreichsten Sprachnormautoritäten eine wichtige Rolle als Entscheidungsträger:innen über richtig und falsch zu. Oftmals wird der Kodex zur Orientierung herangezogen. Dennoch wird das Empfinden darüber, was ein Fehler ist und was nicht, weitgehend selbstständig gebildet (vgl. Ammon 2005: 36). Im schlechtesten Fall erfolgen Korrekturen durch die Lehrperson sogar strenger, als es der Kodex besagen würde. Dies ist insofern bedenklich, als dass Lehrende oft Unsicherheit bezüglich der Korrektur verspüren, da manchmal unklar ist, welche Norm die Richtige ist und inwieweit die Schüler:innen innerhalb dieser Norm variieren dürfen (vgl. Davies 2017: 138).

Unter einem *Sprachkodex* wird eine Beschreibung oder Normierung einer Sprache in einem schriftlichen Regelwerk verstanden. Dabei handelt es sich vor allem um metasprachliche Schriftstücke wie Wörterbücher und Grammatiken (vgl. Ammon 2005:33). In diesen ist also der Kodex einer Sprache vermerkt, welcher sich auf eine funktional prestigebehaftete Varietät, also die oben bereits erwähnte Standardsprache, bezieht (z.B. Metzler Lexikon Sprache 2010, Artikel Kodifizierung). Hierbei zwischen gebrauchorientiert und norm-

beziehungsweise kodexorientiert zu unterscheiden, ist eine wichtige Bedingung für die Einordnung der deutschen Standardsprache. Dabei stellt sich die Frage, welche Wörter dazugehören und welche nicht (vgl. Klein 2013: 26ff.).

Laut Ammon (1995: 74ff.) wird durch die Kodifizierung einer Sprache die Standardisierung derer erst vorangebracht, die Standardvarietät bildet sich dadurch heraus. Dennoch geraten Sprachkodizes immer wieder unter Kritik unter den Sprachwissenschaftler:innen. Während Laien sich auf den zeitlos „gültigen“ Anspruch von Kodizes berufen, sind Experten:innen der Meinung, dass Kodifizierungen keine relevanten Instanzen der Standardisierung einer Sprache darstellen (vgl. Ammon 2005: 35). Demnach müsste von Fall zu Fall die Kraft der Kodizes relativiert und auch im Hinblick des Sprachwandels kritisch gesehen werden. Trotz gegenteiliger Meinungen gilt für beide Seiten, dass klar sein muss, von welcher Kodex-Begrifflichkeit ausgegangen wird. Dazu gilt es, eine sprachwirkliche und realistische Präzisierung des Begriffs „Sprachkodex“ anzuführen. Aus diesem Grund wird folgende Definition vorgeschlagen:

Zum Sprachkodex einer Sprache gehören alle metasprachlichen Schriften, die für eine Sprachgemeinschaft zu einem bestimmten Zeitpunkt als Normautoritäten zur Verfügung stehen und von ihr auch als Normautorität wahrgenommen werden. Der Sprachkodex kann sich auf unterschiedliche Ebenen und Instanzen der Sprache beziehen: Aussprache, Schreibung, Grammatik (Wort- und Wortformenbildung, Syntax), Lexik, Semantik, Pragmatik. (Klein 2014: 222)

Das Zitat lässt erkennen, dass ein Sprachkodex normalerweise nicht nur aus einer einzigen Schrift besteht, sondern durch mehrere Texte geregelt werden kann. Diese Beobachtung lässt sich zwar auf den Bereich der Orthografie übertragen, da hierbei der Kodex mit einer einzigen rechtsförmigen Schrift geregelt ist. Alle anderen Sprachebenen sind jedoch für die deutsche Sprache nicht „staatlich geregelt.“ (Klein 2014: 222) Während es beispielsweise den Duden oder das österreichische Wörterbuch gibt, existiert keine Bibel der Syntax oder der Grammatik (vgl. Schäfer 2018: 18f.). Schäfer (2018: 19) schlägt daher die Duden-Bände als Norm vor. Ammon (2005: 35) hält dagegen, dass der tatsächliche Gebrauch der Schriften und die faktische Reichweite innerhalb der Sprachgemeinschaft von entscheidender Bedeutung ist und somit keine offizielle Vorgabe für die Verwendung eines Kodex notwendig ist. Es bildet sich zwar bewusst oder unbewusst ein Konsens darüber, aber welche Texte eine maßgebliche Rolle für eine Sprachgemeinschaft spielen und somit zum zeitgenössischen Kodex gerechnet werden können, kann nicht immer empirisch überprüft werden. Für den gesamten deutschen Sprachraum wird in der Regel der Duden (2006) oder das Regelwerk des Rates der deutschen Rechtschreibung (Rat für deutsche Rechtschreibung 2018) herangezogen. Durch das Anführen der korrekten standardsprachlichen Variante wird auch implizit deutlich gemacht,

was als nicht normgerecht gilt. Im Hinblick auf Österreich bezeichnet Ammon (1995: 96) das „Österreichische Wörterbuch“ als „zentralen und unzweifelhaftesten“ Sprachkodex.

Als weitere normsetzende Instanz dienen die *Modellsprecher:innen* oder *Modellschreiber:innen*. Sie gelten als sprachliche Vorbilder, an welchen sich Normautoritäten orientieren beziehungsweise auch auf sie verweisen, wenn es darum geht, sprachlich nach der Norm zu handeln (vgl. Ammon 2005: 33). Beispiele dafür wären Autor:innen, Nachrichtensprecher:innen, Journalist:innen und Schauspieler:innen (vgl. Ammon 2005: 33). Diese verfassen meistens sogenannte *Modelltexte*, worunter Prosa- oder Sachtexte verstanden werden, welche von den anderen Instanzen anerkannt werden (vgl. Ammon 2005: 33).

Die vierte und letzte Instanz machen die *Sprachexpert:innen* aus, unter welche beispielsweise Linguist:innen vom Fach oder Laien fallen, auf die sich in Zweifelsfällen auch entgegen der Norm in den Kodizes bezogen werden kann (vgl. Ammon 2005: 36).

Das Wirken aller vier Instanzen aufeinander zeigt sich dadurch, dass sich Sprachnormautoritäten, wie beispielsweise Lehrpersonen, aber auch die anderen Instanzen auf den Sprachkodex beziehen. Die Kodifizierer:innen achten auf Kritiken der Sprachexpert:innen und beobachten Modellschreiber:innen und –sprecher:innen, welche wiederum von Verlagslektor:innen und Radiodirektor:innen kontrolliert werden. Dieses System vollzieht sich nicht ständig, aber beispielsweise deutlich bei Versuchen einer Normänderung (vgl. Ammon 2005: 37).

Um verstehen zu können, wie Korrekturen, welche sich an der Standardsprache als Norm orientieren, aussehen dürfen, muss zunächst geklärt werden, was ein sprachlicher Fehler ist.

2.1.4 Der sprachliche Fehler

„Ein Fehler ist eine Abweichung von der geltenden linguistischen Norm.“ (Kleppin 2004: 19) Mit diesem Zitat setzt Kleppin bei der vorher genannten Erklärung an, die Standardsprache als linguistisches Regelsystem zu betrachten und jegliche Abweichungen davon als Fehler. Grundsätzlich lässt sich der linguistische Fehlerbegriff in Norm- und Systemfehler unterteilen (vgl. Coseriu 1988: 52-53). So können in unterschiedlichen Kommunikationssituationen verschiedene Normen gebraucht werden, sodass die Verwendung sprachsystematisch zwar angemessen ist, aber für den schriftsprachlichen Standard hingegen als Normfehler angesehen

werden würde (vgl. Dürscheid 2012: 109). Darüber hinaus können auch normalerweise nicht gebrauchte Wörter vorkommen, die zwar nach dem Regelwerk korrekt gebildet werden, jedoch in den Kodizes nicht vorzufinden sind (vgl. Dürscheid 2012: 109). Obwohl diese nach den Grundregeln verwendet werden, sind es dennoch Normfehler. Zusammenfassend kann somit gesagt werden, dass es sich trotz Normverstößen dabei nicht um inkorrekte, sondern höchstens um unangemessene Formulierungen handelt (vgl. Abel & Glaznieks 2017: 242).

Diesen Normfehlern stehen die Systemfehler gegenüber, welche laut Eisenberg und Voigt (2007: 212) „unter allen Umständen Fehler“ bleiben, da sie grammatikalische Regeln brechen (vgl. Ágel 2008: 65; Hundt 2009: 133). So wird aus einer Entscheidung über richtig und falsch eine differenzierte Einschätzung der verwendeten Sprache angedacht. In der modernen Linguistik wird deshalb versucht, den Produktionskontext zu beleuchten und erst dann eine Feststellung darüber zu treffen, ob eine Korrektur notwendig ist.

Darüber hinaus wird zwischen Kompetenzfehlern oder sogenannten strukturellen Verstößen und Performanzfehlern unterschieden (vgl. Corder 1967). Zu Letzterem zählen beispielsweise Flüchtigkeits- oder Tippfehler. Kompetenzfehler geschehen meistens durch die Unsicherheit und Unwissenheit der Lernenden (vgl. Hennig 2012: 134). Dies stellt gerade für die Lehrpersonen einen wichtigen Hinweis dar, welche Inhalte im Unterricht noch geübt werden sollten (vgl. Steuer 2014: 21). Außerdem kann die Unterscheidung der Fehlerarten bei der Fehlerkorrektur und der Gewichtung wichtig sein (vgl. Hennig 2012: 142). Diese scheint sich in der Praxis jedoch als schwierig zu gestalten. So merkt auch Nussbaumer (1991: 75) an: „Fehler, die systematisch auftreten, deuten eher auf Kompetenz-Fehler hin. Fehler, die unsystematisch auftreten, sind noch lange kein Indiz für bloße Performanz-Fehler; bekanntlich gibt es bei Unsicherheit über eine Regel die Strategie des Ausprobierens.“ Die Unterscheidung genau zu treffen stellt sich auch als schwierig heraus, wenn beispielsweise ein Buchstabe in einem Wort fehlt. Dabei bleibt fraglich, ob das Regelbewusstsein vorhanden ist und dieser lediglich vergessen wurde, oder ob der Lernende die korrekte Schreibweise nicht kennt. Ohne den Schreibprozess des Schülers oder der Schülerin zu kennen, kann dies nicht genau beurteilt werden. Deshalb kann bezüglich Performanz- und Kompetenzfehlern bei einem einzigen Schreibprodukt lediglich spekuliert werden. Des Weiteren müssen sich Lehrpersonen auch über sprachliche Zweifelsfälle bewusst werden, wo aufgrund Tendenzen des Sprachwandels Uneinigkeit darüber herrscht, was als angemessen und nicht angemessen gilt, oder ob es sich überhaupt wirklich um einen Fehler handelt (vgl. Klein 2003: 7). Es gibt Hinweise darauf, dass sie nach ihren eigenen Norm- und Sprachvorstellungen korrigieren

(vgl. Peter 2011: 280). Dies bedeutet, dass das als korrekt eingestuft wird, was sie auch selbst als richtig erlernt haben (BIFIE 2013: 4) und es deshalb oftmals zu genannten Schwierigkeiten kommt.

Als Veranschaulichung der genannten Ausführungen soll der Fehlerkreislauf nach Hennig (2012) aufgezeigt und kurz erläutert werden.

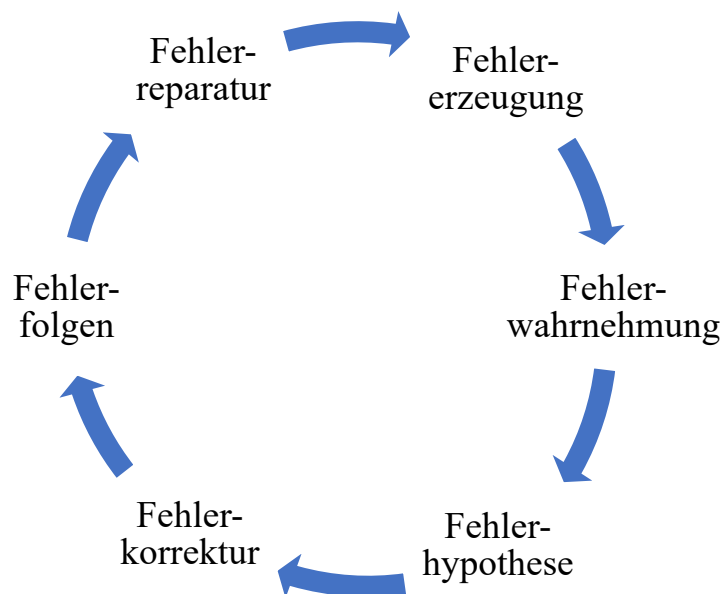


Abbildung 2: Fehlerkreislauf, eigene Darstellung nach Hennig (2012: 125)

Die *Fehlererzeugung* geschieht beim Produzieren eines Textes eines Schülers oder einer Schülerin. Zu diesem Zeitpunkt ist ihnen nicht klar, dass es sich dabei um eine fehlerhafte Konstruktion handelt (vgl. Hennig 2012: 129).

Ob sprachliche Abweichungen als Fehler gesehen werden, hängt von der bewussten *Fehlerwahrnehmung* der Lehrkraft ab (Hennig 2012: 129). Die subjektive Ansicht und Einstellung von Normen kommen hierbei einerseits zu Tragen und andererseits die Verfassung der Lehrperson während des Bewertungsprozesses und der Bezugsrahmen (vgl. Davies et al. 2017).

Besonders die nächsten drei Schritte des Fehlerkreislaufes spielen für vorliegende Masterarbeit eine Rolle und verdeutlichen nochmals die oben genannten Ausführungen zum sprachlichen Fehler.

Beim Schritt der *Fehlerhypothese* klassifiziert die Lehrperson den Fehler und bildet eine Annahme über die korrekte Zieläußerung. Die Einordnung in die verschiedenen Klassifikationssysteme hängt von der Interpretation über die Intention des Schülers oder der Schülerin ab (Hennig 2012: 130). Beispielsweise kann es sich bei „Ich schenke ihm ein Ball“

um einen Fehler des Artikels des Akkusativobjekts handeln oder aber um eine Verwechslung des Wortes wie „Ich schenke ihm ein Buch.“ An diesem Beispiel lässt sich erkennen, dass Fehler anhand der von der Lehrperson gebildeten Hypothese bestimmten Fehlerkategorien zugeordnet (vgl. Hennig 2012: 126) und dementsprechend im nächsten Schritt der *Fehlerkorrektur* korrigiert werden können. Hierbei wird also der Fehler zum ersten Mal sanktioniert. Die meisten Lehrpersonen handhaben dies durch das Anstreichen im Text und/oder der Markierung eines Korrekturzeichens. Dadurch werden die Wahrnehmung und die Fehlerklassifikation des oder der Beurteilenden erkennbar (vgl. Hennig 2012: 130). Die Fehlertoleranz spielt hierbei eine entscheidende Rolle. Laut Hennig (2012: 127) ist diese bei Lehrenden niedriger ausgeprägt als bei den Schüler:innen, was unter anderem institutionell bedingt ist.

Anhand der gemachten Fehler und weiterer qualitativer Merkmale wird der Schüler:innentext schließlich benotet (vgl. Hennig 2012: 130). Der Schritt der *Fehlerfolgen* stellt also das Benoten dar, wo „die Bewertung in eine Ziffer kodiert wird“ (Decker & Kaplan 2019: 122).

Für die Lehrperson ist die Beschäftigung mit dem Schüler:innentext hiermit abgeschlossen. Auf der letzten Ebene der *Fehlerreparatur* sollen die Schüler:innen dazu veranlasst werden, den Fehler aufzuarbeiten und ihn somit in Zukunft zu vermeiden (vgl. Hennig 2012: 131). Für vorliegende Masterarbeit ist dieser letzte Schritt jedoch zu vernachlässigen, da es sich bei der verwendeten Schularbeit um einen abschließenden schriftlichen Text handelt und sich die Forschung lediglich mit dessen Korrektur und Beurteilung beschäftigt. Das nächste Kapitel beschreibt daher zunächst eingehender die Formen der Korrektur.

2.2 Korrektur

Lange Zeit wurden Korrekturen so durchgeführt, dass Fehler gewichtet und quantifiziert wurden (vgl. Kleppin 2010: 1064). Obwohl dies in der Praxis teilweise häufig immer noch so gehandhabt werden dürfte, wird heutzutage vermehrt auf Kriterienkataloge gesetzt, um ein standardisiertes Verfahren durch qualitative Gesichtspunkte statt lediglich quantitative Berechnungen gewährleisten zu können. Kleppin (2010: 1064) betont dahingehend, dass auch bei dieser Korrekturpraxis Fehlermarkierungen von Nöten sind, um den Schüler:innen ihre Normverstöße aufzeigen zu können. Für Lehrpersonen bieten sie eine Übersicht beziehungsweise eine Orientierung über die erbrachte Leistung.

2.2.1 Korrekturstrategien

In der Literatur werden häufig vier verschiedene Strategien erwähnt, welche für eine Korrektur von schriftlichen Leistungen angewandt werden können.

Bei der „einfachen Fehlermarkierung“ (Kleppin 2010: 1064) oder der indirekten Fehlerkorrektur („indirect correct feedback“, Ellis 2009: 100) wird den Schüler:innen zwar anhand von Unterstreichungen oder Auslassungssymbolen angezeigt, dass ein Fehler besteht, dieser wird jedoch nicht korrigiert (vgl. Ellis 2009: 100; Kleppin 2010: 1064). Besonders bei Performanzfehlern scheint diese Strategie geeignet, da die Lernenden eigentlich über die Korrektheit der Aussage Bescheid wissen.

Demgegenüber steht die direkte Korrektur („direct corrective feedback“, Ellis 2009: 99) oder nach Kleppin (2010: 1064) die „Berichtigung durch den Lehrer“. Dabei wird bei der Korrektur auch die richtige Form der fehlerhaften Formulierung aufgezeigt. Sinnlose beziehungsweise inkorrekte Satzteile können hierbei ergänzt oder gänzlich gestrichen werden oder die korrekte Phrase wird darüberschrieben (vgl. Ellis 2009: 99; Kleppin 2010: 1064). Diese Vorgehensweise ist vor allem dann sinnvoll, wenn es sich nicht um Performanzfehler handelt, sondern die Schüler:innen nicht das sprachliche Wissen besitzen. Durch diese Strategie können die Lernenden sofort die korrekte Version einsehen (vgl. Ellis 2009: 99).

Bei der „Fehlermarkierung mit Korrekturzeichen“ (Kleppin 2010: 1064) oder der metalinguistischen Korrektur („metalinguistic corrective feedback“, Ellis 2009: 100) geht es um die Kategorienbildung der verbesserten Fehler (vgl. Ellis 2009: 100; Kleppin 2010: 1064). Dabei werden die fehlerhaften Phrasen durch Korrekturzeichen in linguistische Kategorien eingeteilt. Vor allem im Hinblick auf Abschlussarbeiten macht diese Strategie Sinn, da die Schüler:innen hierbei keine Gelegenheiten mehr bekommen, den Text zu überarbeiten. Für Lehrpersonen stellt es dennoch eine Erleichterung bei der Einschätzung von Fehlerkategorien und für ihre Beurteilung dar, denn sie können darauf bei der Begründung der Note für Eltern und Schüler:innen Bezug nehmen (vgl. Kapitel 2.1.4). Ellis (2009: 100-101) schreibt diesbezüglich über die Frage der Einschränkung der Kategorien und merkt an: „Most of the error codes used in research and language pedagogy employ relatively broad categories“. Wie grob oder detailliert dies von jeder Lehrperson gehandhabt wird, wird in der qualitativen Forschung näher untersucht.

Bei der vierten Strategie handelt es sich um eine weitere Stufe der metalinguistischen Korrektur, bei welcher zusätzlich noch Erklärungen zu den gemachten Fehlern geliefert werden (vgl. Ellis 2009: 101). Diese vor allem für den Lernprozess der Schüler:innen sinnvolle Strategie kann in vorliegender Masterarbeit vernachlässigt werden, da die Studie

lediglich eine punktuelle Überprüfung darstellt und nicht der gesamte Prozess eines Kindes untersucht wird.

An der Korrekturpraxis in österreichischen Schulen kritisiert Hennig (2012: 133) „die tradierten Korrekturzeichen (Gr = Grammatik, R = Rechtschreibung, Z = Zeichensetzung, T = Tempus, SB = Satzbau etc.)“ aufgrund fehlender Standardisierung und Begründung. Dahingehend postuliert sie:

Das Fehlen eines Regelwerks zur Fehlerkorrektur und -klassifikation lässt darauf schließen, dass die Fehlertypen tatsächlich auf Tradierung und Konvention beruhen und nicht auf linguistisch und/oder didaktisch motivierten Überlegungen. (Hennig 2012: 134)

Trotz alldem wird an konventionellen Methoden festgehalten, da es sich um ein vertrautes und verständliches Einteilungssystem innerhalb der Schulgemeinschaft handelt (vgl. Hennig 2012: 134).

2.2.2 Kriterienkataloge

Von zentraler Bedeutung für die Leistungsbeurteilung im Fach Deutsch nach der Korrektur von Fehlern ist auch die Frage nach der Textqualität, also was denn überhaupt ein guter Text ist und ab wann er als gelungen gilt. Wie sich an den bereits genannten Ausführungen sehen lässt, ist diese Frage nicht leicht zu beantworten. Laut Becker-Mrotzek und Böttcher (2006: 89) unterliegt die Bewertung von Texten unterschiedlichen Kriterien. Diese sind einerseits abhängig vom Vorwissen des Adressaten beziehungsweise der Adressatin und andererseits von der Kommunikationssituation. Darüber hinaus spielt die Textfunktion eine wesentliche Rolle (vgl. Becker-Mrotzek & Böttcher 2006: 90).

Da in der Schulpraxis vermehrt Kriterienkataloge zum Einsatz kommen und auch die Reife- und Diplomprüfung mit einem Raster vonstattengeht, soll im Folgenden näher auf die beiden bekanntesten Kriterienkataloge eingegangen werden, welche für schreibdidaktische und unterrichtliche Zwecke stetig weiterentwickelt wurden.

Das *Zürcher Textanalyseraster* von Nussbaumer (1996) entstand aus einer fachdidaktischen Forschung, mithilfe dessen 200 Schweizer Maturant:innen untersucht wurden. Damit fand es den Einzug in die Schreibdidaktik, obwohl es nicht für den Unterricht gedacht war. Es diente nämlich dazu, „zu Einsichten in die Sprachfähigkeiten der jungen Schreiberinnen und Schreiber“ zu gelangen (Nussbaumer 1996: 97f.). Die verfassten Texte sollten dabei als ein „Spiegel dahinterliegender Fähigkeiten“ (Nussbaumer 1996: 97) gesehen werden und diese „möglichst umfassend und systematisch, explizit und reflektiert“ (Sieber 1994: 149)

analysieren. Für die Normvorstellungen des Unterrichts, welche noch immer in richtig und falsch kategorisiert werden, eignet sich das Raster laut Nussbaumer (1996: 103) deshalb nicht. Dennoch wurde es von vielen Fachdidaktiker:innen übernommen und das Konzept für verschiedenste Kriterienkataloge weiterentwickelt (z.B. Becker-Mrotzek 2003, Birkel 2003, Baurmann 2006, Nutz 2006, Becker-Mrotzek & Böttcher 2006).

Bei der Arbeit mit dem Zürcher Textanalyseraster geht es primär darum, den vorliegenden Text vor der Bewertung so zu sehen, wie er ist. Mit dieser neutralen Verwendungsart ist es möglich, den Blick nicht direkt auf die Mängel zu richten, sondern den Text mit seinen globalen und ins Einzelne gehenden Kennzeichen zu erkennen (vgl. Sieber 1994: 150). Um zu solch einer differenzierten Sichtweise zu gelangen, müssen bestimmte Fragen an einen Text gestellt werden. Diese werden daraufhin in einem systematisch gruppierten Katalog bereitgestellt (vgl. Sieber 1994: 142). Dazu müssen zunächst alle Auffälligkeiten in einem Text erfasst werden, um in einem weiteren Schritt die vorerst „neutral“ erfassten Merkmale in den Raster einzuordnen. Hierbei lässt sich bereits der Unterschied zum schulischen Kontext erkennen, da die Orientierung an Mängeln unterbrochen wird. Danach sollen Ersteindrücke notiert und begründet werden (vgl. Binder & Harrer 1998: 9), um die Konzentration auf die Gewichtungen von einzelnen Auffälligkeiten zu lenken. „Mit Hilfe des Rasters scheinen verborgene Zusammenhänge zwischen bestimmten Stärken oder Schwächen in einem Text auf, und es werden umgekehrt Vorannahmen über bestehende Zusammenhänge überprüfbar“. (Sieber 1994: 150)

Der Katalog des Zürcher Analyserasters gliedert sich in drei Teile, den Null-Abschnitt und einen A- und B-Abschnitt, wobei es mit B1, B2 und B3 insgesamt fünf Dimensionen der Bewertung gibt. Der Null-Abschnitt beinhaltet die Grundgrößen, also objektive Eigenschaften des Textes, wie beispielsweise die Länge, syntaktische Muster, den Wortschatz oder die Kohäsionsstiftung, also inwiefern ein roter Faden gegeben ist (vgl. Nussbaumer 1996: 99f.).

Der A-Abschnitt beinhaltet die sprachformale und orthographische Richtigkeit einschließlich Syntax und Semantik (vgl. Nussbaumer 1996: 99f.).

Der erste Teil des B-Abschnittes, also B1, beschreibt die funktionale Angemessenheit des Textes. Darunter fallen die Verständlichkeit und Kohärenz des Textes. Fragen nach der Gesamtidee und dem Aufbau des Textes werden hier genauso analysiert wie die thematische Entfaltung und die Angemessenheit der Sprachmittel. Darüber hinaus werden auch Ausdrücke der Rezipient:innenführung und Textmusternorme, sowie der Grad an Explizitheit begutachtet (vgl. Nussbaumer 1996: 100f.).

Unter B2 fällt die ästhetische Angemessenheit. Darunter werden besondere formale Qualitäten, wie das sprachlich-formale Wagnis und die Attraktivität der Sprachmittel bezüglich der Wortwahl, des Satz- und Textbaus, des Rhythmus und der Tonwahl verstanden. Innerhalb B3 wird auf die inhaltliche Relevanz des Textes geachtet. Darunter fallen das inhaltliche Wagnis und die inhaltliche Wegqualität (vgl. Nussbaumer 1996: 100f.).

Wie an den Ausführungen erkennbar ist, kann dieses Raster für die Analyse von sprachlich-textuellen Qualitäten eines Textes dienen, unabhängig davon, ob negative oder positive Aspekte ermittelt werden. Um es jedoch für den Unterricht brauchbar zu machen, entwickelten es Becker-Mrotzek und Böttcher (2006: 92) weiter, indem sie entwicklungssensitive Kriterien zufügten, die den Entwicklungsstand und den Lernfortschritt berücksichtigen. Dadurch ergibt sich der sogenannte „Basiskatalog“, welcher aus insgesamt zwölf Kriterien besteht. Diese verteilen sich auf fünf Basisdimensionen, nämlich Aufbau, Inhalt, Sprachrichtigkeit, Sprachangemessenheit und Schreibprozess (vgl. Becker-Mrotzek & Böttcher 2006: 92). Die Kriterien können in der Regel auf jede Textsorte angewandt werden, jedoch müssen sie bezüglich der Schreibaufgaben und der Bewertungsziele nachjustiert werden. Dazu zählt das Konkretisieren der allgemeinen Formulierungen im Hinblick auf die verwendete Textsorte und darüber hinaus die Berücksichtigung der didaktischen Ziele, welche damit verfolgt werden. Somit können sich Verschiebungen der einzelnen Dimensionen ergeben.

Mit diesen Hilfsmitteln kann laut Becker-Mrotzek und Böttcher (2006: 92) die Textqualität gemessen werden. Vor allem für das Benoten von Schüler:innentexten ist die Erstellung eines Kriterienkatalogs laut ihnen unabdingbar, da damit Beurteilungsfehler vermieden werden können. Dies kann durch die dadurch erzielte Verbesserung der Gütekriterien der Objektivität, Reliabilität und Validität gewährleistet werden. Astrid Neumann merkt im Forschungshandbuch der empirischen Schreibdidaktik diesbezüglich ebenfalls an, dass die spezifischen Einzelkategorien eines differenzierten Kriterienkataloges für eine Stabilisierung der Beurteilungsmaße sorgen, da die dadurch zusammengeordnete Gesamtskala deutlich haltbarer ist (vgl. Becker-Mrotzek et al. 2016: 210). Im Gegensatz dazu liefert bereits eine Studie aus dem Jahr 1975 Hinweise, dass dem nicht so ist. Dabei hat sich in einer Untersuchung Schröters (1975: 49) gezeigt, dass die Notenstreuungen durch Kriterienkataloge nicht geringer werden. Laut Mayr und Schubert (1985: 13) können differenzierte Kriterien zur Beurteilung „zwar die Gültigkeit der Beurteilung verbessern, nicht aber die Objektivität.“ (Mayr & Schubert 1985 zitiert nach Schmidinger & Vierlinger 2012: 129).

Trotz dieser Hinweise und der ressourcen- und zeitaufwändigen Handhabung des Basiskataloges, bringt er dennoch einige weitere Vorteile mit sich. Dieses Schema sichert eine gute und schnelle Anwendbarkeit für die Bewertenden, auch und vor allem für die Unerfahrenen. Darüber hinaus kann anhand der Darstellung der Ergebnisse ein umfassendes Feedback gegeben werden (vgl. Becker-Mrotzek et al. 2016: 211). Außerdem lassen sich durch die Beurteilung anhand des Basiskataloges die systematischen Fehler beziehungsweise auch gelungene Merkmalsausprägungen sichtbar machen (vgl. Hattie et al. 2013).

Die angeführten Kriterienkataloge stellen die Weichen für Entwicklungen von weiteren Rastern. Im Schulkontext bleibt es jeder Lehrkraft selbst überlassen, ob und welchen Kriterienraster diese verwendet. Darüber hinaus können auch eigene Raster mit eigens gesetzten Schwerpunkten kreiert werden. Eine Studie von Decker und Kolleg:innen (2016: 8) konnte aufzeigen, dass die häufigste Leistungsbeurteilungsform Kriterienkataloge sind. Begründet wird dies dadurch, dass auch bei der Matura beziehungsweise dem Abitur mit Beurteilungsrastern bewertet wird (vgl. Decker et al. 2016: 8) und die Schüler:innen als auch die Lehrer:innen somit früher oder später damit konfrontiert werden.

Da sich vorliegende Masterarbeit mit der Leistungsbeurteilung von Schüler:innentexten auseinandersetzt und im vorigen Absatz bereits von Leistungsfeststellungen gesprochen wurde, sollen diese Begrifflichkeiten im folgenden Kapitel näher definiert und auch rechtliche Erläuterungen dazu angeführt werden.

2.3 Leistungsbeurteilung

Um sich an den Begriff der Leistungsbeurteilung anzunähern, soll zunächst der allgemeine Leistungsbegriff definiert werden.

2.3.1 Leistungsbegriff

Der Leistungsbegriff spielt für die heutige Gesellschaft eine zentrale Rolle. Vor allem in Österreich wird häufig von einer Leistungsgesellschaft gesprochen, was schon anhand der viel verwendeten Begriffe wie beispielsweise Leistungsniveau, Leistungserfolg, Leistungserwartung oder Höchstleistung zu Tragen kommt (vgl. Fuss 2002: 12).

Auch Hartfiel (2000: 377) formuliert, dass Leistung in Österreich einen hohen Stellenwert hat und gewissermaßen als Ranginstrument funktioniert:

Wir leben in einer Leistungsgesellschaft, d.h. Leistung dient unter diesen sozialen Gegebenheiten zentral und prinzipiell als Verteilungsschlüssel für gesellschaftliche Positionen, für Verdienst, Macht und Einfluss, also als umfassender „Verkehrsregelungsgrundsatz unseres Zusammenlebens. (Hartfiel zitiert nach Böhm 2000: 377)

In der Literatur finden sich verschiedene Definitionen zum Begriff Leistung, da es sehr viele Erklärungs- und Interpretationsmöglichkeiten gibt. Im Folgenden soll ein Grundcharakteristikum dieses Begriffs anhand verschiedener Definitionen erstellt und im Detail besonders auf die psychologische und pädagogische Komponente eingegangen werden.

Mit der Definition von Florek (1999: 13f.) wird ein Versuch gewagt, den Begriff mit einem allgemeingültigen Verständnis einzugrenzen, dass Leistung immer durch eine selbstständige Tätigkeit entsteht. Demnach kann Leistung nichts sein, was jemandem widerfährt (vgl. Florek 1999: 13f.). Laut Bartnitzky (2000: 114) wohnt es grundsätzlich jedem Menschen inne, lernen zu wollen und etwas zu leisten. Diese Annahme wird beispielsweise auch von Sacher (2014: 13) gestützt, welcher postuliert, dass Leistungsstreben ein grundlegendes Bedürfnis aller Menschen darstellt, „von dessen Befriedigung seine Selbstverwirklichung ganz entscheidend abhängt.“ Vor allem im Schulkontext wird jedoch eine gewisse Leistungserbringung gefordert und stellt somit laut Böhm (2000: 377) „Anstrengung und Aktivitäten, die aufgrund von Leistungsanforderungen vollzogen werden als auch deren Ergebnis“ dar. Klafki (1976: 90) ergänzt diesen Leistungsbegriff um die Gütemaßstäbe, welche im Kapitel 2.3.2 näher betrachtet werden und postuliert:

Versteht man Leistung als Ergebnis und Vollzug einer Tätigkeit, die mit Anstrengung und gegebenenfalls mit Selbstüberwindung verbunden ist und für die Gütemaßstäbe anerkannt werden, die also beurteilt wird, so erfordern die genannten Zielsetzungen vom jungen Menschen ein hohes Maß an Anstrengung und spezifischem Können. (Klafki 1976: 90)

Dabei wird auch gleichzeitig der erzieherische Anspruch deutlich. Bereits an diesen Zitaten lässt sich ein großer Interpretationsspielraum bezüglich des Begriffs der Leistung erkennen, die Unterscheidung zwischen einem psychologischen und pädagogischen Leistungsbegriff fällt dennoch auf. Vor allem bei letzterem wird deutlich, dass immer eine Art von Anstrengung oder Anforderung beziehungsweise Gütemaßstäben und damit verbundenen Bewertungen einbezogen wird. Obwohl das Leisten also ein menschliches Grundbedürfnis darstellt, bedeutet dies grundsätzlich nicht, dass danach verlangt wird, Leistungsanforderungen zu erfüllen (vgl. Jürgens 2006: 6). Dem Schulkontext liegt das

Leistungsprinzip jedoch so weit zugrunde, dass Schüler:innen lediglich aufgrund ihrer Leistungserbringung in der Schule beurteilt werden (vgl. Criblez 2002a: 85). Inwiefern dies nun kritisch zu betrachten ist, sei dahingestellt, aber aufgrund des Status quo der Leistungsschule ist es wichtig, die Leistungen der Schüler:innen angemessen feststellen zu können, um diese überhaupt in einem weiteren Schritt zu beurteilen.

2.3.2 Leistungsfeststellung

Bei der Leistungsfeststellung geht es darum, einen Sachverhalt zu erheben, bevor diese Leistung dann benotet werden kann. Konkret geht es also um die Messung einer Leistung und die Frage, inwiefern diese von den Vorgaben oder dem Lernziel abweicht (vgl. Benischek 2006: 92). Die Grundlage dafür bieten die Lehrpläne mit den seit der Bildungsreform verfassten Kompetenzen und den Zielen und Inhalten für eine Leistungsbeurteilung (vgl. Criblez 2002a: 85). Die Problematik daran ist jedoch, dass Lehrpläne diesbezüglich allgemein gehalten sind und die Lehrperson damit weder das „Messobjekt“ noch das Messinstrument vorgegeben bekommt. Darüber hinaus müssen zwar unter Einhaltung der Rechtsvorschriften, aber dennoch bis zu einem gewissen Grad selbst Kriterien entwickelt werden, was als richtig und als falsch zu werten ist. Umso komplexer die zu beurteilende beziehungsweise messende Leistung ist, desto schwieriger wird ein konkretes Ablesen (vgl. Neuweg 2014: 9ff.). Gerade im Bereich der Textproduktion trifft diese Aussage Neuwegs zu, denn wie schon Baumann (2006: 120) erläutert, ist ein Text aufgrund der Vielfalt an relevanten Aspekten eine komplex zu erfassende Leistung. Deshalb ist es wichtig, dass Leistungsmessungen den Gütekriterien der Objektivität, Reliabilität und Validität entsprechen.

Objektiv ist eine Erhebung dann, wenn mehrere beurteilende Personen zum gleichen Ergebnis kommen (vgl. Tortik 2016: 41). Vor allem dieser Aspekt wird in vorliegender Masterarbeit untersucht, deshalb soll die weitere Unterscheidung zwischen der Durchführungsobjektivität, Auswertungsobjektivität und Interpretationsobjektivität im Folgenden angeführt werden.

Erstere beschreibt die Reglementierung und Standardisierung der Durchführungsbedingungen, unter welchen die diagnostischen Informationen eruiert werden. Grundvoraussetzungen dafür sind beispielsweise die Festlegung von erlaubten Hilfsmitteln, der Bearbeitungszeit und der Prüfungsgestaltung (vgl. Sacher 2014: 35).

Bei der Auswertungsobjektivität geht es um Maßnahmen, welche für eine Feststellung von erhobenen Leistungen notwendig sind, sodass die Korrektur unabhängig zu der zu beurteilenden Person erfolgt (vgl. Sacher 2014: 35-36). Neuweg (2014: 120) schlägt vor, dass

sich Lehrpersonen in der Gemeinschaft über das Niveau von Aufgaben, über Vorgaben und Gewichtungen zur Korrektur austauschen und einigen. Somit sollen verschiedene auswertende Personen zu einem identen Ergebnis kommen (vgl. Sacher 2014: 36).

Bei der Interpretationsobjektivität geht es um die schlussendliche Bewertung der eruierten und ausgewerteten Information. Wenn eine gleiche Beurteilung einer Leistung durch unterschiedliche Lehrende erfolgt, kann von hoher Interpretationsobjektivität gesprochen werden (vgl. Sacher 2014: 36).

Kollegiumsinterne, einheitliche Bewertungskriterien, unabhängig des Leistungsniveaus einer Klasse und gemeinsam festgelegte Mindestanforderungen stellen eine unerlässliche Grundlage für eine übereinstimmende Interpretation der Prüfungsleistung dar (vgl. Tent & Stelzl 1993: 44). Der Einschätzung der Beurteilenden liegt also der gleiche Maßstab zugrunde (vgl. Sacher 2014: 36).

Da das Hauptanliegen vorliegender Arbeit lediglich die Objektivität bei Beurteilungen betrifft, werden die beiden weiteren Gütemaßstäbe nur kurz erläutert, um einen Überblick zu geben.

Reliabel ist eine Erhebung dann, wenn sie auch bei mehrmaliger Wiederholung zuverlässig misst. Die Messung sollte also so präzise sein, dass die Ergebnisse auch bei einer erneuten Anwendung der Testung dieselben sind (vgl. Sacher 2014: 36). Reliabilität ist deshalb bei Überlegungen wichtig, ob einem einmaligen Messergebnis vertraut und es aufgrund dessen als stabil angesehen beziehungsweise das Ergebnis verallgemeinert werden kann (vgl. Becker-Mrotzek et al. 2016: 206).

Das dritte Gütekriterium ist die *Validität*. Diese ist dann gegeben, wenn eine Erhebung wirklich nur das misst, was gemessen werden soll. Dabei wird auch von der Gültigkeit einer Testung gesprochen (vgl. Becker-Mrotzek et al. 2016: 206). Für eine Kontrolle der Validität muss danach gefragt werden, was beispielsweise alles mitgemessen wird. Bei Textproduktionen, wie auch im konkreten Fall vorliegender Masterarbeit, kann dies zum Beispiel der Ausdruck, der Inhalt, die Sprachrichtigkeit, die Textlänge und die Orthografie sein (vgl. Neuweg 2014: 120). Die Handschrift stellt ein Beispiel dar, welches eigentlich nicht in die Leistungsfeststellung miteinbezogen werden dürfte. Dennoch konnte das Gegenteil bereits anhand einiger Studien aufgezeigt werden (siehe Kapitel 2.4).

Laut einiger Meinungen und Untersuchungen (siehe etwa Haas 2008, Stern 2010, Neuweg 2014) werden diese Gütekriterien bei Leistungsfeststellungen und –beurteilungen in der Schule nicht erfüllt.

Das Ziel einer jeden Leistungsfeststellung wäre es demnach, den Leistungsstand eines Schülers oder einer Schülerin möglichst objektiv, reliabel und valide zu erheben. Aufbauend auf diesem Schritt kann die Leistungsbeurteilung erfolgen.

2.3.3 Leistungsbeurteilung

Bei der Leistungsbeurteilung wird der zuvor erhobene Sachverhalt anhand von Noten bewertet. Sie ist also ein Sachverständigengutachten durch die Lehrperson und diese ordnet die Leistung einer von mehreren Beurteilungsstufen zu. In Österreich reichen diese von „Sehr gut“ bis „Nicht Genügend“ (vgl. Tortik 2016: 200), welche im Anschluss an dieses Kapitel näher ausgeführt werden. Dies bedeutet, dass eine einzige Ziffer reicht, um das Ergebnis einer Leistungsüberprüfung oder den Leistungsstand eines Schülers oder einer Schülerin am Jahresende abzubilden (vgl. Wild & Krapp 2001: 541). Die Lehrperson ist an diesen Prozessen maßgeblich beteiligt, stellt sie doch die zu messende Größe her, fertigt das Überprüfungsmaterial, also das Messinstrument, an und misst beziehungsweise beurteilt das Ergebnis (vgl. Kiper 2001: 22). Dadurch können Noten laut Winter (2004: 40) auch als Schätzurteile der Lehrenden angesehen werden. Die Qualität der Leistungsbeurteilung liegt oftmals nicht im Messinstrument selbst begründet, sondern in dem Menschen, welcher eine Bewertung abgibt. Dies bedeutet, dass die Genauigkeit der Messung, das Vorgehen, die Kriterien und die Vergleiche die Qualität des Urteils maßgeblich bestimmen (vgl. Winter 2004: 40).

In der Theorie lässt sich der Unterschied beider Begriffe erkennen, in der Praxis führt dies jedoch häufig zu Missverständnissen, da Leistungsfeststellung und Leistungsbeurteilung nicht immer voneinander getrennt werden können. Eine Leistung zu messen, ohne sie zu bewerten, funktioniert im Kontext Schule nicht, da dies allein schon aufgrund der Rechtsgrundlage nicht möglich wäre (vgl. Benischek 2006: 96). Im dritten Paragraphen der Leistungsbeurteilungsverordnung wird demnach festgehalten, welche Formen der Leistungsfeststellungen für eine Beurteilung zulässig sind: die Mitarbeit der Schüler:innen im Unterricht, besondere mündliche Leistungsfeststellungen wie beispielsweise Prüfungen oder Übungen, besondere schriftliche Leistungsfeststellungen, wie beispielsweise Schularbeiten,

Tests, Diktate und ähnliches, besondere praktische Leistungsfeststellungen und besondere graphische Leistungsfeststellungen (vgl. Neuweg 2000: 10).

Für vorliegende Masterarbeit wird nur folgender aus den vielen allgemeinen Grundlagen der Leistungsbeurteilung genannt: „Die Leistungen der Schüler sind sachlich und gerecht zu beurteilen; es ist größtmögliche Objektivierung anzustreben.“ (vgl. Leistungsbeurteilungsverordnung 2022: Paragraph 11)

Um diese Aussage zu untersuchen, beschäftigt sich die Forschung vorliegender Arbeit genauer mit verschiedenen Leistungsbeurteilungen derselben Schularbeit. Da auch NMS-Lehrpersonen in der Stichprobe enthalten sind, soll die Leistungsbeurteilung an der Neuen Mittelschule im Folgenden thematisiert werden.

2.3.4 Leistungsbeurteilung an der Neuen Mittelschule

Die NMS startete mit einem vierjährigen Modellversuch im Jahr 2008/2009 und löste jede Hauptschule ab dem Schuljahr 2018/19 ab. Bereits seit März 2015 liegen erste Studien vor, welche die Ziele der Einführung der Neuen Mittelschule im Gegensatz zur Hauptschule überprüft haben. Dabei kam heraus, dass das Leistungsniveau nicht höher und meistens sogar niedriger ist als an Hauptschulen (vgl. Tortik 2016: 84). Tortik (2016: 84) berichtet auch von ersten Evaluierungen, dass leistungsschwache Schüler:innen nicht besser gefördert werden und das Ziel für mehr soziale Chancengleichheit und Bildungsgerechtigkeit nicht erreicht wurde. Was als Idee einer Gesamtschule begann, endete auch wie bei den Hauptschulen als Konkurrenzkampf mit der AHS. Wesentliche Weichenstellungen seitens der Schüler:innen sind zu dieser Zeit bereits geschehen und sie wurden in ihrem Lernverhalten beziehungsweise ihrem Vorwissen bereits geprägt (vgl. Tortik 2016: 84). Es gelang also nicht wie erhofft, die Schüler:innenströme von der AHS in die NMS umzulenken und eine Aufwertung der HS durch die Umstellung auf eine NMS zu erfahren. Laut der breiten Meinung des Lehrpersonals fehlt dadurch die soziale Durchmischung und damit geht unter anderem auch die Abnahme des Leistungsniveaus einher, wobei gleichzeitig gefordert wird, keine „Nicht genügend“ zu vergeben (vgl. Tortik 2016: 85). Aufgrund der genannten Punkte gibt es an den Neuen Mittelschulen ein abweichendes Beurteilungsschema. So erfolgt die Leistungsbeurteilung in der fünften und sechsten Schulstufe zwar wie in der AHS anhand des gewohnten Beurteilungsspektrums, welches von „Sehr gut“ bis „Nicht genügend“ reicht. In der 7. und 8. Schulstufe findet die Leistungsbeurteilung jedoch in abgewandelter Form statt (vgl. Tortik

2016: 250). Die Leistungsfeststellung und -beurteilung in den differenzierten Pflichtgegenständen erfolgt „nach den Anforderungen des Lehrplans nach grundlegenden und vertieften Gesichtspunkten.“ (Paragraf 18 Abs 2a SchUG) Anhand eines Zusatzes bei der Notenvergabe muss ersichtlich sein, ob die Leistung nach der grundlegenden oder vertieften Allgemeinbildung beurteilt worden ist (vgl. Kremser 2017: 137). Für eine bessere Übersichtlichkeit wird im Folgenden eine Abbildung von der Beurteilungsskala in der 7. und 8. Schulstufe der NMS dargestellt:

5.-6. Schulstufe (s. LBVO)	7.-8. Schulstufe
Sehr gut Gut Befriedigend Genügend	Sehr gut / vertiefte Allgemeinbildung Gut / vertiefte Allgemeinbildung Befriedigend / vertiefte Allgemeinbildung Genügend / vertiefte Allgemeinbildung
Nicht genügend	Befriedigend / grundlegende Allgemeinbildung Genügend / grundlegende Allgemeinbildung Nicht genügend / grundlegende Allgemeinbildung

Abbildung 3: Beurteilungsskala Neue Mittelschule, eigene Darstellung nach Bildungsdirektion Wien (Quelle: <https://www.nms.ssr-wien.at/index.php/neue-leistungsrueckmeldung-in-der-nms/leistungsbeurteilung-und-berechtigungen>)

Wie sich hier erkennen lässt, gibt es in der vertieften Allgemeinbildung kein „Nicht genügend“. In diesem Fall müssten die Anforderungen der grundlegenden Allgemeinbildung herangezogen und nach ihnen beurteilt werden. In dieser gibt es lediglich die Beurteilungsskala von „Befriedigend“ bis „Nicht Genügend“. Ein „Sehr gut“ oder „Gut“ würde demnach wiederum in den Bereich der vertieften Allgemeinbildung verortet und nach diesen Anforderungen beurteilt werden. Somit ergibt sich eine siebenstufige Beurteilungsskala für die 7. und 8. Schulstufe der Neuen Mittelschulen (vgl. Kremser 2017: 138). Darüber hinaus lässt sich keine parallele Existenz an der Abbildung erkennen. Die grundlegende Allgemeinbildung bildet die Basis und die vertiefte Bildung baut darauf auf. Obwohl sieben Beurteilungsstufen verwendet werden, sind die Benennungen und Definitionen rechtlich geregelt und somit bleiben weiterhin die gleichen fünf Noten wie für

alle anderen Schularten und -formen bestehen. Dies führt oftmals zu rechtlichen Problemen in der Leistungsbeurteilung (vgl. Kremser 2017: 139).

Die vertiefte Bildung geht dabei über die Grundanforderungen hinaus und weist einen höheren Komplexitätsgrad auf. Damit entspricht sie denselben Anforderungen wie in den entsprechenden Schulstufen der Unterstufe der AHS (vgl. Kremser 2017: 139).

Problematisch dabei ist, dass die Lehrpläne der vertieften und grundlegenden Allgemeinbildung ident sind und deshalb auch nicht ersichtlich ist, was denn nun unter einem „höheren Komplexitätsgrad“ zu verstehen ist (vgl. Westfall-Greiter 2013: 804). Werden die Notendefinitionen betrachtet, bleibt auch fraglich, inwiefern ein Schüler oder eine Schülerin einerseits gleichzeitig beispielsweise in einem „über das Wesentliche hinausgehende Ausmaß“ arbeitet und dieses andererseits aber lediglich „überwiegend erfüllt“. Die angeführte Kritik zeigt auf, dass die verwendete siebenstufige Skala der 7. und 8. Schulstufe in den Neuen Mittelschulen den Notendefinitionen der Leistungsbeurteilungsverordnung nicht gerecht wird und die so vorgenommenen Beurteilungen somit in rechtlicher Hinsicht nicht gedeckt sind (vgl. Kremser 2017: 143). Die gängige Praxis als auch die genannte Fachliteratur zeigen auf, dass vor allem auch aufgrund dieser Kritikpunkte Unsicherheiten bezüglich Korrekturen aufkommen, welche somit zu fehlender Nachvollziehbarkeit führen können (siehe Kapitel 2.1).

Die Leistungsbeurteilung schließt im Schulkontext in der Regel mit einer Note ab. Im Folgenden soll daher das österreichische Notensystem näher erläutert werden.

2.3.5 Notensystem

Wie bereits erwähnt, erfolgt die Leistungsbeurteilung an österreichischen Schulen durch Ziffernoten und diese werden vom Schulgesetz geregelt. Im Paragraphen 15 Absatz 2 der Leistungsbeurteilungsverordnung (LBVO) steht geschrieben, dass sogenannte Beurteilungsstufen von „Sehr gut“ bis zu einem „Nicht genügend“ verwendet werden müssen und diese „in Worten einzusetzen“ sind. Mit den einzelnen Notenstufen sind „die Selbstständigkeit der Arbeit, die Erfassung und die Anwendung des Lehrstoffes, die Durchführung der Aufgaben und die Eigenständigkeit des Schülers zu beurteilen.“ (Paragraph 18 Abs 3 SchUG)

Im Folgenden werden die einzelnen Stufen mit der zusätzlichen Erklärung aus der Leistungsbeurteilungsverordnung § 14 aufgelistet:

Beurteilungstufen (§ 14 LBVO)

- (1) Mit „Sehr gut“ sind Leistungen zu beurteilen, mit denen der Schüler die nach Maßgabe des Lehrplanes gestellten Anforderungen in der Erfassung und in der Anwendung des Lehrstoffes sowie in der Durchführung der Aufgaben in weit über das Wesentliche hinausgehendem Ausmaß erfüllt und, wo dies möglich ist, deutliche Eigenständigkeit beziehungsweise die Fähigkeit zur selbständigen Anwendung seines Wissens und Könnens auf für ihn neuartige Aufgaben zeigt.
- (2) Mit „Gut“ sind Leistungen zu beurteilen, mit denen der Schüler die nach Maßgabe des Lehrplanes gestellten Anforderungen in der Erfassung und in der Anwendung des Lehrstoffes sowie in der Durchführung der Aufgaben in über das Wesentliche hinausgehendem Ausmaß erfüllt und, wo dies möglich ist, merkbare Ansätze zur Eigenständigkeit beziehungsweise bei entsprechender Anleitung die Fähigkeit zur Anwendung seines Wissens und Könnens auf für ihn neuartige Aufgaben zeigt.
- (3) Mit „Befriedigend“ sind Leistungen zu beurteilen, mit denen der Schüler die nach Maßgabe des Lehrplanes gestellten Anforderungen in der Erfassung und in der Anwendung des Lehrstoffes sowie in der Durchführung der Aufgaben in den wesentlichen Bereichen zur Gänze erfüllt; dabei werden Mängel in der Durchführung durch merkbare Ansätze zur Eigenständigkeit ausgeglichen.
- (4) Mit „Genügend“ sind Leistungen zu beurteilen, mit denen der Schüler die nach Maßgabe des Lehrplanes gestellten Anforderungen in der Erfassung und in der Anwendung des Lehrstoffes sowie in der Durchführung der Aufgaben in den wesentlichen Bereichen überwiegend erfüllt.
- (5) Mit „Nicht genügend“ sind Leistungen zu beurteilen, mit denen der Schüler nicht einmal alle Erfordernisse für die Beurteilung mit „Genügend“ (Abs. 5) erfüllt.
(Leistungsbeurteilungsverordnung Paragraph 14)

Werden nun die Ausführungen der Leistungsfeststellung mit diesen Erläuterungen verglichen, lässt sich erkennen, dass erstere sehr detailliert beschrieben wird und letztere viele unbestimmbare Rechtsbegriffe und ungenaue Floskeln enthält. Dies wird damit erklärt, „dass die Definition, will sie alle Schularten und Unterrichtsgegenstände erfassen, so abstrakt gehalten sein muss, dass sie notwendig an Aussagekraft verliert und damit auch ein normativer Gehalt problematisch wird.“ (Leistungsbeurteilungsverordnung 2022: RV 345 BlgNR 13) Aufgrund des hohen Interpretationsspielraums wird die Lehrperson gezwungen, eigenständig Maßstäbe zu definieren (vgl. Neuweg 2000: 72f.). Besonders bei den Leistungsfeststellungen bezüglich „Genügend“ und „Nicht genügend“ ist oftmals nicht klar, was denn genau die wesentlichen Bereiche eigentlich sind. Durch die Lehrpläne ist hierbei zwar ein gewisser Rahmen gegeben, was die Lehrpersonen jedoch genau unterrichten oder worauf sie ihre Schwerpunkte legen, bleibt ihnen überlassen (vgl. Neuweg 2000: 74). Laut Kremser (2016: 131) wird die dabei getroffene Gewichtung, in welchem Ausmaß was als

wesentlich gilt, in der Praxis oftmals nicht den Schüler:innen mitgeteilt. Dadurch entstehen Vorwürfe fehlender Intransparenz und Nachvollziehbarkeit (vgl. Kremser 2016: 131).

Warum vor allem die Ziffernoten zum Einsatz von Leistungsbeurteilungen kommen, abgesehen von den rechtlichen Pflichten, liegt an der leichten Erteilung und Verrechnung dieser. Darüber hinaus erfüllen Noten weitere Funktionen, welche im Folgenden kurz vorgestellt werden.

Zunächst ist vor allem die *Selektionsfunktion* für weiterführende Schulen und gesellschaftliche und berufliche Positionen zu nennen, welche anhand der Noten erbracht wird (vgl. Becker-Mrotzek et al. 2006: 86; Sacher 2014: 22). Beispielsweise werden für eine Zulassung in eine AHS gute Noten benötigt. In einer Vorarlberger Studie von Böheim-Galehr und Engleitner (2014) wurde diesbezüglich beobachtet, dass ein „Sehr gut“ in den letzten Jahren öfter vergeben wurde. Dies zeigte sich deutlich in Orten, wo sich mehrere Allgemeinbildende Höhere Schulen befanden. Bei einem Vergleich von Volksschulbeurteilungen und den Orientierungsarbeiten der fünften Klassen sowohl in der AHS als auch an der NMS konnte aufgezeigt werden, dass es Überschneidungen der Leistungen und Kompetenzausprägungen bei Kindern mit und ohne AHS-Reife gab. Dazu merken Böheim-Galehr und Engleitner (2014) an, dass „bei der Beurteilung und bei der Zuerkennung der AHS-Reife in vielen Fällen die erforderliche Objektivität nicht gesichert ist und auch andere Faktoren als die bisherigen Leistungen eine Rolle spielen dürften.“ Neben dieser Kritik steht ebenso die viel zu früh beginnende und endgültige Selektionsentscheidung (vgl. Sacher 2014: 23). Außerdem ist Selektion auch immer mit Stigmatisierung verbunden (vgl. Sacher 2014: 24).

Eine weitere Funktion von Noten ist die der *Sozialisation*. So werden Schüler:innen bereits in jungen Jahren mit der Leistungsgesellschaft beziehungsweise mit der Leistungsorientierung vertraut gemacht. Durch die Leistungsbeurteilung anhand von Noten werden sie verglichen und vergleichen sich in weiterer Folge auch mit anderen (vgl. Sacher 2014: 24). Darüber hinaus werden die Kinder auch an das Prinzip der Bürokratie herangeführt, indem ihre individuelle Persönlichkeit und tatsächliche Kompetenz Nachrang gegenüber formalen Kriterien in amtlichen Dokumenten hat (vgl. Sacher 2014: 24). Sacher (2014: 27) nimmt an, dass die beharrliche Verteidigung des Ziffernotensystems im österreichischen Schulwesen vor allem auf der Sozialisationsfunktion basiert, da diese am meisten von einer Abschaffung beeinträchtigt werden würde.

Leistungsbeurteilungen durch Noten dienen aufgrund ihres Zahlencharakters auch der *Legitimation* von administrativen, politischen und unterrichtlichen Entscheidungen. Sie sind

leicht zu verstehen und zu verrechnen und wirken objektiv. Daher scheint es einleuchtend, Noten zu verwenden, um Entscheidungen zu rechtfertigen (vgl. Sacher 2014: 27).

Eine weitere Funktion von Noten ist die *Kontrolle* des Systems Schule. Lehrpersonen, Lehrpläne, Schulen, verschiedene Schularten und ganze Schulsysteme werden über die Notenvergabe kontrolliert (vgl. Sacher 2014: 27). Sacher (2014: 27) merkt hierbei an, dass ein widersprüchliches Verhältnis zwischen dem Notenniveau und der angenommenen pädagogischen Qualität besteht. So stehen gute Noten im Allgemeinen für erfolgreiches Lehren und Lernen, schlechte Noten jedoch für Überforderung aufgrund eines zu hohen Anforderungsniveaus. Insofern führen sehr gute als auch schlechte Noten zur Fraglichkeit der Lehrqualität, weswegen in der Praxis häufig die Tendenz zur Mitte beobachtet wird (vgl. Sacher 2014: 27).

Noten dienen darüber hinaus der Funktion der *Prognose*. Sowohl Schüler:innen, Lehrpersonen, Eltern als auch gesellschaftliche Organisationen und Institutionen interpretieren Noten prognostisch und leiten damit Erwartungen über zukünftige Leistungen ab beziehungsweise wollen verfügbares Potenzial erkennen (vgl. Sacher 2014: 27). Durch die Prognosefunktion von Noten werden Konsequenzen für die Bildungs- und Arbeitsmarktpolitik gezogen (vgl. Sacher 2014: 28). Eine genaue Vorhersage über die weiteren Entwicklungen der Leistungen zu treffen, kritisieren jedoch beispielsweise Baumert und Kolleg:innen (2010) und Perleth und Sen (2010) aufgrund der fehlenden Objektivität von Noten.

Zusätzlich kann Noten eine *Informations- und Rückmeldungsfunktion* zugeschrieben werden, indem sie über den erreichten Lernstand beziehungsweise die gemachten Fortschritte in der Leistung informieren (vgl. Sacher 2014: 28). Schüler:innen können somit erfahren, wo sie sich im Leistungsspektrum der Klasse befinden, ob sie den Anforderungen entsprechen oder ob der Arbeitseinsatz intensiviert werden sollte (vgl. Sacher 2014: 28). Sacher (2014: 28) fordert, dass Rückmeldungen bei Misserfolgen einfühlsamer mitgeteilt werden sollten, da ansonsten Demotivation und Resignation folgen können, was sich in weiterer Folge auf die Persönlichkeitsentwicklung auswirkt.

Eine weitere Funktion ist die der *Disziplinierung*. Durch Noten kann aufgezeigt werden, dass weitere Anstrengungen oder eine hohe Lernbereitschaft maßgeblich für den Erfolg in der Schule notwendig sind (vgl. Sacher 2014: 28). In der Praxis werden Noten aber auch als Druck- und Machtmittel eingesetzt, um den Schüler:innen zu demonstrieren, dass die Lehrperson die Autorität besitzt (vgl. Sacher 2014: 29). Die vierte Kinderstudie aus dem Jahr

2018 konnte diesbezüglich aufzeigen, dass Noten einen der verbreitetsten Auslöser von Angst darstellen (vgl. Kinderstudie World Vision 2018).

Des Weiteren erfüllen Noten eine Funktion als *Lehr- und Lerndiagnose*. Dadurch diagnostizieren Lehrpersonen den Lernstand der Schüler:innen, um Überlegungen für die weitere Unterrichtsgestaltung und Lernberatung machen zu können (vgl. Sacher 2014: 29). Diese Funktion stellt auch eine Qualitätsüberprüfung der Lehrperson selbst dar.

Die letzte hier zu nennende Funktion von Noten ist die *Lern- und Leistungserziehung*. Die Beurteilungen tragen zur Erziehung der Schüler:innen bei, indem sie lernen, ihr Lernverhalten selbst zu kontrollieren und Verantwortung für eigenständiges Arbeiten zu entwickeln (vgl. Sacher 2014: 30).

Laut Becker-Mrotzek et al. (2006) ist es insofern problematisch, Noten als Begründung und Funktion für die genannten Bereiche zu sehen, als dass sie einerseits nur bedingt valide sind, da sie nicht nur die Textqualität per se erfassen, sondern beispielsweise auch das Erscheinungsbild des Textes oder lediglich die Sprachrichtigkeit begutachtet wird (vgl. Becker-Mrotzek et al. 2006: 87).

Andererseits sind Noten auch nicht reliabel, denn sie führen bei Wiederholungen der Testungen nicht immer zum gleichen Ergebnis. Ebenso ist laut aktueller Studienlage auch wenig bis keine Objektivität gegeben, da die Leistungsbeurteilungen oftmals damit zusammenhängen, wer benotet (siehe Kapitel 2.4).

Gründe dafür können sogenannte Bezugsnormen sein, welche bei Lehrpersonen dazu beitragen, dass Unklarheiten bezüglich der Leistungsbeurteilung aufkommen. Daher sollen diese im nächsten Kapitel näher erläutert werden.

2.3.6 Bezugsnormen

Leistungsergebnisse werden bewusst oder unbewusst immer in Relation zu anderen Leistungen gesehen (vgl. Sacher 2014: 84). Es wird sich also an einem bestimmten Bezugssystem orientiert, in welchem sogenannte Bezugsnormen herrschen. Diese bezeichnen die Art des Maßstabes, nach welchen eine Leistung bewertet wird (vgl. Sacher 2014: 84). Beispielsweise kann dies über das Lehrziel, den Durchschnitt der Klasse oder über die Schüler:innen selbst erfolgen (vgl. Neuweg 2014: 76ff.). Im Folgenden sollen diese drei Bezugsnormen genauer erläutert werden.

Sachliche Bezugsnorm

Die sachliche Bezugsnorm wird in der Literatur auch objektive, curriculare oder lehrziel- und kriteriumsorientierte Bezugsnorm genannt und stellt die Idealnorm dar. Wie der Name bereits erahnen lässt, zielt diese auf ein vorher definiertes Lernziel oder Anforderungen ab, an welchen die Leistungen gemessen werden, ohne dabei die individuellen Anstrengungen zu berücksichtigen (vgl. Olechowski 2003: 223). Die Lehrperson muss daher einerseits ein Lernziel und andererseits das Ausmaß der Leistung definieren, ab wann das Ziel erreicht ist, also beispielsweise die Anzahl an Punkten (vgl. Benischek 2006: 117). Trotz einiger Vorgaben des Lehrplans wirkt es so, als ob es hierbei große Variationen zwischen Lehrpersonen und auch zwischen den Schulen gibt. Dennoch scheint die Anwendung der sachlichen Norm tendenziell positiv auszufallen, vor allem wenn differenzierte Aufgabenstellungen zur Verfügung gestellt werden, um so allen Schüler:innen eine faire Chance geben zu können (vgl. Benischek 2006: 118).

Individuelle Bezugsnorm

Bei der individuellen Bezugsnorm werden frühere Leistungen einer einzelnen Person herangezogen und mit der aktuellen verglichen. Dies bedeutet, dass der Schüler oder die Schülerin nicht wie bei der sachlichen Bezugsnorm an einem erreichten Lehr- oder Lernziel gemessen wird, sondern lediglich an seiner oder ihrer Person selbst. Dadurch werden Lernprozesse, sowohl Verschlechterungen als auch Verbesserungen realisiert (vgl. Neuweg 2014: 91, Benischek 2006: 116). Ein Lernzuwachs verzeichnet demnach eine gute Leistung, während Stagnation oder sogar ein Rückschritt eine schlechte Leistung darstellt.

Vorteil dieser Bezugsnorm ist vor allem die Lernmotivation seitens der Schüler:innen, da das Ausgangsniveau in den Hintergrund rückt und auch benachteiligte Kinder den positiven Zusammenhang zwischen Bemühungen und Fortschritt wahrnehmen können (vgl. Rheinberg 2001: 6).

Der Nachteil dieser Bezugsnorm ist der erhöhte Arbeitsaufwand für die Lehrpersonen (vgl. Thonhauser 2002: 53-54).

Soziale Bezugsnorm

Die soziale Bezugsnorm bezeichnet den Leistungsvergleich eines Schülers oder einer Schülerin mit einer bestimmten Bezugsgruppe, in dem Fall mit der Schule beziehungsweise der Schulklasse. Dazu werden die Durchschnittswerte der Gruppe errechnet und die

Streuungen davon berücksichtigt (vgl. Tortik 2016: 35). Die individuelle Leistung eines jeden Schülers oder einer jeden Schülerin wird dann an der Gruppenleistung gemessen und stellt demnach eine gute Leistung dar, wenn der Mittelwert übertroffen wird und eine schlechte Leistung, wenn die Mitte nicht erreicht wird. So entsteht gewissermaßen eine Rangfolge aller Schüler:innenleistungen (vgl. Benischek 2006: 105).

Benischek (2006: 105) sieht dies als fragwürdig an, da eine Normalverteilung der Leistungen nur für größere Stichproben angewendet werden kann. Darüber hinaus erscheint es pädagogisch nicht sehr wertvoll, Leistungen anhand des zufälligen Durchschnittswerts von Schüler:innenleistungen zu beurteilen (vgl. Persy 1999: 205). Die Problematik dabei ergibt sich vor allem an den Enden der Beurteilungsskala, wo beispielsweise die schlechtesten Schüler:innen einer Klasse mit einem „Nicht Genügend“ beurteilt werden, obwohl die Leistungen gar nicht unzureichend waren und die besten Schüler:innen mit einem „Sehr gut“, trotzdem die Arbeit unter anderen Bedingungen keinen 1er rechtfertigen würde (vgl. Olechowski 2003: 221f.). Dadurch werden individuelle Lernfortschritte ausgeblendet und Leistungsunterschiede übersehen und somit ist es für leistungsschwächere Schüler:innen im ständigen Vergleich zu den Leistungsstarken schwierig, dem unteren Ende der Verteilung zu entfliehen (vgl. Benischek 2006: 118).

Die Anwendung dieser Bezugsnorm macht lediglich in Settings Sinn, wo die dauerhaft Besten herauskristallisiert werden sollen, also beispielsweise bei (mehrstufigen) Bewerbungsverfahren (vgl. Rheinberg 2001: 5).

Rechtslage der Bezugsnormen

Die Leistungsbeurteilungsverordnung regelt im § 11 Abs 1 im zweiten Satz ausdrücklich, dass die Leistungen eines Schülers beziehungsweise einer Schülerin mit der sachlichen Bezugsnorm zu messen sind, da sich die Vorgaben und Leistungsfeststellungen an dem Lehrplan orientieren müssen. Es muss dennoch darauf geachtet werden, dass der Stand des Unterrichtes mit den Anforderungen des Lehrplanes übereinstimmt.

Angenommen die Lehrpersonen würden die soziale Bezugsnorm wählen, würde dies nicht mehr dem Anspruch genügen, sachlich, objektiv und gerecht beurteilen zu können, da vom Klassendurchschnitt ausgegangen wird. Dies merkt auch der österreichische Verwaltungsgerichtshof an:

Der Maßstab der Leistungsbeurteilung ist daher kraft Gesetzes ein von der Beurteilung anderer Schüler oder von der durchschnittlichen Beurteilung von Schülern gleicher Schulart und Schulstufe, sei es im ganzen Bundesgebiet oder in bestimmten Teilen desselben, unabhängiger. (§§ 18 Abs 1 SchuG)

2.4 Studienlage zu Korrekturverhalten und Beurteilungen von Deutschlehrer:innen

Grundsätzlich gibt es in der aktuellen Literatur wenig zum Korrekturverhalten von Lehrpersonen zu finden. Den größten Umfang liefern dabei Forschungen über den Bereich der Grammatik. Ein Beispiel einer solchen Studie ist die von Hennig (2012), welche es sich zum Ziel gemacht hat, Datenmaterial zum Korrekturverhalten von Deutschlehrer:innen zusammenzutragen. Dabei wurde beobachtet, was genau von 40 Deutschlehrenden, aber auch Germanistikstudent:innen als Fehler in einem fiktiven Schüler:innenaufsatz angestrichen wird. Dieser Aufsatz war mit 25 potenziellen Fehlern versehen, darunter sowohl System- als auch Normfehler, aber auch Flüchtigkeitsfehler und Zweifelsfälle. Die Proband:innen mussten daraufhin den Text korrigieren, was zu insgesamt 57 unterschiedlichen Fehlermarkierungen führte. Dies macht 32 Korrekturen mehr aus, als potenziell vorhanden waren. Ebenso die Kategorisierungen der Fehler wurden verschieden gehandhabt (vgl. Hennig 2012).

Auch Davies (2017) beschäftigte sich mit Korrekturen im Bereich der Grammatik, indem er eine Umfrage mit 65 Lehrpersonen durchführte, um deren Toleranz hinsichtlich bestimmter Phänomene in der Grammatik zu beleuchten. Die Ergebnisse zeigen, dass Phänomene im Vergleich zur kodifizierten Norm viel strenger beurteilt werden (vgl. Davies 2017). Bereits an den genannten Studien lässt sich die Problematik an Korrekturen, vor allem im Fach Deutsch, durch die Komplexität von Texten erkennen.

Bezüglich Beurteilungen konnten bereits die Studie von Schröter (1981) und die Replikationsstudie von Birkel (2003) zeigen, dass das Beurteilungsspektrum bei Schüler:innentexten, vor allem bei den Messungen im mittleren Leistungsbereich, äußerst groß ist. Dabei wurde auch bemerkt, dass ein Text mit vielen Rechtschreibfehlern schlechter abschneidet, da die Note hier vermehrt die orthografische Korrektheit anstatt der eigentlich zu bewertenden Textqualität als Gesamtes ausdrückt (vgl. Becker-Mrotzek 2006: 87).

Einen weiteren interessanten Aspekt zum Korrekturverhalten von Deutschlehrenden liefert die Studie von Ransmayr (2020), in welcher untersucht wurde, ob das Schreibmedium Auswirkungen auf die Benotung hat. Dabei wurde, wie auch bereits aus anderen Forschungen (Graham & Morphy 2011; Russell & Tao 2004) hervorgehend festgestellt, dass handgeschriebene Texte im Vergleich zu computergeschriebenen besser beurteilt werden.

Dies verwundert zunächst, da am Computer bereits viele Fehler automatisch ausgebessert werden und alles auch viel übersichtlicher aussieht. Jedoch könnte dies auf die Erwartungshaltungen der Lehrenden zurückzuführen sein oder damit zusammenhängen, dass eine digitale Arbeit übersichtlicher wirkt und somit die Schwächen des Textes offensichtlicher werden (vgl. Ransmayr 2020: 70). Die Beurteilungsraster der untersuchten Maturaarbeiten hätten jedenfalls bessere Noten erwarten lassen, als dies schlussendlich der Fall war (vgl. Ransmayr 2020: 70).

Bezüglich formaler Kriterien lieferten einige Untersuchungen Hinweise, dass beispielsweise eine unschöne Handschrift als auch schlechte Grammatik- und Rechtschreibleistungen zu ungerechtfertigt negativen Bewertungen führen (vgl. Baurmann 2006: 144; siehe auch Ingenkamp 1977).

Wenn Untersuchungsergebnisse hinsichtlich der Notenvergabe aus dem 20. Jahrhundert mit Aktuelleren verglichen werden, lassen sich laut Brügelmann (2014: 72) keine großen Unterschiede vernehmen. Die Datenlage hat sich diesbezüglich nicht stark verändert, weswegen in diesem Kapitel auch ältere Studien erwähnt werden.

Beispielsweise konnten Studien zeigen, dass in den Unterrichtsfächern, welche Leistungsüberprüfungen anhand schriftlicher Arbeiten durchführen und wo Schulleistungen quantifizierbar gemacht werden können, strenger beurteilt wird (vgl. Ziegenspeck 1979: 40). Diese unterschiedliche Beurteilung und Härte in der Notengebung führen dazu, dass Unterrichtsfächer in unterschiedlichem Maß dazu beitragen können, wie aussortiert wird. Dabei konnte auch ein Zusammenhang mit dem jeweiligen Anteil an der Gesamtwochenstundenanzahl festgestellt werden (vgl. Sacher 2014: 44).

Der Ort der Schule spielt bei der Notenvergabe ebenso eine Rolle, da verschiedene Regionen und Bundesländer die Strenge der Benotungen beeinflussen (vgl. Sacher 2014: 44). Laut Sacher (2014: 44) spielen auch Schulstufen und Schularten eine Rolle. Beim Übertritt von der Volksschule ins Gymnasium lässt sich beispielsweise ein sogenannter „Notenknicke“ beobachten, welcher trotz stattfindender Selektion in den höheren Schulstufen nicht besser wird (vgl. Sacher 2014: 44).

Eine weitere Problematik ergibt sich aus dem klasseninternen Bezugsmaßstab, welchen Lehrende bei Beurteilungen heranziehen. So verteilen Lehrpersonen

die Zensuren in ihrer Klasse nach dem von ihnen für richtig gehaltenen Schlüssel, ohne hinreichende Informationen über den Leistungsstand ihrer Klasse im Vergleich zu dem anderer Klassen zu besitzen. Das führt dazu, dass die nach objektivem Vergleichstest gleiche Leistung in der einen Klasse mit „1“, in der anderen Klasse mit „6“ beurteilt werden kann.“ (Ingenkamp 1985: 179)

Dies zieht die Tatsache nach sich, dass die gleichen Noten unterschiedliche Kompetenzgrade in den verschiedenen Klassen darstellen und repräsentieren (vgl. Sacher 2014: 44).

Darüber hinaus werden auch geschlechterspezifische Mängel an der Notengebung deutlich. Einige Studien lieferten Hinweise, dass Mädchen bessere Noten für erbrachte Leistungen erhalten als Jungen (siehe etwa Ingenkamp 1995, Maaz et al. 2011) und dass Lehrerinnen milder beurteilen als Lehrer (vgl. Rittberger 2000: 100). Diese bereits weiter zurückliegenden Untersuchungen konnten anhand einer groß angelegten Metastudie von Voyer und Voyer (2014) bestätigt werden. Dabei wurden 369 Primärstudien aus den Jahren 1914 und 2012 zusammengefasst (vgl. Voyer & Voyer 2014).

Des Weiteren zeigten Untersuchungen bezüglich sozialer Herkunft, dass sozial benachteiligte Schüler:innen im Durchschnitt schlechter benotet werden als Kinder und Jugendliche aus eher gehobenen Sozialschichten (vgl. Maatz et al. 2011). Diese Benachteiligung schlägt sich vor allem in sprachlichen Fächern und am stärksten in Deutsch nieder (vgl. Wolf 1980, Ingenkamp 1995).

Auch sogenannte Erwartungseffekte schwächen die Begründung der Notenvergabe. So konnte bereits Zillig (1928) in einer Untersuchung aufzeigen, dass bei rechtschreibstarken Schüler:innen 39 Prozent der Fehler übersehen wurden, während dieser Wert bei den „schlechten“ Schüler:innen bei nur 12 Prozent lag.

Ebenso konnte Weiss bereits im Jahr 1965 aufzeigen, dass die Bewertung eines Schülers, welchem positive Eigenschaften zugeschrieben worden waren, deutlich besser ausfiel als der Aufsatz eines durchschnittlichen Schülers. Was die urteilenden Lehrpersonen jedoch nicht wussten, war, dass die Aufsätze vertauscht worden waren.

Einige weitere Studien, wie die von Hofer (1969), Jürgens (1977), Schwarzer und Lange (1979) konnten auch den Zusammenhang zwischen Lehrendenpersönlichkeit und Schüler:innenleistung beziehungsweise -charakter aufzeigen. Persönlichkeitsmerkmale wie das Alter, das Geschlecht, die Ausbildung und der Status können Beurteilungsunterschiede erklären (vgl. Sacher 2014: 42). Es wird angenommen, dass implizite Beurteilungs- und auch Unterrichtstheorien dafür verantwortlich sind und Lehrpersonen somit ein Bild von „guten“

und „schlechten“ Schüler:innen generieren, was in weiterer Folge auch in Unterrichtsprozesse und Interaktionen miteinfließt. Dies führt zu Urteilsverzerrungen bezüglich der Notengebung und somit auch zwangsläufig zu Schwierigkeiten für den weiteren Schulerfolg (vgl. Jürgens 2005: 70). Die damit verbundenen Erwartungshaltungen bezüglich Erfolges oder Misserfolges können sich sowohl auf Seiten der Lehrenden als auch der Lernenden und deren Eltern verfestigen (vgl. Jürgens 1977).

Die Beeinflussung der Notengebung durch Sympathie und Antipathie konnte ebenso in verschiedensten Studien aufgezeigt werden (vgl. Bergin & Bergin 2009; Birkel 2009; Rausch 2013).

Aufgrund der eben genannten und weiteren Untersuchungen kann angenommen werden, dass Noten weder den Ansprüchen der Objektivität, Reliabilität und Validität gerecht werden (vgl. Deimel 2015), noch eine hinreichende pädagogische und gesellschaftliche Leistungskontrolle gewährleisten können (vgl. Ingenkamp 1985: 179). Dahingehend betont auch Frigge (2000: 426) einen weiteren Kritikpunkt der Notengebung, indem er postuliert, dass Schüler:innen zwar die Note entgegennehmen, aber nichts dadurch lernen. Darüber hinaus werden die Beurteilenden tendenziell dazu angehalten, Fehler und Mängel in den Texten zu finden und lediglich anhand dieser zu beurteilen. Vor allem bei der Matura oder schriftlichen Abschlussprüfungen wird nur der fertige Text beurteilt, der gesamte Lernprozess der Schüler:innen wird dabei außer Acht gelassen. An dieser Stelle muss außerdem angemerkt werden, dass die Bewertung anhand einer Ziffernzahl keineswegs die komplexe Gesamtheit eines geschriebenen Textes erfassen kann (vgl. Baumann 2006: 120). Laut Baumann könnte diese Vorgehensweise lediglich einen groben Vergleich zwischen den einzelnen Schüler:innenarbeiten darstellen. Somit ist der Prozess der Notengebung bestenfalls ein Schätzverfahren und ein subjektives Schätzurteil, wie bereits erwähnt wurde. Anstatt der vorgegebenen Notenskala wird die urteilende Lehrperson somit zum eigentlichen Messinstrument (vgl. Baumann 2006: 120). Noten als Unterstützung und Hilfe für individuelle Lernförderungen und -prozesse anzusehen ist deshalb problematisch. Um curriculare und didaktische Entscheidungen treffen zu können, ist es laut Jürgens (2005: 71) unabdingbar, differenzierte Beurteilungsformen zu etablieren.

Im Hinblick auf die Thematik vorliegender Masterarbeit soll auch folgende Untersuchung aufgezeigt werden: Vierlinger (2012) führte im Rahmen eines Vortrages ein ad-hoc-Experiment durch, um aufzeigen zu können, dass Ziffernnoten nicht objektiv sind. Dazu

projizierte er einen Aufsatz eines elfjährigen Bauernkindes auf die Leinwand und ließ diesem von über 200 Expert:innen des österreichischen Schulwesens im Raum eine Note hinsichtlich Inhalt und Sprachgestaltung geben, welche per Handzeichen eruiert wurden. Während sich ungefähr 90 Personen für ein „Befriedigend“ aussprachen, wurden dennoch von den restlichen Besucher:innen auch alle anderen Noten für denselben Text vergeben (vgl. Schmidinger et al. 2012: 120). Wie Vierlinger aufzeigen konnte, kann ein:e Schüler:in durch einen bloßen Lehrer:innenwechsel bei gleichbleibender Leistung somit von einem „Sehr gut“ auf ein „Genügend“ springen. Ein Grund dafür dürfte im ungleichen Niveau erklärt sein, welches aufgrund der Klassenzusammensetzung als auch wegen der, durch die unterrichtenden Lehrenden erreichten, unterschiedlichen Kompetenzen und Engagements vorhanden ist (vgl. Schmidinger et al. 2012: 123). Wenn aber bedacht wird, dass für eine objektiv gesehen gleiche Leistung verschiedene Lehrende unterschiedliche Noten vergeben, dann kann im Umkehrschluss auch gesagt werden, dass differierende Leistungen hinter gleichen Noten stehen.

Angesichts der Tatsache, dass Noten in der heutigen Gesellschaft und dem vorherrschenden Bildungssystem weitreichende Auswirkungen haben können, als Richtwerte für Zukunftschancen und im schlechtesten Fall als Messung für die Intelligenz oder Eignung angesehen werden, sind sowohl formale Mindeststandards notwendig als auch eine möglichst objektive Korrektur durch die Lehrpersonen. Viele Aspekte eines Textes sind aber nicht operationalisierbar und selbst diejenigen, die es eventuell anhand einer Zuhilfenahme des Kodex wären, wie beispielsweise die Grammatik und die Orthografie, können erfahrungsgemäß dennoch nicht zu einer objektiven Beurteilung beitragen. Dies geht auch aus der Literatur hervor. Diese Tatsache allein stellt bereits einen Grund dafür dar, vertiefend auf dem Gebiet zu forschen. Deshalb und aufgrund der zu dieser Thematik relativ veralteten erforschten Erkenntnisse hat es sich vorliegende Masterarbeit zum Ziel gemacht, eine Schularbeit aus der dritten Klasse der Sekundarstufe 1, von Lehrenden sowohl aus dem AHS und NMS-Bereich korrigieren und benoten zu lassen und diese Ergebnisse dann zu untersuchen. Dabei soll erforscht werden, inwiefern das Korrekturverhalten und die Beurteilung zwischen den einzelnen Lehrpersonen als auch zwischen den beiden Schularten divergiert. Im nächsten Kapitel soll deshalb die methodische Vorgehensweise für vorliegende Forschung vorgestellt werden.

3. Methodik

In vorliegender Masterarbeit geht es um Unterschiede im Korrekturverhalten und in der Leistungsbeurteilung von AHS- und NMS-Lehrenden für Schularbeiten im Fach Deutsch. Basierend auf den theoretischen Ausführungen sollte die Untersuchung sowohl quantitativ als auch qualitativ erfolgen, um die Forschungsfrage umfassend beantworten zu können. Neben der Beschreibung der methodischen Vorgehensweise wird auch die Datengrundlage und die Stichprobe näher erläutert. Außerdem werden aufgetretene Schwierigkeiten und Herausforderungen thematisiert.

3.1 Methodentriangulation

Vorliegende Untersuchung wurde anhand einer Between-Methods-Triangulation durchgeführt. Im Gegensatz zur Within-Methods-Triangulation, wo die Kombination innerhalb einer Methodik erfolgt, wurden hierbei unterschiedliche Methoden für die Beantwortung desselben Gegenstands verwendet (vgl. Hussy et al. 2013: 288). Die Erhebung der quantitativen und der qualitativen Daten erfolgte dabei einerseits gleichzeitig und andererseits fließen diese auch gleichwertig in die Auswertung mit ein. Die dadurch gewonnenen Ergebnisse und Erkenntnisse wurden daraufhin einem Vergleich unterzogen, um ein umfassenderes und vertiefendes Bild über das Forschungsinteresse zu bekommen (vgl. Creswell & Plano Clark 2007).

Zugleich handelte es sich auch um eine Mixed-Methods-Studie, da sowohl eine quantitative als auch qualitative Methode verwendet wurde:

Mixed methods research [...] combines elements of qualitative and quantitative research approaches (e.g., use of qualitative and quantitative viewpoints, data collection, analysis, inference techniques) for the broad purposes of breadth and depth of understanding and corroboration. (Johnson, Onwuegbuzie & Turner 2007: 123)

3.1.1 Quantitative Methode

Die quantitative Methode vorliegender Forschung beinhaltete das Zählen der Korrekturzeichen aller miteinander zu vergleichenden Korrekturen des Schularbeitstextes. Darüber hinaus wurden die Benotungen der einzelnen Lehrpersonen begutachtet. Von diesen Ergebnissen ausgehend wurde der Mittelwert gebildet, um den Durchschnitt und die sich daraus ergebenden Abweichungen sichtbar zu machen.

Die Vorteile von quantitativen Erhebungsmethoden liegen vor allem in der hohen Präzision, die durch die Auswertung durch Zahlen gegeben ist. Dadurch kann auch eine gute Vergleichbarkeit gewährleistet werden. Des Weiteren können die Ergebnisse übersichtlich dargestellt und einfach verknüpft werden (vgl. Hussy et al. 2013: 53).

Der größte Nachteil einer quantitativen Forschung ist vermutlich der Informationsverlust, welcher durch die Darstellung in Zahlenform vonstattengeht (vgl. Hussy et al. 2013: 52). Diesem Umstand soll in vorliegender Masterarbeit die qualitative Analyse Rechnung tragen.

3.1.2 Qualitative Methode

Um eine umfassende Conclusio zu erhalten, erfolgte zusätzlich zur quantitativen eine qualitative Analyse, wo die korrigierten Schularbeitstexte und die qualitativen Aspekte des Fragebogens untersucht wurden. Dazu wurden die Schularbeiten hinsichtlich des Korrekturverhaltens und der dazugehörigen Beurteilung der Lehrpersonen analysiert, indem Korrekturzeichen und markierte Fehler miteinander verglichen wurden.

Die Vorteile von qualitativen Methoden liegen in der Vertiefung der Ergebnisse und in der Möglichkeit, Individualität abzubilden. Außerdem ist eine flexible Anpassung an den Forschungsgegenstand gegeben, was die qualitativen Methoden sehr vielfältig macht (vgl. Hussy et al. 2013: 53).

Im Folgenden sollen jedoch ebenso die negativen Aspekte erläutert werden.

Zum einen ist diese Methode schwer objektivierbar, da es sich um subjektive Erfahrungen und Meinungen der Teilnehmenden handelt, daher sollte dies in der Interpretation der Ergebnisse auch berücksichtigt werden. Damit einhergehend lassen sich die Ergebnisse auch schwer verallgemeinern, da diese in der Regel aus wenigen „Fällen“ stammen und somit nicht auf eine Grundgesamtheit geschlossen werden kann. Außerdem ist auch der höhere Arbeitsaufwand von qualitativen Erhebungs- und Auswertungsmethoden zu nennen (vgl. Hussy et al. 2013: 53; 192).

3.2 Datengrundlage

Der für die Untersuchung verwendete Schularbeitstext stammt von einer Schülerin oder einem Schüler aus der dritten Klasse einer Allgemeinbildenden Höheren Schule in Wien. Die Einwilligung des Direktors, der Eltern und des Kindes wurde mündlich eingeholt und danach wurde die handschriftlich verfasste Schularbeit per Fotodatei an die Autorin vorliegender Masterarbeit weitergeleitet. Dadurch, dass alles anonymisiert durchgeführt wurde, sind ausführlichere Informationen dazu nicht bekannt. In qualitativer Hinsicht wird von einem durchschnittlichen Text ausgegangen. Bei der Textsorte handelt es sich um eine Kriminal-Kurzgeschichte. Zuerst wurde überlegt den Schularbeitstext zu digitalisieren, um so dem Problem zu entgehen, dass eventuelle Ungereimtheiten bezüglich einer Korrektur und Benotung auftauchen, die lediglich dem Umstand der Handschrift geschuldet sind. Um das Material so authentisch wie möglich zu belassen, entschied sich die Autorin vorliegender Masterarbeit jedoch dagegen. Auf der Angabe der Schularbeit sind vier Auswahlmöglichkeiten vermerkt. Der oder die Verfasser:in des Schularbeitstextes entschied sich für die erste Möglichkeit. Der geschriebene Text umfasst 423 Wörter und kann mitsamt der Schularbeitsangabe im Kapitel 3.4 eingesehen werden. Die Schularbeit wurde in den Fragebogen eingebettet, welcher somit die Datengrundlage für die Forschung vorliegender Masterarbeit darstellte und für die Datenerhebung herangezogen wurde.

3.3 Datenerhebung

3.3.1 Pretest

Vor der eigentlichen Datenerhebung mittels eines Fragebogens wurde ein sogenannter Pretest durchgeführt, um diesen an Gewährspersonen erproben zu können, bevor er dann für die finale Erhebung verwendet wurde. Dazu wurde der vorläufige Fragebogen an zwei Lehrpersonen sowohl aus der AHS als auch der NMS geschickt. Nach der Durchführung wurde überprüft, ob die Fragen einerseits verständlich und sinnvoll waren, also ob ein roter Faden erkennbar war und andererseits ob die Rahmenbedingungen wie der zeitliche Aufwand gut bemessen war. Durch die Ergebnisse des Pretests wurde klar, dass der ausgewählte Schularbeitstext in qualitativer Hinsicht sehr schlecht abschnitt und dies für vorliegende Forschung zu einseitigen Ergebnissen führen hätte können. Um eine eventuell breitere Streuung zu bekommen, wurde daraufhin ein anderer Schularbeitstext besorgt, welcher als durchschnittlich eingestuft worden war und im Kapitel 3.4 eingesehen werden kann.

3.3.2 Durchführung der Erhebung

Der verwendete Schularbeitstext wurde Lehrenden zugeschickt, welche einerseits durch Kooperationen mit ausgewählten Schulen rekrutiert und andererseits auch durch bekannte Lehrpersonen animiert wurden. Die Aufgabe der Lehrer:innen war es, die zugeschickte Schularbeit mit ihrem eigenen Beurteilungssystem zu korrigieren und zu benoten. Zusätzlich wurde ein Online-Fragebogen mithilfe der Software mysurveio.com erstellt, welchen sie ebenfalls ausfüllten. Dieser kann im Anhang unter dem Kapitel 10.2 eingesehen werden. Die Mehrheit der Gewährspersonen bearbeitete den Fragebogen online, welchen sie anhand eines Links zugeschickt bekommen hatten. Für eine neue Mittelschule wurde dieser jedoch ausgedruckt und mitsamt dem Schularbeitstext ausgelegt. Der schriftliche Fragebogen enthielt geschlossene Fragen, beispielsweise über soziodemografische Aspekte, aber auch offene Fragen über die Lehrer:innentätigkeit und das Korrekturverhalten.

Der Fragebogen wurde für vorliegende Forschung ausgewählt, da er einige Vorteile bietet, wie beispielsweise eine sehr ökonomische und anonyme Datenerhebung. Obwohl ein Fragebogen eine kostengünstige Methode ist, kann gleichzeitig eine hohe Stichprobe untersucht werden (vgl. Hug & Poscheschnik 2015: 123). In vorliegendem Fall kann dieser Pool der Teilnehmenden eine bestimmte Grundgesamtheit der Lehrer:innen in Österreich repräsentieren. Dadurch sind die Ergebnisse dieser Befragung belastbarer als Erkenntnisse, die aus einer Forschung stammen, die lediglich qualitativ arbeitet. Außerdem können bestimmte Zufallsfehler, welche bei kleineren Stichproben auftauchen könnten, durch die im Regelfall große Anzahl an Teilnehmenden ausgeschlossen werden (vgl. Hug & Poscheschnik 2015: 123). In dem Fall vorliegender Masterarbeit muss jedoch hinzugefügt werden, dass die erhoffte Größe der Stichprobe nicht erreicht werden konnte.

Im Folgenden sollen auch die Nachteile einer Fragebogenerhebung erläutert werden. Die Beantwortung der Teilnehmer:innen ist hierbei schwer zu kontrollieren, da die Forscher:innen im Regelfall dabei nicht anwesend sind. Dadurch kann es auch passieren, dass der Fragebogen nicht eigenständig oder wahrheitsgetreu ausgefüllt wird. Außerdem können Unklarheiten bezüglich des Inhalts gegeben sein und die Teilnehmenden können daraufhin keine Nachfragen stellen (vgl. Hug & Poscheschnik 2015: 123f.), jedoch wurde diesem Aspekt mit dem Pretest entgegengewirkt.

3.4 Schularbeitsangabe und -text

Nachfolgend lässt sich die Angabe der Schularbeit in digitalisierter Form einsehen:

Thema: Kriminal-Kurzgeschichten

Wähle ein Thema aus und verfasse eine spannende Krimi-Kurzgeschichte. Achte dabei auf die im Unterricht erlernten Kriterien! (W-FRAGEN BEARBEITEN!)

Vergiss nicht am Ende die Wörter zu zählen. (absolutes Minimum: 300 Wörter) und deiner Geschichte einen kreativen Titel zu geben!

1. Erzählanfang: Setze fort!

Die letzten weichen Strahlen der Herbstsonne versuchten vergeblich das blasse Gesicht des Mädchens, das unter dem Fenster lag, zu erwärmen,....

2. Informationen ausbauen:

Todesopfer: klassenbesteR Schüler:in

Todesursache: Sturz vom Dach des Schulgebäudes

Verdächtige: Klassenkamerad:innen, Stiefvater, ...

Ermittler:in: Kommissar:in Müller

Motiv, Täter:in etc. FREI WÄHLBAR!

3. „Der Tote im Dorfteich“: Schreibe die Geschichte so um, dass es doch eine:n Mörder:in gibt!

4. Zeitungsbericht: Baue aus!

Im Folgenden findet sich der Schularbeitstext in digitalisierter Form, welcher von dem Schüler oder der Schülerin zu dem ersten Thema in der Angabe „Erzählanfang: Setze fort!“ verfasst wurde. Einen Screenshot des Originaltextes lässt sich im Kapitel 4.2.1 und nochmals im Anhang einsehen.

blutige Herbstferien

Die letzten weichen Strahlen der Herbstsonne versuchten vergeblich das blasse Gesicht des Mädchens, das unter dem Fenster lag, zu erwärmen, doch sie blieb kühl. Der Täter wusste, das er geschnappt werden würde, wenn er nicht die Spuren beseitigen und weglaufen würde. Ein Handy klingelte. Der Täter ging sofort ran. Es stellte sich heraus, dass dies sein Komplize sei, der wissen wollte, ob die Tat vollbracht war, doch er schwieg, legte auf und ging davon.

Es war ein typischer Samstag. Kommissar Brot verspeißte seinen Apfel und dachte sich nichts dabei, als auf einmal sein Handy ertönte. „Kommissar Brot zu ihren Diensten!“ wie kann ich heute behilflich sein?“, sagte er mit einem höflichem Umgangston. Die nervöse Frau am Handy schrie, dass er doch schnell zur Algostraße 12 kommen sollte da es ein Notfall sei. Kommissar Brot versuchte die Frau zu beruhigen, doch es half nicht. Daraufhin hoppte er in seinen Wagen und machte sich auf den schnellsten Weg zur Algostraße 12. Er sah 2 junge Damen. Eine kullerten die Tränen herunter, wie ein Wasserfall. Die andere war jedoch voll mit Blut und Glassplittern als näher kam, sah er, dass das junge Mädchen 2 Schüsse in den Kopf und einen in die Brust bekommen hatte. Der Mann trat näher und befragte die Dame. Es kam heraus, dass sie Anna hieß und das anscheinend tote Mädchen Marie. Marie war 17 Jahre alt und ging noch zur Schule. Eine Sirene ertönt. „Das wird wohl die Polizei sein.“, sagte Anna mit Tränen in ihren Augen. Kommissar Brot erzählte den Polizisten über das Verbrechen. Dabei fiel ihm ein mitte 60 Jähriger Mann, der das Mädchen leblos anstarrt. Es kam ihm so vor, dass der Mann seinen Blick nicht von ihr abwenden könne. Dies gab Kommissar Brot ein merkwürdiges Gefühl. Der Mann seufzte und sagte mit einer leisen und ruhigen Stimme: „2 unnötige Schüsse.“ Da bekam Kommissar Brot ein Geistesblitz. Herr Brot starrte ihn mit großen Augen an und sagte: „Sie waren es.“ Der Mann starrte zurück. Alle Augen waren auf die beiden gerichtet. Man sah dem Mann die Furcht an, als hätte man es ihm auf die Stirn geschrieben. Er musste sich die Tränen unterdrücken und gab anschließend ein Geständnis ab. Es kam heraus, dass er ihr Freund sei und das nur getan hat, da sie alles über seine illegalen Geschäfte herausgefunden hatte und das sein Komplize noch auf freiem Fuß sei. Doch er wollte noch wissen, wie er darauf gekommen ist, dass er der Mörder sei? Kommissar Brot meinte: „Woher wussten sie, dass 2 Schüsse unnötig waren?“

3.5 Datenauswertung

Für die Datenauswertung wurde eine Methodentriangulation angewandt, da hierbei verschiedene Zugänge zur Beantwortung der Forschungsfrage forciert werden. Diese wurde bereits zu Beginn des Methodik-Kapitels genauer beschrieben. Für die Zählungen und Berechnungen der quantitativen Untersuchungen wurde dabei das Programm Microsoft Excel verwendet. Mithilfe dessen wurden ebenso die Diagramme und Tabellen, welche im nächsten Kapitel der Ergebnisse dargestellt werden, angefertigt. Für die qualitative Untersuchung wurden die korrigierten Schularbeitstexte der teilnehmenden Lehrer:innen herangezogen und

das Korrekturverhalten auf Unterschiede und Gemeinsamkeiten analysiert. Darüber hinaus wurden die Antworten der offenen Fragen des Fragebogens beleuchtet und zusammenfassend dargestellt.

Sowohl die demografischen Daten des Fragebogens als auch schulbezogene Aspekte der Proband:innen sollen im nächsten Kapitel aufgezeigt werden, wo die Stichprobe vorliegender Untersuchungen beschrieben wird.

3.6 Stichprobe

Der Fragebogen wurde mitsamt des Schularbeitstextes an zwölf Schulen verschickt und die Daten im Zeitraum von 20. März bis 20. April 2022 erhoben. Dabei konnten insgesamt 21 Lehrpersonen erreicht werden. Die Zusammensetzung dieser Stichprobe bestand aus 17 weiblichen und vier männlichen Teilnehmer:innen. Wie sich aus folgendem Kreisdiagramm entnehmen lässt, beträgt die Verteilung somit 81% Frauen und 19% Männer. Dies entspricht beinahe dem Geschlechterverhältnis an den Schulen Österreichs, wo laut Statistik Austria im Schuljahr 2020/21 nahezu drei Viertel des Lehrpersonals Frauen sind (vgl. Statistik Austria 2020/21). Diesbezüglich kann gesagt werden, dass die Beteiligung von Frauen am Fragebogen über dem statistischen Durchschnitt liegt.

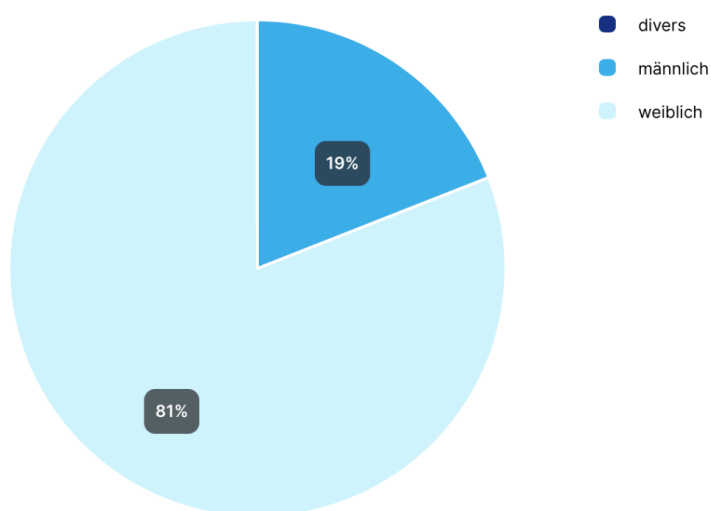


Abbildung 4: Geschlecht der Stichprobe

Davon unterrichten zwölf Personen in einer Neuen Mittelschule und neun in einer Allgemeinbildenden Höheren Schule, was 57,1% zu 42,9% ausmacht.

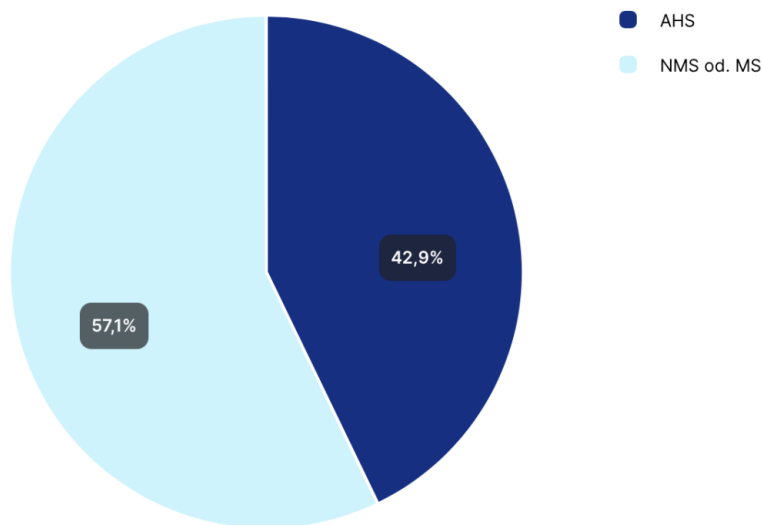


Abbildung 5: Schulart

Über 90 Prozent von ihnen mit insgesamt 19 Lehrpersonen unterrichten in der Stadt, während die restlichen beiden Lehrer:innen ihrem Job auf dem Land nachgehen.

Das Alter der Proband:innen lässt sich in folgendem Diagramm einsehen.

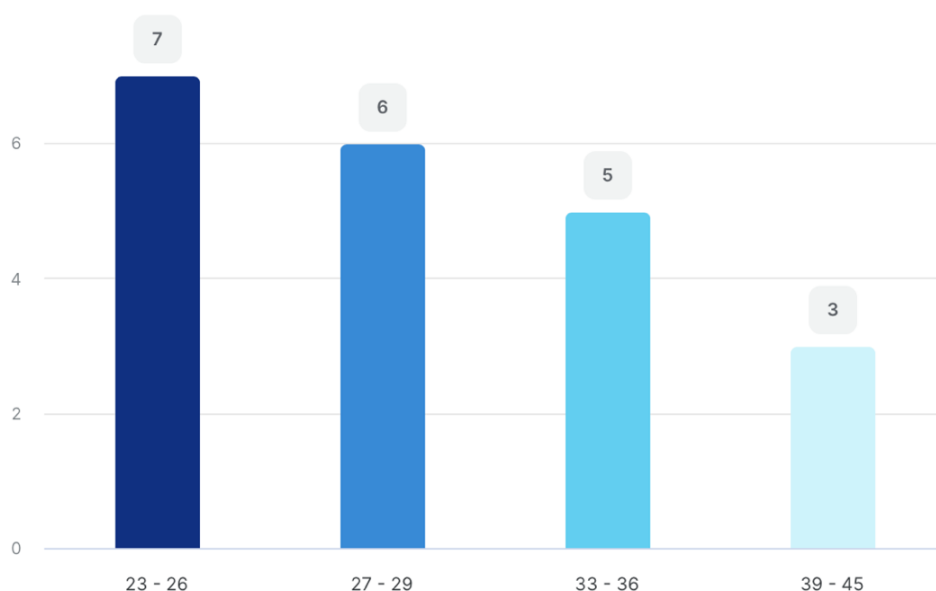


Abbildung 6: Alter der Proband:innen

Die meisten von ihnen sind junge Lehrer:innen. Insgesamt 13 Lehrpersonen sind unter 30 Jahre alt, während die restlichen acht Personen ein Alter von 33-45 Jahren aufweisen. Nach

der Erhebung des genauen Alters eines und einer jeden Proband:in kann gesagt werden, dass der Altersdurchschnitt der Stichprobe bei 30 Jahren liegt.

Bezüglich der bereits absolvierten Dienstjahre der Gewährspersonen konnte ausgemacht werden, dass die Mehrheit Junglehrer:innen sind. Dieser Begriff findet in vorliegender Arbeit eine eigene Definition. Damit ist eine Lehrer:innentätigkeit und -erfahrung von bis zu fünf Jahren gemeint. Ein Drittel der Stichprobe arbeitet bereits seit über fünf Jahren im Lehrer:innenberuf, wobei die Person mit der meisten Unterrichtserfahrung eine Lehrer:innentätigkeit von 22 Jahren aufweist.

Ebenso wurde die Anzahl an Wochenstunden abgefragt, in welchen die Lehrpersonen Deutsch unterrichten. Dabei geben 15 Personen maximal 12 Stunden Deutschunterricht, während der Rest bis zu 20 Wochenstunden aufweisen kann.

Um weitere Informationen bezüglich Erfahrungen der Lehrpersonen in der Stichprobe zu bekommen, wurde auch abgefragt, in welchen Schulstufen sie bereits unterrichtet haben oder aktuell eingesetzt werden. Die für vorliegende Masterarbeit relevanten Antworten betreffen die 3. Klasse Unterstufe beziehungsweise die 7. Schulstufe, da es sich bei dem verwendeten Schularbeitstext um eine Arbeit aus dieser Klassenstufe handelt. Dahingehend konnte eruiert werden, dass drei Personen aus der Stichprobe noch nie in einer dritten Unterstufenklasse unterrichtet haben. Die übrigen 18 haben bereits dort unterrichtet beziehungsweise unterrichten elf Lehrkräfte aktuell in einer 7. Schulstufe entweder in der AHS oder NMS.

4. Ergebnisse

4.1 Quantitative Aspekte: Fehlermarkierungen der Lehrpersonen

In diesem Kapitel werden die wichtigsten quantitativen Ergebnisse zu den Fehlermarkierungen der einzelnen Lehrpersonen in Diagrammen dargestellt. Für eine bessere Übersicht kann zunächst eine Tabelle eingesehen werden, in welcher die Anzahl an gezählten und ausgewiesenen Korrekturzeichen jeder einzelnen Lehrperson vermerkt sind. Textstellen, die lediglich im Text verbessert oder angestrichen wurden, werden hierbei nicht berücksichtigt, aber in der qualitativen Analyse beleuchtet. Anzumerken ist auch, dass Folge- und Wiederholungfehler aufgrund der Vergleichbarkeit in der quantitativen Analyse nicht miteinberechnet wurden, da nicht ersichtlich war, ob diese in die Wertung miteingeflossen sind oder nicht. Dennoch wird auch dieser Punkt bei den qualitativen Aspekten untersucht und beschrieben.

Für die vorliegende Untersuchung wurde den Lehrenden beziehungsweise den zugehörigen Schularbeitstexten die Nummern 1 bis 21 zugewiesen, wobei die besser Beurteilten vorne lagen und nach schlechter werdenden Benotungen weitergereiht wurden. Die Bereiche in der Tabelle wurden unterschieden in Sprachrichtigkeit mit den Kategorien Rechtschreibung, Zeichensetzung, Grammatik, Satzbau, Zeitfehler und in sonstige Fehler mit Inhalt und Ausdruck. Grammatik wurde hierbei unterstrichen, da dies als übergeordnete Kategorie gezählt wurde, welche Satzbau- und Zeitenfehler miteinschließt und auch in der qualitativen Analyse so vorgegangen wurde. Am Ende der Spalten wurde für jede Lehrperson sowohl die Gesamtfehleranzahl als auch die gegebene Schularbeitsnote angeführt. Dabei wurde auch mit einem plus (+) oder minus (-) neben der Note vermerkt, wenn eine Lehrkraft explizit angegeben hatte, ob es beispielsweise noch beinahe ein „Genügend“ war oder ob sie nahezu eine bessere Note gegeben hätte. Angemerkt werden muss auch, dass die Lehrperson Nummer 2 kein einziges Korrekturzeichen verwendete, sondern lediglich direkt im Text ausgebessert hatte. Daher wurde in dieser Spalte auch immer die Zahl „0“ vermerkt. Im Kapitel der qualitativen Aspekte wird die Korrekturpraxis aber genau beschrieben, sodass auch die Fehlerwahrnehmung und -verbesserung der Lehrperson Nummer 2 ersichtlich wird. Bei den restlichen Lehrpersonen stimmten die markierten Stellen weitgehend mit den ausgewiesenen Korrekturzeichen am Rand überein. Detailliertere Aussagen dazu lassen sich im Kapitel 4.2.1 einsehen.

Lehrperson	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
Anzahl Korrekturzeichen (Sprachrichtigkeit)																					
Rechtschreibung	13	0	15	16	5	2	7	15	19	19	15	24	14	12	11	9	18	20	15	14	21
Zeichensetzung	2	0	3	5	6	1	4	5	8	7	3	6	0	6		3	0	10	6	3	4
<u>Grammatik</u>	7	0	2	10	21	11	10	8	6	6	14	10	7	9	7	9	11	23	9	8	10
Satzbau		0			2			1			2			2		1			3		
Zeitfehler	1	0	3	7	5			1	5	2	3	4	1	7	4	1			5		
Summe der Korrekturzeichen (Sprachrichtigkeit)	24	0	23	38	39	14	21	30	38	34	37	44	22	36	22		29	53	38	26	34
Anzahl Korrekturzeichen (Sonstiges)																					
Inhalt			2		2	6	2											2		1	3
Ausdruck	1		2	3	5	9	11	7	15	3	2	9	5	12	6	1	2	17	5	17	27
Summe der Korrekturzeichen	24	0	27	41	46	29	34	37	53	37	39	53	27	48	28	24	31	72	43	44	65
Benotung der Schularbeit	2	2	2	3	3	3-	3	3	3	3	3-	3	3+	3	4	4+	4-	4	4	4	5

Tabelle 1: Übersicht Korrekturzeichen & Benotungen

In einem ersten Schritt der Untersuchung wurden die Leistungsbeurteilungen der verschiedenen Lehrpersonen erhoben, welche in folgendem Balkendiagramm eingesehen werden können. Über 50 Prozent der Teilnehmer:innen beziehungsweise elf Lehrpersonen bewerteten den Schularbeitstext mit einem „Befriedigend“. Weitere sechs Personen oder 29 % befanden ein „Genügend“ als passend, während drei Lehrkräfte beziehungsweise 14 Prozent der Stichprobe davon überzeugt waren, der Text würde einem „Gut“ entsprechen. Eine einzige Lehrkraft beurteilte die untersuchte Schularbeit mit einem „Nicht genügend“. Ein „Sehr gut“ war für alle teilnehmenden Personen ausgeschlossen.

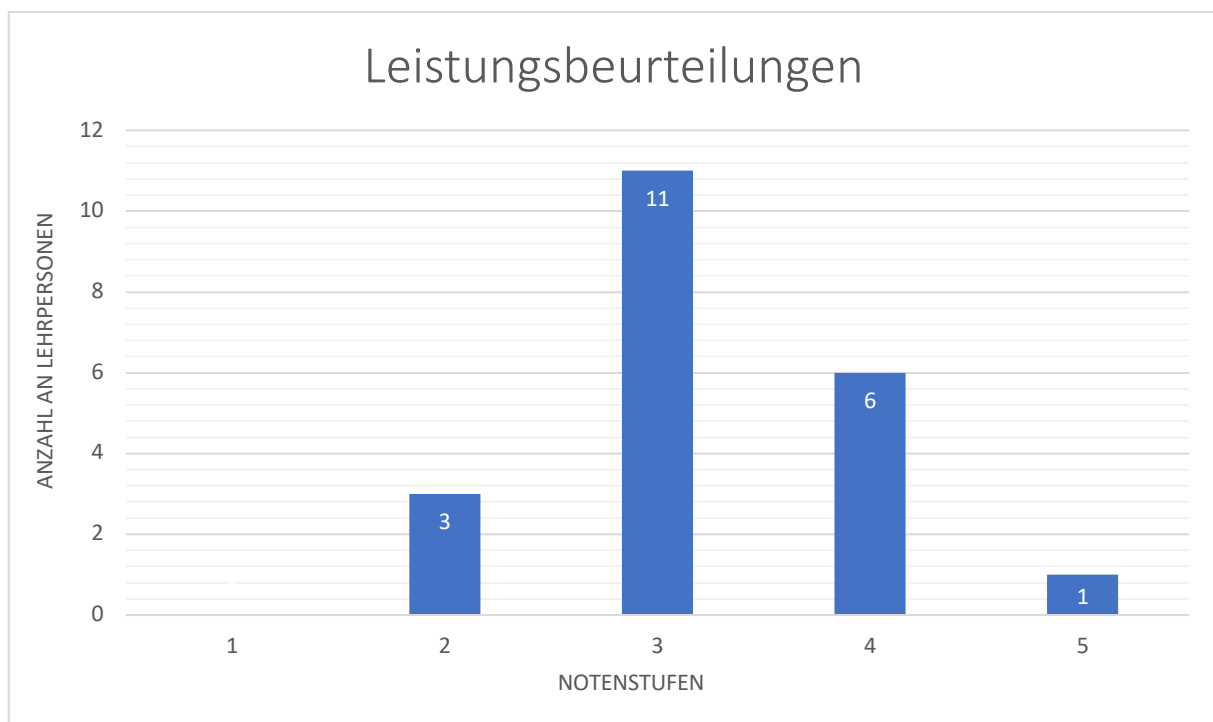


Abbildung 7: Leistungsbeurteilungen aller Teilnehmer:innen

Im Folgenden werden die Arbeiten mit den meisten und den wenigsten Korrekturzeichen angeführt. Dabei wurden, wie oben bereits erwähnt, Fehler des Satzbaus und Zeitfehler unter der Kategorie Grammatik zusammengefasst, um die Einteilung übersichtlicher zu gestalten. Darüber hinaus fallen diese Aspekte in der Theorie im Regelfall auch unter den Punkt der Grammatik.

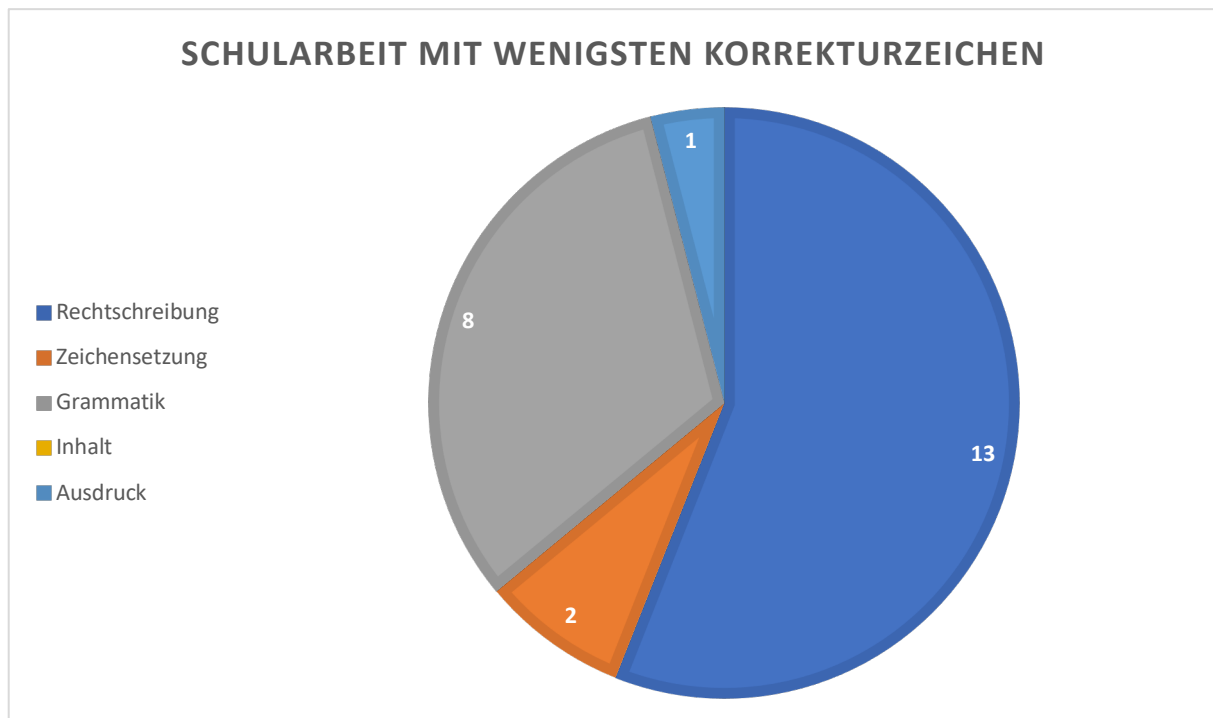


Abbildung 8: Korrektur Lehrperson Nr.1

Das Diagramm zeigt die Korrektur der Lehrperson Nummer 1 auf, welche die niedrigste Fehleranzahl mit insgesamt 24 korrigierten Fehlern in den Bereichen Rechtschreibung, Zeichensetzung, Grammatik und Ausdruck aufweist. Inhaltsfehler wurden keine markiert. Die Mehrheit der Normverstöße wurde im Bereich der Rechtschreibung mit 13 Fehlern, einsehbar in der dunkelblauen Fläche, markiert. Bezüglich der Grammatik, hier im Diagramm grau hinterlegt, wurden acht Fehler von der Lehrperson gefunden. Beistrich- oder Zeichensetzungsfehler im orangenen Areal wurden zwei angestrichen und im Bereich des Ausdrucks, hier hellblau hinterlegt, wurde ein Fehler verbessert. Diese Lehrperson beurteilte die Schularbeit mit der Note „Gut“.

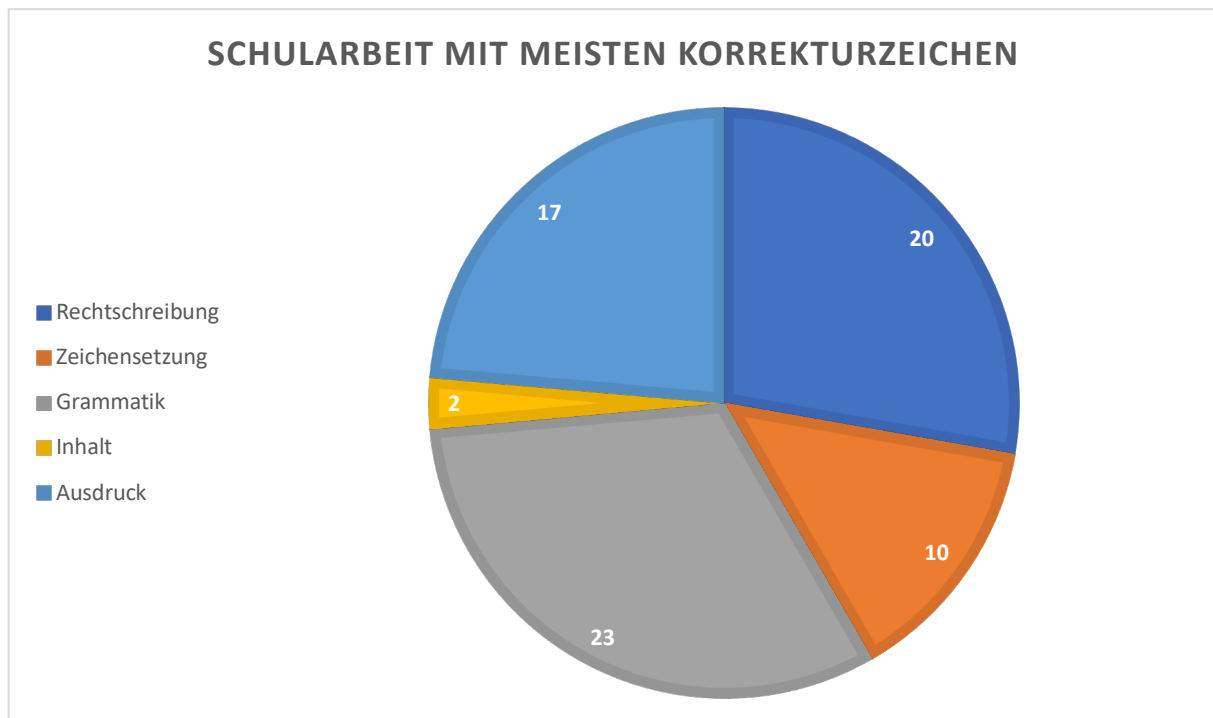


Abbildung 9: Korrektur Lehrperson Nr.18

Im Gegensatz dazu zeigt dieses Diagramm die Aufteilung der Fehler bei der Schularbeit mit den meisten Korrekturzeichen. Die mit insgesamt 72 Markierungen versehene Schularbeit wurde von der Lehrperson Nummer 18 korrigiert. Die Mehrheit der Korrekturen, nämlich 23 Stellen, wurde im Bereich der Grammatik gefunden. Auch bei der Rechtschreibung wurden 20 Normverstöße markiert. Bezüglich der Zeichensetzung wurden 10 Fehler verbessert. Darüber hinaus machte die Lehrperson 17 Anmerkungen im Bereich Ausdruck, während inhaltlich nur zwei Fehler gekennzeichnet wurden. Diese Schularbeit wurde mit einem „Genügend“ beurteilt, obwohl bei der Bewertung ersichtlich wird, dass nur aufgrund eines Punktes Unterschied kein „Nicht Genügend“ vergeben wurde.

Darüber hinaus soll auch die Schularbeit gezeigt werden, welche als einzige mit einem „Nicht Genügend“ beurteilt wurde.

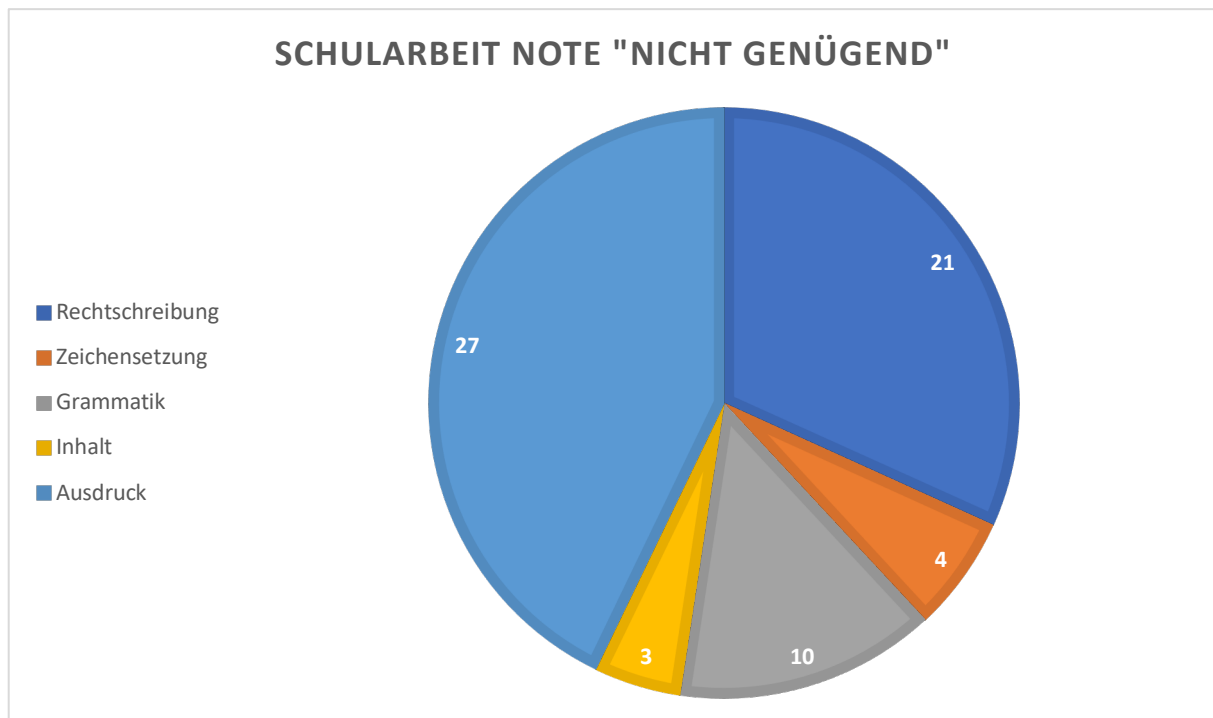


Abbildung 10: Korrektur Lehrperson Nr.21

Dieses Diagramm zeigt die Aufteilung der Korrekturen der Lehrperson Nummer 21. Der Text ist mit insgesamt 64 markierten Fehlern neben der zuvor genannten Schularbeit jener mit den zweitmeisten Korrekturen. Dabei entfällt die Mehrheit mit 27 Fehlern auf den Bereich Ausdruck, während 21 Normverstöße im Feld der Rechtschreibung markiert wurden. In grammatikalischer Hinsicht bemängelte die Lehrperson zehn Fehler. Bezüglich der Zeichensetzung wurden vier Fehler gefunden und im Bereich Inhalt wurden drei Verstöße markiert.

Im Folgenden wird die Anzahl an Korrekturzeichen je nach Benotung der einzelnen Arbeiten aufgezeigt. Die ersten drei Schularbeiten, welche mit einem „Gut“ beurteilt wurden, weisen in einem Fall 24 markierte Gesamtfehler und in einem anderen Fall 27 Korrekturzeichen auf, während die Lehrperson Nr. 2 lediglich direkt im Text ausgebessert hat ohne ein einziges Korrekturzeichen am Rand zu vermerken.

Für die insgesamt elf Schularbeiten, welche mit einem „Befriedigend“ beurteilt wurden, können die Werte für eine bessere Übersicht in folgendem Diagramm eingesehen werden.

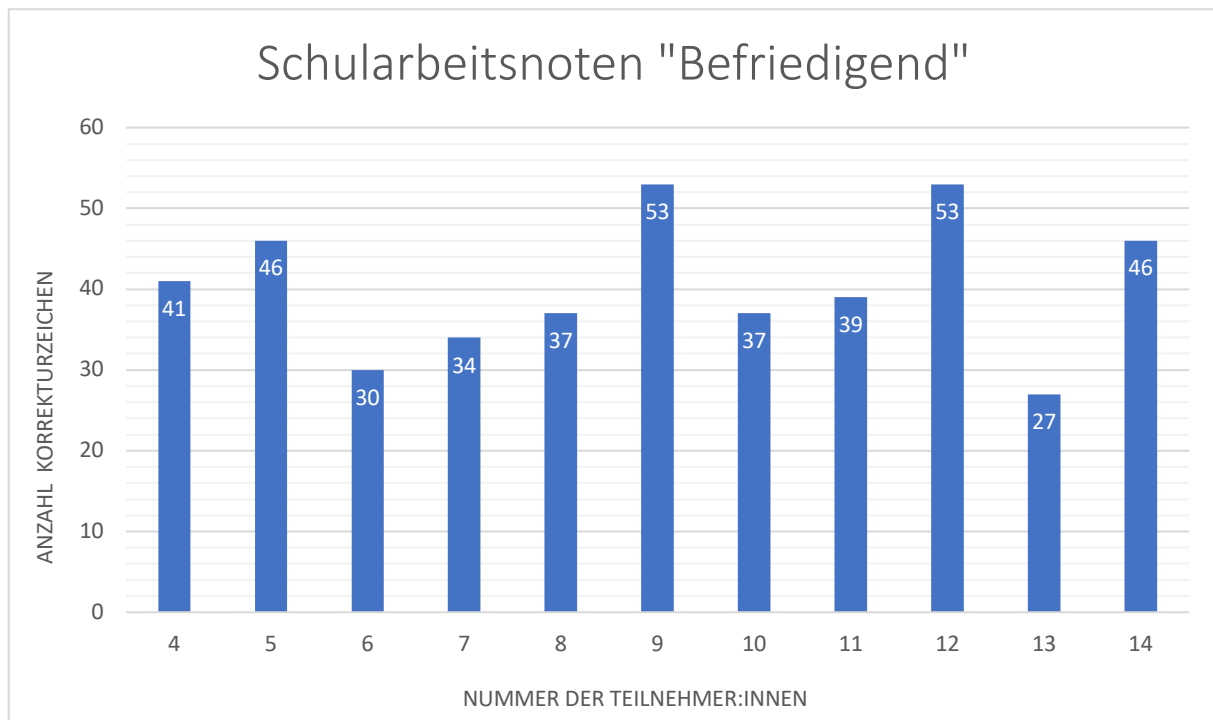


Abbildung 11: Übersicht Korrekturzeichen bei Arbeiten mit Note "Befriedigend"

Diese reichen dabei von 27 markierten Gesamtfehlern bei Lehrperson 13 bis zu insgesamt 53 gezählten Korrekturzeichen bei den Lehrpersonen 9 und 12. Bei der Schularbeit Nummer 13 mit den wenigsten Korrekturzeichen kann hinzugefügt werden, dass die Lehrperson die Arbeit eigenen Angaben zufolge zunächst mit einem „Gut“ bewerten wollte, dies dann aber bei einem zweiten Durchgang doch verworfen hat. Der Rest der Schularbeiten verteilt sich auf den Mittelbereich und liegt somit bei 30 und 46 gezählten Korrekturzeichen.

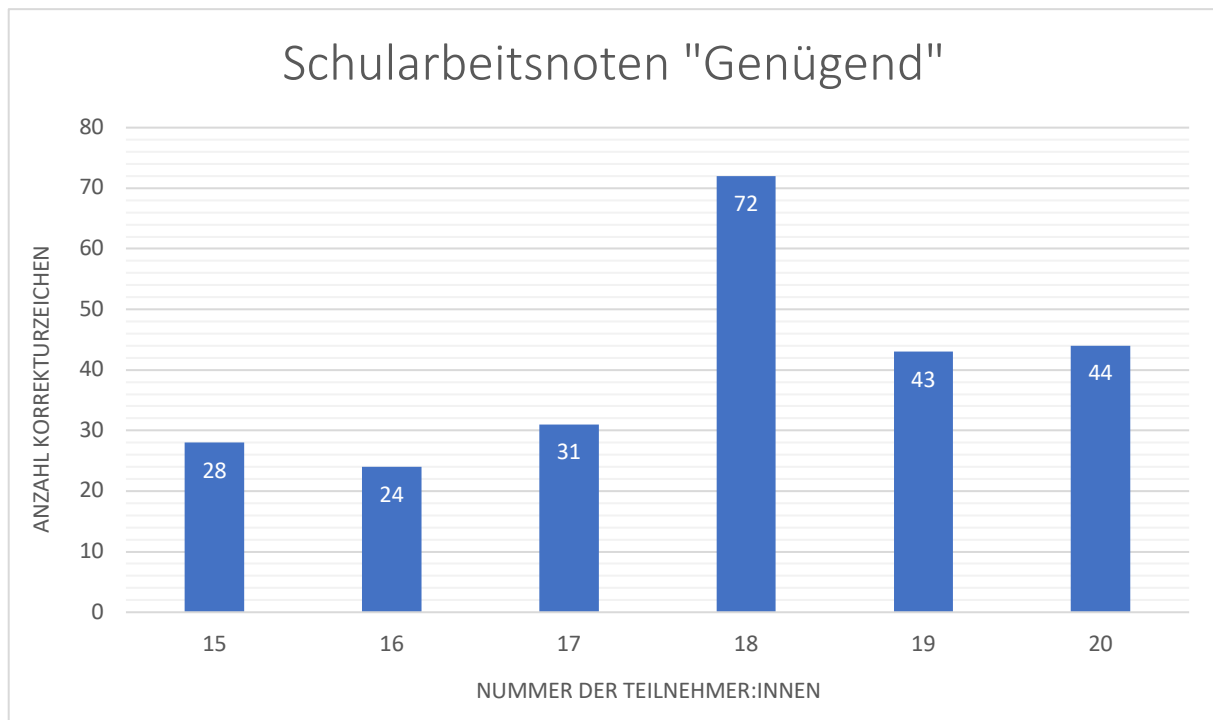


Abbildung 12: Übersicht Korrekturzeichen bei Arbeiten mit Note "Genügend"

Dieses Diagramm zeigt die Korrekturzeichenanzahl bei den insgesamt sechs Schularbeiten mit der Note „Genügend“ auf. Während die geringste Anzahl mit 24 Markierungen beim Schularbeitstext 16 liegt, weist der Text Nummer 18 72 gezählte Korrekturzeichen auf. Die restlichen Markierungen in den Schularbeiten reichen dabei von 28 bis 44 Verbesserungen.

Auf der nächsten Seite kann eine Übersicht über die Daten der einzelnen Proband:innen eingesehen werden. Hierbei wurden die wichtigsten Erkenntnisse des Fragebogens noch einmal zusammengefasst dargestellt und darüber hinaus für die Fragestellung relevante Mittelwerte berechnet.

Lehrperson	Note	Schulart	Ort	Geschlecht	Unterrichtserfahrung	Erfahrung 7. Klasse	Anstellung
1	2	NMS	Stadt	weiblich	unter 5 Jahre	ja	>10h
2	2	NMS	Stadt	weiblich	über 5 Jahre	ja	>10h
3	2	NMS	Stadt	weiblich	unter 5 Jahre	ja	>10h
4	3	NMS	Stadt	weiblich	unter 5 Jahre	ja	<10h
5	3	AHS	Stadt	weiblich	unter 5 Jahre	nein	>10h
6	3	AHS	Stadt	weiblich	über 5 Jahre	ja	<10h
7	3	AHS	Stadt	weiblich	unter 5 Jahre	nein	<10h
8	3	AHS	Stadt	weiblich	über 5 Jahre	ja	<10h
9	3	NMS	Stadt	weiblich	gleich 5 Jahre	ja	>10h
10	3	AHS	Stadt	männlich	über 5 Jahre	ja	<10h
11	3	AHS	Stadt	weiblich	über 5 Jahre	ja	<10h
12	3	NMS	Stadt	weiblich	gleich 5 Jahre	ja	>10h
13	3	NMS	Stadt	weiblich	gleich 5 Jahre	ja	<10h
14	3	NMS	Stadt	weiblich	unter 5 Jahre	ja	>10h
15	4	NMS	Stadt	männlich	unter 5 Jahre	ja	<10h
16	4	AHS	Stadt	männlich	unter 5 Jahre	nein	<10h
17	4	NMS	Stadt	männlich	unter 5 Jahre	ja	>10h
18	4	AHS	Land	weiblich	über 5 Jahre	ja	>10h
19	4	NMS	Stadt	weiblich	unter 5 Jahre	ja	>10h
20	4	NMS	Stadt	weiblich	unter 5 Jahre	ja	<10h
21	5	AHS	Land	weiblich	über 5 Jahre	nein	>10h

Tabelle 2: Übersicht Fragebogendaten der Proband:innen

Der Mittelwert aller Noten beläuft sich auf 3,2. Die Arbeiten der männlichen Probanden weist dabei einen Wert von 3,8 auf, der Mittelwert bei der Notengebung der Frauen beträgt 3,1. Für die Interpretation der Ergebnisse muss hierbei jedoch auf die ungleiche Verteilung der Proband:innen hingewiesen werden.

Der Mittelwert der Beurteilungen von AHS-Lehrkräften beträgt 3,5, während NMS-Lehrende im Durchschnitt mit 3,2 beurteilten.

Junglehrende, welche mit bis zu fünf Jahren Berufserfahrung definiert werden, benoteten im Durchschnitt mit 3,2 und die restlichen Proband:innen, welche zu den erfahreneren Lehrkräften zählen, wiesen einen Durchschnittswert von 3,3 auf.

4.2 Qualitative Aspekte: Vergleich des Korrekturverhaltens und der Beurteilungen

In diesem Kapitel erfolgt eine qualitative Analyse der korrigierten Schularbeiten und der Antworten aus dem Fragebogen. Dazu werden das Korrekturverhalten und das Beurteilungsschema der teilnehmenden Lehrpersonen aufgezeigt. Hierfür werden die fehlerhaften Sätze und Phrasen der Schularbeit extrahiert und die Unterschiede und Gemeinsamkeiten im Korrekturverhalten der 21 Lehrenden aufgezeigt. Darüber hinaus wird die Frage nach den üblichen Beurteilungsschemata der Proband:innen ausgewertet.

4.2.1 Darstellung der Korrekturen

Bevor im weiteren Verlauf die Beschreibung des Korrekturverhaltens erfolgt, sollen zunächst die in der Schularbeit vorgefundenen Korrekturzeichen und die damit verbundene, vermutete Kategorisierung, welche sich aus dem Kontext erschließen ließ, aufgelistet werden.

Kategorie	Korrekturzeichen
Rechtschreibung	R RS
Grammatik	G Gr SB (Satzbau) SG (Satzgefüge)
Zeit	Z t
Zeichensetzung	B BS x Sz (Satzzeichen)
Ausdruck	A SB (Satzbau) Wh (Wiederholung) ? Ugs
Inhalt	I i

Tabelle 3: Übersicht über verwendete Korrekturzeichen

Zu den folgenden Ausführungen zur Korrekturanalyse soll angemerkt werden, dass die Sätze aus der Schularbeit 1:1 mit den von dem Schüler oder der Schülerin gemachten Fehlern übernommen und extrahiert wurden. Bei den Beschreibungen des Korrekturverhaltens werden die Nummern der Gewährspersonen immer in Klammern angeführt. Somit kann bei näherem Interesse nachgesehen werden, um welche Lehrpersonen es sich handelt.

Insgesamt gab es 33 Sätze im Schüler:innentext, welche fehlerhaft waren. Dabei konnten aber unterschiedlichste Fehlerstellen von den verschiedenen Lehrpersonen ausgemacht werden. Im Folgenden kann zunächst ein Screenshot der Original-Schularbeit eingesehen werden und im Anschluss daran werden die einzelnen Fehlerstellen betrachtet und Schritt für Schritt analysiert.

blutige Herbstferien

Die letzten weichen Strahlen der Herbstsonne versuchten das blasse Gesicht des Mädchens, das unter dem Fenster lag, zu erwärmen, doch sie blieb kühl. Der Täter wusste, dass er geschnappt werden würde, ~~falls~~ wenn er nicht die Spuren beseitigen ~~würde~~ und weglaufen würde. Ein Handy klingelte. Der Täter ging sofort ran. Es stellte sich heraus, dass dies sein Komplize sei, der wissen wollte, ob die Tat vollbracht war, doch er schwieg, legte auf, und ging davon.

Es war ~~ein~~ Ein typischer Samstag. Kommissar Brot verspeißte seinen Apfel und dachte sich nichts dabei, als auf einmal sein Handy ertönte. „Kommissar Brot zu Ihren Diensten! wie kann ich heute behilflich sein?“, sagte er mit einem höflichem Umgangston. Die nervöse Frau am Handy schrie, dass er doch schnell zur Algostraße 12 kommen sollte da es ein Notfall gab. Kommissar Brot versuchte die ~~Frau~~ Frau zu beruhigen, doch es half nix. Daraufhin hoppte er in seinen Wagen und machte sich auf den schnellsten Weg zur Algostraße 12. Er sah 2 Junge Damen. Eine kullerten die Tränen herunter, wie ein Wasserfall. Die andere, war jedoch voll mit ~~blut~~ blut und Glassplittern als er näher kam, sah er dass das junge Mädchen 2 Schüsse in den Kopf und einen in die Brust bekommen hatte.

Der Mann trat näher und befragte die Dame. Es kam heraus, dass sie Anna hieß und das anscheinend tote Mädchen Marie. Marie war 17 Jahre alt und ging noch zur Schule. Eine Sirene ertönt „Das wird wohl die Polizei sein.“, sagte Anna mit Tränen in ihren Augen. ~~Kommissar~~ Kommissar Brot erzählte den Polizisten über das Verbrechen. Dabei fiel ihm ein mitte 60 Jähriger Mann auf, der das Mädchen leblos anstarrt. Es kam ihm so vor, dass der Mann seinen Blick nicht von ihr abwenden könne. Dies gab Kommissar Brot ein merkwürdiges Gefühl. ~~Der~~ Der Mann seufzte und sagte mit einer leisen und ruhigen Stimme: „2 unnötige Schüsse.“ Da bekam Kommissar Brot ein Geistesblitz. Herr Brot starrte ihm mit großen Augen an und sagte: „Sie waren es.“ Der Mann starrte zurück. Alle Augen waren auf die Beiden ~~gerichtet~~ gerichtet. Man sah dem Mann die ~~Furcht~~ Furcht an als hätte man es ihm auf die Stirn geschrieben. Er musste sich die Tränen unterdrücken und gab anschließend ein Geständnis ab. Es kam heraus, dass er, ihr Freund sei und das nur getan hat, da sie ~~er~~ über seine illegalen Geschäfte heraufgefunden hatte und das sein Komplize noch auf freien Füße sei. Doch er wollte noch wissen, wie er darauf gekommen ist, dass er der Mörder sei? Kommissar Brot meinte: „Woher wissen sie, dass 2 Schüsse unnötig ~~zu~~ waren?“

Abbildung 13: Original-Schularbeit

Fehlerstelle:

1. blutige Herbstferien

Die Verfehlung der Groß- und Kleinschreibung in der Überschrift der Schularbeit wurde von fünf Lehrpersonen, nämlich von Nummer 3, 5, 16, 17 und 18 weder im Text verbessert noch mit Randkommentaren versehen. Vier Lehrkräfte (Nummer 2, 8, 14, 19) haben das Wort „blutige“ lediglich im Text ausgebessert, indem sie ein großes „B“ über der fehlerhaften Phrase vermerkten. Über 50 Prozent der Stichprobe, nämlich 12 Lehrpersonen mit der Nummer 1, 4, 6, 7, 9, 10, 11, 12, 13, 15, 20 und 21 verbesserten sowohl im Text und markierten den Fehler auch mit einem Korrekturzeichen am Rand. Dabei wurde der Normverstoß von allen bis auf die Nummer 6 als Rechtschreibfehler aufgefasst und mehrheitlich mit einem „R“ ausgewiesen, während diese einen Grammatikfehler mit einem „G“ beschrieb.

2. Die letzten weichen Strahlen der Herbstsonne versuchten das blasse Gesicht des Mädchens, das unter dem Fenster lag, zu erwärmen, doch sie blieb kühl.

Der zweite Satz des Schularbeitstextes rief bei den Lehrpersonen ebenso eine Frage der Korrektheit hervor, jedoch wurde der Fehler von 43% (Nummer 1, 2, 3, 6, 8, 10, 13, 17, 19) entweder nicht gefunden und somit weder im Text verbessert noch mit Randkommentaren versehen, oder er wurde schlichtweg nicht als Normverstoß empfunden. Lehrperson Nummer 20 verbesserte lediglich im Text. Sie empfand die Formulierung unpassend, besserte den Beistrich nach „erwärmen“ zu einem Punkt aus und empfahl den Satz „Doch sie blieb kühl.“ als eigenes Gefüge.

Insgesamt 38 Prozent beziehungsweise acht Lehrpersonen korrigierten sowohl direkt im Text als auch mit einer Notiz von Korrekturzeichen am Rand. Die Einordnung der Fehler wies hierbei jedoch einige Unterschiede auf.

Nummer 4, 12, 15 und 18 unterstrichen beziehungsweise unterwellten das Pronomen „sie“ und wiesen es am Rand als Grammatikfehler aus. Bis auf Lehrperson Nummer 18 schrieben dabei auch alle die korrekte Form von „es“ im Text darüber. Nummer 15 unterwellte darüber hinaus das Adjektiv „kühl“ und markierte einen Ausdrucksfehler mit einem „A“. Dieser Meinung waren auch Nummer 7 und 14. Die Lehrperson mit der Nummer 5 unterwellte die gesamte Phrase „doch sie blieb kühl“ und befand einen Inhaltsfehler als passend, welcher mit einem „I“ am Rand markiert wurde.

3. Der Täter wusste, dass er geschnappt werden würde, wenn er nicht die Spuren beseitigen und weglaufen würde.

71 Prozent der Stichprobe mit insgesamt 15 Lehrpersonen (1, 2, 4, 5, 6, 7, 8, 9, 10, 12, 13, 14, 16, 17, 18) empfand diesen Satz als korrekt. Hierbei wurde weder im Text noch mit Randkommentaren eine Korrektur ausgewiesen.

Lehrperson Nummer 15 unterwellte „Der Täter“ und stellte darüber die Frage „Wer?“. Lehrperson Nummer 20 strich das Partikel „nicht“ durch und fügte ein Fehlzeichen zwischen „Spuren“ und „beseitigen“ ein, um den Satzbau neu zu ordnen. Beide Lehrpersonen nahmen lediglich Verbesserungen im Text vor, ohne dies mit einem Korrekturzeichen am Rand zu vermerken.

Lehrperson Nummer 19 teilte dieselbe Meinung über den Normverstoß wie Nummer 20, aber notierte zusätzlich das Kürzel „SB“ am Rand. Lehrperson Nummer 3 und 21 markierten ein „A“ und korrigierten die Wortwiederholung „würde“ im Text durch Einkreisen oder Unterwellungen. Dabei hat Lehrperson Nummer 3 zweimal einen Ausdrucksfehler nebenbei geschrieben und in Klammer mit „WH“ detaillierter den Fehler kategorisiert. Durch das Einklammern einer Markierung wird jedoch angenommen, dass sie dies nur als einen statt zwei Fehlern zählt.

4. Ein Handy klingelte.

Dieser grundsätzlich korrekte Satz wurde von drei Lehrpersonen (3, 20, 21) sowohl im Text verbessert als auch mit Randkommentaren versehen, während die restlichen 86 Prozent (Nummer 1, 2, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19) nichts zu bemängeln hatten. Lehrperson Nummer 3 kreiste den unbestimmten Artikel „Ein“ ein, schrieb „Wessen?“ darüber und vermerkte dies mit dem Korrekturzeichen „I“ als Inhaltsfehler am Rand. Lehrperson Nummer 20 und 21 unterwellten den Artikel „Ein“ und besserten zu „Sein“ aus. Am Rand wurde dies je einmal mit einem großen „I“ und einem kleinen „i“ vermerkt.

5. Der Täter ging sofort ran.

71 Prozent der Stichprobe (1, 2, 3, 4, 5, 7, 8, 10, 11, 12, 13, 14, 15, 16, 17) stufte diesen Satz als korrekt ein und markierte weder im Text noch am Rand durch Korrekturzeichen. Sechs Lehrpersonen (6, 9, 18, 19, 20, 21) bemängelten jedoch den Ausdruck, unterstrichen beziehungsweise unterwellten das Wort „ran“ und markierten das Kürzel „A“ am Rand.

Lehrperson 6, 18 und 20 machten darüber hinaus Korrekturvorschläge mit „hob“ oder „nahm sofort ab“.

6. Es stellte sich heraus, dass dies sein Komplize sei, der Wissen wollte, ob die Tat vollbracht war, doch er schwieg, legte auf, und ging davon.

Bei diesem Satz waren sich alle Lehrpersonen einig, dass hierbei Korrekturen notwendig sind. Lediglich im Text verbesserte Lehrperson Nummer 2, indem sie den „dass“-Fehler, „Komplize“ und „sei“ unterwellte. Alle anderen Lehrpersonen verbesserten sowohl im Text als auch mit Korrekturzeichenvermerk am Rand, jedoch unterschieden sich die erkannten Fehler und Fehlerkategorien.

Die Lehrperson 15 wies lediglich einen Rechtschreibfehler bezüglich des Substantivs „Komplize“ aus, zusätzlich wurden jedoch „sei“ durchgestrichen und das Pronomen „er“ unterwellt. Dabei wurde auch „wer genau?“ als Frage darüberschrieben. Der letzte Beistrich im Satz wurde eingeklammert, ohne dies als Fehler zu werten.

Lehrperson 1 korrigierte den „dass“-Fehler und „Komplize“ und markierte dies mit jeweils einem „R“ am Rand. Lehrperson 7 bemängelte dieselben Fehler, wies den ersten jedoch als Grammatikfehler mit einem „G“ aus. Derselben Meinung war Lehrperson 20 und korrigierte darüber hinaus noch die Phrase „Doch er schwieg, legte auf und ging davon.“ als eigenen Satz, wies dies jedoch nicht mit einem Korrekturzeichen aus.

Lehrperson 3 markierte zwar auch zwei Rechtschreibfehler mit dem Korrekturzeichen „R“, diese bezogen sich jedoch auf den „dass“-Fehler und das Wort „Wissen“, welches im Text groß geschrieben wurde. Hierbei muss jedoch angemerkt werden, dass aus der Schrift des Schülers oder der Schülerin nicht klar ersichtlich war, ob es sich um eine Groß- oder Kleinschreibung handelte (siehe Original-Schularbeit im Anhang).

Vier Lehrpersonen (6, 10, 13, 17) markierten jeweils drei Fehler in diesem Satz, welche sich jedoch unterscheiden. Lehrperson 6 vermerkte bezüglich des Wortes „Komplize“ und „sei“ zwei inhaltliche Fehler mit „I“ und verbesserte eine überflüssige Beistrichsetzung am Ende des Satzes mit einem „B“ am Rand. Die restlichen drei Lehrpersonen besserten die Phänomene „dass“, „Komplize“ und „Wissen“ aus und markierten alles mit jeweils einem Rechtschreibfehler. Zusätzlich strich Lehrperson 10 den letzten Beistrich im Satz, merkte jedoch keinen Beistrichfehler wie Lehrperson 6 an.

Von insgesamt 6 Lehrpersonen (4, 5, 8, 12, 19, 21) wurden vier Fehler in diesem Satz markiert. Während Nummer 4, 8 und 12 die Verstöße bei „dass“, „Komplize“, „sei“ und „wissen“ mit Rechtschreib- und Grammatikfehler vermerkten, wobei der Grammatikfehler

„sei“ bei der Lehrperson Nummer 4 differenzierter mit einem „Z“ markiert wurde, waren für die Lehrpersonen 5, 19 und 21 andere Verstöße relevant. Lehrkraft 5 verbesserte zwar auch das Wort „wissen“ im Text, vermerkte jedoch keinen Fehler am Rand. Anstelle dessen besserte diese jedoch einen Beistrichfehler aus. Dasselbe galt für Proband:in Nummer 19. Die Lehrperson 21 vermerkte als vierten Verstoß einen Ausdrucksfehler im letzten Absatz und verbesserte zu „Doch er schwieg, legte auf, und ging davon.“

Lehrperson Nummer 14 korrigierte die bereits genannten Fehler zusammengefasst und kam somit auf insgesamt fünf markierte Korrekturen, wobei statt des Grammatikfehlers abermals ein „Z“ verwendet wurde.

Zwei Lehrkräfte (9, 18) konnten sogar sechs Fehler ausmachen. Die Lehrperson 9 markierte dabei jeweils ein „R“ für „dass“, „Komplize“ und „doch“, da es großgeschrieben werden sollte. Darüber hinaus ein „Z“ für „sei“, ein „SZ“ für einen Punkt statt einem Beistrich vor dem Wort „doch“ und schlussendlich das Korrekturzeichen „SZ“ für den letzten Beistrich, welcher laut der Proband:in überflüssig war. Das Wort „wissen“ wurde darüber hinaus auch im Text angezeichnet, jedoch ohne Fehlermarkierung am Rand. Die Lehrperson Nummer 18 markierte jeweils ein „R“ für „dass“, „Komplize“ und „wissen“, ein „G“ für „sei“. Den letzten Absatz als neuen Satz zu sehen war auch für diese Person klar, sie markierte dies jedoch mit einem Ausdrucksfehler mit dem Korrekturzeichen „A: SB“. Das Wegstreichen des letzten Beistriches wurde genauso mit einem „SZ“ markiert.

7. Es war Ein typischer Samstag.

95 Prozent oder 20 Personen aus der Stichprobe erkannten den Fehler. Während Lehrkraft 2 wie immer lediglich im Text ausgebessert hat, vermerkten 15 Lehrende den großgeschriebenen Artikel „Ein“ als Rechtschreibfehler mit Korrekturzeichen am Rand und der Rest markierte einen Grammatikfehler.

8. Kommissar Brot verspeißte seinen Apfel. und dachte sich nichts dabei, als auf einmal sein Handy ertönte.

Drei Lehrkräfte verbesserten lediglich im Text (ohne Korrekturzeichen am Rand), wobei von Nummer 2 und 6 nur der Punkt nach dem Substantiv „Apfel“ weggestrichen wurde. Lehrkraft 7 strich das Verb „verspeißte“ an.

Sowohl im Text als auch mit Randkommentaren verbesserten 17 Lehrpersonen. Dabei wurde von den Lehrpersonen 1 und 3 das Substantiv „Kommissar“ als Rechtschreibfehler markiert,

während Nummer 11, 15 und 17 lediglich „verspeißte“ mit einem „R“ ausbesserten. Lehrperson 11 korrigierte zwar auch das Wort „Kommissar“ und schrieb ein „R“ an den Rand, klammerte dieses jedoch ein. Fehlermarkierungen wie diese wurden in der Analyse nicht zu den Gesamtkorrekturzeichen gezählt, da sie vermutlich einen Folgefehler darstellen sollten. Beide Fehler wurden von den restlichen 57 Prozent oder 12 Proband:innen erkannt und jeder von ihnen vermerkte dies als Rechtschreibfehler. Zusätzlich dazu markierten Nummer 9, 12 und 21 einen Ausdrucksfehler, wobei Lehrperson 9 und 12 das Verb „ertönte“ unterwellten. Für die Lehrkraft Nummer 21 stellte aber die letzte Phrase ein Problem dar. Dahingehend verbesserte sie einen Punkt statt des Beistrichs und begann den Satz neu, indem sie „Plötzlich“ darüberschrieb. Lehrperson Nummer 9 befand den Punkt darüber hinaus als Satzzeichenfehler und markierte diesen mit „SZ“. Andere Proband:innen kreisten den Punkt ein, strichen ihn durch oder schrieben ein Fragezeichen darüber, werteten diesen aber nicht als Fehler.

9. „Kommissar Brot zu Ihren Diensten! wie kann ich heute behilflich sein?“, sagte er mit einem höflichem Umgangston.

Lehrperson Nummer 7 konnte in diesem Satz keine Fehler ausmachen. Daher wurde weder im Text verbessert noch Randzeichen vermerkt. Lehrperson 2 und 10 korrigierten das kleingeschriebene „wie“ lediglich im Text. Außerdem wurde von Lehrperson 2 das Wort „mit“ unterwellt und „in“ darüberschrieben und auch der Fehler in „höflichem“ wurde bemerkt. Die Verbesserungen der anderen Proband:innen fielen sehr unterschiedlich aus. Während die Lehrpersonen Nummer 3, 11 und 15 lediglich das „wie“ als Rechtschreibfehler deklarierten, konnten 67 Prozent der Stichprobe, nämlich die Proband:innen 1, 4, 5, 6, 8, 9, 12, 13, 14, 17, 18, 19, 20 und 21 auch den Fehler bei dem Adjektiv „höflich“ erkennen. Die Mehrheit vermerkte dies als einen Grammatikfehler, dennoch wurde dies von einigen auch als Ausdrucksfehler markiert. Beispielsweise strich Lehrperson Nummer 6 die Phrase „mit einem höflichem Umgangston“ durch und schrieb „fragte höflich“ darüber. Ebenso die Lehrkräfte Nummer 20 und 21 besserten im Text aus und schlugen „antwortete mit höflicher Stimme“ und „meldete er sich mit einem höflichen Umgangston“ vor. Lehrperson 21 markierte zusätzlich zu einem „A“ auch einen Rechtschreibfehler in diesem Satz. Die beiden Proband:innen 9 und 12 verbesserten das Wort „Diensten“ mit einem „R“ am Rand, wobei dies nicht gänzlich aus dem Originaltext ersichtlich war. Person 11 und 18 markierten darüber hinaus einen Beistrich nach der Phrase „Kommissar Brot, zu Ihren Diensten!“ und kennzeichneten dies mit den Korrekturzeichen „SZ“ und „B“.

10. Die nervöse Frau am Handy schrie, dass er doch schnell zur Algostraße 12 kommen sollte da es ein Notfall gab.

Lehrperson Nummer 3 bemängelte an diesem Satz nichts. Verbesserungen lediglich im Text wurden von Proband:in Nummer 2 und 6 getätigt, wobei nur ein Beistrich markiert wurde. Die Lehrpersonen 1, 13, 15, 17 und 19 verbesserten den Artikel „ein“ vor „Notfall“ und markierten einen Grammatikfehler. Zusätzlich zu diesem Fehler wurde von den Lehrpersonen 5, 7, 8, 10, 11, 12, 18 und 20 ein Beistrichfehler verbessert und am Rand notiert, um den letzten Teil des Satzes einzuleiten. Drei Personen korrigierten zwar auch den Beistrich im Text, wiesen diesen jedoch nicht als Fehler aus. Statt eines Grammatikfehlers kennzeichneten Nummer 9 und 21 die letzte Phrase als Ausdrucksfehler.

11. Kommissar Brot versuchte die Frau zu beruhigen, doch es half nix.

24 Prozent der Stichprobe (1, 3, 8, 11, 15) konnten an diesem Satz nichts Unzureichendes feststellen. Der Rest der Proband:innen korrigierte das Phänomen „nix“. Während die Mehrheit der Stichprobe einen Ausdrucksfehler vermerkte, markierten vier Lehrpersonen (5, 13, 17, 21) bei diesem Wort einen Rechtschreibfehler. Die Lehrkraft Nummer 20 fügte am Rand den Vermerk „Ugs“ hinzu.

12. Daraufhin hoppte er in seinen Wagen und machte sich auf den schnellstem weg zur Algostraße 12.

Jede einzelne Lehrperson der Stichprobe hatte an diesem Satz bestimmte Phänomene zu bemängeln. Während Nummer 2 im Text das Wort „hoppte“ unterwellte und die Phrase „den schnellstem weg“ ausbesserte, korrigierten die anderen mit Vermerken am Rand. Bezüglich des inkorrekten Satzteils „auf den schnellstem Weg“ konnten verschiedene Korrekturen beobachtet werden. Die Lehrpersonen 1, 3, 8, 9, 11 und 21 verbesserten lediglich das Substantiv „Weg“ als Rechtschreibfehler. Der Rest kennzeichnete meistens einen Grammatik- und Rechtschreibfehler für die gesamte Phrase, aber es gab auch Beispiele, die einen Ausdrucksfehler werteten. 48 Prozent der Lehrpersonen (4, 6, 9, 10, 12, 14, 15, 18, 19, 20) bemängelten zusätzlich den Ausdruck von „hoppte“ und gaben teilweise auch Verbesserungsvorschläge wie beispielsweise „hüpfte“ und „sprang“ an. Nummer 6 und 9 beanstandeten zusätzlich das Verb „machen“ mit dem Korrekturzeichen „A“ und besserten auf „begab“ und „fuhr“ aus.

13. Er sah 2 Junge Damen.

Lehrperson Nummer 1 war die Einzige, die an diesem Satz nicht Ungewöhnliches feststellen konnte oder absichtlich nicht ausbesserte. 71 Prozent der Restlichen (insgesamt 15 Lehrpersonen) markierten einen Rechtschreibfehler für das großgeschriebene Adjektiv „junge“. Die Lehrkräfte Nummer 5, 6 und 11 empfanden dies als Grammatikfehler. Während einige Proband:innen die Nummer „2“ durchstrichen und die Zahl ausschrieben, aber dies dennoch nicht als Fehler werteten, vermerkten die Lehrenden 9, 12, 13 und 21 einen Ausdrucksfehler. Lehrperson Nummer 18 markierte hierbei sogar einen Rechtschreibfehler.

14. Eine kullerten die Tränen herunter, wie ein Wasserfall.

Hierbei waren sich nahezu alle Lehrpersonen einig, „Eine“ zu korrigieren. Lediglich Nummer 2 und 6 haben zwar im Text ausgebessert, vermerkten aber kein Korrekturzeichen. Für 71 Prozent der Stichprobe stellte das Fehlen des letzten Buchstabes „R“ einen Grammatikfehler dar (Nummer 1, 4, 5, 7, 8, 9, 10, 11, 13, 14, 15, 16, 17, 18, 19). Die Lehrkräfte 3, 12 und 21 korrigierten hierbei einen Rechtschreibfehler. Proband:in Nummer 20 unterwühlte den kompletten Satz und wies dies mit einem „A“ für Ausdrucksfehler am Rand aus. Nummer 7 und 18 unterwühlten lediglich den Zusatz „wie ein Wasserfall“, aber vermerkten dadurch einen zusätzlichen Ausdrucksfehler zum Grammatikverstoß. Varianten wie „Sie weinte wie ein Wasserfall.“ und „wie einem Wasserfall“ wurden über dem Satz vermerkt. Darüber hinaus wurde der Beistrich von Nummer 10 und 18 gestrichen und als Beistrichfehler am Rand ausgewiesen.

15. Die andere war, jedoch voll mit blut und Glassplittern als er näher kam, sah er dass das junge Mädchen 2 Schüsse in den Kopf und einen in die Brust bekommen hatte.

An diesem Satz hatten alle Lehrpersonen bestimmte Phänomene zu bemängeln. Während Proband:in Nummer 2 wieder nur im Text korrigierte, verbesserte der Rest sowohl Inhalts-, Ausdrucks-, Grammatik-, Rechtschreib- und Zeichensetzungsfehler in unterschiedlichem Ausmaß. Der erste falsch gesetzte Beistrich wurde von einigen direkt im Text gestrichen, aber nur von 57 Prozent oder 12 Personen auch als Fehler gewertet.

Zwei Lehrpersonen (Nummer 5 und 6) markierten den Fehler der Klein- und Großschreibung des Substantivs „Blut“ mit einem Grammatikverstoß, während der Rest einen Rechtschreibfehler für angebracht hielt. Der Satzbau wurde von Nummer 3 und 6 mit einem

Inhaltsfehler markiert, wobei viele andere dies als Ausdrucksfehler wahrnahmen. Hierbei waren sich alle einig über den Vorschlag, einen neuen Satz zu beginnen: „Als er näher kam, ...“. Lehrperson Nummer 17 markierte für diesen als inkorrekt eingestuften Satzbau einen Rechtschreibfehler und Lehrperson 5 einen Grammatikfehler.

Die verwendete Mehrzahl von „Glassplittern“ wurde fälschlicherweise von den Lehrpersonen Nummer 9, 10 und 21 als Fehler am Rand mit einem Korrekturzeichen ausgewiesen, wobei Nummer 21 diesen als einen Grammatikfehler auffasste und die anderen beiden als Rechtschreibfehler. Die Lehrperson Nummer 11 verbesserte „Glassplittern“ zwar im Text, markierte aber keinen Fehler dazu am Rand. Den Vorschlag, einen Beistrich vor dem Satzgefüge „...dass das junge Mädchen...“ zu setzen, kam von 57 Prozent der Lehrpersonen (4, 5, 7, 8, 9, 10, 12, 16, 18, 19, 20, 21). Diese markierten daher auch einen Beistrichfehler. Einige wenige, wie beispielsweise Nummer 17, verbesserten den Beistrich zwar, wiesen diesen aber nicht als Fehler am Rand aus und der Rest bemerkte oder korrigierte den Fehler nicht. Die nicht ausgeschriebene Zahl „2“ wurde zwar von sehr vielen durchgestrichen und ausgebessert auf „zwei“, als Fehler gewertet wurde dieser jedoch nur von Nummer 9, 12, 13 und 21. Dabei waren sich alle vier einig, dass es sich um einen Ausdrucksfehler handelte.

16. Der Mann trat näher und befragte die Dame.

Bis auf die Lehrpersonen Nummer 15, 20 und 21 wurde dieser Satz von niemandem korrigiert. Die genannten Proband:innen bemängelten alle den Ausdruck des Substantivs „Mann“. Während Lehrperson Nummer 15 einen Ausdrucksfehler markierte und „Der Polizist“ darüberschrieb, stellte dies für Lehrkraft Nummer 21 ein Inhaltsfehler dar, welchen sie im Text noch zusätzlich mit einem Fragezeichen kennzeichnete. Proband:in Nummer 20 machte zwar auch Vorschläge wie „Der Inspektor“ und „Der Kommissar“, vermerkte dies aber lediglich mit einem Fragezeichen am Rand.

17. Es kam heraus, dass sie Anna hieß und das anscheinend tote Mädchen Marie.

Mehr als die Hälfte der Proband:innen kritisierte nichts im ersten Teil des Satzes. Lediglich 33 Prozent beziehungsweise sieben Personen (5, 6, 13, 14, 18, 20, 21) waren sich einig, hierbei einen Ausdrucksfehler anzustreichen. Teilweise wurden Vorschläge wie „Es stellte sich heraus“ und „Er erfuhr“ darüberschrieben (Nummer 5, 6, 14, 20). 76 Prozent (1, 3, 4, 5, 6, 8, 9, 10, 11, 12, 14, 16, 17, 18, 19, 21) markierten „anscheinend“ mit einem Rechtschreibfehler, während Lehrperson 7 und 15 dieses Adverb mit einem Ausdrucksfehler

korrigierten. Die Nummer 15 fügte dabei als Randkommentar „Mit solchen Verletzungen ist man eindeutig tot.“ hinzu. Proband:in Nummer 20 schrieb diesbezüglich ein Fragezeichen an den Rand. Die Lehrpersonen 5, 9 und 12 korrigierten den Satzbau zu „dass sie Anna und das anscheinend tote Mädchen Marie hieß.“ Während Nummer 5 hierbei einen Grammatikfehler kennzeichnete, stellte dies für Nummer 9 einen Ausdrucksverstoß dar. Nummer 12 besserte dies lediglich im Text ohne Markierung am Rand aus. Die Lehrpersonen 4 und 7 fügten das Verb „war“ am Satzende hinzu und markierten dies mit einem Ausdruck- (Nr. 4) und einem Inhaltsfehler (Nr. 7).

18. Marie war 17 Jahre alt und ging noch zur Schule.

Der Satz wurde von allen Lehrpersonen außer der Nummer 18 als korrekt eingestuft. Durch dessen Einklammerung und einer Markierung mit „I“ am Rand wurde ein Inhaltsfehler gekennzeichnet.

19. Eine Sirene ertönt: „Das wird wohl die Polizei sein.“, sagte Anna mit Tränen in ihren Auge.

Lehrperson Nummer 1 bemängelte nichts an diesem Satz. Lehrperson Nummer 2 besserte lediglich im Text auf: „Eine Sirene ertönte und sie sagte mit Tränen:...“ aus. Auch Nummer 6 korrigierte sehr gründlich, aber wies kein einziges Korrekturzeichen am Rand aus. Der Zeitenfehler fiel 76 Prozent der Personen auf, jedoch korrigierte ihn die Lehrperson Nummer 12 als Rechtschreibfehler, während der Rest ein „G“ oder ein „Z“ am Rand vermerkte. Den Punkt am Ende der direkten Rede markierten Nummer 10, 12, 14, 18 und 19 als Satzzeichenfehler. Die Lehrperson 21 korrigierte zwar auch mit dem Vermerk „SZ“ am Rand, dieser bezog sich jedoch auf einen fehlenden Punkt nach „ertönte“, welcher das Ende des Satzes markieren sollte. Die Korrektur der fehlenden Endung bei dem Substantiv „Auge“ wurde sehr unterschiedlich gehandhabt. Von den 67 Prozent, die es bemerkten und mit einem Korrekturzeichen versahen, markierten sieben Lehrkräfte einen Rechtschreibfehler (3, 4, 9, 12, 13, 20, 21) und sieben (5, 7, 8, 11, 14, 17, 18) einen Grammatikfehler. Die Lehrperson Nummer 20 markierte darüber hinaus einen Ausdrucksfehler aufgrund des Pronomens „ihren“ und korrigierte zu „den Augen“. Insgesamt vier Personen konnten den Fehler bei dem Wort „Augen“ nicht erkennen oder werteten ihn absichtlich nicht.

20. Kommissar Brot erzählte den Polizisten über das Verbrechen.

Drei Lehrpersonen (18, 20, 21) korrigierten einen Ausdrucksfehler für das Verb „erzählte“ und schrieben alle „berichtete“ darüber. Der Satzteil „über das Verbrechen“ wurde von 57 Prozent der Lehrpersonen ausgebessert, wobei sechs (1, 7, 8, 9, 17, 19, 21) von einem Ausdrucksfehler und sechs (11, 12, 13, 14, 15, 18) von einem Grammatikfehler ausgingen. Nahezu von jeder dieser Lehrkräfte wurde in Klammer „von dem“ als Korrekturvorschlag vermerkt. Lehrperson Nummer 18 notierte darüber hinaus „Inhalt: Polizist? Passant?“. Die sechs Proband:innen mit den Nummern 2, 3, 4, 5, 10 und 16 hatten nichts an diesem Satz zu bemängeln.

21. Dabei fiel ihm ein mitte 60 Jähriger Mann auf, der das Mädchen leblos anstarrt.

Jede einzelne Lehrperson hatte zumindest einen Aspekt an diesem Satz zu bemängeln. Während Nummer 1 und 2 den ersten Teil nur im Text verbesserten, konnten bei der Schularbeitsanalyse der restlichen Proband:innen die unterschiedlichsten Korrekturen mit Randvermerk beobachtet werden. Nummer 3, 9 und 13 markierten einen Rechtschreibfehler für „mitte“ und Nummer 10, 11, 12 und 17 für „60 Jähriger“, welches sie mit „60-jähriger“ ausbesserten. Nummer 6, 8 und 17 korrigierten beide Wörter zu „Mitte 60-Jähriger“ und „Mitte-60-jähriger“, aber vermerkten dennoch nur einen Rechtschreibfehler. Gleich zwei Rechtschreibfehler wurden für diese fehlerhafte Kombination von den Lehrpersonen Nummer 4, 14, 18 und 19 vergeben. Das Adjektiv „leblos“ wurde von der Lehrkraft Nummer 5 als Inhaltsfehler und von Nummer 18 als Ausdrucksfehler ausgewiesen. Der Zeitenfehler in dem Verb „anstarren“ wurde von 67 Prozent der Lehrpersonen (4, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19) erkannt und entweder mit einem „Z“ oder „G“ am Rand vermerkt. Zwei Proband:innen (15, 17) markierten einen Rechtschreibfehler und änderten den Satzbau zu „der das leblose Mädchen“. Die Lehrkräfte Nummer 20 und 21 notierten beide ein Fragezeichen für den gesamten Satz und einen Ausdrucksfehler am Rand.

22. Es kam ihn so vor, dass der Mann seinen Blick nicht von ihr abwenden könne.

Der Hauptsatz in diesem Satzgefüge wurde bis auf Lehrperson Nummer 3 von allen korrigiert, wobei Nummer 2 und 15 den falschen Fall bei dem Pronomen „ihn“ lediglich im Text ausgebessert haben. Die restlichen 86 Prozent der Proband:innen (1, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 16, 17, 18, 19, 20, 21) waren sich einig, dass es sich dabei um einen Grammatikfehler handelte.

Die Phrase „dass der Mann“ wurde von den sechs Lehrkräften 1, 4, 12, 14, 20 und 21 als Ausdrucksfehler markiert. Während zwei davon zu „als könnte der Mann..“ korrigierten, waren sich die anderen einig über „als ob...“.

Der Zeitenfehler in „abwenden könne“ wurde von sieben Lehrpersonen korrigiert, wobei Nummer 2 und 6 lediglich im Text zu „könnte“ ausbesserten und der Rest (5, 11, 14, 18, 19) einheitlich einen Grammatik- bzw. Zeitenfehler markierte. Lehrperson Nummer 11 war jedoch der Meinung, dass „konnte“ die richtige Form darstellen würde und vermerkte zusätzlich am Rand „direkte Rede“.

23. Dies gab Kommissar Brot ein merkwürdiges Gefühl.

Während 71 Prozent der Stichprobe (15 Personen) diesen Satz als korrekt einstufen, waren sich die restlichen sechs Lehrpersonen (6, 8, 14, 18, 20, 21) einig, den ersten Teil oder in dem Fall von Person Nummer 20 und 21 den gesamten Satz als Ausdrucksfehler zu markieren. Einen Vorschlag zur Korrektur gab es von Nummer 6: „Dies verursachte bei Kommissar Brot ein merkwürdiges Gefühl.“

24. Der Mann seufzte und sagte mit einer leisen und ruhigen stimme: „2 unnötige Schüsse.“

Lehrperson Nummer 21 korrigierte das Verb „seufzte“ fälschlicherweise mit einem Rechtschreibfehler und den Rest des Hauptsatzes mit einem Ausdrucksfehler. Die Lehrperson 20 war dergleichen Meinung, markierte dafür ein „A“ am Rand und schlug als Verbesserung „Der merkwürdige Mann seufzte und sagte mit leiser und ruhiger Stimme“ vor.

Bis auf Nummer 5 und 10 wurde das Substantiv „Stimme“ von jeder Lehrperson korrigiert, wobei Nummer 2 und 14 nur im Text ohne einen Vermerk am Rand verbesserten, während alle anderen einen Rechtschreibfehler markierten. Die Nummer 6 war dabei die einzige Ausnahme. Sie befand diesen Verstoß als einen Grammatikfehler.

Bezüglich der direkten Rede besserte zwar mehr als die Hälfte mit elf Lehrpersonen die Zahl „2“ aus und schrieb diese aus, jedoch wurde dies nur von den Lehrkräften Nummer 9, 12 und 13 als Ausdrucksfehler vermerkt. Letztere gibt aber zusätzlich zu dem Korrekturzeichen ein „(WH)“ an, womit somit (Wiederholung) für einen Folgefehler gemeint sein dürfte.

25. Da bekam Kommissar Brot ein Geistesblitz.

Der Satzanfang wurde von 33 Prozent der Lehrpersonen (6, 7, 8, 13, 14, 15, 18) als Ausdrucksfehler markiert. Einige schlugen „Da hatte Kommissar Brot“ oder „Da kam Kommissar Brot ein Gedanke“ vor. Der Artikel „ein“ vor Geistesblitz wurde von nahezu allen verbessert. Bis auf eine Rechtschreibfehlermarkierung der Nummer 9 bemängelten hierbei alle die Grammatik. Lediglich Nummer 3, 8 und 15 haben den Fehler übersehen. Die Lehrkräfte 2 und 6 korrigierten den Verstoß ausschließlich direkt im Text.

26. Herr Brot starrte ihm mit großen Augen an und sagte: „Sie waren es.“

Die Lehrperson 1 markierte einen Zeitenfehler für das erste Prädikat im Satz und besserte auf „starrt“ aus.

Die Verwendung des dritten Falles bei dem Pronomen „ihm“ wurde von allen bis auf die Nummern 3 und 13 korrigiert. Während die Lehrpersonen 2 und 15 lediglich im Text verbesserten, ohne ein Korrekturzeichen zu verwenden, war sich der Rest darüber einig, hierbei einen Grammatikfehler vorgefunden zu haben. Nummer 8 markierte darüber hinaus das Verb „sagen“ als Wortwiederholung zu einem vorangegangenen Satz.

27. Der Mann starrte zurück.

Vier Personen der Stichprobe (3, 18, 20, 21) markierten in diesem Fall einen Ausdrucksfehler und notierten detaillierter, dass es sich bei „starrte“ um eine Wortwiederholung handelte. Die Lehrperson Nummer 20 unterwellte dabei sogar den ganzen Satz. Alle anderen hatten nichts zu bemängeln.

28. Alle Augen waren auf die Beiden gerichtet.

Hierbei wurde die Großschreibung von „die Beiden“ kritisiert und 48 Prozent der Lehrpersonen (4, 8, 10, 12, 15, 17, 18, 19, 20, 21) markierten dies als Rechtschreibfehler. Im Gegensatz dazu hielten dies die Personen Nummer 5 und 6 für einen Grammatikfehler. Nummer 9 verbesserte lediglich im Text, ohne ein Korrekturzeichen am Rand zu vermerken.

29. Man sah dem Mann die Furcht an als hätte man es ihm auf die Stirn geschrieben.

Der fehlende Beistrich für die Einleitung des Nebensatzes wurde von drei Lehrpersonen (2, 6, 13) erkannt und direkt im Text ausgebessert. Darüber hinaus wurde er von neun Lehrkräften

(3, 4, 5, 8, 9, 12, 14, 16, 18) als Beistrichfehler markiert. Die Lehrperson Nummer 19 verbesserte einen Beistrich nach dem Substantiv „Furcht“. Außerdem unterwellte diese Person das Pronomen „es“ und korrigierte, wie auch die Lehrkraft Nummer 18, einen Grammatikfehler.

Die Proband:innen Nummer 9 und 12 kreisten den Anfangsbuchstaben des Wortes „auf“ ein und markierten einen Rechtschreibfehler.

Die Lehrpersonen Nummer 20 und 21 unterwellten den gesamten Satz und markierten einen Ausdrucksfehler sowie bei Nummer 20 den Vermerk „Ugs“ für Umgangssprache. Als Korrektur schlug die Person „Dem Mann stand die Angst sprichwörtlich ins Gesicht geschrieben.“ vor.

30. Er musste sich die Tränen unterdrücken und gab anschließend ein Geständnis ab.

Das Pronomen „sich“ wurde von 19 Prozent der Lehrpersonen (2, 4, 11, 12) im Text durchgestrichen, ohne einen Fehler am Rand zu vermerken. Die Proband:innen 6, 9 und 14 markierten diesbezüglich einen Ausdrucksfehler, während Nummer 5, 18, 20 und 21 dies als einen Grammatikfehler empfanden. Die Person Nummer 19 vermerkte einen Ausdrucksfehler für das Verb „unterdrücken“. Lediglich Nummer 11 kritisierte den Ausdruck von „ein Geständnis abgeben“ und korrigierte zu „legte anschließend ein Geständnis ab“. Der Lehrperson Nummer 14 sagte der Ausdruck „schließlich“ nicht zu und daher tauschte sie dies durch „anschließend“ aus. Beide vermerkten ein „A“ am Rand.

31. Es kam heraus, dass er, ihr Freund sei und das nur getan hat, da sie über seine illigalen Geschäfte heraufgefunden hatte und das sein Komplize noch auf freien Füße sei.

Da dieser Satz einerseits eine hohe Länge und einige Fehler aufweist und andererseits die Korrekturen sehr unterschiedlich ausfielen, soll er zunächst aufgeteilt werden, um eine bessere Übersicht zu gewährleisten.

Während für die Lehrperson Nummer 14 der gesamte Satz mit einem „A“ markiert wurde, befanden die Lehrpersonen 5, 7, 18 und 21 den ersten Teil des Satzes „Es kam heraus“ als Schwäche des Ausdrucks. Zwei davon schlugen „Es stellte sich heraus“ vor und notierten dies oberhalb des Satzes.

Der falsch gesetzte Beistrich hinter dem Pronomen „er“ wurde von 48 Prozent der Lehrpersonen (1, 3, 4, 5, 7, 8, 12, 14, 18, 20) gewertet und von weiteren zwei (13, 17) lediglich im Text gestrichen. Somit wurde dieser Fehler fast von der Hälfte der Stichprobe

übersehen oder nicht korrigiert. Darüber hinaus fiel zehn Personen das Wort „sei“ ins Auge. 80 Prozent davon (4, 5, 11, 12, 14, 15, 18, 19) kennzeichneten dies als Grammatik- bzw. detaillierter als Zeitenfehler mit „Z“, während Nummer 8 und 9 das Wort lediglich im Text unterwellten. Die falsche Zeit bei der Phrase „getan hat“ fiel den genannten Proband:innen nicht zur Gänze auf, aber anderen wiederum schon. So markierten 33 Prozent der Stichprobe (6, 7, 11, 12, 15, 17, 18) einen Grammatik- bzw. Zeitenfehler. Die Lehrperson Nummer 5 verbesserte hierbei einen Ausdrucksfehler. Die Lehrperson Nummer 10 vermerkte zwar kein Randkommentar aber schrieb „habe“ im Text über das falsche Wort, während sich die anderen einig waren über die Form „getan hatte“.

An der Phrase „da sie über seine illigalen Geschäfte“ bekrittelten fünf Personen (7, 8, 17, 20, 21) den Ausdruck und besserten auf „von seinen illegalen Geschäften erfahren hatte/Bescheid wusste“ aus. Der Schreibfehler im Wort „illigalen“ fiel fünf Personen nicht auf, der Rest markierte einheitlich mit einem Rechtschreibfehler bzw. Nummer 8 als Ausnahme mit einem Ausdrucksfehler.

Den „dass“-Fehler bemerkten lediglich 14 Proband:innen von 21. Und selbst von dieser Anzahl wiesen ihn nur neun Lehrpersonen als Fehler aus. Dabei wurde unterschieden in Grammatik- und Rechtschreibfehler. Die restlichen Lehrkräfte verbesserten den Verstoß lediglich im Text. Die Person 19 markierte einen Wiederholungsfehler bezüglich der Rechtschreibung.

Bei dem Substantiv „Komplieze“ war der Wiederholungsfehler deutlicher, denn sechs Personen kennzeichneten hierbei einen Rechtschreibfehler und in Klammer einen Vermerk mit „WH“ oder „FF“. Lediglich drei Lehrkräfte (3, 4, 10) werteten hierbei abermals einen Fehler wie zu Beginn des Schularbeitstextes.

Die letzte Phrase „noch auf freien Fuße sei“ wurde von einigen Lehrpersonen hinsichtlich des fehlerhaften Falles kritisiert, als auch aufgrund der falschen Zeit. So besserten fünf Personen (1, 8, 10, 15, 18) das Adjektiv „freien“ zu „freiem Fuße“ aus und markierten einen Grammatikfehler, außer Lehrperson Nummer 10. Diese entschied sich für einen Rechtschreibfehler. Lehrperson Nummer 6 korrigierte dies direkt im Text, aber vermerkte kein Korrekturzeichen. Die Lehrkräfte 4, 5, 11 und 18 waren darüber hinaus unzufrieden mit der Zeit bezüglich des Wortes „sei“ und markierten dies mit einem Grammatikfehler.

32. Doch er wollte noch wissen, wie er darauf gekommen ist, dass er der Mörder sei?

Die Lehrpersonen 5 und 14 schlugen andere Satzanfänge wie „Außerdem wollte er noch wissen“ und „Danach wollte er noch wissen“ vor und wiesen dies auch mit einem

Ausdrucksfehler am Rand aus. Die Lehrpersonen 20 und 21 unterwellten dahingehend den kompletten Satz, setzten ein Fragezeichen dazu und werteten einen Ausdrucksfehler.

Im zweiten Teil des Satzes wurde von 38 Prozent der Stichprobe (4, 5, 9, 11, 12, 15, 18, 19) ein Zeitenfehler gekennzeichnet und darüber hinaus von drei Lehrkräften lediglich im Text angestrichen. Während Nummer 4 beispielsweise der Meinung war, die Phrase „wie er darauf kam“ auszubessern, gab es auch die Meinungen für „gekommen sei“ und „gekommen war“.

Das letzte Wort „sei“ wurde von 33 Prozent (2, 3, 4, 5, 11, 14, 18) angestrichen, wobei nur Nummer 3, 4, 5 und 14 auch einen Grammatikfehler dafür markierten. Lehrperson Nummer 18 unterstrich das letzte „dass“ und kennzeichnete einen Ausdrucksfehler mit detaillierterer Beschreibung in Klammer „(WW)“ für Wortwiederholung.

33. Kommissar Brot meinte: „Woher wissen sie, dass 2 Schüsse unnötig waren?“

Der erste Teil der direkten Rede wurde von 29 Prozent der Proband:innen (1, 5, 8, 10, 11, 19) verbessert, wobei hierbei unterschiedliche Fehler und Kategorien wahrgenommen wurden. Während vier Personen einen Rechtschreibfehler für das kleingeschriebene „sie“ vermerkten, notierte Lehrperson Nummer 5 ein „A/G“. Hierbei konnte sie sich offensichtlich nicht zwischen Ausdruck- oder Grammatikfehler entscheiden. Die Lehrkraft Nummer 19 übersah zwar diesen Fehler, besserte aber „wissen“ zu „wussten“ aus und markierte ein „Z“ in Klammer, um so einen Zeitenfehler zu verdeutlichen.

Bezüglich der nicht ausgeschriebenen Zahl „2“ markierte die Lehrperson 3 einen Rechtschreibfehler und die Personen 12, 13 und 21 einen Ausdrucksfehler. Weitere 43 Prozent beziehungsweise 9 Proband:innen besserten lediglich im Text aus und die Lehrperson Nummer 9 kennzeichnete diesbezüglich einen Wiederholungsfehler des Ausdrucks.

Um Unterschiede in den Korrekturen auch bildlich mit den Original-Verbesserungen darzulegen, wird auf der nächsten Seite ein Screenshot aus der Arbeit mit den wenigsten markierten Korrekturen mit der Note „Gut“ von der Lehrperson Nummer 1 abgebildet. Im Vergleich dazu wird außerdem die am schlechtesten bewertete Schularbeit von der Lehrkraft Nummer 21 aufgezeigt.

Blutige Herbstferien

R

Die letzten weichen Strahlen der Herbstsonne versuchten das blasser Gesicht des Mädchens, das unter dem Fenster lag, zu erwärmen, doch sie blieb kühl. Der Täter wusste, dass er geschnappt werden würde, ~~falls~~ wenn er nicht die Spuren beseitigen ~~würde~~ und weglaufen würde. Ein Handy klingelte. Der Täter ging sofort ran. Es stellte sich heraus, dass dies sein Komplize sei, der wissen wollte, ob die Tat vollbracht war, doch er schwieg, legte auf, und ging davon.

R, R

Es war ~~ein~~ ^REin typischer Samstag. Kommissar Brot verspeiste seinen Apfel und dachte sich nichts dabei, als auf einmal sein Handy ertönte. „Kommissar Brot zu Ihren Diensten! Wie kann ich heute behilflich sein?“, sagte er mit einem höflichem ⁿUmgangston. Die nervöse Frau am Handy schrie, dass er doch schnell zur Algostraße 12 kommen sollte da es ~~einen~~ Notfall gab. Kommissar Brot versuchte die ~~Frau~~ Frau zu beruhigen, doch es half nix. Daraufhin hoppte er in seinen Wagen und machte sich auf den schnellsten ~~Weg~~ Weg zur Algostraße 12. Er sah 2 junge Damen. Eine kullerten die Tränen herunter, wie ein Wasserfall. Die andere war jedoch voll mit ~~Blut~~ ^{und} Blut und Glassplittern. ^{und} Als er näher kam, sah er dass das junge Mädchen 2 Schüsse in den Kopf und einen in die Brust bekommen hatte.

R, R

R

G

G

R

G

B

R

Der Mann trat näher und befragte die Dame. Es
 kam heraus, dass sie Anna hieß und das anscheinend
 tote Mädchen Marie. Marie war 17 Jahre alt und
 ging noch zur Schule. Eine Sirene ertönt „Das wird
 wohl die Polizei sein“, sagte Anna mit Tränen in
 ihren Augen. ~~Kommissar~~ Brot erzählte den Polizeisten
 über das Verbrechen. Dabei fiel ihm ein mitte
 60-jähriger Mann auf, der das Mädchen leblos
 anstarrt. Es kam ihm so vor, ~~als ob~~ dass der Mann seinen
 Blick nicht von ihr abwenden könne. Dies gab
 Kommissar Brot ein merkwürdiges Gefühl. Der
 Mann seufzte und sagte mit einer leisen und ruhigen
 Stimme: „2 unnötige Schüsse.“ Da bekam Kommissar
 Brot einen Geistesblitz. Herr Brot starrte ihn mit
 großen Augen an und sagte: „Sie waren es.“ Der
 Mann starrte zurück. Alle Augen waren auf die
 Beiden ~~gerichtet~~ gerichtet. Man sah dem Mann die
 Furcht an als hätte man es ihm auf die Stirn ge-
 schrieben. Er musste sich die Tränen unterdrücken
 und gab anschließend ein Geständnis ab. Es kam
 heraus, dass er ihr Freund sei und das nur getan
 hat, da sie ~~etw.~~ über seine illegalen Geschäfte
 heraufgefunden hatte und dass sein Komplize
 noch auf freiem Fuße sei. Doch er wollte noch
 wissen, wie er darauf gekommen ist, dass er der
 Mörder sei? Kommissar Brot meinte: „Woher wissen
 Sie, dass 2 Schüsse unnötig ~~es~~ waren?“

Abbildung 14: Schularbeit Note "Gut"

^Bblutige Herbstferien

R

Die letzten weichen Strahlen der Herbstsonne versuchten das blasser Gesicht des Mädchens, das unter dem Fenster lag, zu erwärmen, doch sie blieb kühl. Der Täter wusste, dass er geschnappt werden würde, ~~falls~~ wenn er nicht die Spuren beseitigen ~~würde~~ und weglaufer würde.
^{SEID}Ein Handy klingelte. Der Täter ging sofort ran. Es stellte sich heraus, dass dies sein Komplize sei, der ^WWissen wollte, ob die Tat vollbracht war. ^DDoch er schwieg, legte auf, und ging davon.

G

A

IA

RR

R

A

Es war ~~ein~~ ^EEin typischer Samstag. Kommissar Brot verspeiste ^Sseinen Apfel und dachte sich nichts dabei, als ^{plötzlich}auf einmal sein Handy ertönte. „Kommissar Brot zu Ihren Diensten! ^WWie kann ich heute behilflich sein?“, ^{melodisch}sagte er ^{mit}mit einem höflichem ^hUmgangston. Die nervöse Frau am Handy schrie, dass er doch schnell zur Algostraße 12 kommen sollte, da es ein Notfall gab. Kommissar Brot versuchte die ~~Frau~~ Frau zu beruhigen, doch es half ^{nichts}nix. Daraufhin hoppte er in seinen Wagen und machte sich auf den schnellsten ^Wweg zur Algostraße 12. Er sah ^{zwei}zwei junge Damen. Einer kullerten die Tränen herunter, wie ein Wasserfall. Die andere ~~war~~ war jedoch voll mit ~~Blut~~ ^BBlut und Glassplittern. Als er näher kam, sah er, dass das junge Mädchen ^{zwei}zwei Schüsse in den Kopf und einen in die Brust bekommen hatte.

RR

R

AA

R

A

R

BA

RA

R

AR

R

ABA

RBG

A

Der [?]Mann trat näher und befragte die Dame. Es
 kam heraus, dass sie Anna hieß und das anscheinend ^{R?}
 tote Mädchen Marie. Marie war 17 Jahre alt und ^{Warum anscheinend?}
 ging noch zur Schule. Eine Sirene ertönte. Das wird
 wohl die Polizei sein. sagte Anna mit Tränen in
 ihren Augen. Kommissar Brot ^{berichtete} erzählte den Polizisten ^{2 A}
^{von dem} über das Verbrechen. Dabei fiel ihm ein mitte ^{1 A}
 60 Jähriger Mann auf, der das Mädchen leblos
 anstarrt. Es kam ihm ^m so vor, ^{gls gl} dass der Mann seinen ^{GA}
 Blick nicht von ihr abwenden konnte. Dies gab
 Kommissar Brot ein merkwürdiges Gefühl. Der ^A
 Mann seufzte und sagte mit ~~einer~~ leiser und ruhiger ^{RA}
 Stimme. „^{zwei} unnötige Schüsse.“ Da bekam Kommissar ^R
 Brot einen Geistesblitz. Herr Brot starrte ihm ^{GA}
 großen Augen an und sagte: „Sie waren es.“ Der ^{WH/A}
 Mann starrte zurück. Alle Augen waren auf die
 Beiden ~~gerichtet~~ gerichtet. Man sah dem Mann die ^R
 Furcht an als hätte man es ihm auf die Stirn ge-
 schrieben. Er musste sich die Tränen unterdrücken ^A
 und gab anschließend ein Geständnis ab. Es kam ^A
 heraus, dass [?]er, ihr Freund sei und das nur getan ^{2 A}
 hatte da sie ~~erf~~ ^{von} über seinen illegalen Geschäften ^{AGRG}
^{erfahren} herausgefunden hatte und das sein Komplize ^{RA}
 noch auf freien Füße sei. Doch er wollte noch ^{2 A 1}
 wissen, wie er darauf gekommen ist, dass er der
 Mörder sei? Kommissar Brot meinte: „Woher wissen
 sie, dass ^{zwei} Schüsse unnötig ~~se~~ waren?“ ^A

Nicht Genügend

Keine Gliederung, schwere Fehler in Bereich
 Rechtschreibung und Grammatik. Sprachrichtigkeit!
 Ausdruck lässt auch zu wünschen übrig.

Abbildung 15: Schularbeit Note "Nicht genügend"

4.2.2 Beurteilungen

In einem ersten Schritt wurde im Fragebogen abgefragt, wie die gegebene Note auf den Schularbeitstext begründet wird. Die Antworten dabei fielen unterschiedlich aus und sollen im Folgenden zusammengefasst werden.

Während die drei Lehrpersonen mit der Nummer 1, 2 und 3 ihr „Gut“ mit lediglich positiven Aspekten begründeten, lässt sich im Mittelbereich der Noten klar erkennen, dass Verbesserungsvorschläge und ein wenig Kritik angebracht wurden. Die Lehrperson, welche ein „Nicht Genügend“ vergeben hat, fand keine positiven Worte in der Begründung der Notengebung. So beantwortete beispielsweise Lehrperson Nummer 2 die Frage mit „bemüht, gutes Sprachgefühl, ausdrucksstarke Wortwahl, bildhafte Ausdrucksweise“ (LP 2) und im Gegensatz dazu die Lehrperson Nummer 21 mit „keine Gliederung, sehr viele schwere Fehler im Bereich der Sprachrichtigkeit, unlogischer Schluss, Ausdruck lässt auch zu wünschen übrig“ (LP 21).

Die Mehrheit vermerkte die vielen Fehler im Bereich der Sprachrichtigkeit und meinte, dass diese eine bessere Note verhindert hätten.

Bezüglich des Ausdrucks hielten sich die Ansichten der Lehrenden die Waage. Einige waren der Meinung, dass der Text „eine schöne Sprache“ (LP 3, 4, 6) beinhaltete, die anderen sprachen sich dagegen aus. Der teils umgangssprachliche Ausdruck war lediglich für die Lehrpersonen 14 und 20 nicht passend.

Der Inhalt wurde von nahezu allen gelobt, nur die Lehrenden 12 und 21 kritisierten einen unlogischen Schluss. Proband:in Nummer 20 merkte an, dass der Höhepunkt der Geschichte und „das gewisse Etwas“ (LP 20) fehlte. Nummer 15 war darüber hinaus der Meinung, dass die Einleitung nicht zum Text passen würde. Besonders die Lehrpersonen, welche ein „Genügend“ gegeben hatten, wiesen darauf hin, dass der gelungene Inhalt ein „Nicht Genügend“ verhindert hätte.

Lehrperson Nummer 15 und 19 lobten auch den guten Umfang beziehungsweise die Länge des Textes.

Die Lehrperson Nummer 10 betonte darüber hinaus, dass sich der oder die Schüler:in trotz der Fehler Mühe gegeben zu haben schien.

Die Lehrpersonen 13 und 19 beschrieben das unterschiedliche Vorgehen an einer NMS und dass dies einen Grund für Verunsicherung bei der Notengebung darstellen würde. Beispielsweise meint erstere, dass an ihrer Schule mehrteilige Schularbeiten zusammengestellt werden, da das Verfassen von Texten eine der größten Schwachstellen der

Schüler:innen darstellt. Aufgrund der Einschätzung eines inhaltlichen guten Produktes schnitt der Text nach seiner/ihrer Meinung auch besser ab.

Eine weitere Frage im Fragebogen lautete: „Wie gelangen Sie im Allgemeinen zu Ihrer Beurteilung bei Deutschtexten?“. Die Ergebnisse sollen im Folgenden angeführt werden.

Beinahe die Hälfte der Lehrpersonen (3, 4, 5, 7, 9, 12, 14, 18, 19) gab im Fragebogen an, mithilfe eines Beurteilungsrasters zu arbeiten. Auch die restlichen Lehrkräfte hatten konkrete Vorstellungen einer Beurteilung. Hierbei waren sich die meisten einig, Inhalt, Ausdruck, Struktur und Sprachrichtigkeit zu differenzieren.

Darüber hinaus achteten die Lehrpersonen Nummer 2 und 19 auch auf den Umfang des Textes. Während der Auswertung des Fragebogens und der qualitativen Analyse wird aber deutlich, dass auch weitere Lehrpersonen (z.B. 14, 15) auf die Länge des Textes Wert legten. Die Lehrkraft Nummer 1 merkte an, dass sie besonderes Augenmerk darauflegt, ob das, was im Unterricht gelernt wurde, auch angewendet wird. Dieser Aspekt ist im Hinblick auf vorliegende Forschung für die genannte Person nicht überprüfbar.

Der oder die Proband:in Nummer 3 gab an, zwischen Hausübungen und Schularbeiten zu unterscheiden. Wie bereits oben erwähnt, wird zwar ein Raster für Letztere verwendet, aber bei Hausübungen zählt der Gesamteindruck.

Von der Lehrperson Nummer 7 wird bei der Frage nach der Beurteilung zusätzlich der Austausch mit Kolleg:innen genannt. Außerdem wiegt bei ihr der Inhalt am meisten. Sobald dieser Teilbereich negativ ist, kann die Gesamtnote nicht mehr positiv werden. Dieser Meinung war auch die Lehrkraft Nummer 13.

Die Gewährsperson Nummer 8 beschrieb die unterschiedlichen Bereiche, nach welchen sie beurteilt, mit Rechtschreibung, Grammatik, Zeichensetzung, direkte Rede, Ausdruck, Aufbau und Spannung.

Die Lehrperson Nummer 11 gab an, in einer ersten Runde die Fehler zu korrigieren und Noten zu verteilen. In einem zweiten Schritt betrachtet sie genauer den Aufbau und den Inhalt und vergleicht dann mit anderen Schularbeiten aus derselben Klasse: „Jede Notenebene wird miteinander verglichen. Deshalb variiert es auch von Klasse zu Klasse. Dieses Befriedigend wäre in einer anderen Klasse ein Genügend.“ (LP 11)

Eigenen Angaben zufolge händigt Nummer 14 den Schüler:innen vor der Schularbeit einen Raster aus, auf welchem die gewünschten und geforderten Merkmale der Textsorte vermerkt sind. Die Rechtschreibfehler werden hierbei aber mit einer Formel auf die Wortzahl des Textes aufgerechnet.

Ebenso gab Lehrperson 15 an, die Anforderungen an die Textsorte im Vorherein zu kommunizieren und anhand der Erfüllung dieser und der Kompetenzvorgaben zu beurteilen. Lehrkraft Nummer 19 beschrieb die Vorgehensweise bei der Beurteilung sehr detailliert mit dem erstmaligen Lesen der Texte und mithilfe eines eigens erstellten Rasters. Dabei werden Punkte hinsichtlich Länge und Ausdruck und mehreren nicht weiter ausformulierten Kriterien, welche erfüllt sein müssen, vergeben. Von dieser Gesamtpunkteanzahl werden dann die Punkte der Fehler abgezogen.

Der Auskunft im Fragebogen zufolge war der Gewährsperson Nummer 20 sehr wichtig, dass Rechtschreib- und Grammatikregeln beherrscht werden und daher wird verstärkt auf diese Komponenten geachtet. Darüber hinaus merkte sie auch an, dass der strukturierte Aufbau eines Textes gegeben sein sollte, da dies die Verständlichkeit des Geschriebenen erleichtert.

Nummer 21 beschreibt die Vorgangsweise bei Beurteilungen durch das Verwenden eines Punkteschlüssels und das Benoten durch langjährige Erfahrung.

Im nächsten Kapitel *Interpretation und Diskussion* werden die Ergebnisse und die daraus resultierenden, wichtigsten Erkenntnisse der quantitativen als auch der qualitativen Untersuchung interpretiert. Darüber hinaus werden das Korrekturverhalten und Verbesserungsvorschläge zu diesem diskutiert.

5. Interpretation und Diskussion der Ergebnisse

An den Ergebnissen der quantitativen Analyse lässt sich erkennen, dass die Streuung der Fehleranzahl in beide Richtungen vom Mittelwert ausgehend, also von 38,1 gesamten Fehlern, sehr breit war. Während die geringste Gesamtfehleranzahl bei 24 lag, wies die Schularbeit mit den meisten Fehlern 72 gekennzeichnete Verstöße auf. Dies spiegelte sich auch in den Benotungen wider, da von einem „Gut“ bis zu einem „Nicht genügend“ jede Notenstufe vertreten war. Diese Ergebnisse decken sich mit den im Kapitel 2.4 erwähnten Untersuchungen von Birkel (2003), welcher beobachten konnte, dass das Beurteilungsspektrum, vor allem im mittleren Leistungsbereich, sehr groß ist. Die Untersuchung gab Hinweise darauf, dass die markierten Fehler mit den vergebenen Noten korrelieren, jedoch lediglich an den äußeren Enden der Notenskala. Beispielsweise wurden zwei Schularbeiten, einmal mit 24 und das andere Mal mit 27 gekennzeichneten Fehler mit einem „Gut“ beurteilt, welches die beste Note in den Untersuchungen darstellte. Der Text mit den meisten Kennzeichnungen mit 72 Fehlern wurde mit einem knappen Genügend („4-“) benotet. Die mit einem „Nicht genügend“ bewertete Schularbeit weist die zweitmeisten Fehler mit 63 markierten Stellen auf. Im Mittelbereich mit der Note „Befriedigend“ reichten die markierten Stellen von 27 bis hin zu 53 Gesamtfehlern mit einem Mittelwert von 40,2 Verstößen. Bei den Schularbeiten mit der Note „Genügend“ konnte eine Spanne von mindestens 24 bis hin zu maximal 72 Fehlern ausgemacht werden. Der Mittelwert der Korrekturen beträgt hierbei 40,3, also nur ein minimaler Unterschied, weswegen diesbezüglich keine gesicherten Aussagen über den Zusammenhang zwischen den markierten Fehlern und der ausgewiesenen Note gemacht werden können. Werden jedoch die Mittelwerte von allen Notensprüngen betrachtet, lässt sich ein Ansatz bezüglich eines Zusammenhanges zwischen Notenvergabe und Fehlermarkierungen erkennen. Auch im Hinblick auf die normative Sprachrichtigkeit, also Grammatik-, Rechtschreib- und Zeichensetzungsfehler konnten deutliche Sprünge in der Korrektur ausgemacht werden. Die wenigsten Fehler korrigierte Lehrperson Nummer 6 mit 14 Markierungen, während die Lehrkraft Nummer 18 insgesamt 53 Fehler im Bereich der Sprachrichtigkeit ausmachen konnte. An diesen Untersuchungen lässt sich bereits erkennen, dass Texte oftmals äußerst unterschiedlich korrigiert und benotet werden und dies auch im Bereich der normativen Sprachrichtigkeit geschieht, welcher eigentlich eine höhere Objektivität als den anderen Bereichen, wie beispielsweise dem Ausdruck oder dem Inhalt, beigemessen und zugeschrieben wird.

Diese Ansicht konnte durch die qualitative Analyse gestützt werden, da auch hier Unterschiede festgestellt wurden. Die mehrheitlich beobachteten Erkenntnisse zeigten, dass, auch wenn wenige Fehler in den Bereichen des Ausdrucks und des Inhalts markiert wurden und anhand der gängigen Meinungen der Lehrpersonen im Fragebogen relativ gut ausfielen, bei mehreren normativen Regelverstößen die Lehrkräfte „strenger“ agierten als bei Arbeiten mit weniger Fehlern bezüglich der Sprachrichtigkeit. Diese Beobachtungen decken sich mit den Annahmen von Becker-Mrotzek et al. (2006: 87), laut welchen die orthografische Korrektheit stärker gewichtet wird als die gesamte Textqualität. Über die möglichen Gründe für dieses beobachtete Phänomen kann nur spekuliert werden. Eine Hypothese ist, dass früher der Stellenwert von Rechtschreibung und Grammatik höhergestellt war und Fehler in diesen Kategorien vermehrt für die Notengebung herangezogen wurden (vgl. Kleppin 2010: 1064). Dabei musste bereits ab einer bestimmten Anzahl mit einem Fünfer beziehungsweise einem schlechteren Notengrad gerechnet werden. Die Lehrperson Nummer 20 gab im Fragebogen sogar an, dass Sprachrichtigkeit den höchsten Stellenwert bei der Beurteilung hätte. In kompetenzorientierten Zeiten wie heute gibt das Bundesministerium aber vor, darauf zu achten „was die Schülerin/der Schüler kann, nicht darauf, was sie/er nicht kann.“ (vgl. Bundesministerium Bildung, Wissenschaft und Forschung - Handreichung zum Beurteilungsraster 2017). Hierbei wäre es interessant zu beobachten, ob das Alter der beurteilenden Lehrer:innen eine Rolle spielt und ob sie diese Reform miterlebt haben. Dies könnte ein Grund dafür sein, dass eventuell noch das alte Bewusstsein für das Korrekturverhalten verinnerlicht ist.

Diese Erkenntnisse überraschen zunächst, da sowohl in der Theorie beschrieben (vgl. Decker et al. 2016: 8) als auch durch die Ergebnisse vorliegender Masterarbeit ein vermehrter Einsatz von Kriterienkatalogen erkennbar ist und dies vermuten lassen würde, dass alle Aspekte der Textqualität bei der Beurteilung beachtet werden.

Kleppin (2010: 1064) spricht sich trotz Verwendung von Rastern für Fehlermarkierungen im Text aus (siehe Kapitel 2.3.1). In der qualitativen Analyse konnte beobachtet werden, dass dies von nahezu jeder Lehrperson umgesetzt wurde. Lediglich die Lehrkraft Nummer 2 korrigierte anhand der „einfachen Fehlermarkierung“ (Kleppin 2010: 1064). Die restlichen Proband:innen verbesserten mit der „metalinguistischen Korrektur“ (Ellis 2009: 100), wobei auch hierbei, wie von Ellis zitiert wurde, Ungereimtheiten im Text auffielen, welche für die Schüler:innen Schwierigkeiten bereiten könnten. Aufgrund dieser Untersuchung liegt die Vermutung nahe, dass selbst trotz Zuhilfenahme von Kriterienkatalogen große Unklarheiten bezüglich des Korrigierens seitens der Lehrpersonen herrschen, was sich mit den Annahmen

von Davies (2017: 138) in den theoretischen Ausführungen deckt. Wie bereits in der Theorie erwähnt wurde, ist ein Text bis zu einem gewissen Grad individuell und somit nicht durch ein Raster objektiv beurteilbar (vgl. Baurmann 2006: 120), was sich auch in den Ergebnissen vorliegender Untersuchung zeigt. Die Lehrpersonen, welcheangaben, mithilfe eines Rasters zu arbeiten, gelangten nämlich nicht zu denselben Leistungsbeurteilungen. Diese Erkenntnis scheint sich auch mit der Annahme von Mayr und Schubert (1985:13) zu decken, laut denen die Objektivität der Beurteilung durch Kriterienkataloge nicht verbessert werden kann. Dennoch muss hinzugefügt werden, dass nicht ersichtlich ist, ob es ohne Raster vielleicht nicht noch höhere Sprünge in der Leistungsbeurteilung gegeben hätte, wie auch von Becker-Mrotzek et al. (2016: 210) angenommen wird. Hierbei könnte als erster Schritt angedacht werden, einen standardisierten und einheitlichen Raster für alle Schulen und Lehrenden zu etablieren. Dies würde vermutlich sowohl eine höhere Objektivität als auch Reliabilität gewährleisten.

In der qualitativen Analyse konnte die Aussage von Hennig (2012:134) bestätigt werden, dass in der Schulpraxis an den gängigen Korrekturzeichen wie „G“ für Grammatik, „R“ für Rechtschreibung und „Z“ für Zeichensetzung festgehalten wird. Nahezu jede Lehrperson orientierte sich an diesen tradierten Fehlermarkierungen. Darüber hinaus wurden Beistrichfehler meistens mit einem „B“ markiert und Ausdrucksfehler mit einem „A“. Einige wenige kennzeichneten zusätzlich inhaltliche Fehler mit einem „I“. Eher ungewöhnliche und selten verwendete Markierungen waren „x“ und „?“.

Zu beobachten war, dass eine Abweichung von diesen verwendeten Zeichen unübersichtlich wirkte und zu einem Unverständnis bei der Autorin vorliegender Masterarbeit führte. Dadurch lässt sich erahnen, dass dies ebenso ein Problem für Schüler:innen, Eltern und andere Lehrer:innen darstellen könnte. Obwohl Hennig (2012: 134) die tradierten Korrekturzeichen aufgrund fehlender Begründung kritisiert, könnte mit einer Standardisierung der Markierungen zumindest der Unübersichtlichkeit entgegengewirkt werden, da in vorliegender Untersuchung die Erfahrung gemacht wurde, dass klassische Fehlermarkierungen, wie beispielsweise „A“ für Ausdruck, „R“ für Rechtschreibung, „G“ für Grammatik und „B“ für Beistrich am ehesten und leichtesten eingeordnet werden konnten. Dies würde nicht nur der Vergleichbarkeit zwischen den Korrekturen zugutekommen, sondern auch den Schüler:innen und Eltern zu einer erhöhten Nachvollziehbarkeit verhelfen. Damit könnte wiederum der Vorwurf von Kremser (2016: 131) bezüglich Intransparenz entkräftet werden, vorausgesetzt die Korrekturen korrelieren mit der gegebenen Note, wie es in den Ergebnissen vorliegender

Untersuchung größtenteils zu beobachten war. Eine solche Vorgabe würde auch vor allem bei einem Lehrer:innen- oder Schulwechsel Sinn machen.

Aufgrund der unterschiedlichen Korrekturen kann angenommen werden, dass Lehrkräfte nach ihren eigenen Sprachvorstellungen korrigieren, wie bereits im Kapitel 2.1.4 erwähnt wurde. Sprachliche Zweifelsfälle scheinen bei den meisten Lehrpersonen nicht angestrichen worden zu sein, manche andere markierten diese jedoch mit Unterwellungen und Fragezeichen. Die Wörter „rangehen“, „hoppte“ und „nix“ wurden von den meisten als Ausdrucksfehler markiert. Eine einzige Lehrperson vermerkte, dass es sich dabei um umgangssprachliche Begriffe handelt. Das Phänomen „hoppen“ ist im Duden nicht vorzufinden. Hierbei handelt es sich laut Kodex nicht um ein umgangssprachliches Wort. Die anderen beiden Wörter stehen im Duden und sind als Umgangssprache ausgewiesen. Bei Betrachtung der theoretischen Grundlagen bezüglich der Dynamik von Sprache (siehe Kapitel 2.1) lässt sich hierbei diskutieren, ob und wie in einem solchen Fall korrigiert werden sollte.

Tatsache ist aber, dass vor allem im Sprachenunterricht eine Standardisierung von Korrekturen und Bewertungen nur teilweise Sinn macht, denn Sprache kann niemals in richtig und falsch, schwarz und weiß gegliedert werden. Hierbei bewegt man sich als Lehrperson oftmals in einer Grauzone, wie bereits in der Theorie ausführlicher erläutert wurde (vgl. Ammon 2005: 36). Sprache ist ein dynamisches System und deshalb stetig im Wandel (vgl. Wiesinger 1988: 27), was auch bei Bewertungen berücksichtigt werden sollte.

Darüber hinaus gab es auch Phänomene, bei welchen ein beziehungsweise der letzte Buchstabe vergessen wurde, wie beispielsweise bei „Augen“. Einige Lehrer:innen schienen dies als Performanzfehler (siehe Kapitel 2.1.4) zu werten, da dieser nicht korrigiert wurde. Dabei kann spekuliert werden, ob die Lehrpersonen somit erkannt haben, dass normalerweise das Regelbewusstsein über den Inhalt vorhanden sein müsste (vgl. Nussbaum 1991: 76-77). Eine andere Möglichkeit könnte aber auch sein, dass der Fehler übersehen wurde.

Ein weiteres Beispiel stellt das Adverb „anscheinend“ dar, welches beim ersten Mal falsch als „antscheinend“ und danach von dem Schüler oder der Schülerin korrekt geschrieben wurde. Mögliche Erklärungen dafür könnten sein, dass der Fehler entweder nicht bemerkt oder nicht korrigiert wurde, da das Bewusstsein für die richtige Schreibung bei dem Schüler oder der Schülerin angenommen wurde.

Interessant war auch der Fall im Schüler:innentext, dass „2“ nicht ausgeschrieben wurde. Lehrer:innen, welche eine gute Note vergaben und auch weniger im Text korrigierten, markierten zwar die Stelle aber wiesen keinen Fehler aus, während bei Lehrkräften, die

strenger beurteilten, eine Tendenz zur Fehlermarkierung zu erkennen war. Hierbei wird wieder das subjektive Empfinden deutlich, was als Fehler gilt und gewertet wird.

Teilweise wurden auch grammatikalisch richtige Sätze ausgebessert, da sie dem Anspruch eines guten Ausdrucks nicht genügten. Der Satz „Der Mann seufzte und sagte mit einer leisen und ruhigen Stimme:“ wurde beispielsweise von der Mehrheit nicht korrigiert. In den unteren Notenstufen wurde hierbei jedoch der Ausdruck bemängelt, obwohl der Satz eigentlich korrekt ist und für diese Schulstufe auch angemessen erscheint. Wenn Fehler gesucht werden, werden auch welche gefunden beziehungsweise macht es auch oftmals den Anschein, als werden diese sogar „kreiert“. Dazu konnte in den Ergebnissen bereits ein Auszug von der am besten und der am schlechtesten bewerteten Schularbeit eingesehen werden. Die Unterschiede in den Korrekturen zeigen diese Annahmen deutlich auf.

Ein und derselbe Fehler wurde oftmals sehr verschieden korrigiert. Beispielsweise wurde in der fünfzehnten extrahierten Phrase der fehlerhafte Satzbau sowohl als Inhalts-, Ausdrucks-, Grammatik- und Rechtschreibfehler gewertet. Dies ist nur eines von vielen Beispielen, welche den subjektiven Charakter bezüglich Korrekturen aufweist. Obwohl das Rechtschreib- und Grammatikverständnis der Lehrpersonen am sichersten wirkte, fiel das Einordnen in die Kategorien dennoch schwer. Dahingehend wurden des Öfteren Unstimmigkeiten beobachtet. Während beispielsweise Groß- und Kleinschreibung für viele zur Rechtschreibung gezählt wird, markierten auch einige einen Grammatikfehler. Über Fehler hinsichtlich der Zeitform waren sich die Lehrkräfte meistens einig, diesen entweder gesondert auszuweisen oder als Grammatikfehler zu zählen.

Im letzten Kapitel werden schlussendlich alle Ergebnisse und die wichtigsten Erkenntnisse der Untersuchungen vorliegender Masterarbeit zusammengefasst, Schlussfolgerungen für die Praxis gezogen und ein Ausblick auf mögliche weiterführende oder vertiefende Forschungen geliefert.

6. Zusammenfassung und Ausblick

Es gibt Hinweise darauf, dass die Beurteilungen von Deutschtexten in der Praxis nicht gänzlich frei von Subjektivität erfolgen (siehe etwa Baurmann 2006). Da die bestehende Literatur sehr gering ausfällt und wenn, den Blick lediglich auf den Bereich der Grammatik wirft, stellt sich in vorliegender Arbeit die Frage, ob diese Annahmen auch auf alle Bereiche eines Textes zutreffen beziehungsweise welche Unterschiede es generell im Korrekturverhalten von Lehrenden für ein und dieselbe Schularbeit gibt.

Von Interesse in vorliegender Arbeit war es demnach, das Korrekturverhalten und die Leistungsbeurteilungen von AHS und NMS Lehrenden anhand eines *mixed method designs*, also sowohl einer quantitativen als auch qualitativen Analyse, zu untersuchen. Deshalb sollen im Folgenden die eingangs erwähnte Forschungsfrage anhand der gewonnenen Ergebnisse beantwortet und die Hypothesen überprüft werden.

6.1 Zentrale Ergebnisse

Forschungsfrage:

Gibt es Unterschiede im Korrekturverhalten und in der Leistungsbeurteilung verschiedener Lehrer:innen für Schularbeiten im Fach Deutsch sowie bei Lehrenden zweier verschiedener Schultypen?

Um diese Fragestellung beantworten zu können war einerseits die quantitative und andererseits auch die qualitative Untersuchung der Schularbeit und des Fragebogens zielführend. Sowohl im Korrekturverhalten als auch in den Leistungsbeurteilungen gab es Differenzen. Keine einzige der verbesserten Schularbeiten wies eine idente Korrektur mit einer anderen auf. Daran lässt sich bereits der deutliche Unterschied bei Korrekturen im Fach Deutsch erkennen. Während von einigen Lehrer:innen Fehler erkannt wurden, markierten andere diese nicht. Selbst wenn sich die Lehrenden bezüglich eines Verstoßes einig waren, wurde dieser auf unterschiedliche Art und Weise gekennzeichnet. Von einem Rechtschreib- über Grammatikfehler bis hin zu einem Ausdruck- oder Inhaltsverstoß kam in einigen Phrasen alles vor. Am sichersten schienen sich die Lehrkräfte bezüglich Rechtschreib- und Grammatikfehler zu sein. Dabei wurden die meisten Übereinstimmungen erkannt. Dennoch kam es sogar hierbei zu Unstimmigkeiten. Das Unwissen darüber, welche Phänomene zu

einem Grammatikfehler zählen oder welche zur Rechtschreibung scheint bei einigen Lehrkräften trotz Zuhilfenahme eines Kodex nicht gesichert zu sein.

Unterschiede zwischen AHS- und NMS- Lehrenden wurden auch deutlich. Im Allgemeinen kann gesagt werden, dass Lehrpersonen aus Allgemeinbildenden Höheren Schulen mehr korrigierten und auch strenger beurteilten als Lehrende der NMS. Detaillierter können diese Erkenntnisse nachfolgend bei den Ausführungen der zweiten Hypothese nachgelesen werden.

1. Hypothese:

Junglehrende markieren im Durchschnitt mehr Fehler als Lehrer:innen mit mehr Unterrichtserfahrung.

Für die Überprüfung dieser Hypothese wurde in vorliegender Masterarbeit eine eigene Definition und Einteilung für Junglehrer:innen herangezogen. Die Stichprobe wurde somit in Proband:innen mit bis zu fünf Jahren Unterrichtserfahrung und über fünf Jahre Erfahrung geteilt. Anhand dieser Kategorisierung wurde dann das Korrekturverhalten quantitativ verglichen.

Die Markierungen der Fehler weisen deutliche Sprünge auf. Während jede:r Junglehrer:in im Durchschnitt 43,8 Fehler kennzeichnete, markierten erfahrene Lehrende durchschnittlich 32,2 Verstöße. Insofern kann die Hypothese, dass Junglehrende im Durchschnitt mehr Fehler als Lehrer:innen mit mehr Erfahrung markieren, bestätigt werden. Auswirkungen auf die Noten scheint dies jedoch nicht zu haben. Der Mittelwert der Benotungen beläuft sich bei den Junglehrer:innen auf 3,2 und bei den erfahrenen Lehrkräften auf 3,3. Demnach kann hierbei kein gravierender Unterschied bezüglich der Leistungsbeurteilungen ausgemacht werden.

Die vermehrte Fehlermarkierung könnte im Engagement und der anfänglichen Sorgfalt begründet liegen, mit welcher Junglehrende ihren Job antreten. Direkt nach der Universität sind die theoretischen Ausführungen des Berufes noch sehr verinnerlicht. Nichts falsch und möglichst alles genau zu machen, um etwaigen Missverständnissen vorzubeugen, scheint oftmals die Devise zu sein. Mit zunehmender Erfahrung scheinen Lehrpersonen eingespielter bezüglich Korrekturen zu werden und es wird eine gewisse Routine entwickelt. Eventuell werden erfahrene Lehrende bei den Mengen an zu korrigierenden Texten auch „arbeitsfauler“ und kennzeichnen somit nur das Nötigste.

2. Hypothese:

AHS-Lehrende beurteilen strenger als NMS-Lehrende.

Um diese Hypothese überprüfen zu können, wurden einerseits die Benotungen verglichen und andererseits das Korrekturverhalten analysiert, indem die Fehlermarkierungen gezählt wurden. Insgesamt gibt es zwölf Personen, welche an einer Neuen Mittelschule arbeiten und neun Lehrkräfte sind an einer AHS beschäftigt. Der Mittelwert der Noten mit 3,5 zu 3,2 bestätigt die Hypothese, dass AHS-Lehrende im Durchschnitt strenger beurteilen als NMS-Lehrer:innen. Auch bezüglich der Fehlermarkierungen konnte ausgemacht werden, dass AHS-Lehrkräfte im Durchschnitt insgesamt 44,1 Fehler im Text korrigierten. Beim NMS-Lehrpersonal hingegen wurden durchschnittlich nur 34,9 Verstöße beobachtet. Ein Grund dafür könnte darin liegen, dass das Leistungsniveau an den Neuen Mittelschulen geringer ist (siehe Kapitel 2.3.4) und den Lehrer:innen deshalb oftmals Druck gemacht wird, besser zu beurteilen (vgl. Tortik 2016: 85). Wie auch Lehrperson Nummer 11 erwähnte (siehe Kapitel 4.2.2), wird die Benotung an das Niveau angepasst, da der schul- und klasseninterne Bezugsmaßstab für die Beurteilungen herangezogen wird (vgl. Ingenkamp 1985: 179).

Zusammenfassend kann gesagt werden, dass große Unterschiede in der Leistungsbeurteilung und im Korrekturverhalten von AHS- und NMS-Lehrenden festzustellen waren. Mit der vermehrten Verwendung von Rastern hätte angenommen werden können, dass Subjektivität und Intransparenz bei den Benotungen verringert wird, jedoch lässt sich an den Ergebnissen vorliegender Arbeit gut erkennen, dass die Komplexität eines deutschen Textes und die dazu erforderlichen Schreibkompetenzen nicht leicht zu standardisieren sind. Dies ging bereits aus der Theorie hervor (siehe etwa Baurmann 2006: 120). Darüber hinaus lässt sich vermuten, dass gerade im Bereich der Sprachrichtigkeit am ehesten Objektivität geübt und Übereinstimmung gewährleistet werden könnte. Diese Arbeit liefert jedoch Hinweise, dass dem nicht so ist, da bei der qualitativen Analyse deutlich wurde, wie unterschiedlich die Wahrnehmung bezüglich Fehlerkorrekturen ist. Durch diesen neuen Kontext, welcher in vorliegendem Format noch nicht untersucht wurde, sind die Ergebnisse auch einigermaßen generalisierbar.

Trotz pädagogischer Beurteilungsspielräume und zumindest ein wenig eingefärbten subjektiven Empfindens, welches auch keiner Lehrperson zugeschrieben werden kann, muss gesagt werden, dass eine gewisse Objektivität anzustreben versucht werden sollte. Einige Möglichkeiten, wie dies gelingen kann, sollen im nächsten Kapitel thematisiert werden.

6.2 Allgemeine Schlussfolgerungen für die Praxis

Die gemachten Beobachtungen laden dazu ein, die Ausbildung von Lehramtsstudent:innen zu überdenken. Während diese kaum Inhalte über Grammatik, Rechtschreibung und im Allgemeinen über die fachdidaktische Sprachrichtigkeit vermittelt bekommen, werden Themen wie beispielsweise Alt- und Mittelhochdeutsch gelernt, die defacto nicht in einer Schule gelehrt werden. Hierbei sollte deutlich werden, dass besonders im Bereich der Sprachrichtigkeit eine einheitliche Vorgehensweise über eine dementsprechende Lehre an der Hochschule erreicht werden kann. Zentrale Fragen, wie beispielsweise, ob Groß- und Kleinschreibung oder das/dass-Schreibung zu Rechtschreibung oder Grammatik gezählt werden soll, könnten somit leicht vermittelt werden. Dass bei der Bewertung von Ausdruck und Inhalt immer auch ein subjektives Empfinden mitschwingt, bleibt bis zu einem gewissen Grad verständlich und kann eventuell anhand standardisierter Raster teilweise ausgeglichen werden. Voraussetzung dafür ist jedoch eine präzise Anleitung beziehungsweise Beschreibung und eine für jede Textsorte angepasste Kriterienliste (vgl. Becker-Mrotzek & Böttcher 2006: 92).

Wünschenswert wäre ein alternativer Ansatz, wo mehrere Lehrkräfte eine Schularbeit korrigieren und daraus eine Durchschnittsnote vergeben wird. Dies lässt sich jedoch praktisch aufgrund des Aufwands wohl kaum umsetzen.

Darüber hinaus könnte angedacht werden, Prüfungsleistungen von externen Prüfer:innen bewerten zu lassen, wie es beispielsweise in Schulversuchen wie der Privatschule Walz geschieht. Dadurch gelten Lehrer:innen eher als Trainer:innen und nicht als Schiedsrichter:innen. Sie verfügen somit nicht über das Machtinstrument der Noten (vgl. Winter 2004: 50), wodurch auch der Unterricht positiv beeinflusst wird.

Im bildungspolitischen Kontext gesehen sollte darüber hinaus eine rechtliche Reform erfolgen, um den hohen Interpretationsspielraum in der Beschreibung von Leistungsfeststellungen und –beurteilungen zu verringern, sodass einheitliche Maßstäbe definiert werden können (vgl. Neuweg 2000: 72f.).

Im letzten Kapitel *Ausblick* sollen weitere Forschungsmöglichkeiten überlegt werden, welche an vorliegende Untersuchung anknüpfen und die Thematik weiter ausführen könnten.

6.3 Ausblick

Um die Ergebnisse zu vertiefen, erscheint eine Wiederholung der Studie mit einer größeren Stichprobe sinnvoll. Die Studie vorliegender Masterarbeit wurde durch die geringe Anzahl an Proband:innen limitiert und könnte im gegensätzlichen Fall eine gewisse Repräsentation der Grundgesamtheit gewährleisten. Die Schwierigkeiten bei der Rekrutierung der Sample lagen laut Aussagen der Proband:innen und bei Nachfragen von weiteren Lehrpersonen in dem Corona-Chaos begründet. Insofern wäre es zu einem anderen Zeitpunkt wahrscheinlich gut möglich, eine größere Stichprobe erreichen zu können.

Darüber hinaus wären Untersuchungen anzudenken, welche sich mit den einzelnen Aspekten eines Textes beschäftigen und wie diese Auswirkungen auf die Beurteilung der Gesamtnote haben.

Eine weitere Überlegung wäre, dass verschiedene Lehrkräfte anhand eines standardisierten Rasters einen Text beurteilen sollen. Damit würde eine höhere Vergleichbarkeit bei den Ergebnissen gewährleistet werden. Des Weiteren wäre es interessant, eine Kontrollgruppe zu rekrutieren, welche ohne Raster beurteilt. Anhand der Notensprünge könnte thematisiert werden, ob Kriterienkataloge für eine höhere Objektivität bei der Beurteilung sorgen.

Ebenso anzudenken wäre eine Studie, in welcher dieselben Lehrer:innen eine Schularbeit korrigieren und dies zu einem späteren Zeitpunkt nochmals wiederholen. Somit könnte der Frage der Reliabilität nachgegangen werden, also ob ein und dieselbe Person einen Text zu unterschiedlichen Zeitpunkten gleich beurteilen würde.

Wünschenswert wäre ebenso eine Studie hinsichtlich des Stadt-Land-Gefälles, da eine strengere Beurteilung auf dem Land angenommen wird, wenn einigen Zeitungsartikeln der letzten Jahre Glauben geschenkt werden möchte. Dieser Aspekt sollte bereits in vorliegender Masterarbeit mituntersucht werden, jedoch musste dies anhand der Zusammensetzung der Stichprobe wieder verworfen werden.

Abschließend bleibt zu sagen, dass sowohl die Einführung der Bildungsstandards als auch die Umstellung auf die Zentralmatura nicht die notwendige Sicherung der Objektivität gewährleisten konnten. Einige Lösungsvorschläge wurden bereits in vorliegender Arbeit thematisiert, weitere Vorgehensweisen müssen jedoch von der Bildungspolitik überlegt werden. Dennoch muss gesagt werden, dass der subjektive Charakter in der Arbeit mit Menschen niemals zur Gänze verhindert werden kann.

7. Literaturverzeichnis

- Abel, Andrea/Aivars Glaznieks (2017): KoKo: Bildungssprache im Vergleich: korpusunterstützte Analyse der Sprachkompetenz bei Lernenden im deutschen Sprachraum—ein Ergebnisbericht, Korpus-Südtirol, [online] http://www.korpus-suedtirol.it/KoKo/Documents/Ergebnisse_Dokumentation_gesamt_FINAL.pdf [abgerufen am 03.04.2022].
- Ágel, Vilmos (2008): Bastian Sick und die Grammatik. Ein ungleiches Duell, in: *Informationen Deutsch als Fremdsprache*, Bd. 35, Nr. 1, S. 64–84, [online] doi:10.1515/infodaf-2008-0107.
- Ammon, Ulrich (2005): Standard und Variation: Norm, Autorität, Legitimation, in: Ludwig M. Eichinger/Kallmeyer Werner (Hrsg.), *Standardvariation. Wie viel Variation verträgt die deutsche Sprache?* Berlin, Deutschland: De Gruyter, S. 28–40.
- Ammon, Ulrich/Hans Bickel/Jakob Ebner (2004): *Variantenwörterbuch des Deutschen: Die Standardsprache in Österreich, der Schweiz und Deutschland sowie in Liechtenstein, Luxemburg, Ostbelgien und Luxemburg, Ostbelgien und Südtirol*, 1. Aufl., Berlin, Deutschland: De Gruyter.
- Bábosik, István/Richard Olechowski (2003): *Lehren – Lernen – Prüfen (Schule - Wissenschaft - Politik, Band 16)*, Neuausg., Bern, Schweiz: Peter Lang.
- Bartnitzky, Horst (2000): Leistung und Leistungsbeurteilung. Leistung der Kinder? Leistung der Schule?, in: Dieter Haarmann (Hrsg.), *Handbuch Grundschule*, 4. Aufl., Weinheim, Deutschland: Beltz Verlag, S. 114–128.
- Bartsch, Renate (1985): *Sprachnormen. Theorie und Praxis*, Tübingen, Deutschland: M. Niemeyer.
- Baumert, Jürgen/Kai Maaz/Cornelia Gresch/Nele McElvany/Yvonne Anders/Kathrin Jonkmann/Marko Neumann/Rainer Watermann (2010): *Der Übergang von der Grundschule in die weiterführende Schule. Leistungsgerechtigkeit und regionale, soziale und ethnisch-kulturelle Disparitäten: Zusammenfassung der zentralen Befunde*, Bonn, Deutschland: Bundesministerium für Bildung und Forschung, Referat Bildungsforschung.
- Baurmann, Jürgen (2006): *Schreiben, Überarbeiten, Beurteilen*, Seelze, Deutschland: Kallmeyer.
- Becker-Mrotzek, Michael (2003): Mündlichkeit - Schriftlichkeit - Neue Medien, in: Ursula Bredel/Hartmut Günther/Peter Klotz/Jakob Ossner/Gesa Siebert-Ott (Hrsg.), *Didaktik*

- der deutschen Sprache - ein Handbuch. 1. Teilband*, Stuttgart, Deutschland: UTB, S. 69–89.
- Becker-Mrotzek, Michael/Ingrid Böttcher (2006): *Schreibkompetenz entwickeln und beurteilen, Praxishandbuch für die Sekundarstufe I und II*, Berlin, Deutschland: Cornelsen.
- Becker-Mrotzek, Michael/Joachim Grabowski/Torsten Steinhoff/Nadine Anskeit/Thomas Bachmann/Ulrike Behrens/Katrin Böhme/Albert Bremerich-Vos/Thomas Canz/Helmuth Feilke/Cornelia Glaser/Jörg Jost/Matthias Knopp/Michael Krelle/Katrin Lehnen/Markus Linnemann/Nicole Marx/Debora Meyer/Astrid Neumann/Guido Nottbusch/Maik Philipp/Thorsten Pohl/Kirsten Schindler/Stefan Schipolowski/Markus Schmitt/Hansjakob Schneider/Afra Sturm/Christian Weinzierl/Arne Wrobel (2016): *Forschungshandbuch empirische Schreibdidaktik*, 1. Aufl., Münster, Deutschland: Waxmann.
- Benischek, Isabella (2006): *Leistungsbeurteilung im österreichischen Schulsystem*, Berlin, Deutschland: Lit Verlag.
- Bergin, Christi/David Bergin (2009): Attachment in the Classroom, in: *Educational Psychology Review*, Bd. 21, Nr. 2, S. 141–170, [online] doi:10.1007/s10648-009-9104-0.
- Beuge, Patrick (2021): *Was ist gutes Deutsch?: Eine qualitative Analyse laienlinguistischen Sprachnormwissens (Sprache und Wissen (SuW), 40, Band 40)*, 1. Aufl., Berlin, Deutschland: De Gruyter.
- BIFIE (2013): Erläuterungen zur Bewertung der normativen Sprachrichtigkeit im Rahmen der schriftlichen Reife- und Diplomprüfung in Deutsch, Internationale Arbeitsgruppe - Schriftliche standardisierte kompetenzorientierte Reife- und Diplomprüfung Deutsch, [online] <https://docplayer.org/21504820-Erlaeuterungen-zur-bewertung-der-normativen-sprachrichtigkeit-im-rahmen-der-schriftlichen-reife-und-diplompruefung-in-deutsch.html> [abgerufen am 13.05.2022].
- Binder, Eva/Irmgard Harrer (1998): *Untersuchung von Maturaaufätzen mit dem Zürcher Textanalyseraster.*, Diplomarbeit, [online] http://germanistik.univie.ac.at/fileadmin/user_upload/inst_germanistik/Wiss_Arbeiten/matura.rtf.
- Birkel, Peter (2003): Aufsatzbeurteilung–ein altes Problem neu untersucht, in: *Didaktik Deutsch*, Bd. 9.15, S. 46–63.

- Birkel, Peter (2009): Rechtschreibleistung im Diktat - eine objektiv beurteilbare Leistung?, in: *Didaktik Deutsch*, Bd. 15.27, S. 5-32.
- Böheim-Galehr, Gabriele/Johann Engleitner (2014): *Schule der 10- bis 14-Jährigen in Vorarlberg: Entwicklungen, Bildungshaltungen und Bildungserwartungen. Projektbericht Band 1 (FokusBildungSchule)*, Innsbruck, Österreich: Studien Verlag.
- Böhm, Winfried (2000): *Wörterbuch der Pädagogik*, Stuttgart, Deutschland: UTB.
- Brügelmann, Hans (2014): *Sind Noten nützlich - und nötig? Ziffernzensuren und ihre Alternativen im empirischen Vergleich. Eine wissenschaftliche Expertise des Grundschulverbandes*, Frankfurt am Main, Deutschland: Grundschulverband e.V., S. 72, [online] doi: 10.25656/01:18828
- Bundesministerium Bildung, Wissenschaft und Forschung (2022): Bildungsstandards in der Allgemeinbildung, [online]
<https://www.bmbwf.gv.at/Themen/schule/bef/bist/bildungsstandards.html> [abgerufen am 23.02.2022].
- Corder, Stephen Pit (1967): The significance of learner's errors, in: *IRAL - International Review of Applied Linguistics in Language Teaching*, Bd. 5, Nr. 1-4, S. 161-170, [online] doi:10.1515/iral.1967.5.1-4.161.
- Coseriu, Eugenio (1988): *Einführung in die allgemeine Sprachwissenschaft*, Stuttgart, Deutschland: UTB.
- Creswell, John/Vicki Plano Clark (2007): *Designing and Conducting Mixed Methods Research*, Thousand Oaks, Kanada: SAGE Publications.
- Criblez, Lucien (2002a): Beurteilen als Grundlage von Promotion, Selektion und Berechtigung im Bildungswesen, in: Heinz Rhy (Hrsg.), *Beurteilung macht Schule: Leistungsbeurteilung von Kindern, Lehrpersonen und Schule*, 1., Aufl., Bern, Schweiz: Haupt Verlag, S. 83-94.
- Criblez, Lucien (2002b): Wozu Pädagogik? Zum Funktionswandel der Pädagogik in der Lehrerbildung, in: *Beiträge zur Lehrerbildung* 20, S. 300-318, [online] doi:10.25656/01:13499.
- Davies, Winifred V. (2017): Gymnasiallehrkräfte als SprachnormvermittlerInnen und Sprachnormautoritäten, in: Davies Winifred (u.a.) (Hrsg.): *Standardsprache zwischen Norm und Praxis. Theoretische Betrachtungen, empirische Studien und sprachdidaktische Ausblicke*, Tübingen, Deutschland: Narr Francke Attempto, S. 123-146.

- de Cillia, Rudolf/Jutta Ransmayr (2019): *Österreichisches Deutsch macht Schule. Bildung und Deutschunterricht im Spannungsfeld von sprachlicher Variation und Norm*, Wien, Österreich: Böhlau Verlag GmbH & Co. KG.
- Decker, Lena/Ina Kaplan (2019): Beurteilung schriftlicher Schülerleistungen mithilfe der Textsorten Lehrerkommentar und Kriterienkatalog - ein Vergleich, in: Ina Kaplan/Inger Petersen (Hrsg.), *Schreibkompetenzen messen, beurteilen und fördern (Sprachliche Bildung)*, 1. Aufl., Münster, Deutschland: Waxmann, S. 121–139.
- Decker, Lena/Nina Gersdorf/Ina Kaplan (2016): *Beurteilungspraxis von Lehrkräften und ihre Erwartungen zu Kompetenzen angehender Lehrkräfte im Praxissemester - Eine Experteninterviewstudie*, Ako Working Papers, [online] [https://www.uni-siegen.de/phil/ako/publikationen_vortraege/ako_working_papers_2_\(2016\).pdf](https://www.uni-siegen.de/phil/ako/publikationen_vortraege/ako_working_papers_2_(2016).pdf). [abgerufen am 24.03.2022].
- Deimel, Rainer (2015): Über die Unmöglichkeit, objektiv zu urteilen – Zur Klärung eines Paradoxons, ABA-Fachverband, [online] <https://aba-fachverband.info/ueber-die-unmoeglichkeit-objektiv-zu-urteilen-zur-klaerung-eines-paradoxons/> [abgerufen am 10.03.2022].
- Deutsche Rechtschreibung - Regeln und Wörterverzeichnis (2018): Rat für deutsche Rechtschreibung, [online] https://www.rechtschreibrat.com/DOX/rfdr_Regeln_2016_redigiert_2018.pdf [abgerufen am 18.02.2022].
- Duden (2005): *Die Grammatik*, 7. Aufl., Mannheim, Deutschland: Bibliographisches Institut & F.A. Brockhaus AG.
- Duden (2006): *Die deutsche Rechtschreibung*, 24. Aufl., Mannheim, Deutschland: Bibliographisches Institut & F.A. Brockhaus AG.
- Dudenredaktion (2022): Nix, in: *Duden*, [Wörterbucheintrag].
- Dudenredaktion (2022): rangehen, in: *Duden*, [Wörterbucheintrag].
- Dürscheid, Christa (2012a): Reich der Regeln, Reich der Freiheit. System, Norm und Normenreflexion in der Schule, in: *Kommunikation und Öffentlichkeit*, S. 105–120, [online] doi:10.1515/9783110289916.105.
- Dürscheid, Christa (2012b): *Syntax: Grundlagen und Theorien*, 6. Aufl., Stuttgart, Deutschland: UTB.
- Ebner, Jakob (2008): *Duden. Österreichisches Deutsch. Eine Einführung von Jakob Ebner*. Mannheim, Deutschland: Dudenverlag.

- Echterhoff, Gerald/Margit Schreier/Walter Hussy (2013): *Forschungsmethoden in Psychologie und Sozialwissenschaften für Bachelor*, 2., überarbeitete Aufl., Berlin, Deutschland: Inprint Springer.
- Eder, Ferdinand (2002): *Qualitätsentwicklung und Qualitätssicherung im österreichischen Schulwesen*, Weinheim, Deutschland: Beltz Verlag.
- Eisenberg, Peter & Gerhard Voigt (1990): Grammatikfehler, in: *Praxis Deutsch* 102, S. 10-15.
- Ellis, Rod (2009): A typology of written corrective feedback types, in: *ELT Journal*, Bd. 63, Nr. 2, S. 97–107, [online] doi:10.1093/elt/ccn023.
- Engelbrecht, Helmut (2013): Historische Entwicklung schulischer Leistungsbewertung, VCL-OE, [online] https://www.vcl-oe.at/images/stories/vcl/artikel_aus_vcl_news/vcl_zerlegt/vcl_zerlegt_20131/vcl-news%20maerz%202013%20artikel%20helmut%20engelbrecht.pdf [abgerufen am 17.05.2022].
- Ernst, Peter (2012): *Deutsche Sprachgeschichte*, 2. Aufl., Wien, Österreich: UTB.
- Bayrhammer, Bernadette (2014): Migration und Schule: Von verschenkten Noten und Versagen, Die Presse, [online] <https://www.diepresse.com/3853938/migration-und-schule-von-verschenkten-noten-und-versagen> [abgerufen am 17.05.2022].
- Feilke, Helmuth (2012): Bildungssprachliche Kompetenzen fördern und entwickeln. Basisartikel, in: *Praxis Deutsch*, Nr. 233, S. 4–13.
- Feilke, Helmuth/Mathilde Hennig (2018): *Zur Karriere von ›Nähe und Distanz‹: Rezeption und Diskussion des Koch-Oesterreicher-Modells (Reihe Germanistische Linguistik, 306, Band 306)*, 1. Aufl., Berlin, Deutschland: De Gruyter.
- Florek, Hans-Christian (1999): *Leistungsbegriff und pädagogische Praxis*, Münster, Deutschland: LIT.
- Frigge, Reinhold (2000): Wir wollen besser werden. Perspektiven, Verfahren und Instrumente einer reflexiven Beurteilungspraxis, in: *Deutschunterricht* 53, Heft 6, S. 426-435.
- Fuss, Elisabeth (2002): Begründung und Intention des Jahresthemas, in: Konrad Lemnitzer/Werner Wiater (Hrsg.), *Die Schule fördert und fordert: Leistungsbereitschaft und Leistungsfähigkeit in pädagogischer Verantwortung*, Seelze, Deutschland: Kallmeyer, S. 11–19.
- Glück, Helmut (2010): *Metzler Lexikon Sprache*, Heidelberg, Deutschland: J.B. Metzler.

- Goldgruber, Barbara Elisabeth (2011): *Einstellungen zu Dialekt und Standardsprache in Österreich. Eine empirische Untersuchung in Graz und Wien*, Diplomarbeit, Hauptuniversität Wien.
- Haas, Viktoria Maria (2008): *Alternative Leistungsbeurteilung in der Sekundarstufe I*, Diplomarbeit, [online] <https://unipub.uni-graz.at/obvugrhs/content/titleinfo/205568/full.pdf>. [abgerufen am 12.02.2022].
- Hattie, John/Wolfgang Beywl/Klaus Zierer (2013): *Lernen sichtbar machen*, 1. Aufl., Baltmannsweiler, Deutschland: Schneider Verlag Hohengehren.
- Hennig, Mathilde (2012): Was ist ein Grammatikfehler, in: *Kommunikation und Öffentlichkeit: Sprachwissenschaftliche Potenziale zwischen Empirie und Norm* 296, S. 121.
- Hofer, Manfred (1969): *Die Schülerpersönlichkeit im Urteil des Lehrers*, Weinheim, Deutschland: Beltz Verlag.
- Höhle, Mandy (2010): Standardsprache, in: Barkowski, Hans/ Hans-Jürgen Krumm (Hrsg): *Fachlexikon Deutsch als Fremd- und Zweitsprache*. Tübingen, Deutschland: Francke Verlag.
- Huesmann, Anette (1998): *Zwischen Dialekt und Standard: Empirische Untersuchung zur Soziolinguistik des Varietätenspektrums im Deutschen (Reihe Germanistische Linguistik, 199, Band 199)*, Reprint 2017, Berlin, Deutschland: De Gruyter.
- Hundt, Markus (2009): Normverletzungen und neue Normen, in: *Deutsche Grammatik - Regeln, Normen, Sprachgebrauch*, S. 117–140, [online] doi:10.1515/9783110217360.3.117.
- Ingenkamp, Karl-Heinz/Urban Lissmann (2008): *Lehrbuch der Pädagogischen Diagnostik (Beltz Pädagogik)*, 6. Aufl., Weinheim, Deutschland: Beltz Verlag.
- Ingenkamp, Karlheinz (1977): *Die Fragwürdigkeit der Zensurengebung. Texte und Untersuchungsgebiete*, Weinheim, Deutschland: Beltz.
- Ingenkamp, Karlheinz (1977): *Schüler- und Lehrerbeurteilung.: Empirische Untersuchungen zur pädagogischen Diagnostik. (Theorie und Praxis der Schulpsychologie)*, Weinheim, Deutschland: Beltz Verlag.
- Ingenkamp, Karlheinz (1985): *Lehrbuch der Pädagogischen Diagnostik*, Weinheim, Deutschland: Beltz.
- Ingenkamp, Karlheinz (1995): *Die Fragwürdigkeit der Zensurengebung: Texte und Untersuchungsberichte*, Weinheim, Deutschland: Beltz Verlag.

- Jachmann, Michael (2003): *Noten oder Berichte?: Die schulische Beurteilungspraxis aus der Sicht von Schülern, Lehrern und Eltern*, Reihe Schule und Gesellschaft, 29, Opladen, Deutschland: Leske + Budrich.
- Johnson, R. Burke/Anthony J. Onwuegbuzie (2004): Mixed Methods Research: A Research Paradigm Whose Time Has Come, in: *Educational Researcher*, Bd. 33, Nr. 7, S. 14–26, [online] doi:10.3102/0013189x033007014.
- Joseph, John (1987): *Eloquence And Power: The Rise Of Language Standards And Standard Languages*, New York, USA: B. Blackwell.
- Jürgens, Eiko (2005): *Leistung und Beurteilung in der Schule: Eine Einführung in Leistungs- und Bewertungsfragen aus pädagogischer Sicht*, 6. Aufl., Sankt Augustin, Deutschland: Academia.
- Jürgens, Eiko (2006): Beurteilen und Fördern ist kein Widerspruch – Begründungsstruktur bedeutungsbezogener Lernformierung, in: *Erziehungswissenschaft und Beruf*. 54. Jahrgang. 1/06. Rinteln, Deutschland, S. 3-16
- Kellermeier-Rehbein, Birte (2013): Standard oder Nonstandard? Ungelöste Probleme der Abgrenzung, in: Karina Schneider-Wiejowski/Birte Kellermeier-Rehbein/Jakob Haselhuber (Hrsg.), *Vielfalt, Variation und Stellung der deutschen Sprache*, 1. Aufl., Berlin, Deutschland: De Gruyter, S. 3–22.
- Kiper, Hanna (2001): *Einführung in die Schulpädagogik*, Berlin, Deutschland: Cornelsen.
- Klafki, Wolfgang (1976): Erziehungswissenschaft als kritisch-konstruktive Theorie. Hermeneutik–Empirie–Ideologiekritik, in: Wolfgang Klafki (Hrsg.), *Aspekte kritisch-konstruktiver Erziehungswissenschaft. Gesammelte Beiträge zur Theorie-Praxis-Diskussion*, Weinheim, Deutschland: Beltz Verlag, S. 13–49.
- Klein, Wolf-Peter (2013): Warum brauchen wir einen klaren Begriff von Standardsprachlichkeit und wie könnte er gefasst werden?, in: Jörg Hagemann/Klein Wolf-Peter/Staffeldt, Sven (Hrsg.), *Pragmatischer Standard*. Tübingen, Deutschland: Stauffenburg, S. 15–33.
- Klein, Wolf-Peter (2014): Gibt es einen Kodex für die Grammatik des Neuhochdeutschen und, wenn ja, wie viele? Oder: Ein Plädoyer für Sprachkodexforschung, in: Albrecht Plewnia/Witt Andreas (Hrsg.), *Sprachverfall? Dynamik – Wandel – Variation*. Berlin, Deutschland: De Gruyter, S. 219–242.
- Kleppin, Karin (1998): *Fehler und Fehlerkorrektur*, Berlin, Deutschland: Langenscheidt.

- Konopka, Marek/Bruno Strecker (2009): Normverletzungen und neue Normen, in: *Deutsche Grammatik - Regeln, Normen, Sprachgebrauch*, S. 117–140, [online] doi:10.1515/9783110217360.3.117.
- Krapp, Klaus-Peter (2001): Pädagogisch-psychologische Diagnostik, in: Andreas Krapp/Bernd Weidenmann (Hrsg.), *Pädagogische Psychologie*, 4. Aufl., Weinheim, Deutschland: Beltz Verlag, S. 513–564.
- Kremser, Martin/NWV Verlag GmbH (2017): *Die Leistungsbeurteilung im österreichischen Schulrecht: Notengebung aus rechtlicher Sicht*, Wien, Österreich: NWV.
- Linke, Angelika/Markus Nussbaumer/Paul Portmann (2004): *Reihe Germanistische Linguistik 121 Kollegbuch: Studienbuch Linguistik*, 5. Aufl., Berlin, Deutschland: De Gruyter.
- Maaz, Kai/Franz Baeriswyl/Ulrich Trautwein (2013): II. Studie: „Herkunft zensiert?“ Leistungsdiagnostik und soziale Ungleichheiten in der Schule, in: *Chancen bilden*, S. 185–188, [online] doi:10.1007/978-3-531-19393-9_13.
- Michael, Russell/Wei Tao (2004): Effects of handwriting and computer-print on composition scores: A follow-up to Powers, Fowles, Farnum, & Ramsey, in: *Practical Assessment, Research, and Evaluation* 9.1, S. 1.
- Morphy, Paul/Steve Graham (2012): Word processing programs and weaker writers/readers: A meta-analysis of research findings, in: *Reading and Writing* 25.3, S. 641-678.
- Mroczynski, Robert (2013): „Ja nein, ich meine...“ – zur ja nein-Konstruktion im gesprochenen Deutsch, in: Jörg Hagemann/Wolf Peter Klein/Sven Staffeldt (Hrsg.), *Pragmatischer Standard*, 1. Aufl., Tübingen, Deutschland: Stauffenburg, S. 245–263.
- Neuweg, Georg Hans (2014): *Schulische Leistungsbeurteilung*, 5. akt. und erw. Auflage 2014, Linz, Österreich: Trauner Verlag.
- Neuweg, Georg Hans (2018): *Kompetenzorientierte Leistungsbeurteilung. Pädagogische und rechtliche Hilfestellungen für die Schulpraxis*, 1. Auflage 2018, Linz, Österreich: Trauner Verlag.
- Nussbaumer, Markus (1991): *Was Texte sind und wie sie sein sollen: Ansätze zu einer sprachwissenschaftlichen Begründung eines Kriterienrasters zur Beurteilung von schriftlichen Schülertexten (Germanistische Linguistik, 119, Band 119)*, Reprint 2010, Berlin, Deutschland: De Gruyter.
- Nussbaumer, Markus/Peter Sieber (1994): Texte analysieren mit dem Zürcher Textanalyseraster, in: Peter Sieber (Hrsg.), *Sprachfähigkeiten - Besser als ihr Ruf und nötiger denn je!*, Aarau, Schweiz: Sauerländer, S. 141–186.

- Nutz, Maximilian (2006): Beurteilung sprachlicher Leistungen, in: Ursula Bredel/Hartmut Günther/Peter Klotz/Gesa Siebert-Ott (Hrsg.), *Didaktik der deutschen Sprache, Bd. 2: Ein Handbuch*, 2. Aufl., Stuttgart, Deutschland: UTB, S. 924–937.
- Olechowski, Richard (2003): Alternative Leistungsbeurteilung – Humanisierung der Schule, in: István Bábosik/Richard Olechowski (Hrsg.), *Lehren – Lernen – Prüfen (Schule - Wissenschaft - Politik, Band 16)*, Neuausg., Frankfurt am Main, Deutschland: Peter Lang GmbH, Internationaler Verlag der Wissenschaften, S. 215–232.
- Perleth, Christoph/Mitra-Anne Sen (2010): Zuverlässigkeit von Schulnoten, kognitiven Fähigkeitstests und Begabungseinschätzung von Eltern für die weitere Schullaufbahn, in: Susanne Lin-Klitzing/David di Fuccia/Gerhard Müller-Frerich (Hrsg.), *Übergänge im Schulwesen: Chancen und Probleme aus sozialwissenschaftlicher Sicht (Gymnasium - Bildung - Gesellschaft)*, 1. Auflage, Bad Heilbrunn, Deutschland: Klinkhardt, S. 105–126.
- Persy, Elisabeth (1999): Ermutigende Leistungsbeurteilung in einer humanen Schule, in: Richard Olechowski/Karl Garnitschnig (Hrsg.), *Humane Schule (Schule - Wissenschaft - Politik, Band 14)*, Neuausg., Frankfurt, Deutschland: Peter Lang GmbH, Internationaler Verlag der Wissenschaften, S. 202–220.
- Peter, Klaus (2011): *Textbewertungen: Eine empirische Untersuchung zu Sprachbewusstheit und Spracheinstellungen*, Weinheim, Deutschland: Beltz Verlag.
- Poelchau, Heinz-Werner (2004): Veränderte Schülerwelten-Herausforderungen für die Motivation, in: *Schülermotivation. Konzepte und Anregungen für die Praxis 2*, S. 3–12.
- Ransmayr, Jutta (2020): Eine Frage des Schreibmediums. Deutschmatura mit dem Stift oder am Computer schreiben?, in: Helen Bito/Ursula Esterl/Elisabeth Schabus-Kant (Hrsg.): *Schrift und Schriftlichkeit*, Innsbruck, Österreich: StudienVerlag, S. 61–70.
- Rausch, Tobias (2013): Wie Sympathie und Ähnlichkeit Leistungsbeurteilungen beeinflussen kann, Researchgate, [online]
https://www.researchgate.net/publication/259382244_Wie_Sympathie_und_Ahnlichkeit_Leistungsbeurteilungen_beeinflussen_kann [abgerufen am 02.04.2022].
- Rheinberg, Falko (2001): Leistungsbeurteilung im Schulalltag: Wozu vergleicht man was womit?, in: Weinert, Franz (Hrsg.), *Leistungsmessungen in Schulen (Beltz Pädagogik)*, 2. Aufl., S. 59–73, [online] <http://www.psych.uni-potsdam.de/people/rheinberg/messverfahren/Leistbeurteilung.pdf>. [abgerufen am 12.03.2022].

- Rinas, Karsten (2017): *Theorie der Punkte und Striche: die Geschichte der deutschen Interpunktionslehre*, Norderstedt, Deutschland: BoD–Books on Demand.
- RIS - Leistungsbeurteilungsverordnung - Bundesrecht konsolidiert, Fassung vom 04.03.22 (2022): Rechtsinformationssystem des Bundes, [online]
<https://www.ris.bka.gv.at/GeltendeFassung.wxe?Abfrage=Bundesnormen&Gesetzesnummer=10009375> [abgerufen am 04.03.2022].
- Rittberger, Michael (2000): Zensurengebung in der Sekundarstufe, in: *Erziehung und Unterricht 1–2*, S. 96–107.
- Sacher, Werner (2014): *Leistungen entwickeln, überprüfen und beurteilen: Bewährte und neue Wege für die Primar- und Sekundarstufe*, 6., überarb und erw. Auflage., Berlin, Deutschland: Klinkhardt Julius.
- Schäfer, Roland (2018): *Einführung in die grammatische Beschreibung des Deutschen: Dritte, überarbeitete und erweiterte Auflage (Textbooks in Language Science, Band 2)*, Berlin, Deutschland: Language Science Press.
- Schmidinger, Rupert/Elfriede Vierlinger/Österreich Bundesministerium Für Unterricht (2012): *Zeitgemäße Leistungsbeurteilung*, Weinheim, Deutschland: Beltz Verlag.
- Schmidlin, Regula (2011): *Die Vielfalt des Deutschen. Standard und Variation. Gebrauch, Einschätzung und Kodifizierung einer plurizentrischen Sprache*, Berlin, Deutschland: De Gruyter.
- Schröter, Gottfried (1975): Von der Theorie zur Praxis der Aufsatzbeurteilung, in: *Praxis Deutsch*, 10, Weinheim, Deutschland: Beltz Verlag.
- Schröter, Gottfried (1981): *Zensuren? Zensuren!: Allgemeine und fachspezifische Probleme : Grunderkenntnisse und neue Forschungsergebnisse für Lehrer, Eltern und interessierte Schüler*, Weinheim, Deutschland: Beltz Verlag.
- Schwarzer, Ralf/Bernward Lange (1979): Implizite Unterrichtstheorie von Lehrern, in: Dietmar Bolscho/Christine Schwarzer (Hrsg.), *Beurteilen in der Grundschule*, München, Deutschland: Urban und Schwarzenberg, S. 67–84.
- Seneca, Lucius Annaeus/Heinz Gunermann, H. (1998). *Epistulae morales ad Lucilium*, Liber XVII et XVIII. Stuttgart, Deutschland: Reclam.
- Sieber, Peter (1994): *Sprachfähigkeiten, Besser als ihr Ruf und nötiger denn je!*, Aarau, Schweiz: Sauerländer.
- Statistik Austria (2021): Statistik Austria - Die Informationsmanager, [online]
<https://www.statistik.at/statistiken/bevoelkerung-und-soziales/bildung/lehrpersonen> [abgerufen am 27.03.2022].

- Stern, Thomas (2010): *Förderliche Leistungsbewertung*, Wien, Österreich: Österreichisches Zentrum für Persönlichkeitsbildung und soziales Lernen.
- Steuer, Gabriele (2014): *Fehlerklima in der Klasse: zum Umgang mit Fehlern im Mathematikunterricht*, Berlin, Deutschland: Springer-Verlag.
- Tent, Lothar/Ingeborg Stelzl (1993): *Pädagogisch-psychologische Diagnostik: Theoretische und methodische Grundlagen*, Göttingen, Deutschland: Hogrefe Verlag.
- Thonhauser, Josef (2002): Die Beurteilung der Leistungen von Schülerinnen und Schülern, in: Ferdinand Eder (Hrsg.), *Qualitätsentwicklung und Qualitätssicherung im österreichischen Schulwesen*, Weinheim, Deutschland: Beltz Verlag, S. 3–16.
- Tortik, Evelyn/NWV Verlag GmbH (2016): *Rechtsfragen der Leistungsbeurteilung: Allgemeine Fragestellungen und Besonderheiten der Sekundarstufe I im Kontext bildungspolitischer Reformprozesse*, Weinheim, Deutschland: Beltz Verlag.
- Voyer, Daniel/Susan D. Voyer (2014): Gender differences in scholastic achievement: A meta-analysis., in: *Psychological Bulletin*, Bd. 140, Nr. 4, S. 1174–1204, [online] doi:10.1037/a0036620.
- Weiss, Rudolf (1965): *Zensur und Zeugnis*, Linz, Österreich: Haslinger.
- Wiesinger, Peter (1988): *Die deutsche Sprache in Österreich. Das österreichische Deutsch*, Wien/Köln/Graz, Österreich: Böhlau Verlag.
- Winter, Felix (2004): *Leistungsbewertung: Eine neue Lernkultur braucht einen anderen Umgang mit den Schülerleistungen*, Weinheim, Deutschland: Beltz Verlag.
- World Vision Deutschland e.V. (2018): *Kinder in Deutschland 2018: 4. World Vision Kinderstudie*, 1. Aufl., [online] <https://www.worldvision.de/sites/worldvision.de/files/pdf/World-Vision-Zusammenfassung-vierte-Kinderstudie.pdf>.
- Ziegenspeck, Jörg (1979): Zensur und Zeugnis - Ein Mängelbericht, in: Dietmar Bolscho/Christine Schwarzer (Hrsg.), *Beurteilung in der Grundschule*, München, Deutschland: Urban & Schwarzenberg, S. 36–54.
- Zillig, Maria (1928): Einstellung und Aussage [Attitude and Aussage], in: *Zeitschrift für Psychologie und Physiologie der Sinnesorgane. Abt. 1. Zeitschrift für Psychologie*, 106, S. 58–106, [online] URL: <https://psycnet.apa.org/record/1929-01862-001> [abgerufen am 15.03.2022].

8. Abbildungsverzeichnis

Abbildung 1: Soziales Kräftefeld einer Standardvarietät, eigene Darstellung nach Ammon (1995: 80)	16
Abbildung 2: Fehlerkreislauf, eigene Darstellung nach Hennig (2012: 125)	20
Abbildung 3: Beurteilungsskala Neue Mittelschule, eigene Darstellung nach Bildungsdirektion Wien (Quelle: Leistungsbeurteilung und Berechtigungen (ssr-wien.at)).....	32
Abbildung 4: Geschlecht der Stichprobe.....	51
Abbildung 5: Schulart	52
Abbildung 6: Alter der Proband:innen	52
Abbildung 7: Leistungsbeurteilungen aller Teilnehmer:innen.....	56
Abbildung 8: Korrektur Lehrperson Nr.1.....	57
Abbildung 9: Korrektur Lehrperson Nr.18.....	58
Abbildung 10: Korrektur Lehrperson Nr.21.....	59
Abbildung 11: Übersicht Korrekturzeichen bei Arbeiten mit Note "Befriedigend"	60
Abbildung 12: Übersicht Korrekturzeichen bei Arbeiten mit Note "Genügend"	61
Abbildung 13: Original-Schularbeit.....	66
Abbildung 14: Schularbeit Note "Gut"	83
Abbildung 15: Schularbeit Note "Nicht genügend"	85

9. Tabellenverzeichnis

Tabelle 1: Übersicht Korrekturzeichen & Benotungen.....	55
Tabelle 2: Übersicht Fragebogendaten der Proband:innen.....	61
Tabelle 3: Übersicht über verwendete Korrekturzeichen	62

10. Anhang

10.1 Original-Schularbeitstext

blutige Herbstferien

Die letzten weichen Strahlen der Herbstsonne versuchten das blasser Gesicht des Mädchens, das unter dem Fenster lag, zu erwärmen, doch sie blieb kühl. Der Täter wusste, dass er geschnappt werden würde, ~~falls~~ wenn er nicht die Spuren beseitigen ~~würde~~ und weglaufen würde. Ein Handy klingelte. Der Täter ging sofort ran. Es stellte sich heraus, dass dies sein Komplize sei, der wissen wollte, ob die Tat vollbracht war, doch er schwieg, legte auf, und ging davon.

Es war ~~ein~~ Ein typischer Samstag. Kommissar Brot verspeißte seinen Apfel und dachte sich nichts dabei, als auf einmal sein Handy ertönte. „Kommissar Brot zu Ihren Diensten! wie kann ich heute behilflich sein?“, sagte er mit einem höflichem Umgangston. Die nervöse Frau am Handy schrie, dass er doch schnell zur Algostraße 12 kommen sollte da es ein Notfall gab. Kommissar Brot versuchte die ~~Frau~~ Frau zu beruhigen, doch es half nix. Daraufhin hoppte er in seinen Wagen und machte sich auf den schnellsten Weg zur Algostraße 12. Er sah 2 Junge Damen. Eine kullerten die Tränen herunter, wie ein Wasserfall. Die andere, war jedoch voll mit ~~blut~~ blut und Glassplittern als er näher kam, sah er dass das junge Mädchen 2 Schüsse in den Kopf und einen in die Brust bekommen hatte.

Der Mann trat näher und befragte die Dame. Es kam heraus, dass sie Anna hieß und das anscheinend tote Mädchen Marie. Marie war 17 Jahre alt und ging noch zur Schule. Eine Sirene ertönt „Das wird wohl die Polizei sein.“, sagte Anna mit Tränen in ihren Augen. ~~Kommissar~~ ^{Kommissar} Brot erzählte den Polizisten über das Verbrechen. Dabei fiel ihm ein mitte 60 Jähriger Mann auf, der das Mädchen leblos anstarrt. Es kam ihm so vor, dass der Mann seinen Blick nicht von ihr abwenden könne. Dies gab Kommissar Brot ein merkwürdiges Gefühl. ~~Der~~ Der Mann seufzte und sagte mit einer leisen und ruhigen Stimme: „2 unnötige Schüsse.“ Da bekam Kommissar Brot ein Geistesblitz. Herr Brot starrte ihm mit großen Augen an und sagte: „Sie waren es.“ Der Mann starrte zurück. Alle Augen waren auf die Beiden ~~gerichtet~~ gerichtet. Man sah dem Mann die ~~Furcht~~ Furcht an als hätte man es ihm ~~auf~~ auf die Stirn geschrieben. Er musste sich die Tränen unterdrücken und gab anschließend ein Geständnis ab. Es kam heraus, dass er, ihr Freund sei und das nur getan hat, da sie ~~hat~~ über seine illegalen Geschäfte heraufgefunden hatte und das sein Komplize noch auf freien Füße sei. Doch er wollte noch wissen, wie er darauf gekommen ist, dass er der Mörder sei? Kommissar Brot meinte: „Woher wissen sie, dass 2 Schüsse unnötig ~~zu~~ waren?“

10.2 Fragebogen

Umfrage: Umgang mit Schüler*innentexten

Liebe*r Teilnehmer*in,

mein Name ist Nina Straussberger, ich studiere Psychologie/Philosophie, Sport und Deutsch auf Lehramt und verfasse derzeit meine Masterarbeit am Institut der Germanistik. Dabei beschäftige ich mich mit dem Umgang von Schularbeitstexten im Fach Deutsch. Dazu würde ich Sie bitten, die folgenden Fragen wahrheitsgemäß zu beantworten und die Ihnen zugesandte Deutsch-Schularbeit nach Ihren herkömmlichen Kriterien zu beurteilen. Geben Sie bitte des Weiteren auch den von Ihnen verwendeten Notenschlüssel an. Die Korrektur sollte direkt auf der Schularbeit erfolgen. Die Bearbeitung dieses Fragebogens dauert etwa 15 Minuten, für die Korrektur der Schularbeit ist die Dauer sehr individuell. Für einen erfolgreichen Abschluss dieser Studie ist es wichtig, dass Sie den Fragebogen vollständig ausfüllen und keine der Fragen auslassen. Alle Daten werden selbstverständlich anonym erhoben und streng vertraulich behandelt. Somit können sie Ihrer Person nicht zugeordnet werden.

Vielen Dank für Ihre Teilnahme!

1 Welches Geschlecht haben Sie?

Fragesupport: *Wählen Sie eine Antwort*

☐ weiblich ☐ männlich ☐ divers

2 Wie alt sind Sie?

3 In welcher Schule unterrichten Sie?

Fragesupport: *Wählen Sie eine Antwort*

☐ NMS od. MS ☐ AHS

4 Unterrichten Sie in einer Stadt oder auf dem Land?

Fragesupport: *Wählen Sie eine Antwort*

☐ Stadt ☐ Land

5 Wie lange unterrichten Sie bereits das Fach Deutsch?

6 Wieviele Stunden pro Woche unterrichten Sie das Fach Deutsch?

7 In welchen Schulstufen haben Sie bereits unterrichtet?

Fragesupport: Wählen Sie eine oder mehr Antworten

- | | | | | |
|---|---|---|---|--|
| ☐ 5. Schulstufe (1. Klasse Unterstufe) | ☐ 6. Schulstufe | ☐ 7. Schulstufe | ☐ 8. Schulstufe | ☐ 9. Schulstufe |
| ☐ 10. Schulstufe | ☐ 11. Schulstufe | ☐ 12. Schulstufe | ☐ 13. Schulstufe | |

8 In welchen Schulstufen unterrichten Sie aktuell?

Fragesupport: Wählen Sie eine oder mehr Antworten

- | | | | | |
|---|---|---|---|--|
| ☐ 5. Schulstufe (1. Klasse Unterstufe) | ☐ 6. Schulstufe | ☐ 7. Schulstufe | ☐ 8. Schulstufe | ☐ 9. Schulstufe |
| ☐ 10. Schulstufe | ☐ 11. Schulstufe | ☐ 12. Schulstufe | ☐ 13. Schulstufe | |

Bitte drucken Sie nun die Ihnen zugeschickte Schularbeit aus, korrigieren diese direkt auf dem Zettel und gehen Sie dann weiter zur nächsten Frage.

9 Geben Sie hier bitte einen selbstausedachten, fünfstelligen numerischen Code ein (zB. 38639). Dieser Code soll auch auf der Schularbeit oben rechts vermerkt werden.

Fragesupport: Dieser Code dient lediglich Zuordnungszwecken, all Ihre Angaben bleiben selbstverständlich weiterhin anonym.

Bitte vermerken Sie den Code jetzt auf der Schularbeit oben rechts!

10 Welche Note würden Sie dem Ihnen zugesandten Schularbeitstext aus einer 3. Klasse Unterstufe (7. Schulstufe) geben?

Fragesupport: Bitte korrigieren Sie den Schularbeitstext und beurteilen Sie nach Ihrem üblichen Vorgehen.

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

11 Bitte begründen Sie Ihre Entscheidung in ein paar Sätzen.

12 Wie gelangen Sie im Allgemeinen zu Ihrer Beurteilung bei Deutschtexten?

13 Wie empfinden Sie die Leistungsanforderungen an Ihrer Schule?

Fragesupport: Wählen Sie eine Antwort

☐ zu niedrig ☐ passend ☐ zu hoch

14 Wie empfinden Sie das Leistungsniveau an Ihrer Schule?

Fragesupport: Wählen Sie eine Antwort

☐ zu niedrig ☐ passend ☐ zu hoch

Bitte schicken Sie zuletzt noch den korrigierten Schularbeitstext per Foto- oder pdf-Datei an die Email-Adresse
01448241@unet.univie.ac.at.

Vielen Dank für Ihre Teilnahme an der Studie!

10.3 Abstract

Abstract Deutsch

Die vorliegende Masterarbeit beschäftigt sich mit Unterschieden in Korrekturen und Leistungsbeurteilungen im Fach Deutsch. Das Ziel ist es, Differenzen und Gemeinsamkeiten verschiedener Lehrer:innen sowohl aus Allgemeinbildenden Höheren Schulen als auch Neuen Mittelschulen zu beleuchten. Dazu wurde mittels eines mixed-methods-designs quantitativ und qualitativ vorgegangen. Anhand einer anonymen Schularbeit im Fach Deutsch und eines Online-Fragebogens wurden das Korrekturverhalten und die Notengebungen von 21 Lehrer:innen untersucht. Es konnten deutliche Unterschiede in der Korrekturpraxis aufgezeigt werden. Darüber hinaus waren hohe Sprünge in den Benotungen zu beobachten. Die Spanne reichte dabei von einem „Gut“ bis zu einem „Nicht genügend“ für ein und denselben Schularbeitstext. Diese Erkenntnisse decken sich mit ähnlichen Untersuchungen zur Objektivität bei Bewertungen (vgl. Davies 2007; Hennig 2012; Vierlinger 2012), wenngleich bis dato keine vergleichbaren Studien durchgeführt wurden, welche die Thematiken Korrekturverhalten und Leistungsbeurteilungen kombinieren.

Abstract English

This master's thesis deals with differences in correcting and performance assessments in the school subject German. The aim is to shed light on differences and similarities between different teachers from both general secondary schools and new secondary schools. For this purpose, a mixed-methods-design was used, both quantitatively and qualitatively. The correction behavior and grading of 21 teachers was investigated by means of anonymous school work in German and an online questionnaire. Significant differences in correction practices could be shown. In addition, high jumps in the grades were observed. The range extended from a "good" to an "insufficient" for one and the same text. These findings are in line with similar studies on objectivity in assessments (cf. Davies 2007; Hennig 2012; Vierlinger 2012), although no comparable studies combining the topics of corrective behavior and performance assessments have been conducted to date.