

Radboud Universiteit Nijmegen
Radboud Docenten Academie
Faculteit der Letteren

Masterarbeit

Der Einfluss der *Self-Questioning*-Methode auf
die Interpretationsfähigkeiten von Schülern
und wie man diese bewerten kann

Betreuerin: dr. Stefanie Ramachers

Begutachterin: dr. Eva Knopp

Abgabedatum: 02.08.2023

Ymke Martine Brummel

S4541375

Tweejarige Educatieve Master Duits

ymke.brummel@ru.nl

Zusammenfassung

Die vorliegende Arbeit hat sich mit schriftlichen Dialogen, die Schüler vor und nach dem Lesen einer Kurzgeschichte anlässlich einer Unterrichtsintervention, die sich auf *Self-Questioning* richtete, schrieben, auseinandergesetzt. Die Intervention war Teil des Fremdsprachenunterrichts Deutsch an niederländischen Schulen in der vierten Klasse (= 10. Klasse in Deutschland) der Niveaus *havo* und *vwo*. Die Studie befasste sich mit den Fragen (1) wie Schüler literarische Kurzgeschichten im Fremdsprachenunterricht vor und nach einer Unterrichtsintervention, die sich auf *Self-Questioning* richtet, interpretieren; (2) inwiefern die *Self-Questioning*-Unterrichtsmethode beim Lesen von Kurzgeschichten einen Effekt auf die Interpretationen der Schüler hat; (3) wie ein Bewertungsschema für Interpretationsfähigkeiten von Schülern, zu einer Aufgabe, bei der von Schülern einen schriftlichen Dialog zu einer Kurzgeschichte geschrieben wurde, aussehen könnte. Um diese Fragen zu beantworten, wurden die schriftlichen Dialoge codiert und analysiert. Dabei stellte sich heraus, dass die Schüler vielfältige Aussagen zum Text treffen und dass deren Interpretationsfähigkeiten sich nach der Unterrichtsintervention verbessert hatten. Daraufhin wurde ein Bewertungsschema anhand der Codiererergebnisse und unter Berücksichtigung der wissenschaftlichen Literatur zu Interpretation und Prüfungen konzipiert. Das Bewertungsschema besteht aus den Kernkategorien ‚Beschreiben‘, ‚Beurteilen‘, ‚Interpretieren: hypothetisch‘, ‚Interpretieren: Charaktere‘, ‚Interpretieren: Geschichte‘ und ‚Referieren‘. Diese Kernkategorien enthalten alle vier Abstufungen, in denen erläutert wird, was Schüler machen müssen, um die jeweilige Abstufung zu erreichen. Schließlich werden Empfehlungen für den Einsatz des Bewertungsschema gegeben und wird die ganze Studie kritisch diskutiert.

Samenvatting

Voor deze scriptie werden schriftelijke dialogen, die leerlingen voor en na het lezen van een kort verhaal schreven naar aanleiding van een onderwijsinterventie die zich op *self-questioning* richtte, onderzocht. De interventie maakte deel uit van het vreemdetalenonderwijs Duits in Nederland in vierde klassen van de niveaus *havo* en *vwo*. Het onderzoek richtte zich op de vragen (1) hoe leerlingen korte literaire verhalen in het vreemdetalenonderwijs voor en na een onderwijsinterventie, die zich op *self-questioning* richt, interpreteren; (2) in hoeverre de *self-questioning*-onderwijsmethode bij het lezen van korte verhalen een effect op de interpretaties van leerlingen heeft; (3) hoe een beoordelingsschema voor interpretatievaardigheden van leerlingen, bij een opdracht, waarbij leerlingen een schriftelijke dialoog over een kort verhaal moesten schrijven, eruit zou kunnen zien. Om deze vragen te beantwoorden, werden de schriftelijke dialogen gecodeerd en geanalyseerd. Uit analyse van de gecodeerde data bleek, dat leerlingen veelzijdige uitspraken over de tekst doen en dat zich hun interpretatievaardigheden na de onderwijsinterventie verbeterd hadden. Vervolgens werd een beoordelingsschema ontworpen aan de hand van de codeerresultaten en rekening houdend met de wetenschappelijke literatuur over interpretatie en toetsing. Het beoordelingsschema bestaat uit de kerncategorieën ‚beschrijven‘, ‚beoordelen‘, ‚interpreteren: hypothetisch‘, ‚interpreteren: personages‘, ‚interpreteren: verhaal‘ en ‚refereren‘. Deze kerncategorieën bevatten allemaal vier niveaus, waarbij duidelijk beschreven wordt wat leerlingen moeten laten zien om aan een bepaald niveau te voldoen. Tenslotte worden aanbevelingen over het gebruik van het beoordelingsschema gedaan en wordt het hele onderzoek kritisch bediscussieerd.

Inhaltsverzeichnis

Zusammenfassung.....	2
Samenvatting.....	2
1. Einleitung.....	5
2. Theoretischer Rahmen	8
2.1 Literarisches Lesen in einer Fremdsprache.....	8
2.2 Interpretation.....	14
2.3 Die <i>Self-Questioning</i> -Methode	19
2.4 Bewertung von Klausuren.....	22
2.4.1 <i>Eindtermen</i>	22
2.4.2 Bestehende Modelle zu literarischen Kompetenzen	25
2.4.3 Qualitätskriterien eines guten Bewertungsschemas.....	27
3. Datenbestand.....	29
3.1 Zusammenfassungen der Kurzgeschichten.....	31
3.1.1 Helm auf.....	31
3.1.2 Stinkefisch.....	32
4. Methodik.....	32
4.1 Datenanalyse	32
4.1.1 Entscheidungen bei der Erstellung des Codebuches.....	32
4.1.2 Codiervorgang.....	36
4.2 Interrater-Reliabilität.....	37
4.2.1 Auswahl der Dialoge zur Berechnung der Interrater-Reliabilität	37
4.2.2 Berechnung der Werte	38
5. Ergebnisse des Codierens	40
6. Diskussion zu den Ergebnissen des Codierens	46
7. Entwicklung eines Bewertungsschemas	51
7.1 Auswahl der Kategorien	51
7.2 Literarischer Wortgebrauch und Textstruktur	53
7.3 Beschreiben.....	55
7.4 Beurteilen.....	57
7.5 Interpretieren.....	59
7.5.1 Hypothetisch	59
7.5.2 Charaktere.....	61
7.5.3 Geschichte.....	64
7.6 Referieren.....	66
7.7 Das gesamte Bewertungsschema	67

7.7.1 Vergleich der verschiedenen Kategorien und ihrer Abstufungen.....	68
7.7.2 Empfehlungen für den Einsatz des Bewertungsschemas.....	70
8. Diskussion.....	74
8.1 Erste Forschungsfrage.....	74
8.2 Zweite Forschungsfrage.....	75
8.3 Dritte Forschungsfrage.....	75
Literaturverzeichnis	81
Anhang.....	85
Ausgewählte Dialoge inkl. Mengenangabe der Codes zur Berechnung der Interrater- Reliabilität.....	85
Codebuch	86
Descriptive: default.....	87
Descriptive: summary	87
Essential question.....	88
Evaluative: character.....	89
Evaluative: text	89
Literary words.....	89
Problem understanding: self assessed.....	90
Problem understanding: without noticing.....	90
Reasoning: evaluative character	90
Reasoning: evaluative text	91
Reasoning: hypothetical.....	91
Reasoning: moral – in story	92
Reasoning: moral – beyond story	92
Reasoning: understanding character	92
Reasoning: understanding story events.....	93
Referential: real world	93
Referential: self.....	94
Transformative reading.....	94
Off topic	95
Voorbeelddialog.....	96
Format schriftlicher Dialoge	98

1. Einleitung

Im Fremdsprachenunterricht Deutsch an niederländischen Sekundarschulen wird seit jeher (Jugend-)Literatur gelesen. Literarisches Lesen hat somit eine gewisse Tradition im niederländischen DaF-Unterricht. Allerdings ist es nicht selbstverständlich, dass Schüler¹ auch die Fähigkeiten besitzen, einen literarischen Text sinngemäß zu lesen. Die Kompetenz, literarische Texte lesen zu können und dabei selbst zu einer Interpretation des Textes zu kommen, muss von Schülern an Sekundarschulen größtenteils noch erworben werden.

In den letzten Jahren gab es mehrmals besorgniserregende Berichte darüber, dass Schüler in den Niederlanden immer weniger lesen. Eine Studie des SCPs zeigte, dass Jugendliche im Alter von dreizehn bis neunzehn Jahren in 2018 durchschnittlich nur noch vierzehn Minuten pro Tag lasen.² Im Vergleich zu einer gleichen Studie aus 2013 war dies eine Abnahme von fast 40%.³ Auch in ‚De Staat van het Onderwijs 2023‘ wurde neulich festgestellt, dass die Basisfähigkeiten, darunter Lesekompetenz, von Schülern sich verschlechtern.⁴ Deswegen läutet De Leescoalitie die Alarmglocken; Es soll auf ganzer Linie in der Bildung von Schülern auf Spaß beim Lesen gesetzt werden.⁵ Diese Entwicklung der Verringerung der Lesezeit von Jugendlichen in den letzten Jahren hat selbstverständlich auch Auswirkungen auf die Lesekompetenz der Schüler.⁶ Somit wird das Fördern der Lesekompetenz eine immer wichtigere Aufgabe für die Schule.

Für die Entwicklung von Lesekompetenz ist es essenziell, dass Schüler verschiedene und vor allem viele Texte lesen.⁷ An den Sekundarschulen in den Niederlanden ist die Lesekompetenz deswegen in allen Sprachen Teil des Lehrplans. Das kann man auf der einen Seite daran erkennen, dass in allen Abiturprüfungen (*centraal schriftelijke examens*) aller modernen Sprachen auf Niveau *havo* und *vwo* Leseverstehen von vor allem Sachtexte, gemäß ‚Domein A: Leesvaardigheid‘ der *Eindtermen* des niederländischen Bildungsministeriums⁸, geprüft wird. Andererseits gibt es zudem in den *Eindtermen* aller modernen Sprachen eine

¹ In dieser Arbeit wird das generische Maskulinum verwendet. Dadurch sollen keinesfalls Mädchen, Frauen und diverse Personen ausgeschlossen werden. Diese Wahl wurde lediglich zur Verbesserung der Leserlichkeit des Textes getroffen.

² Vgl. Schaper, Wennekers, und De Haan, „Trends in Media:Tijd“.

³ Vgl. ebd.

⁴ Vgl. Inspectie van het onderwijs, „De Staat van het Onderwijs 2023“, S. 15.

⁵ Vgl. De Leescoalitie, „Oproep tot een ambitieus leesoffensief“, S. 6.

⁶ Vgl. Inspectie van het onderwijs, „Peil.Leesvaardigheid einde (speciaal) basisonderwijs 2020-2021“, S. 14.

⁷ Vgl. Van der Knaap, *Literatuur en film in het vreemdetalenonderwijs*, S. 62.

⁸ College voor Toetsen en Examens, „Examenprogramma moderne vreemde talen en literatuur havo/vwo“, S. 7.

separate Domäne für Literatur. Kurz zusammengefasst umfasst der Begriff ‚Lesekompetenz‘ also sowohl das Leseverstehen als auch Interpretationsfähigkeiten von literarischen Texten. Die Förderung der Lesekompetenz, insbesondere im Bereich der literarischen Texte, ist dementsprechend ein zentraler Bestandteil des niederländischen Lehrplans für Deutsch als Fremdsprache. Schüler sollen dabei nicht nur in der Lage sein, Texte zu verstehen, sondern auch, diese zu interpretieren und eigene Meinungen dazu zu bilden und begründen zu können. Die vorliegende Studie befasst sich mit den literarischen Aspekten der Lesekompetenz von Schülern, insbesondere mit Blick auf Interpretationsfähigkeiten und -Strategien.

Eine Methode, die die Interpretationsfähigkeiten von Schülern fördern kann, ist die *Self-Questioning*-Methode. Diese Methode erfordert ein bewusstes und aktives Nachdenken über den Text durch das systematische Formulieren von Fragen an den Text und an sich als Leser. Diese Fragen beziehen sich auf den Inhalt, die Charaktere, den Kontext und andere Aspekte des Textes. Die Methode soll den Schülern helfen, ihre eigenen Gedanken und Ideen bezüglich eines Textes zu entwickeln, anstatt, so wie die Schüler das bei Sachtexten als Vorbereitung auf die Abiturprüfungen in den modernen Sprachen lernen, einfach nur Informationen aus dem Text zu extrahieren. Bisherige Studien haben gezeigt, dass die *Self-Questioning*-Methode dazu beitragen kann, die Interpretationsfähigkeiten von Schülern zu verbessern.⁹ Allerdings wurde die Wirkung dieser Lesestrategie auf die Interpretationsfähigkeiten von Schülern in den Niederlanden bisher nur im muttersprachlichen Unterricht untersucht. Es ist also noch unklar, ob die *Self-Questioning*-Methode auch im Fremdsprachenunterricht einen positiven Einfluss auf die Interpretationsfähigkeiten von Schülern hat.

Wenn man von *Eindtermen*, Anforderungen des niederländischen Bildungsministeriums an in diesem Fall den Literaturunterricht, redet, kann man logischerweise davon ausgehen, dass diese auch auf irgendeine Art und Weise geprüft werden müssen. Die Bewertung von Schülerinterpretationen von literarischen Texten stellt daher einen wichtigen Bestandteil des DaF-Curriculums der Sekundarstufe II dar. In Wirklichkeit stellt sich jedoch heraus, dass das Bewerten von Schülerinterpretationen Schwierigkeiten bereitet. Zum Beispiel werden zu literarischen Texten auch Fragen, die das Leseverstehen prüfen, gestellt. Eine Möglichkeit, um Interpretationsfähigkeiten von Schülern zu prüfen, könnten schriftliche Dialoge sein. Mit schriftlichen Dialogen werden in diesem Fall Dialoge, die Schüler nach dem Lesen eines Textes allein schreiben, gemeint. Schüler tun dabei so, als ob sie zwei Leute, die zusammen

⁹ Vgl. Janssen, Braaksma, und Couzijn, „Self-Questioning in the Literature Classroom: Effects on Students’ Interpretation and Appreciation of Short Stories“; Urlaub, „Reading Strategies and Literature Instruction: Teaching Learners to Generate Questions to Foster Literary Reading in the Second Language“.

über den Text diskutieren, sind. Ramachers hat für ihre Forschung Schüler schriftliche Dialoge schreiben lassen und stellt fest, dass “deze taak [wellicht] zelfs kansen [biedt] voor het toetsen van interpretatievaardigheid in het VO.”¹⁰ Folgend auf diese Erkenntnis wird im Rahmen dieser Masterarbeit untersucht, wie ein Bewertungsschema für die Interpretationsfähigkeiten von Schülern in Bezug auf schriftliche Dialoge aussehen könnte.

Die Daten, mit denen im Rahmen dieser Masterarbeit gearbeitet wird, stammen von der Forschung von Ramachers.¹¹ Sie bestehen aus schriftlichen Dialogen, die Schüler aus den Klassen *4havo* und *4vwo* anlässlich zweier Kurzgeschichten geschrieben haben. Einen Dialog schrieben sie vor der Intervention, die unter anderem die *Self-Questioning*-Methode vermittelt hat, und den anderen schrieben sie danach. Diese Daten werden codiert, um herauszufinden, was Schüler in den schriftlichen Dialogen eigentlich überhaupt beim Vorgang des Interpretierens machen. Die Ergebnisse, die daraus folgen, werden anschließend verwendet, um, unter Berücksichtigung der wissenschaftlichen Literatur und der Anforderungen des Bildungsministeriums, ein Bewertungsschema zu erstellen.

Im Rahmen dieser Masterarbeit werde ich mich also mit den Fragen beschäftigen, wie Schüler der Klassen *4havo* und *4vwo* im niederländischen DaF-Unterricht literarische Kurzgeschichten interpretieren und wie man ihre Interpretationen bewerten kann. Konkret lauten die Forschungsfragen für diese Masterarbeit wie folgt:

- Wie interpretieren Schüler literarische Kurzgeschichten im Fremdsprachenunterricht vor und nach einer Unterrichtsintervention, die sich auf *Self-Questioning* richtet?
- Inwiefern hat die *Self-Questioning*-Unterrichtsmethode beim Lesen von Kurzgeschichten einen Effekt auf die Interpretationen der Schüler?
- Wie könnte ein Bewertungsschema für Interpretationsfähigkeiten von Schülern, zu einer Aufgabe, bei der von Schülern einen schriftlichen Dialog zu einer Kurzgeschichte geschrieben wurde, aussehen?

Ziel dieser Arbeit ist es also, einen Beitrag zur aktuellen Diskussion über die Förderung der Lesekompetenz von Schülern im niederländischen DaF-Unterricht zu leisten und dabei gleichzeitig neue Erkenntnisse über die Wirksamkeit der *Self-Questioning*-Methode und schriftlicher Dialoge als Bewertungsmethode zu gewinnen. Im Rahmen dieser Arbeit wird ein Bewertungsschema konzipiert, das von DaF-Lehrern konkret in ihrer Unterrichtspraxis verwendet werden kann.

¹⁰ Ramachers, „Wanderlust, dialogisch leren met korte verhalen binnen het schoolvak Duits“, S. 5. Übersetzung: diese Aufgabe vielleicht Möglichkeiten bietet, Interpretationsfähigkeiten zu prüfen.

¹¹ Ebd.

2. Theoretischer Rahmen

2.1 Literarisches Lesen in einer Fremdsprache

Wenn man von literarischem Lesen in einer Fremdsprache spricht, stellen sich zunächst die Fragen, wozu man überhaupt literarische Texte mit Schülern in einer Fremdsprache lesen würde und welche Arten von Texten eigentlich genau mit Literatur gemeint werden.

Obwohl die meisten Menschen schon ein Gespür dafür haben, was Literatur sein soll, ist die Frage gar nicht so einfach zu beantworten. Schaut man sich den Begriff in einem Lexikon an, stellt sich heraus, dass dort schon keine eindeutige Definition von Literatur vorzufinden ist. Das Metzler Lexikon Literatur führt zum Beispiel drei unterschiedliche Definitionen auf, die von den Herausgebern als die wichtigsten Verwendungsweisen des Ausdrucks Literatur angemerkt werden:¹² „1. L[iteratur] als die Gesamtheit des Geschriebenen [...] 2. L[iteratur] als Gegenstand der L[iteratur]wissenschaft [...] 3. L[iteratur] als Gesamtheit aller Texte von bestimmtem Wert.“¹³ Die erste Definition kann man im schulischen Kontext sofort verwerfen, denn es wird an niederländischen Schulen in allen Sprachen ein Unterschied zwischen Literatur und Sachtexte gemacht und diesen Unterschied bräuchte man nicht machen, wenn alles Geschriebene Literatur wäre. Die Definitionen zwei und drei sind zwar brauchbarer als die erste, allerdings immer noch ziemlich ungenau. Diese Definitionen sind also nicht ausreichend und dementsprechend suchen wir weiter.

Im *Handboek Vreemdetalendidactiek* ist eine weitere interessante Definition von Literatur vorzufinden, denn dort werden alle kreativen Sprachäußerungen als Literatur bezeichnet.¹⁴ Problematisch an dieser Definition ist jedoch, dass dadurch verschiedene Fertigkeiten, wie zum Beispiel Hören und Lesen, beansprucht werden. Im Fremdsprachenunterricht ist das nicht immer praktisch, denn die Lernziele bezogen auf Literatur werden dann noch vielfältiger und somit schwerer zu prüfen. In dieser Arbeit wurde also dazu entschieden, nur geschriebene Texte als Literatur zu deuten. Allerdings ist die Frage, was genau Literatur ist, damit immer noch nicht geklärt.

Der Frage, was Literatur ist, hat Jonathan Culler ein ganzes Kapitel in seinem Buch *Literaturtheorie* gewidmet.¹⁵ Er stellt dabei zunächst fest, dass einerseits die Frage selbst nicht

¹² Vgl. Schweikle und Schweikle, *Metzler Lexikon Literatur. Begriffe und Definitionen*, S. 445.

¹³ Ebd.

¹⁴ Vgl. Dönszelmann et al., *Handboek vreemdetalendidactiek: vertrekpunten - vaardigheden - vakinhoud*, S. 402.

¹⁵ Culler, *Literaturtheorie. Eine kurze Einführung*, S. 32ff.

eindeutig ist und andererseits, dass die Frage kaum eine Rolle zu spielen scheint. Im Unterrichtskontext ist die Frage allerdings durchaus berechtigt, weil vom Ministerium vorgegeben wird, dass alle Schüler für den Unterricht in einer modernen Fremdsprache mindestens drei literarische Werke gelesen haben sollen.¹⁶ Die Frage, was Literatur überhaupt ist, so sagt Culler, könnte auf mehreren Weisen aufgefasst werden; fragt man beispielsweise nach der Funktion von Literatur, oder aber nach den Unterscheidungsmerkmalen im Vergleich zu nicht-literarischen Texten?¹⁷ Im schulischen Kontext stellen sich die Schüler vielleicht eher die Frage nach der Definition von Literatur mit den Unterscheidungsmerkmalen im Hinterkopf; Sie wollen vom Dozenten nämlich wissen, weswegen sie einen bestimmten Text schon und einen anderen aber nicht lesen dürfen. Für die Dozenten dürfte aber die Funktion von Literatur entscheidend sein; Sie haben eine bestimmte Vorstellung davon, was die Schüler beim Lesen von Literatur lernen oder aber entdecken sollen.

Für den Literaturbegriff an niederländischen Sekundarschulen ist es also weniger wichtig, dass er einer formellen Definition entspricht, eher soll er an einigen Eigenschaften und Unterscheidungsmerkmalen festgemacht werden. Culler macht einen Versuch, einige Merkmale von Literatur, auch wenn das Fehlen eines oder mehrerer dieser Merkmale keinen Text ausschließt, aufzuführen. Ein wichtiges Merkmal sei, dass Literatur ein besonderes Verhältnis, das wir fiktional nennen, zur Wirklichkeit darstellt.¹⁸ Auch sei der Sprachgebrauch in literarischen Werken häufiger ein besonderer; laut Culler schreit einem die Sprache eines Textes förmlich an: „Schau mich an! Ich bin Sprache!“¹⁹ Sprachgebrauch könne auch mehrfach kodiert sein, das heißt, dass die Form der Sprache auch etwas über den Inhalt aussagt.²⁰ Zudem sei die Intertextualität ein wichtiges Merkmal von Literatur; Ein literarischer Text befindet sich immer in einer Beziehung zu seinen Vorgängertexten.²¹ Schließlich sei das wichtigste Merkmal eines literarischen Textes sein Kontext; Ein Text wird zu Literatur, wenn man ihn an einer Stelle auffindet, wo man erwartet Literatur aufzufinden.²² Im schulischen Kontext könnte somit gesagt werden, dass ein Text, in dem Moment, in dem ein Lehrer ihn als Literatur bezeichnet, zu Literatur wird. Das niederländische Bildungsministerium traut es den Lehrern offenbar zu, selbst beurteilen zu können, was Literatur ist, denn in den *Eindtermen*

¹⁶ College voor Toetsen en Examen, „Examenprogramma moderne vreemde talen en literatuur havo/vwo“, S. 8.

¹⁷ Vgl. Culler, *Literaturtheorie. Eine kurze Einführung*, S. 35.

¹⁸ Vgl. ebd., S. 49.

¹⁹ Ebd., S. 45.

²⁰ Vgl. ebd., S. 47f.

²¹ Vgl. ebd., S. 53.

²² Vgl. ebd., S. 44.

steht zwar, dass Schüler am Ende ihrer Schulkarriere drei literarische Werke in der jeweiligen Fremdsprache gelesen haben sollen, aber es steht nirgendwo, welche Werke als literarisch aufgefasst werden beziehungsweise welche Merkmale jene Texte haben sollen. Eine Definition von Literatur, die demzufolge für eine Anwendung an Sekundarschulen absolut vertretbar wäre, könnte wie folgt lauten: Literatur im schulischen Kontext sind geschriebene Texte, die als solche durch eine Fachgruppe oder auch einen einzelnen Lehrer, aufgrund verschiedener Merkmale, deren Gewichtung durch das Lehrpersonal auf Basis ihrer Ausbildung und Erfahrung bestimmt wird, festgelegt werden.

Die Frage, weswegen man im Fremdsprachenunterricht literarische Texte lesen würde, ist damit allerdings noch nicht beantwortet worden. Das Lesen von Literatur im niederländischen DaF-Unterricht an Sekundarschulen hat eine lange Tradition, trotzdem ist es wichtig, der Frage, warum Literatur eingesetzt wird, nachzugehen, denn es gibt auch Argumente gegen literarisches Lesen in der Fremdsprache. Zumal weist die Frage auf mögliche Probleme, die man als Lehrer bei der Didaktisierung vom literarischen Lesen in der Fremdsprache beachten soll, hin. Naheliegend ist der Einwand, dass das Lesen von Texten in einer Fremdsprache generell schwieriger als das Lesen in der Muttersprache ist, schlichtweg weil eine Fremdsprache nun einmal eine Fremdsprache ist und man dementsprechend nicht die gleiche Kompetenz besitzt, Wörter und Inhalte von Texten, geschweige denn ihre mehrfach kodierten Zusammenhänge, zu verstehen. In diesem Zusammenhang wird auch oft aufgeworfen, dass der Sprachgebrauch in literarischen Texten zu komplex ist und weit entfernt vom alltäglichen Sprachgebrauch ist.²³ Es gibt allerdings auch Gegenstimmen dazu, die meinen, dass der literarische Sprachgebrauch nicht fundamental anders als der alltägliche Sprachgebrauch ist;²⁴ „[It is] ordinary language [being used] in extraordinary ways.“²⁵ Dem Argument folgend weisen Naji et al. darauf hin, „that the language of literary texts might actually provide an excellent model for language use since it often expresses meaning in profound, skilful, and multilayered ways.“²⁶ Man kann somit sagen, dass gerade der literarische Sprachgebrauch dafür geeignet ist, eine Sprache zu lernen, weil in der Literatur Sprache so bedeutungsvoll eingesetzt wird. Auch kann der Sprachgebrauch für sich schon ein Grund dafür sein, Literatur zu lesen: die Ästhetik der Sprache allein ist interessant genug.

²³ Vgl. Naji, Subramaniam, und White, *New Approaches to Literature for Language Learning*, S. 31.

²⁴ Vgl. Hall, *Literature in Language Education*, S. 17.

²⁵ Ebd., S. 13.

²⁶ Naji, Subramaniam, und White, *New Approaches to Literature for Language Learning*, S. 32.

Obwohl die (Un)gewohnheit vom literarischen Sprachgebrauch also umstritten ist und das Lesen von Literatur in einer Fremdsprache Probleme aufwirft, gibt es auch viele Gründe, um Literatur für den Fremdsprachenunterricht einzusetzen. Literatur kann im DaF-Unterricht verschiedenen Zwecken, die anderswie nicht erreichbar wären, dienen. Im Büchlein *Literatur im Unterricht* führt das Deutschland Institut Amsterdam (DIA) zehn Gründe auf, weswegen Literatur im DaF-Unterricht von Vorteil ist.²⁷ Einerseits werden viele sprachliche Gründe, Literatur im DaF-Unterricht einzusetzen, aufgeführt – beispielsweise: „Literarische Texte bieten die Möglichkeit, abstrakte Wortschatzbereiche (z.B. Denken und Fühlen) in den Unterricht einzubinden.“²⁸ Andererseits wird die Stärkung (inter)kultureller und literarischer Fähigkeiten als Grund für den Einsatz von Literatur im DaF-Unterricht genannt – beispielsweise: „Interkulturelle Kompetenzen und die Fähigkeit des Perspektivenwechsels werden beim Arbeiten mit Literatur trainiert und gestärkt.“²⁹

Eine genauere Angabe zu den Zwecken von literarischen Texten im Schulgebrauch führt Simone Schiedermaier auf, indem sie die Lernziele in vier Kategorien aufteilt:

[L]iterarische Texte [können eingesetzt werden]

1. als Medien zum Spracherwerb – Grammatische Strukturen in literarischen Texten
2. als Medien zum Erwerb von landeskundlichem Wissen – Literatur als authentischer und gleichzeitig subjektiver und emotionaler Zugang zu landeskundlichen Inhalten
3. als Medien zur Auseinandersetzung mit Interkulturalität – Interkulturelle Kompetenz durch Perspektivenwechsel und Ambiguitätstoleranz
4. als Trigger für die Kreativität der Lernenden – handlungs- und produktionsorientierter Literaturunterricht³⁰

Es gibt also genug Gründe, um literarische Texte im Fremdsprachenunterricht einzusetzen. Dabei fällt auf, dass bei den Zielen für den Einsatz von Literatur im Sprachunterricht eine Dichotomie sprachspezifischer und allgemeiner einsetzbarer Lernziele wahrnehmbar ist. Interessanterweise nennt Schiedermaier aber keine grundlegenden literarischen Ziele für den Einsatz von Literatur im Fremdsprachenunterricht. Paran hat schon einen Fokus auf literarische Ziele und hat zur Verdeutlichung der Zweiteilung zwischen literarischen und sprachlichen Fähigkeiten, die beim Einsatz von Literatur im Fremdsprachenunterricht trainiert werden können, eine Graphik, mit Fokus auf Literatur auf der x-Achse und Fokus auf das Erlernen der

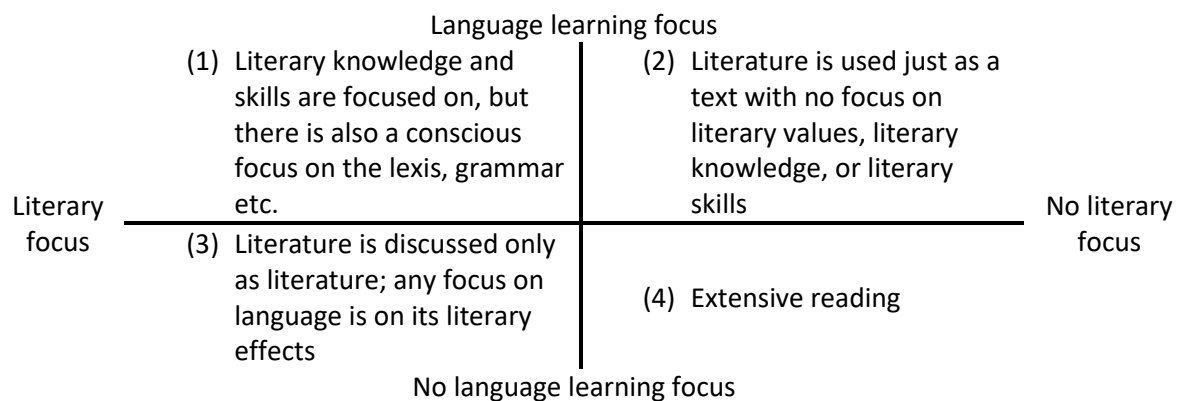
²⁷ Koppenteiner, Jürgen und Schwarz, Eveline. Literatur im DaF/DaZ-Unterricht. Eine Einführung in Theorie und Praxis. Wien 2012, 30-33. Zitiert nach Duitsland Instituut Amsterdam. „Literatur im Unterricht. Unterrichtsmaterial für den Einsatz von deutschsprachiger Literatur am *Dag van de Duitse taal*“, S. 3.

²⁸ Ebd.

²⁹ Ebd.

³⁰ Schiedermaier, „Gegenwartsliteratur im Unterricht Deutsch als Fremdsprache. Ein Vorschlag zur Arbeit mit Textnetzen“, S. 95.

Fremdsprache auf der y-Achse, erstellt (siehe Graphik 1).³¹ Die Einteilung scheint anwendbar, auch wenn die (inter)kulturelle Ebene fehlt. Idealerweise beschäftigt man sich im



Graphik 1: The intersection of literature and language teaching.³²

Fremdsprachenunterricht sowohl mit den sprachlichen als auch mit den literarischen Aspekten der Literatur, entsprechend Nummer (1) in der Graphik von Paran. Literatur ohne Aufmerksamkeit für die Sprache und den Sprachgebrauch (Nummer (3)) im Fremdsprachenunterricht zu lesen wird dem literarischen Text überhaupt nicht gerecht, weil Sprache, wie weiter oben schon besprochen, einen wesentlichen Teil eines literarischen Textes und seiner Bedeutung ausmacht. Genauso wenig Sinn ergibt es, einen literarischen Text ohne Berücksichtigung der literarischen Besonderheiten des Textes und der (nicht) vorhandenen literarischen Fähigkeiten der Schüler zu lesen (Nummer (2)). Wenn nur auf die wortwörtliche Bedeutung eines Textes fokussiert wird, kann man davon ausgehen, dass die Bedeutung des literarischen Textes dabei verloren geht. Wenn das Ziel überhaupt keine literarischen Teilziele umfasst, bräuchte man keine Literatur lesen; Dafür reichen Sachtexte. Wenn man also entscheidet, literarische Texte im Fremdsprachenunterricht einzusetzen, müssen dabei auch literarische Ziele verfolgt werden.

Bloemert et al. haben für den niederländischen Kontext, weiterführend auf die Forschung von unter anderem Paran, untersucht, wie niederländische Dozenten Literatur im Unterricht Englisch als Fremdsprache einsetzen.³³ Die Studie hat eine Graphik, die vier verschiedenen Ansätze (textbezogen, leserbezogen, kontextbezogen und sprachbezogen) für den Literaturunterricht umfasst, ergeben (siehe Graphik 2).³⁴ Obwohl diese Studie unter Dozenten,

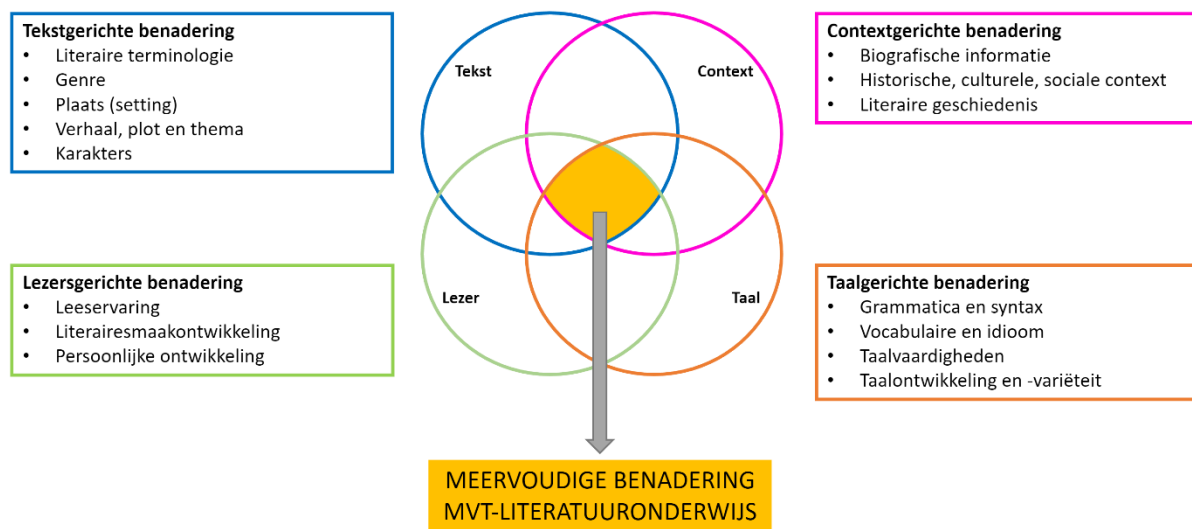
³¹ Paran, „The Role of Literature in Instructed Foreign Language Learning and Teaching: An Evidence-Based Survey“, S. 7.

³² Ebd.

³³ Bloemert, Jansen, und Grift, „Exploring EFL Literature Approaches in Dutch Secondary Education“.

³⁴ Vgl. Bloemert, De Goede, und Goedhard, „Via een meervoudige benadering naar verrijkt mvt-literatuuronderwijs“, S. 13.

die Englisch anstelle von Deutsch in den Niederlanden unterrichten, durchgeführt wurde, sind die Resultate interessant, denn die *Eindtermen* sind in den Niederlanden für alle modernen Fremdsprachen die Gleichen. Bloemert et al. plädieren, basierend auf den Ergebnissen ihrer Studie und der bestehenden wissenschaftlichen Literatur, für die Einbeziehung aller vier der Ansätze im Fremdsprachenunterricht, um zu erreichen, dass die Schüler so viel wie möglich vom Literaturunterricht lernen³⁵ und zudem ein besseres Verständnis von literarischer Prosa zu erlangen.³⁶ Lehrer sollten also idealerweise alle vier Ansätze in den Literaturunterricht mit einbeziehen und dafür auch passende Lernziele konzipieren. Dabei darf auf keinen Fall vergessen werden, dass der Einsatz von Literatur im Fremdsprachenunterricht eine andere Herangehensweise an den Text von den Schülern erfordert.



Graphik 2: Mehrfache Herangehensweise für den Literaturunterricht in den modernen Fremdsprachen in den Niederlanden.³⁷

Vorhin wurde bereits angesprochen, dass Lesekompetenz zwei unterschiedliche Komponenten, nämlich einerseits Leseverstehen von Sachtexten und andererseits Interpretationsfähigkeiten literarischer Texte, umfasst. Rosenblatt beschreibt diese Unterschiede auch, indem sie zwei unterschiedliche Leseweisen, nämlich *effeferent reading* gegenüber *aesthetic reading*, feststellt. Sie sagt, dass man sich beim effeferenten Lesen damit beschäftigt, welche faktischen Informationen man dem Text entnehmen kann, während der Leser beim ästhetischen Lesen sich damit befasst, was während des Lesens zwischen dem Text

³⁵ Vgl. ebd., S. 12.

³⁶ Vgl. Bloemert, Jansen, und Grift, „Exploring EFL Literature Approaches in Dutch Secondary Education“, S. 184.

³⁷ Bloemert, De Goede, und Goedhard, „Via een meervoudige benadering naar verrijkt mvt-literatuuronderwijs“.

und dem Leser passiert.³⁸ Ästhetisches Lesen ist dementsprechend eine subjektivere Weise des Lesens, bei der der Leser seine eigene individuelle Bedeutung oder auch Interpretation des Textes erschafft. Die beiden Komponenten von Lesekompetenz unterscheiden sich somit nicht nur auf Textebene (Sachtext vs. literarischer Text), sondern sie fordern vom Leser eine völlig andere Herangehensweise an den Text. In den Niederlanden lernen Schüler bei verschiedenen Sprachen und im Verlaufe der Jahre mehrfach wiederholt, wie sie Sachtexte lesen sollen, damit sie die Kompetenz für efferentes Lesen erwerben. Die notwendigen Kompetenzen für ästhetisches Lesen werden jedoch oftmals im Rahmen des Leseunterrichts vernachlässigt, wie auch Lazar beobachtet: „[T]oo often when reading literature, students are expected, as if by osmosis, to acquire a kind of competence in reading literature.“³⁹ Lazar stellt zudem fest, dass es extrem schwierig ist, alle Komponenten, die literarische Lesekompetenz umfassen, aufzuzeichnen.⁴⁰ Er erwähnt allerdings direkt danach, dass es für Dozenten äußerst wichtig ist, wenigstens einige dieser Komponenten zu identifizieren, eben weil man sich dessen bewusst sein soll, dass literarische Lesekompetenzen von Schülern zunächst erlernt werden müssen.⁴¹ Interpretation kann als eine dieser Kernkomponenten literarischer Lesekompetenz angesehen werden.⁴² Literarische Texte stellen nämlich an den Leser bestimmte Anforderungen, die eine Interpretation erfordern.

2.2 Interpretation

Schüler werden vom niederländischen Bildungsministerium dazu aufgefordert, literarische Texte interpretieren zu können.⁴³ Zudem würden Schüler auch selbst gerne interpretierend über literarische Texte reden.⁴⁴ Dabei stellen sich die Fragen, was dieses ‚Interpretieren‘ eigentlich genau beinhaltet und inwieweit Schüler diese Kompetenz besitzen müssen. Mit anderen Worten ist es wichtig zu wissen, was genau Schüler zeigen müssen, wenn sie einen Text interpretieren beziehungsweise was dabei von ihnen überhaupt erwartet werden kann.

Interpretation als Begriff ist ziemlich vage, denn welche Handlungen umfasst das Interpretieren, um zu einer Interpretation zu gelangen? Culler beschreibt das Interpretieren als

³⁸ Vgl. Rosenblatt, *The Reader, the Text, the Poem: The Transactional Theory of the Literary Work*, S. 24f.

³⁹ Lazar, *Literature and Language Teaching: A Guide for Teachers and Trainers*, S. 13.

⁴⁰ Vgl. ebd.

⁴¹ Vgl. ebd.

⁴² Vgl. Cornelissen, *‘Maar als je erover nadenkt ...’: een jaar literatuuronderwijs in groepen 7 en 8 van de basisschool*, S. 16.

⁴³ Vgl. College voor Toetsen en Examens, *„Examenprogramma moderne vreemde talen en literatuur havo/vwo“*, S. 8.

⁴⁴ Vgl. De Nood und Van Herten, *„‘Een open einde kan leuk zijn, maar dit is wel heel open’: Waarover leerlingen praten, als zij over verhalen praten“*, S. 36.

ein ‚Worum-Spiel‘.⁴⁵ Beim Interpretieren sollte man immer Worum-Fragen an den Text stellen und dabei beim Beantworten der Frage mit dem Spiel mitspielen wollen⁴⁶ (und können); im Falle der Kurzgeschichte *Helm auf* (einer der für diese Studie verwendeten Kurzgeschichten) würde es für eine Interpretation zum Beispiel nicht reichen, die Frage, worum es in dem Text geht, mit: ‚Es geht in der Geschichte um einen Jungen mit Helm‘ zu beantworten. Dann wird das Spiel des Interpretierens laut Culler nicht gespielt.⁴⁷ Vielmehr sollte diese Worum-Frage beispielsweise mit: ‚In der Geschichte geht es um einen Jungen, der einen persönlichen Schicksalsschlag nicht gut verarbeiten kann‘ beantwortet werden. Eine Interpretation soll also nicht eine faktische, deskriptive Aussage über die Geschehnisse im Text sein, sondern weiter als der eigentliche Text führen. Culler merkt dabei an:

Was im Interpretationsspiel zählt, ist nicht die letztlich gefundene Antwort [...]. Was vielmehr zählt, ist die Art und Weise, wie man dahin kommt, was man mit den einzelnen Details des Texts macht, wenn man sie zur eigenen Antwort in Beziehung bringt, und wie ansonsten rätselhafte Details von der eigenen Interpretation erhellt werden.⁴⁸

Dass beim Interpretieren das Zustandekommen der Interpretation wichtiger als das Endergebnis, die Interpretation, ist, macht es für die Unterrichtspraxis schwierig, Interpretation zu prüfen, denn eigentlich soll dabei der Prozess dahin auch beobachtet und bewertet werden.

Marshall et al. verweisen auf Rabinowitz⁴⁹, der zudem ein anderes Problem, wenn es um das Interpretieren von literarischen Texten im schulischen Kontext geht, feststellt. Dozenten hätten einen Text, wenn sie ihn im Unterricht besprechen, selbst schon mehrmals gelesen und nehmen deswegen einen retrospektiven Blickwinkel ein.⁵⁰ Die Interpretation von Dozenten sieht deswegen zwangsläufig anders als die der Schüler, die den Text zum ersten Mal erblicken, aus. Dozenten würden deswegen im Unterricht und im Dialog mit den Schülern unbewusst auf Interpretationen, die Muster des Textes zu verstehen versuchen, fokussieren, während das ein Interpretationsverfahren ist, das erst stattfinden kann, wenn man den Text schon (mehrmals) gelesen hat.⁵¹ Schüler seien dagegen eher damit beschäftigt, was in einem Text als nächstes passieren wird und warum.⁵² Dabei berufen sie sich auf ihre eigenen Erfahrungen mit der Welt; „When we predict what a character will do, for example, we base that prediction on our

⁴⁵ Vgl. Culler, *Literaturtheorie. Eine kurze Einführung*, S. 95.

⁴⁶ Vgl. ebd.

⁴⁷ Vgl. ebd.

⁴⁸ Ebd., S. 96.

⁴⁹ Vgl. Rabinowitz, Peter. *A thousand times and never like: Re-reading for class*. Chicago 1991. Zitiert nach Marshall, Smagorinsky, und Smith, *The Language of Interpretation: Patterns of Discourse in Discussions of Literature*, S. 126.

⁵⁰ Vgl. Rabinowitz, Peter. *A thousand times and never like: Re-reading for class*. Chicago 1991. Zitiert nach ebd.

⁵¹ Vgl. ebd.

⁵² Vgl. ebd., S. 126f.

understanding of human nature, on what we have done or would do in similar situations, or on what people we know who remind us of the character have done or would do.”⁵³ Marshall et al. kommentieren, dass sehr viel Potenzial für Diskussionen über Literatur und somit für Interpretationen von der Seite der Schüler verloren geht, wenn Dozenten von Schülern fordern, dass sie sich beim Interpretieren vor allem auf die Zusammenhänge des Textes fokussieren.⁵⁴

Was aber darf denn von Schülern als Leistung beim Interpretieren genau erwartet werden? Was können sie vielleicht schon? Was noch nicht? Und wie kann denen geholfen werden, das Interpretieren noch besser zu erlernen? Verschiedene Indizien dafür können in der wissenschaftlichen Literatur zu Studien, die unerfahrene mit erfahrenen Lesern vergleichen, gefunden werden. Die meisten dieser Studien wurden an Muttersprachlern durchgeführt, doch das stellt kein Problem dar; Verschiedene Studien bestätigen, dass Kompetenzen, die Schüler auf dem Gebiet von Interpretation in der Muttersprache entwickelt haben, zumindest zu einem gewissen Maße in die Fremdsprache übertragbar sind. Unter anderem Altinay-Sezen hat genau diese Übertragbarkeit untersucht und stellt fest, dass „[sich] die in der L1 erworbene literarische Kompetenz [...] im Lesen von Texten in der L2 widerspiegeln lässt.“⁵⁵ Auch Fecteau kommt zu einem ähnlichen Ergebnis, denn ihre Studie zur Übertragbarkeit von literarischen Kompetenzen in eine Fremdsprache ergab, dass die Fähigkeiten von Schülern in der Muttersprache mehr voraussagten über literarische Kompetenz in der Fremdsprache als das Leseverständnis von Schülern in der Fremdsprache.⁵⁶ Allerdings muss auch die mehrfach bestätigte *Threshold-Theory*, die besagt, dass Schüler erst ab einem bestimmten Niveau in der Fremdsprache die Fähigkeiten aus der Muttersprache in die Fremdsprache übertragen können, berücksichtigt werden.⁵⁷ Aus den Ergebnissen der Studien von Altinay-Sezen und Fecteau kann jedoch schon geschlossen werden, dass Studien zu Unterschieden zwischen erfahrenen und unerfahreneren Lesern in der Muttersprache ein Indiz dafür sein können, was man von Schülern auf dem Gebiet von Interpretationen in der Fremdsprache erwarten und nicht erwarten darf.

⁵³ Ebd., S. 127.

⁵⁴ Vgl. ebd.

⁵⁵ Altinay-Sezen, „Literarische Kompetenz – ein sprachübergreifendes Phänomen? Eine empirische Untersuchung über die potenzielle Übertragung der muttersprachlichen literarischen Kompetenz in die L2“, S. 58.

⁵⁶ Vgl. Fecteau, „First-and Second-Language Reading Comprehension of Literary Texts“, S. 484.

⁵⁷ Lee und Schallert, „The Relative Contribution of L2 Language Proficiency and L1 Reading Ability to L2 Reading Performance: A Test of the Threshold Hypothesis in an EFL Context“; Schoonen, Hulstijn, und Bossers, „Metacognitive and Language-Specific Knowledge in Native and Foreign Language Reading Comprehension: An Empirical Study Among Dutch Students in Grades 6, 8 and 10“.

In diesem Zusammenhang ist es interessant, eine schon etwas ältere Studie von Elise Ann Earthman zu betrachten. Sie hat in Amerika unter Muttersprachlern anhand eines *Think-Aloud*-Protokolls untersucht, wie *College Freshman* (Hochschulanfänger) im Vergleich zu *Graduate Students* (Studenten kurz vor dem Abschluss) einen Text interpretieren. Die von ihr für die Studie verwendeten Texte waren zwei Kurzgeschichten und zwei Gedichte.⁵⁸ Die Studie von Earthman umfasste folgende Ergebnisse: erstens stellte sich heraus, dass *Graduate Students* sich mehr als *College Freshman* auf Lücken, die schwieriger zu füllen sind, fokussierten.⁵⁹ Zudem war es für die *Graduate Students* einfacher, Informationen aus dem Text zu sammeln und daraus eine Schlussfolgerung dafür, was in die Lücke hinein passen könnte, zu ziehen.⁶⁰ Unerfahrenere Studenten befassten sich also weniger mit Lücken, die für die Story wichtig sind, und fanden es schwieriger, die Lücken mit einer Interpretation zu füllen. Zweitens stellte sich heraus, dass *College Freshman* beim Lesen der Texte schnell eine bestimmte Perspektive einnahmen und diese auch beibehielten, während die *Graduate Students*, ohne Ausnahme, verschiedene Perspektiven des Textes beleuchteten und ihre Meinung während des Lesens dementsprechend änderten.⁶¹ Dass erfahrenere Leser oft mehrere Perspektive einnehmen können, während unerfahrene Leser diese Fähigkeit kaum aufzeigen, wird unterstützt durch eine Studie von Els Andringa, die im niederländischen Kontext im Muttersprachenunterricht durchgeführt wurde.⁶²

Im Gegensatz zu den beiden zuvor aufgeführten Studien, vergleichen Janssen et al. keine unerfahrenen und erfahrenen Leser aus verschiedenen Altersgruppen untereinander, sondern sie vergleichen Schüler, die schlechtere Noten bekommen, mit Schülern, die bessere Noten bekommen, untereinander. Die niederländischen Schüler, die an dieser Studie teilnahmen, gingen in die 10. Klassenstufe an verschiedenen Schulen.⁶³ Dieser Vergleich ist interessant, denn die Teilnehmergruppe der Studie von Janssen et al. ist ungefähr die Gleiche als von der vorliegenden Studie. Die Ergebnisse sagen dementsprechend viel darüber aus, was leistungsschwache und leistungsstarke Schüler in diesem Alter in der Muttersprache auf dem Gebiet von literarischer Kompetenz bereits können und wie diese Unterschiede auch in der Fremdsprache aussehen könnten. Ein wichtiges Ergebnis der Studie von Janssen et al. ist, dass

⁵⁸ Vgl. Earthman, „Creating the Virtual Work: Readers’ Processes in Understanding Literary Texts“, S. 356.

⁵⁹ Vgl. ebd., S. 351.

⁶⁰ Vgl. ebd., S. 362.

⁶¹ Vgl. ebd., S. 373.

⁶² Vgl. Andringa, „Ontwikkelingen in het omgaan met literaire complexiteit“, S. 282.

⁶³ Vgl. Janssen, Braaksma, und Rijlaarsdam, „Literary Reading Activities of Good and Weak Students: A Think Aloud Study“, S. 38.

leistungsstarke Schüler variiertere Leseaktivitäten aufzeigen; ihre Aussagen betreffen mehr unterschiedliche Kategorien (z.B. nacherzählen, Probleme erkennen) als die Aussagen von leistungsschwachen Schülern.⁶⁴ Außerdem fanden die Forscher heraus, dass leistungsschwächere Schüler viel öfter Teile eines Textes nacherzählen, während leistungsstärkere Schüler sich häufiger emotional und evaluativ mit dem Text auseinandersetzen.⁶⁵ Die leistungsstarken Schüler beschäftigen sich also viel persönlicher und subjektiver mit dem Text.⁶⁶

Eine Studie, die sich mit literarischem Lesen in einer Fremdsprache durch Schüler im Alter von sechzehn Jahren auseinandergesetzt hat, ist von Kadir et al.⁶⁷ In dieser Studie wurden Schüler, die die Fremdsprache gut beherrschten, verglichen mit Schülern, die die Fremdsprache schlechter beherrschten. Die Ergebnisse dieser Studie zeigen einerseits, dass Schüler, die die Sprache besser verstanden, variiertere Aussagen zum Text auf ästhetische, narrative und evaluative Ebene trafen, während Schüler, die ein schlechteres Sprachverständnis hatten, sich viel mehr auf einen Aspekt des Textes, nämlich die narrative Ebene, konzentrierten.⁶⁸ Andererseits war es aber so, dass sowohl die Schüler, die ein gutes Sprachverständnis hatten, als auch die Schüler, die ein schlechteres Sprachverständnis hatten, sowieso eher Aussagen zu der narrativen Ebene des Textes als zu der ästhetischen und evaluativen Ebene trafen.⁶⁹ Die Forscher dieser Studie stellen folglich fest, dass die Schüler sich eher oberflächlich als tiefer mit dem Text engagierten.⁷⁰

Ein Faktor, der bisher noch nicht erwähnt wurde, der aber auch eine wichtige Rolle spielt, wenn es darauf ankommt, ob Schüler zu einer Interpretation eines Textes kommen, ist die Aufgabenstellung. Mason et al. führten eine Studie an Schülern in deren Muttersprache durch, bei dem die Schüler in einem Essay ihre Interpretationen zu einer bestimmten literarischen Kurzgeschichte aufschreiben sollten. Ein interessantes Ergebnis dieser Studie ist, dass eine spezifische Aufgabenstellung, die genau sagte, was von den Schülern erwartet wurde und wozu sie sich alles äußern konnten, zu besseren und tieferen Interpretationen als eine allgemeine Aufgabenstellung, die von Schülern forderte, den gelesenen Text zu kommentieren, führte.⁷¹

⁶⁴ Vgl. ebd., S. 42.

⁶⁵ Vgl. ebd.

⁶⁶ Vgl. ebd.

⁶⁷ Kadir, Maasum, und Vengadasamy, „Conceptualising Literary Responses of L2 Readers: An Exploratory Study“.

⁶⁸ Vgl. ebd., S. 567.

⁶⁹ Vgl. ebd.

⁷⁰ Vgl. ebd.

⁷¹ Vgl. Mason, Scirica, und Salvi, „Effects of Beliefs about Meaning Construction and Task Instructions on Interpretation of Narrative Text“, S. 428.

2.3 Die *Self-Questioning*-Methode

Die vorherigen Kapitel haben sich damit beschäftigt, was Literatur und Interpretation ist, wozu man Literatur im Fremdsprachenunterricht lesen würde und welche Aktivitäten (un)erfahrene Leser beim Interpretieren literarischer Texte aufweisen. Wie aber vorhin schon kurz angedeutet wurde, wird zu oft von Schülern erwartet, dass sie aus dem Nichts heraus die Fähigkeit zu interpretieren entwickeln. De Nood und Van Herten stellen fest, dass Schüler von alleine schon interpretierend über literarische Texte versuchen zu reden, allerdings fänden Schüler es schwierig, dabei zu argumentieren und tiefere Interpretationen zu erwerben.⁷² Schüler brauchen also Hilfe dabei, literarische Kompetenz auf dem Gebiet von Interpretieren zu erlangen. Es gibt verschiedene Methoden, die dabei helfen können, interpretieren zu erlernen. Eine dieser Methoden ist die *Self-Questioning*-Methode. In diesem Kapitel wird besprochen, was *Self-Questioning* überhaupt beinhaltet und wie sie zu einer Entwicklung von literarischer Kompetenz beitragen kann.

Ausgangspunkt der *Self-Questioning*-Methode ist, dass Fragen stellen nicht nur zum literarischen Lesen gehört, sondern sogar eine fundamentale Komponente davon ist.⁷³ Allerdings sind es vor allem die erfahrenen Leser, die viele Fragen an den Text stellen und infolgedessen beim Versuch, diese Fragen zu beantworten, anfangen zu interpretieren.⁷⁴ Für unerfahrenere Leser könnte es also vorteilhaft sein, zu lernen, mehr Fragen an einen Text zu stellen. Genau dies ist das Ziel der *Self-Questioning*-Methode. Sie beinhaltet, dass sich Schüler während des Lesens und im Anschluss daran Fragen mit Bezug auf den Inhalt des Textes stellen. Schüler dürfen zwar alle Fragen, die in ihnen aufkommen, stellen, allerdings sind jene Fragen, die im Text nicht beantwortet werden und die zentrale Inhalte des Textes betreffen, die sogenannten essenziellen Fragen, am wichtigsten.⁷⁵ Der Unterschied der *Self-Questioning*-Methode zum herkömmlichen Literaturunterricht besteht darin, dass im herkömmlichen Literaturunterricht Dozenten meistens die Fragen stellen. Das Problem dabei liegt darin, dass auch wenn die Fragen vom Lehrer gute Fragen, die die Interpretation von der Seite der Schüler fördern, sind, sie dazu führen, dass die Schüler versuchen, den Erwartungen des Lehrers zu

⁷² Vgl. De Nood und Van Herten, „‘Een open einde kan leuk zijn, maar dit is wel heel open’: Waarover leerlingen praten, als zij over verhalen praten“, S. 34 und S. 36.

⁷³ Vgl. Janssen, „Instruction in Self-Questioning as a Literary Reading Strategy: An Exploration of Empirical Research“, S. 99.

⁷⁴ Vgl. Janssen, Braaksma, und Couzijn, „Self-Questioning in the Literature Classroom: Effects on Students’ Interpretation and Appreciation of Short Stories“, S. 95.

⁷⁵ Vgl. ebd., S. 100.

entsprechen, statt sich damit auseinanderzusetzen, welche Inhalte des Textes sie selbst besonders zum Nachdenken angeregt haben.⁷⁶

Die *Self-Questioning*-Methode war schon häufiger Gegenstand der Forschung, jedoch wurde in den meisten Studien *Self-Questioning* bei nicht-literarischen Texten untersucht. Die Ergebnisse dieser Studien waren stets positiv; Das Textverständnis der Leser wurde nach einer *Self-Questioning*-Intervention verbessert.⁷⁷ Rosenshine et al. liefern drei Erklärungen dafür, weshalb das Formulieren von Fragen an den Text so erfolgreich ist, denn (1) Leser werden stimuliert, sich aktiv mit ihrem Leseprozess auseinanderzusetzen, (2) bereits existierendes Hintergrundwissen wird aktiviert und (3) Leser überwachen aktiv ihr Verständnis vom Text.⁷⁸ Es wäre also nicht unmöglich, dass eine Methode, die das Stellen von Fragen an den Text fördert, einen ähnlich positiven Effekt auf das Textverständnis von literarischen Texten haben könnte.

Janssen et al. identifizieren die Lücke, dass *Self-Questioning* im Bereiche des literarischen Lesens noch nicht untersucht wurde, und fragen sich, welchen Effekt die *Self-Questioning*-Methode auf die Interpretationsfähigkeit von Schülern im niederländischen Muttersprachenunterricht hat.⁷⁹ Außerdem erforschen sie, welchen Effekt die *Self-Questioning*-Methode auf die Wertschätzung der Texte durch die Schüler hat. In dieser Studie wurden unter anderem zwei Gruppen, jeweils ohne Instruktion (ungeführt) und geführt durch eine Instruktion bezüglich des Generierens von Fragen, untereinander verglichen. In beiden Fällen wurden die Schüler angeregt, selbst Fragen an den Text zu stellen, doch bei der geführten Instruktion wurden mit den Schülern auch noch Beispiele von guten und schlechten Fragen und Antworten besprochen.⁸⁰ Interessanterweise stellte sich heraus, dass Schüler, die sich mit den Kurzgeschichten anhand der *Self-Questioning*-Methode ungeführt auseinandergesetzt hatten, die Kurzgeschichten erfolgreicher interpretierten als jene, die die *Self-Questioning*-Methode anhand von Beispielen zu guten und schlechten Fragen und Antworten angewendet hatten.⁸¹ Zusätzlich konnte die Schlussfolgerung gezogen werden, dass die *Self-Questioning*-Methode zu einer höheren Wertschätzung der Texte führte. Das ist ein

⁷⁶ Vgl. Singer und Donlan, „Active Comprehension: Problem-Solving Schema with Question Generation for Comprehension of Complex Short Stories“, S. 171.

⁷⁷ Vgl. Rosenshine, Meister, und Chapman, „Teaching Students to Generate Questions: A Review of the Intervention Studies“, S. 195.

⁷⁸ Vgl. ebd., S. 183.

⁷⁹ Janssen, Braaksma, und Couzijn, „Self-Questioning in the Literature Classroom: Effects on Students' Interpretation and Appreciation of Short Stories“.

⁸⁰ Vgl. ebd., S. 102.

⁸¹ Vgl. ebd., S. 106.

wichtiges Ergebnis, weil es für Unterrichtsmethoden auch immer von Bedeutung ist, welchen Effekt sie auf den erlebten Spaß am Lesen von Schülern haben. Da die Wertschätzung von Texten durch die *Self-Questioning*-Methode nicht beeinträchtigt wird, sondern sogar noch gesteigert wird, gibt es keinen Grund die Methode nicht zu implementieren. Gerade weil im Moment stark auf Spaß am Lesen gesetzt wird, wie in der Einleitung schon erwähnt wurde.

Zu den Effekten der *Self-Questioning*-Methode im Fremdsprachenunterricht zu literarischen Texten gibt es lediglich eine Studie, nämlich von Urlaub.⁸² Es stellte sich darin heraus, dass *Self-Questioning* einen positiven Effekt auf Interpretationsfähigkeiten hat und dass *Self-Questioning* möglicherweise das Sprachniveau einigermaßen kompensieren kann. Die Studie von Urlaub wurde durchgeführt unter amerikanischen DaF-Studenten, die im Pre- und Posttest eine Kurzgeschichte lasen und dazu einen Essay schrieben. Es wurde eine Kontrollgruppe, die normalen faktisch-historischen Unterricht gerichtet auf die Kurzgeschichten bekamen, verglichen mit einer Experimentgruppe, die gezielt lernte, Fragen an den Text (bezüglich Charaktere, Handlungsort, Geschehen, Handlungsgegenstände und interkulturelle Spannungen zwischen Leser und Text) zu stellen. Überraschenderweise ergab die Studie, dass Studenten der Experimentgruppe, die ein niedrigeres Sprachniveau (*intermediate*) als ihre Kommilitonen (*advanced*) aufwiesen, im Posttest bessere Noten als die *Advanced Students* bekamen.⁸³ Im Pretest waren die Noten der *Intermediate Students* noch, wie erwartet, niedriger als die Noten der *Advanced Students*. Urlaub vermutet aufgrund dessen, dass die *Self-Questioning*-Methode möglicherweise für bestimmte linguistische Defizite kompensiert, weswegen es möglich war, dass die Studenten mit niedrigerem Sprachniveau bessere Noten erhielten.⁸⁴ Inwieweit dieses Ergebnis übertragbar ist und ob es auch für Schüler im Kontext dieser Studie gelten könnte, ist jedoch ungewiss. Es kann dementsprechend nicht mit Sicherheit davon ausgegangen werden. Zudem erwies sich in der Studie von Urlaub, dass die Essays der Experimentgruppe im Posttest signifikant besser als die der Kontrollgruppe punkteten. Aus dem Ergebnis kann folglich geschlossen werden, dass die *Self-Questioning*-Methode auch im Fremdsprachenunterricht einen positiven Einfluss auf die Interpretationsfähigkeiten der Studenten hat.⁸⁵

⁸² Urlaub, „Reading Strategies and Literature Instruction: Teaching Learners to Generate Questions to Foster Literary Reading in the Second Language“.

⁸³ Vgl. ebd., S. 301.

⁸⁴ Vgl. ebd., S. 302.

⁸⁵ Vgl. ebd., S. 301.

2.4 Bewertung von Klausuren

Bisher haben wir uns angeschaut, was Interpretation ist und wie Schülern dabei geholfen werden könnte, das Interpretieren zu erlernen. Als Ergebnis dieser Arbeit soll abschließend ein Bewertungsschema erstellt werden und dafür sind noch zwei Elemente von besonderer Wichtigkeit, nämlich einerseits die Frage, wie im Moment die literarischen Kompetenzen von Schülern bewertet werden beziehungsweise welche Modelle und Schemas es dafür gibt und auf welchen (wissenschaftlichen) Hintergrund diese basieren. Andererseits gibt es natürlich Qualitätskriterien dafür, wie ein gutes Bewertungsschema aussehen soll. In den folgenden Kapiteln sollen diese Elemente besprochen werden.

2.4.1 *Eindtermen*

Das die literarischen Kompetenzen von Schülern bewertet werden sollen, wird in den *Eindtermen* vom niederländischen Bildungsministerium vorgegeben. Wie aber bereits erwähnt wurde, sind diese *Eindtermen* (siehe Tabelle 1) aktuell ziemlich vage, was dazu führt, dass sie sehr unterschiedlich von Dozenten ausgelegt und zudem auf sehr viele Arten und Weisen geprüft werden können. Es wird unter anderem deswegen schon länger darüber gesprochen,

E1	Havo Vwo	Literaire ontwikkeling	De kandidaat kan beargumenteerd verslag uitbrengen van zijn leeservaringen met ten minste drie literaire werken. ⁸⁶
E2	Vwo	Literaire begrippen	De kandidaat kan literaire tekstsoorten herkennen en onderscheiden, en literaire begrippen hanteren in de interpretatie van literaire teksten. ⁸⁷
E3	Vwo	Literatuurgeschiedenis	De kandidaat kan een overzicht geven van de hoofdlijnen van de literatuurgeschiedenis en de gelezen literaire werken plaatsen in dit historisch perspectief. ⁸⁸

Tabelle 1: Domäne E: Literatur der aktuellen *Eindtermen havo/vwo* für die modernen Fremdsprachen.⁸⁹

die *Eindtermen* zu aktualisieren.⁹⁰ Die aktuellen *Eindtermen* stammen aus 2007 und wurden damals zum ersten Mal an den gemeinsamen europäischen Referenzrahmen (GER)

⁸⁶ Übersetzung: „E1: Literarische Entwicklung. Der Kandidat kann argumentierend von seiner Leseerfahrungen mit mindestens drei literarischen Werke berichten.“

⁸⁷ Übersetzung: „E2: Literarische Begriffe. Der Kandidat kann literarische Textsorten wiedererkennen und unterscheiden, und literarische Begriffe bei der Interpretation von literarischen Texten verwenden.“

⁸⁸ Übersetzung: „E3: Literaturgeschichte. Der Kandidat kann einen Überblick der Kernpunkte der Literaturgeschichte geben und die gelesenen literarischen Werke in einem historischen Kontext einordnen.“

⁸⁹ College voor Toetsen en Examens, „Examenprogramma moderne vreemde talen en literatuur havo/vwo“, S. 8.

⁹⁰ Siehe beispielsweise: <https://curriculum.nu> „Voorstellen ontwikkelteam Engels Moderne vreemde talen. Voorstel voor de basis van de herziening van de kerndoelen en eindtermen van de leraren en schoolleiders uit het ontwikkelteam Engels/Moderne Vreemde Talen“; <https://www.actualisatiemvt.nl> „Actualisatie examenprogramma's moderne vreemde talen“.

gekoppelt.⁹¹ Inzwischen wurde der GER im Jahr 2020 unter anderem im Bereiche Literatur erneuert und spätestens mit dieser Erneuerung kann man feststellen, dass die *Eindtermen* aus dem Jahr 2007 veraltet sind.

In dem aus 2001 stammenden GER wird Literatur nur am Rande unter dem Nenner ‚ästhetischer Sprachgebrauch‘ erwähnt.⁹² Im neuen, aktualisierten GER nimmt Literatur einen höheren Stellenwert ein, denn es wurden drei Bereiche, die Literatur mit einbeziehen, zum GER ergänzt, nämlich (1) Lesen als Freizeitaktivität, (2) eine persönliche Meinung zu einem kreativen Text äußern und (3) analysieren und kritisieren kreativer Texte. Für diese Bereiche wurde ausformuliert, was Lerner der Fremdsprache (ab jetzt: Schüler) in der Fremdsprache zu welchem Niveau genau können müssen. Beispielsweise wurde für den Bereich ‚analysieren und kritisieren kreativer Texte‘ festgelegt, dass Schüler auf dem Niveau B1 folgendes können müssen:

- (a) Can point out the most important episodes and events in a clearly structured narrative in everyday language and explain the significance of events and the connections between them.
- (b) Can describe the key themes and characters in short narratives involving familiar situations that contain only high frequency everyday language.⁹³

Der GER verbindet diese Ziele mit einem definierten Sprachniveau. Schüler müssten diese Ziele laut GER also in der Fremdsprache erreichen können.

Für diese Arbeit ist es jedoch nicht relevant, zu weit auf die unterschiedlichen Niveaus der Bereiche, die Literatur umfassen, des GERs einzugehen, aus dem Grund, dass die schriftlichen Dialoge, die die Schüler anlässlich der Intervention geschrieben haben, auf niederländisch geschrieben worden sind. Wohl kann festgestellt werden, dass im aktualisierten GER Literatur als wesentlicher Inhalt des Fremdsprachenunterrichts deutlicher als in der Vorgängerversion verankert und präziser formuliert worden ist. Eines der in der aktualisierten GER formulierten Ziele ist, dass Schüler einen Text analysieren und sich kritisch mit ihm auseinandersetzen können sollten. Der GER kann dementsprechend so gedeutet werden, dass die Interpretation ein Bestandteil des Sprachcurriculums ist. Man kann also erwarten, dass der Bereich Literatur auch in den neuen *Eindtermen* eine wichtige Rolle spielen wird und dass die Ziele beziehungsweise Anforderungen des Bereiches Literatur präziser formuliert werden.

⁹¹ Vgl. Meijer und Fasoglio, „Handreiking schoolexamen moderne vreemde talen havo/vwo“, S. 17.

⁹² Vgl. Europarat, *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*, S. 56.

⁹³ Europarat, *Common European Framework of Reference for Languages: Learning, Teaching, Assessment - Companion Volume*, S. 108.

Der Autorin dieser Arbeit liegt zurzeit des Schreibens der vorliegenden Arbeit eine Konzeptversion⁹⁴ der neuen *Eindtermen* aus März 2023 vor. Aus dieser Konzeptversion wird klar, dass in Zukunft mehr als bisher auf den Inhalt des Faches Deutsch als Fremdsprache fokussiert werden wird. Die *Eindtermen* werden in andere Domänen eingeteilt, was dazu führt, dass die separate Domäne Literatur wegfällt. Stattdessen sind die Domänen stärker nach Inhalt sortiert und darin ist eine Rolle, vor allem in der Domäne ‚Kulturbewusstsein‘, für den Literaturunterricht vorgesehen. Im Moment ist dabei eine Subdomäne ‚Kultur als Äußerungsmöglichkeit‘, die zwei *Eindtermen*, die am deutlichsten auf literarische Kompetenzen eingehen, umfasst (siehe Tabelle 2), angedacht. Dabei wird einerseits ausgegangen von persönlichen Erlebnissen von unter anderem literarischen Texten, die Schüler beschreiben und erklären können sollen. Andererseits sollen Schüler literarische Texte auch analysieren beziehungsweise interpretieren können. Dabei ist nicht ganz klar, ob sie dies in der Muttersprache oder in der fremden Zielsprache können sollen, allerdings ist dies in den aktuellen *Eindtermen* auch nicht klar vorgegeben und es spricht somit nichts dagegen, diese Ziele anhand von Texten in der Fremdsprache mit einer Verarbeitung in der Muttersprache zu verfolgen. Es wird sogar von einigen Wissenschaftlern davon abgeraten, literarische Kompetenzen in der Fremdsprache zu prüfen, weil dann eigentlich zwei Sachen geprüft werden, nämlich einerseits die literarischen Kompetenzen, die man prüfen möchte, aber andererseits auch die Fähigkeit, sich in der Fremdsprache ausdrücken zu können.⁹⁵ Dadurch wird unklar, welche literarischen Kompetenzen die Schüler eigentlich besitzen, weil die Fremdsprache eine Barriere darstellt.

Beleving / reflectie	Persoonlijke beleving van talige creatieve, kunstzinnige uitingen waaronder narratieve en literaire teksten. Beschrijven en uitleggen van persoonlijke reacties op een werk (o.a. personages, motieven, gebeurtenissen). ⁹⁶
Analyse / interpretatie	Verwerving van kennis en inzicht in vorm en inhoud bij verschillende vormen van fictie waaronder literatuur door analyse en kritische waardering (o.a. opbouw, thema's, genres). ⁹⁷

Tabelle 2: Konzept-*Eindtermen*, die sich eindeutig auf literarische Texte beziehen.⁹⁸

⁹⁴ Actualisatiemvt.nl, „Conceptraamwerk, selectie van concepteindtermen en conceptkarakteristieken Moderne Vreemde Talen. Tussenproducten voor Adviesperiode 2“.

⁹⁵ Vgl. Van der Knaap, *Literatuur en film in het vreemdetalenonderwijs*, S. 58.

⁹⁶ Übersetzung: „Erlebnis / Reflektion: Persönliche Erlebnisse von sprachlichen, kreativen, künstlerischen Äußerungen, darunter narrativen und literarischen Texten. Beschreiben und erklären von persönlichen Reaktionen zu einem Werk (u.a. Charaktere, Motive, Geschehen).“

⁹⁷ Übersetzung: „Analyse / Interpretation: Erwerb von Wissen und Einsicht in Form und Inhalt verschiedener Arten von Fiktion, darunter Literatur, durch Analyse und kritische Wertung (u.a. Aufbau, Thematik, Genres).“

⁹⁸ Actualisatiemvt.nl, „Conceptraamwerk, selectie van concepteindtermen en conceptkarakteristieken Moderne Vreemde Talen. Tussenproducten voor Adviesperiode 2“, S. 19.

2.4.2 Bestehende Modelle zu literarischen Kompetenzen

Es darf davon ausgegangen werden, dass Literatur beziehungsweise literarische Kompetenzen in den modernen Fremdsprachen auf jeden Fall in der Sekundarstufe II aufgrund der *Eindtermen* anhand von Klausuren oder mündlichen Prüfungen geprüft werden. Wie diese Prüfungen für das Fach Deutsch als Fremdsprache in den Niederlanden aussehen, wurde aber bisher, soweit der Autorin dieser Arbeit bekannt ist, nicht erforscht. Aus ihren eigenen Erfahrungen kann die Autorin dieser Arbeit feststellen, dass an Schulen zum Teil Klausuren der von ihnen verwendeten Lehrbücher benutzt werden, zum Teil aber auch eigene Klausuren oder Prüfungen konzipiert werden. In diesen Prüfungen fällt auf, so die Erfahrung der Autorin dieser Arbeit, dass auch wenn literarische Texte gelesen werden sollen, die Aufgaben doch wieder ein reines Leseverstehen (wie in allen Prüfungen zu Sachtexten) erfordern. Jedoch gibt es auch einige Forscher, die Modelle für literarische Kompetenzen entwickelt haben, die beim Erstellen eines Bewertungsschemas von Nutzen sein könnten.

In 2008 hat Witte⁹⁹ für das Fach Niederländisch ausführlich untersucht, wie Schüler der Niveaus *havo* und *vwo* sich auf dem Gebiet von literarischen Kompetenzen während ihrer Schulzeit in den beiden Sekundarstufen entwickeln. Aus dieser Studie entstand ein Modell, dessen Verwendung gesetzlich verankert wurde¹⁰⁰, mit sechs Referenzniveaus für narrative und literarische Texte.

1F	Erlebend lesen
2F	Wiedererkennend lesen
3F	Kritisch und reflektierend lesen
4F	Interpretierend und ästhetisch lesen
5F	Literarisch lesen
6F	Akademisch lesen ¹⁰¹

Niveau 4F sollen Schüler in *6vwo* erreicht haben, aber Niveau 5 soll angestrebt werden.¹⁰² Schüler in *5havo* sollen mindestens Niveau 3F erreichen.¹⁰³ Dieses Modell wurde für das Schulfach Niederländisch entworfen, soll aber laut Van der Knaap auch von Dozenten, die eine Fremdsprache unterrichten, beachtet werden, um eine fächerübergreifende Unterrichtsgestaltung auf dem Gebiet von Literatur zu garantieren.¹⁰⁴

Für die Klassen *4havo* und *4vwo* (die Probandengruppe dieser Arbeit) sind die Niveaus 2F, 3F und 4F am interessantesten zu betrachten, weil 1F schon am Ende der Grundschule

⁹⁹ Witte, *Het oog van de meester: een onderzoek naar de literaire ontwikkeling van havo-en vwo-leerlingen in de tweede fase van het voortgezet onderwijs*.

¹⁰⁰ Vgl. Van der Leeuw und Ravesloot, *Boekverslagen analyseren met het referentiekader*, S. 5.

¹⁰¹ Vgl. Van der Knaap, *Literatuur en film in het vreemdetalenonderwijs*, S. 73ff.

¹⁰² Vgl. ebd., S. 75f.

¹⁰³ Vgl. SLO, „Referentiekader taal“.

¹⁰⁴ Vgl. Van der Knaap, *Literatuur en film in het vreemdetalenonderwijs*, S. 24.

erreicht sein soll. Van der Leeuw und Ravesloot haben für die einzelnen Referenzniveaus festgelegt, was Schüler konkret beherrschen sollten. Das Modell ist für diese Arbeit von Interesse, da es dazu dienen soll, dass Dozenten wissen, worauf sie beim Beurteilen von Aufgaben, die als Schwerpunkt Literatur haben, achten sollen.¹⁰⁵ In dem Modell werden für das Lesen von fiktionalen, narrativen, literarischen Texten drei Kategorien, nämlich ‚verstehen‘, ‚interpretieren‘ und ‚evaluieren‘, unterschieden. Da das Modell zu umfangreich ist um hier in seiner Gänze aufzuführen, wird an dieser Stelle nur die Kategorie ‚interpretieren‘ aufgeführt (siehe Tabelle 3). Das Modell ist einerseits interessant, weil bestimmte Vorgaben darüber gemacht werden, was Schüler auf einem bestimmten Niveau beherrschen sollen.

Interpretieren	1F	2F	3F	4F
Beziehung zur Wirklichkeit	Erkennt Zusammenhänge zwischen Text und Realität.	Stellt fest in welchem Maße Charaktere und Geschehnisse wiedererkennbar und realistisch sind.	Wie 2F.	Platziert Texte in eine kulturhistorische Perspektive.
Struktur-elemente	Zeigt auf spannende, humorhafte oder dramatische Textabschnitte.	Wie 1F.	Erläutert die Wirkung von elementaren Erzähl- und Dichttechniken.	Wie 3F.
Charaktere	Erkennt verschiedene Emotionen (z.B. Traurigkeit, Frust, Freude) im Text wieder.	Charakterisiert Charaktere, sowohl innerlich als auch äußerlich.	Benennt implizite Ziele und Motive der Charaktere.	Identifiziert sich empathisch mit verschiedenen Charakteren.
Thematik	-	Benennt das Thema des Textes.	Formuliert das zentrale Problem oder den Grundgedanken des Textes. Verleiht Symbolen Bedeutung.	Formuliert das allgemeine Thema.

Tabelle 3: Merkmale der Aufgabenverarbeitung durch Schüler beim Interpretieren, sortiert nach den Niveaus des niederländischen Referenzrahmens Sprache.¹⁰⁶

¹⁰⁵ Vgl. Van der Leeuw und Ravesloot, *Boekverslagen analyseren met het referentiekader*, S. 120.

¹⁰⁶ Vgl. ebd.

Andererseits ist aber auch die Gestaltung des Modells interessant. Die verschiedenen Subkategorien und ihre Abstufungen geben wichtige Anhaltspunkte für die Entwicklung eines Bewertungsschemas. Allerdings muss beachtet werden, dass Schüler in der Fremdsprache möglicherweise nicht das Niveau erreichen, das sie in der Muttersprache schon besitzen, weil Verständnisprobleme die Verarbeitung des Gelesenen erschweren. Zudem ist der kulturelle Kontext des Landes der Fremdsprache den Schülern relativ unbekannt.

2.4.3 Qualitätskriterien eines guten Bewertungsschemas

Schließlich sollen wir uns anschauen, welche Kriterien eine Prüfung beziehungsweise ein Bewertungsschema erfüllen muss, damit deren Qualität entsprechend gut ist. Dabei ist von vornherein klar, dass eine einzelne Prüfung nicht alle Qualitätskriterien erfüllen kann, sondern immer ein Kompromiss zwischen den unterschiedlichen Qualitätskriterien sein wird.¹⁰⁷ Trotzdem ist es wichtig, zu versuchen, den bestmöglichen Kompromiss zu treffen. Aus diesem Grund wird in diesem Kapitel eine Auswahl der wichtigsten Qualitätskriterien besprochen.

Bei einer umfangreichen Studie für CITO zur wissenschaftlichen Forschung mit dem Schwerpunkt Qualitätsanalyse von Prüfungen stellten Maassen et al. fest, dass der Begriff ‚Zuverlässigkeit‘ am meisten in den verschiedenen Studien genannt wird.¹⁰⁸ Zuverlässigkeit ist somit ein wichtiges Qualitätskriterium einer Prüfung. Ein bedeutsamer Aspekt von Zuverlässigkeit ist die Objektivität¹⁰⁹; Die Antworten von Schülern sollen immer gleich bewertet werden, egal wer der Schüler und wer der Prüfer ist. Ein Weg, um Objektivität und somit auch Zuverlässigkeit zu erreichen, ist der Verwendung eines Bewertungsschemas.¹¹⁰

Weitere Aspekte, die für die Qualität einer Prüfung von Belang sind, sind Validität, Generalisierbarkeit, Selektivität und Transparenz. Um die Validität einer Prüfung sicher zu stellen, ist es wichtig, dass Dozenten wissen, welche Ziele sie mit der Prüfung verfolgen wollen und welche Prüfungsart dazu angemessen sein könnte.¹¹¹ Wenn man die Interpretationsfähigkeiten von Schülern prüfen möchte, soll die Prüfung es den Schülern ermöglichen, ihre Fähigkeiten zu zeigen. Die Interpretationsfähigkeiten sollten also nicht anhand einer Multiple-Choice-Prüfung geprüft werden, sondern eher durch eine Aufgabe, die offen genug ist, damit Schüler ihre eigenen Gedanken über den Text unterbringen können. Ein

¹⁰⁷ Vgl. Sluijsmans et al., „De kwaliteit van toetsen onder de loep“, S. 20.

¹⁰⁸ Vgl. Maassen et al., „Kwaliteit van toetsen binnen handbereik. Een reviewstudie van onderzoek en onderzoeksresultaten naar de kwaliteit van toetsen“, S. 22.

¹⁰⁹ Vgl. Maassen und Den Otter, „Huh, moesten we dat ook leren? Over toetskwaliteit in het onderwijs“, S. 36.

¹¹⁰ Vgl. ebd.

¹¹¹ Vgl. ebd.

weiterer häufig genannter Begriff ist die Generalisierbarkeit. Dieser beinhaltet, dass eine Prüfung so wenig wie möglich situationsabhängig prüft. Das heißt, dass Schüler die Fähigkeiten, die sie in der Prüfung zeigen, auch außerhalb der Prüfungssituation, sofern sich der Wissensstand nicht verändert, zeigen können sollen.¹¹² Weiterhin soll eine Prüfung selektiv oder auch diskriminierend sein, denn eine Prüfung soll dazu führen, dass man den Unterschied zwischen besser und schlechter leistenden Schülern erkennen kann.¹¹³ Schließlich ist, besonders auch für die Schüler, die Transparenz einer Prüfung am wichtigsten. Sie sollen nämlich nachvollziehen können, welche Inhalte und wie geprüft werden. Das heißt, dass sie ein Recht darauf haben, sich auf die bestmögliche Art und Weise vorbereiten zu können.¹¹⁴ Diese Aspekte hängt auch zusammen mit der Handlungspraxis des Gebens von *Feed-Up*; Man sollte seinen Schülern zeigen, was von ihnen verlangt wird, damit sie wissen, was sie noch lernen sollen beziehungsweise, wie sie sich auf die Prüfung vorbereiten können.¹¹⁵ Dies kann man machen, indem man das Bewertungsschema der Prüfung mit den Schülern teilt. Somit ist die Transparenz der Prüfung und wertvoller *Feed-Up* garantiert.

Ein letzter wichtiger Aspekt eines Bewertungsschemas ist die Einteilung. Dafür gibt es keine strikten Vorgaben, aber Empfehlungen durch das CITO. Das CITO empfiehlt beim Erstellen eines Bewertungsschemas maximal fünf Kategorien beziehungsweise Kriterien aufzuführen, damit das Schema übersichtlich und hantierbar bleibt.¹¹⁶ Dabei soll darauf geachtet werden, dass die Unterschiede zwischen den fünf ausgewählten Kategorien relativ groß sind und die Kategorien dementsprechend gut voneinander zu unterscheiden sind. CITO macht zudem eine Empfehlung, wie viele Abstufungen diese Fertigkeiten haben sollten. Es sollten nämlich nicht zu viele sein, aber mindestens drei, damit Schüler auch erkennen können, dass ihr Einsatz belohnt wird.¹¹⁷ Wenn man aber zu viele Abstufungen wählt, dann wird es schwer, die verschiedenen Stufen mit konkreten Handlungsvorgaben von der Seite der Schüler zu versehen.¹¹⁸ Sanders et al. schreiben zu der Menge von Abstufungen, dass sie von der Komplexität der Aufgabe (je komplexer, desto mehr Abstufungen), den Leistungsunterschieden der Kandidaten (je größer die Unterschiede, desto mehr Abstufungen) und der Zeit des Dozenten für die Bewertung (je geringer die Zeit, desto weniger Abstufungen)

¹¹² Vgl. Maassen et al., „Kwaliteit van toetsen binnen handbereik. Een reviewstudie van onderzoek en onderzoeksresultaten naar de kwaliteit van toetsen“, S. 12.

¹¹³ Vgl. Maassen und Den Otter, „Huh, moesten we dat ook leren? Over toetskwaliteit in het onderwijs“, S. 37.

¹¹⁴ Vgl. Sluijsmans et al., „De kwaliteit van toetsen onder de loep“, S. 20.

¹¹⁵ Vgl. Hattie und Timperley, „The Power of Feedback“, S. 88.

¹¹⁶ Vgl. CITOLab, „6 tips voor het maken van een beoordelingsschaal of -schema“.

¹¹⁷ Vgl. ebd.

¹¹⁸ Vgl. ebd.

abhängig sein sollte.¹¹⁹ Alles in allem muss also gründlich darüber nachgedacht werden, wie viele Fertigkeiten und Abstufungen für ein Bewertungsschema gewählt werden und aus welchen Gründen diese Wahl erfolgt.

3. Datenbestand

Die Daten, mit denen in dieser Arbeit gearbeitet wird, wurden von Stefanie Ramachers im Rahmen ihrer Forschung zur dialogischen Didaktik erworben. Dafür wurde eine Untersuchung mit Pretest, Intervention und Posttest in fünfzehn Klassen der Niveaus *4havo* und *4vwo* während des Schuljahrs 2018-2019 durchgeführt. Die Intervention bestand aus einer Unterrichtsserie zu literarischen Kurzgeschichten¹²⁰, die das Ziel hat, „leerlingen stapsgewijs te leren om hun leeservaringen en vragen [...] tijdens het lezen van Duitstalige korte verhalen te verwoorden (interne dialoog) en in groepsverband of klassikaal te bespreken (externe dialoog).“¹²¹ Ein Drittel der Klassen nahm dabei an der Schüler-geführten Teilintervention (Schüler stellten selbst die Fragen zu den Kurzgeschichten) teil, während das zweite Drittel der Klassen an der Dozenten-geführten Teilintervention (Schüler arbeiteten mit vorausgewählten Fragen) teilnahm und das letzte Drittel der Klassen wiederum wurde als Kontrollgruppe benutzt und bekam entsprechend den Unterricht zu den Kurzgeschichten auf die gängige Art und Weise. Alle in dieser Arbeit verwendeten Daten stammen aus der Schüler-geführten Intervention.

Die Pre- und Posttests bestanden aus einem Fragebogen und einem von den Schülern zu schreibenden schriftlichen Dialog zu einer Kurzgeschichte. In dem Fragebogen wurden, nach dem Lesen der Kurzgeschichte, unter anderem allgemeine Daten (Alter, allgemeine Leseerfahrung, Note im Fach Deutsch (als Fremdsprache), usw.) abgefragt und spezifische Fragen unter anderem zu Einfühlungsvermögen und dem erlebten Schwierigkeitsgrad der Kurzgeschichte gestellt.¹²² Die Kurzgeschichten für die Pre- und Posttests wurden darauf ausgesucht, dass sie ungefähr gleich schwierig sind. Dafür wurde eine LIX-Calculator

¹¹⁹ Vgl. Sanders et al., *Toetsen op school*, S. 138.

¹²⁰ Die Unterrichtsmaterialien sind frei verfügbar unter:
<https://stefanieramachers.wordpress.com/lesmateriaal/>.

¹²¹ Ramachers, „Wanderlust, dialogisch lernen met korte verhalen binnen het schoolvak Duits“, S. 2.
Übersetzung: „Schülern schrittweise zu lehren, ihre Leseerfahrungen und Fragen während des Lesens deutschsprachiger Kurzgeschichten zu äußern (interner Dialog) und im Gruppenzusammenhang oder in einem Klassengespräch zu besprechen (externer Dialog).“

¹²² Vgl. Ramachers und Willems, „Reading Short Stories in a Foreign Language: Is It as Pleasurable as in the L1? Absorption and Appreciation of Short Stories in Dutch Secondary School Students Learning German as a Foreign Language“, S. 5.

verwendet, der nach der Formel von Carl-Hugo Björnsson den Schwierigkeitsgrad eines Textes wie folgt berechnet: $\frac{A}{B} + \frac{C \cdot 100}{A}$ in dem A = Anzahl Wörter; B = Anzahl Sätze; C = Anzahl lange Wörter (das sind Wörter länger als sechs Buchstaben).¹²³ Die ausgesuchten Kurzgeschichten sind *Helm auf* von Finn-Ole Heinrich¹²⁴ und *Stinkefisch* von Ferdinand von Schirach¹²⁵ (eine Zusammenfassung der beiden Geschichten steht in Kapitel 3.1). Die LIX-Berechnung ergibt, dass beide Kurzgeschichten eine sehr geringe Komplexität auf Wort- und Satzebene aufweisen (siehe Tabelle 4). Die teilnehmenden Dozenten bestätigten die niedrige Komplexität der Geschichten.¹²⁶

	Helm auf	Stinkefisch
Anzahl Wörter	1159	1103
Anzahl Sätze	81	103
Durchschnittliche Satzlänge (Wörter)	14,31	10,71
Anteil lange Wörter (%)	16,7	20,04
Lesbarkeitsindex - LIX	31,05	31,11
Komplexität	Sehr niedrig	Sehr niedrig

Tabelle 4: Lesbarkeitsindex der beiden Kurzgeschichten *Helm auf* und *Stinkefisch*.¹²⁷

Die Hälfte der Schüler las die Kurzgeschichten im Pre- und Posttest auf Niederländisch und die andere Hälfte las sie auf Deutsch. Dabei wurden die Geschichten im Pre- und Posttest überkreuzt verteilt, damit einige Schüler im Pretest die Geschichte *Helm auf* lasen und im Posttest *Stinkefisch* und andersherum. Somit kann in den Daten garantiert werden, dass kein Effekt durch die Art der Geschichte auftritt. In dieser Arbeit werden lediglich die Pre- und Posttestergebnisse der Schüler, die die Geschichte in der Fremdsprache, also auf Deutsch, lasen, verwendet.

Die Wahl des Genres Kurzgeschichte erfolgte unter anderem aus dem Grund, dass Lücken Kurzgeschichten inhärent sind.¹²⁸ Genau diese Lücken stellen eine Möglichkeit für Interpretationen dar, weil der Leser diese Lücken zu füllen versuchen muss, um den Text verstehen zu können.¹²⁹ Kurzgeschichten eignen sich also gut, um den Schülern Interpretationen zu entlocken.

¹²³ Vgl. Björnsson, *Läsbarhet*, S. 66.

¹²⁴ Heinrich, *Helm auf*.

¹²⁵ Von Schirach, „Stinkefisch“.

¹²⁶ Vgl. Ramachers und Willems, „Lezen in de vreemde taal: Net zo plezierig als in de moedertaal?“, S. 2.

¹²⁷ Lenhard und Lenhard, „Berechnung des Lesbarkeitsindex LIX nach Björnsson“. Verfügbar unter: <http://www.psychometrica.de/lix.html>.

¹²⁸ Vgl. Schweikle und Schweikle, *Metzler Lexikon Literatur. Begriffe und Definitionen*, S. 416.

¹²⁹ Vgl. Iser, *The Act of Reading. A Theory of Aesthetic Response*, S. 167.

Nach dem Lesen der Kurzgeschichte in jeweils Pre- und Posttest schrieben die Schüler einen imaginären schriftlichen Dialog über die gerade gelesene Kurzgeschichte. Sie bekamen dafür ein Format, das aussieht wie ein einfacher Comic und mit dem Satz ‚Oké, welk stukje van het verhaal sprong er voor jou echt uit?‘¹³⁰ anfängt (siehe Anhang: Format schriftlicher Dialoge). Die Aufgabe des Schreibens eines imaginären Dialogs zu einer Kurzgeschichte stammt aus einer Studie von Schrijvers et al.¹³¹ Sie beinhaltet, dass Schüler sich einen Dialog über den Text ausdenken und diesen aufschreiben. Für die Aufgabe müssen sich die Schüler also überlegen, was sie zu dem Text sagen beziehungsweise schreiben wollen und verkörpern dann im schriftlichen Dialog zwei Stimmen, die sich über die gelesene Kurzgeschichte unterhalten. Diese Aufgabe ist für die Prüfung von Interpretationsfähigkeiten besonders geeignet, weil sie sehr offen ist. Laut McCarthy und Goldman ist die offene Struktur einer Aufgabe, bei der Schüler ihre Fähigkeiten im Bereich der Interpretation zeigen sollen, äußerst wichtig, weil eben nicht nur eine, sondern viele Interpretationen möglich sind.¹³² Die Aufgabe darf diese Möglichkeiten nicht einschränken und der schriftliche Dialog macht dies auch nicht, denn die Schüler sind frei darin, zu entscheiden, was sie aufschreiben möchten.

3.1 Zusammenfassungen der Kurzgeschichten

3.1.1 Helm auf

Ein Junge kauft sich einen Helm und besagt, dass er den nicht wieder abnimmt. Er trägt den Helm wirklich ständig und nimmt ihn beispielsweise auch beim Duschen nicht ab. Er berichtet davon, dass seine Dozenten und Klassenkameraden ihn dafür mobben. Seine Mutter weint und fleht ihn an, den Helm abzusetzen. Um den Helm noch besser zu befestigen, kauft er eine Stahlkette und ein Schloss und wirft den Schlüssel anschließend weg. Er deutet an, dass er den Helm erst wieder absetzt, wenn sein Vater wieder zurück ist. Warum sein Vater jetzt nicht da ist, kann dem Text allerdings nicht eindeutig entnommen werden. Am Ende des Textes wird klar, dass der Junge erwartet, den Helm bis in die Ewigkeit zu tragen, denn auch wenn er heiratet und ein Opa ist, bis ins Grab, wird man ihn mit Helm vorfinden können.

¹³⁰ Übersetzung: Okay, welcher Teil der Geschichte stach für dich echt heraus?

¹³¹ Schrijvers et al., „Transformative Dialogic Literature Teaching Fosters Adolescents’ Insight into Human Nature and Motivation“.

¹³² Vgl. McCarthy und Goldman, „Comprehension of Short Stories: Effects of Task Instructions on Literary Interpretation“, S. 6.

3.1.2 Stinkefisch

Tom wohnt in einem Viertel, in dem viele Nationalitäten zusammen wohnen und in dem keiner seine Kinder zur Schule bringt. Die Kinder laufen also selbstständig zur Schule. Tom ist gemeinsam mit seinen Freunden auf dem Weg zur Schule. Tom hat Angst, denn er soll zu einem vernachlässigt aussehenden Mann ‚Stinkefisch‘ rufen. Tom ist nämlich an der ersten Mutprobe, dem Klauen von Zigaretten, gescheitert. Tom schreit Stinkefisch und sieht in dem Moment, dass der Mann auch einfach nur ein normaler Mann ist. Er entschuldigt sich deswegen bei dem Mann. Kurz darauf werfen seine Freunde Steinen zu dem Mann und sie alle rennen danach schnell weg. Die Polizei kommt zur Schule und redet mit den Kindern; der Mann wurde von einem der Steine getroffen und liegt im Krankenhaus. Die Kinder wurden nicht angeklagt, weil sie zu jung sind. Das Haus des Mannes wurde abgerissen und es wurde dort ein Einkaufszentrum gebaut. Was mit dem Mann passiert ist, nachdem er vom Stein getroffen wurde, kann dem Text nicht entnommen werden. Der Name Stinkefisch verschwand allmählich aus dem Gedächtnis der Kinder.

4. Methodik

Dieses Kapitel umfasst die Methodik, die erforderlich ist, um die erste und zweite Forschungsfrage zu beantworten. Die Methodik, die benutzt wurde um die dritte Forschungsfrage zu beantworten, kann in Kapitel 7 vorgefunden werden.

4.1 Datenanalyse

Um sich mit den Daten vertraut zu machen, wurde zunächst eine der beiden Kurzgeschichten, *Helm auf*, anhand der *Self-Questioning*-Methode (das heißt in diesem Fall, dass Fragen zum Text während des Lesens am Seitenrand notiert wurden) gelesen. Dadurch entwickelte sich ein Verständnis für die möglichen Antworten der Schüler. Daraufhin wurden die schriftlichen Dialoge des Posttests zu *Helm auf* zum ersten Mal durchlesen. Diese beiden Schritte wurden folglich für die Kurzgeschichte *Stinkefisch* wiederholt.

4.1.1 Entscheidungen bei der Erstellung des Codebuchs

Für die Erstellung des Codebuchs wurde eine Vorlage aus der Studie von Schrijvers et al. verwendet.¹³³ Die Studie von Schrijvers et al. hat sich nämlich auch mit schriftlichen Dialogen auseinandergesetzt, allerdings wurde die Studie mit dem Schwerpunkt auf *Transformative*

¹³³ Schrijvers et al., „Transformative Dialogic Literature Teaching Fosters Adolescents’ Insight into Human Nature and Motivation“.

Dialogic Literature Teaching (Vorstellungskraft, Einfühlvermögen, sympathische Gefühle bezüglich Charaktere) unter Muttersprachlern durchgeführt. Da der Schwerpunkt dieser Studie ein anderer ist, muss bei der Erstellung des Codebuchs für diese Studie gut durchdacht werden, welche Codes beibehalten und welche noch hinzugefügt werden sollen.

Die Entscheidungen, die dabei getroffen wurden, basieren größtenteils auf den jeweils ersten fünf schriftlichen Dialogen des Posttests pro Kurzgeschichte in der Liste. Diese insgesamt zehn Dialoge wurden quasi wie ein Pilotprojekt verwendet, um die richtigen Codes für das Codebuch und eine Vorgehensweise für das Codieren auszuwählen. In Tabelle 5 kann man sehen, welche Veränderungen am Codebuch von Schrijvers et al. für die Zwecke dieser Studie vorgenommen wurden. Die Entscheidungen dazu werden im Folgenden der Reihenfolge nach von oben nach unten erläutert.

Codes der Studie von Schrijvers et al.	Codes dieser Studie
Descriptive-evaluative	Descriptive: default
	Descriptive: summary
	Evaluative: character
	Evaluative: text
Reasoning: evaluative	Reasoning: evaluative character
	Reasoning: evaluative text
Reasoning: hypothetical	Reasoning: hypothetical
Reasoning: understanding character	Reasoning: understanding character
Reasoning: understanding story events	Reasoning: understanding story events
Reasoning: moral – in story	Reasoning: moral – in story
Reasoning: moral – beyond story	Reasoning: moral – beyond story
Referential: self	Referential: self
Referential: real world	Referential: real world
Transformative reading: imagery	Transformative reading
Transformative reading: identification	
Transformative reading: experience-taking	
Transformative reading: sympathy	
Transformative reading: aesthetic awareness	
Transformative reading: character evaluation	→ siehe: Evaluative: character / Reasoning: evaluative character
–	Essential question
–	Literary words
–	Problem understanding: self assessed
–	Problem understanding: without noticing
–	Off topic
Other	–

Tabelle 5: Veränderungen des Codebuchs im Vergleich zwischen der Studie von Schrijvers et al. und dieser Studie.

Der Code *Descriptive-evaluative* umfasst im Codebuch von Schrijvers et al. Aussagen, die Teile des Textes beschreiben, und beurteilende Aussagen. In der Studie von Schrijvers et al. wurde dieser Code mit Abstand am häufigsten codiert.¹³⁴ Dies allein ist schon ein Grund gewesen, zu überlegen, diesen Code aufzuteilen. Zudem kam die Überlegung, dass deskriptive Aussagen beim Lesen in einer Fremdsprache einen anderen Wert als in der Muttersprache haben und deswegen vielleicht separat codiert werden sollten, hinzu. Dementsprechend wurde der Code aufgeteilt. Folgend auf die Erkenntnis der Taxonomie von Kwakernaak, dass zusammenfassende Aussagen eine bessere Beherrschung des Materials als nur beschreibende Aussagen erfordern¹³⁵, ergaben sich die Codes *Descriptive: default* und *Descriptive: summary*. Auch für den *Evaluative*-Code wurde der Entschluss gefasst, ihn nochmal aufzuteilen. Der Code *Transformative reading: character evaluation* wurde dabei berücksichtigt und es wurde demnach überlegt, ob es wertvoll sein könnte die *Evaluative*-Codes (*Evaluative* und *Reasoning: evaluative*) in eine *Character*- und eine *Story*-Variante aufzuteilen. Es wurde beschlossen, dass dies sinnvoll sein könnte, weil Schüler sich auf eine tiefere Ebene mit dem Text auseinandersetzen müssen, um eine urteilende Aussage zu einem Charakter als zu dem Text zu treffen.

Für alle *Reasoning*-Codes gilt die Voraussetzung, dass Schüler ihre Aussage auf irgendeine Weise erklären. Diese Codes wurden alle als für diese Studie wichtig betrachtet, denn in diesen Codes geht es darum, dass Schüler über die Geschehnisse und Charaktere der Geschichte nachdenken. Deswegen wurden die *Reasoning*-Codes alle übernommen. Das gesamte Codebuch inklusive genauerer Erklärungen der Codes kann im Anhang gefunden werden. An dieser Stelle werden die Codes deswegen nur kurz und knapp erläutert. Der Code *Reasoning: hypothetical* wird codiert, wenn Schüler hypothetische Aussagen (wie ein mögliches Ende) zum Text treffen. *Reasoning: understanding character* wird verwendet, wenn Schüler versuchen zu erklären, weswegen ein Charakter handelt wie er handelt. *Reasoning: understanding character* und *Reasoning: understanding story events*, bei dem die Schüler versuchen zu erklären, weshalb bestimmte Sachen in der Geschichte passieren, sind zusammen die wichtigsten Codes für Interpretationen, weil dabei wirklich versucht werden muss, den Kern der Geschichte zu erklären. Es wurde noch überlegt, ob es sinnvoll wäre, diese beiden Codes zusammenzufügen, aber es wurde dagegen entschieden, weil es interessant wäre zu beobachten, ob Schüler vielleicht mehr Aussagen zu dem einen als zu dem anderen Code

¹³⁴ Vgl. ebd., S. 11.

¹³⁵ Vgl. Kwakernaak, „Een taxonomie voor het vreemdetalenonderwijs“, S. 14.

machen würden. Die beiden *Reasoning: moral*-Codes beziehen sich auf moralische Aussagen von Schülern jeweils innerhalb der Geschichte oder über die Geschichte hinaus.

Es wurde beschlossen, die beiden *Referential*-Codes auch beizubehalten, weil Schüler laut Marshall et al. beim Interpretieren auf ihre Erfahrungen mit der Welt zurückgreifen.¹³⁶ Der Unterschied zwischen *Referential: self* und *Referential: real world* scheint dabei sinnvoll, weil Verweisungen auf die reale Welt über das eigene Bekannte hinaus für Interpretationen aufgrund ihrer Generalisierbarkeit wertvoller sein können.

Die *Transformative reading*-Codes sind im Gegensatz zu der Studie von Schrijvers et al. für diese Studie weniger interessant. Die Codes wurden zusammengefügt, denn es ist zwar interessant, wie oft sich Schüler emotional mit dem Text auseinandersetzen, aber die Aufteilung in verschiedenen *Transformative reading*-Codes ist für diese Studie nicht relevant.

Zusätzlich wurden dem Codebuch noch ein paar Codes für die Zwecke dieser Studie hinzugefügt. Der Code *Essential question* wurde hinzugefügt, weil sich die Unterrichtsintervention unter anderem mit essenziellen Fragen beschäftigt hat und weil die Menge essenzieller Fragen, die einen Weg zu Interpretationen eröffnen, in den schriftlichen Dialogen auch interessant sein könnte. Der Code *Literary words* basiert auf die Anforderungen der *Eindtermen*, dass Schüler einen bestimmten literarischen Fachjargon beherrschen sollten. Demnach wäre es interessant, wie oft Schüler solche Wörter verwenden. Die beiden *Problem understanding*-Codes wurden hinzugefügt, weil Schüler für diese Studie Kurzgeschichten in einer Fremdsprache gelesen haben. Deshalb ist es interessant, festzulegen, wie häufig Schüler den Text merkbar nicht verstanden haben. Wenn Schüler einen Text nämlich nicht richtig verstehen, hat das Auswirkungen auf ihre Interpretationsfähigkeiten. In der Analyse der ersten zehn Dialoge fiel auf, dass Schüler sich manchmal auch selbst dazu äußern, ob sie die Geschichte verstanden haben oder nicht. Deswegen wurde der Code *Problem understanding: self assessed* als Gegenstück zu *Problem understanding: without noticing* entworfen.

Schließlich wurde der Code *Off topic* erschaffen. Dieser Code basiert zum Teil auf dem Code *Other* aus dem Codebuch von Schrijvers et al. Der Code *Off topic* soll an den Stellen, wo Schüler vollständig von der Geschichte abgeschweift sind, codiert werden. Dieser Code wurde hinzugefügt, weil *Off topic*-Aussagen ein Indiz dafür sein könnten, dass die Schüler den Text nicht ausreichend verstanden haben, um die vorgegebenen drei Seiten dazu zu schreiben. Es wurde gegen einen allgemeinen *Other*-Code entschieden, weil diese Restkategorie nicht viel

¹³⁶ Vgl. Marshall, Smagorinsky, und Smith, *The Language of Interpretation: Patterns of Discourse in Discussions of Literature*, S. 127.

Aussagekraft hätte; sie wird alle Aussagen, die sonst nicht codiert worden wären, enthalten. Zudem ist nicht immer eindeutig, wo ein Code anfängt und wieder aufführt. Das könnte dazu führen, dass viele Wörter und Teilsätze dem Restcode zugeordnet werden. Beim Anschauen dieses Restcodes würde dann die Übersicht (des Kontexts) fehlen. Die Aussagen, die keinem der Codes zugeordnet werden können, sollen dementsprechend nicht codiert werden. Wenn aber bestimmte Aussagen, die nicht codiert werden können, während des Codierens besonders auffallen würden, wird dies aufgeschrieben. Sollte sich beim Durchlesen der Dialoge, wenn sie fertig codiert worden sind, herausstellen, dass noch einige interessante Aussagen verpasst wurden, kann immer noch beschlossen werden, diese in einer (Rest)Kategorie zu codieren.

4.1.2 Codiervorgang

Bevor mit dem Codieren angefangen wurde, wurde überlegt, ob die schriftlichen Dialoge vorher in Segmente aufgeteilt werden sollten. Schrijvers et al. haben die schriftlichen Dialoge nämlich zunächst in Segmente nach Thema eingeteilt.¹³⁷ Beim Versuch, dies ebenso zu machen, stellte sich heraus, dass es schwierig ist, zu entscheiden, wo ein Thema aufhört und das nächste anfängt. Auch war der Sinn der Einteilung in Segmente nicht eindeutig, weil Schrijvers et al. den Segmenten trotzdem mehrere Codes zugeordnet hatten. Außerdem werden Dozenten bei der Bewertung solch eines Dialogs diesen nicht auch noch in Segmente aufteilen. Deswegen hat die Einteilung in Segmente keinen Mehrwert und wurde entschieden, keine Segmente einzuteilen, sondern direkt mit dem Codieren anzufangen.

Als mit dem Codieren angefangen worden war, stellte sich relativ schnell heraus, dass die Arbeitsweise des Codierens pro Dialog vielleicht nicht die richtige war. Auf diese Weise mussten nämlich für jeden Dialog erneut alle Codes und ihre Beschreibungen durchgelesen werden. Es war nicht möglich, sich zu merken, wann welche Aussage welchem Code zugeordnet werden sollte. Deswegen wurde der Entschluss gefasst, nicht pro Dialog, sondern pro Code oder manchmal auch pro semantische Domäne¹³⁸ zu codieren. Das hat nämlich den Vorteil, dass man sich besser merken kann, ob man ähnliche Aussagen schon mit diesem einen Code codiert hat oder nicht. Wenn die Umschreibung der Codes im Codebuch klar genug ist, sollte diese Vorgehensweise keine anderen Ergebnisse, als wenn pro Dialog codiert worden wäre, erzeugen.

¹³⁷ Vgl. Schrijvers et al., „Transformative Dialogic Literature Teaching Fosters Adolescents’ Insight into Human Nature and Motivation“, S. 8.

¹³⁸ Eine semantische Domäne umfasst mehrere Codes, die sich auf ein gleiches Thema beziehen. Beispielsweise ist *Reasoning* eine semantische Domäne für die Codes *Reasoning: evaluative text*, *Reasoning: evaluative character*, *Reasoning: hypothetical* usw.

Zunächst wurden für alle Dialoge des Posttests die vier *Evaluative*-Codes codiert, gefolgt von *Literary words*. Daraufhin wurden *Off topic*, die beiden *Problem understanding*-Codes und die beiden *Referential*-Codes codiert. Danach wurden die beiden *Reasoning: moral*-Codes, *Transformative reading* und *Reasoning: hypothetical* codiert. Abschließend wurden *Essential question*, die beiden *Reasoning: understanding*-Codes und die beiden *Descriptive*-Codes codiert. Die Reihenfolge der Codes beim Codieren ist vollkommen zufällig gewählt worden. Die gleiche Reihenfolge wurde für das Codieren des Pretests wiederholt. Im Übrigen soll an dieser Stelle nochmal kurz erwähnt werden, dass codieren auch immer gewissermaßen interpretieren heißt; Wenn schriftliche Dialoge codiert werden, muss man eine Interpretation davon machen, was Schüler gemeint haben.

Beim Codieren des Pretests fiel auf, dass viel häufiger als im Posttest der Code *Descriptive: summary* codiert wurde. Um nachzugehen, ob hier nicht ein Versehen vorlag, wurden alle Dialoge des Posttests nochmal bezüglich dieses Codes angeschaut, die Ergebnisse blieben aber dieselben. Sowie so wurden, nachdem das Codieren erledigt war, alle schriftlichen Dialoge noch einmal geprüft, um sicher zu gehen, dass keine Aussagen vergessen wurden zu codieren.

4.2 Interrater-Reliabilität

4.2.1 Auswahl der Dialoge zur Berechnung der Interrater-Reliabilität

Um die Interrater-Reliabilität berechnen zu können, wurden schriftliche Dialoge ausgewählt, um von einer zweiten Raterin codiert zu werden. Um sicher zu gehen, dass alle Codes mindestens zwei Mal vertreten sein werden, wurde entschieden, ungefähr 15% (= 12 Dialoge) statt die übliche 10% der schriftlichen Dialoge von der zweiten Raterin codieren zu lassen.

Die Auswahl der Dialoge zur Berechnung der Interrater-Reliabilität erfolgte wie folgt; Zunächst wurde darauf geschaut, welcher Code überhaupt am wenigsten codiert wurde. Das ist *Reasoning: moral – beyond story* und deswegen wurden zwei Dialoge (einer aus dem Pre- und einer aus dem Posttest) mit diesem Code ausgewählt. Anschließend wurde von beiden Dialogen in eine Exceldatei eingegeben, welche Codes sie enthalten. Die Tabelle, die auf dieser Datei basiert, ist im Anhang (Ausgewählte Dialoge inkl. Mengenangabe der Codes zur Berechnung der Interrater-Reliabilität) vorzufinden. Danach wurde darauf geschaut, welche Codes überhaupt noch nicht in den Dialogen vorhanden sind und welcher der Codes davon am seltensten ist. Deswegen wurden zwei Dialoge (einer aus dem Pre- und einer aus dem Posttest), die den Code *Evaluative: character* enthalten, ausgewählt. Im Anschluss daran wurden drei

Dialoge (zwei aus dem Pre- und einer aus dem Posttest), die sowohl den Code *Reasoning: hypothetical* als auch *Problem understanding: without noticing* enthielten, ausgewählt, damit alle Codes mindestens zwei Mal vorhanden waren. In der Exceldatei fiel an der Stelle auf, dass der Code *Reasoning: understanding story events* erst zwei Mal in den Dialogen auftauchte. Weil dieser Code für Interpretationen wichtig ist, wurden nochmal drei Dialoge (einer aus dem Pre- und zwei aus dem Posttest), die diesen Code enthielten, gezielt ausgewählt. Auch wurde bei der Wahl dieser drei Dialoge darauf geachtet, dass *Reasoning: understanding character* noch einmal vorkommt. Damit waren fünf Dialoge aus dem Pretest und fünf aus dem Posttest ausgewählt worden. Allerdings waren bisher noch keine Dialoge, die ganz wenig Codes enthalten, vertreten, weswegen davon auch noch zwei (einer aus dem Pre- und einer aus dem Posttest) ausgewählt wurden.

4.2.2 Berechnung der Werte

Um die Validität der Codes und Zitate (angemerkte Teile eines Dialogs, denen ein Code zugewiesen wird) zu überprüfen, wurden die Dialoge von einer zweiten Raterin (Stefanie Ramachers, Betreuerin dieser Arbeit) codiert. Sie bekam das Codebuch, das im Anhang zu finden ist, als Anleitung dafür, wie codiert werden sollte. Weil die Länge der Zitate relativ irrelevant für diese Studie ist, wurde beschlossen, für die zweite Raterin die Zitate ohne Codes in den Dialogen stehen zu lassen. Die zweite Raterin sollte somit nur noch den Zitaten die ihrer Meinung nach passenden Codes zuordnen, war aber unabhängig darin, weitere Zitate, falls ihrer Einsicht nach notwendig, hinzuzufügen.

Die zweite Raterin hat während des Codierens notiert, über welche Zitate sie beim Zuordnen der Codes gezweifelt hat. Diese wenigen Zitate (8 auf einer Gesamtzahl von 179 Zitate), die Probleme bereiteten, wurden besprochen und es wurde in jedem Fall ein Konsensus zwischen den beiden Raterinnen erreicht. Gegebenenfalls wurde die Beschreibung eines Codes im Codebuch etwas angepasst, um die Unsicherheiten zu beseitigen. Die Datei mit den Codes der zweiten Raterin wurde daraufhin der Datei der ersten Raterin in dem Programm Atlas.ti hinzugefügt, um die Interrater-Reliabilität zu berechnen.

Für die Berechnung der Interrater-Reliabilität wurde beschlossen, die beiden Krippendorff Alphas (siehe Tabelle 6) und eine allgemeine prozentuale Übereinstimmung zu berechnen. Der Krippendorffs $c\alpha_{\text{binary}}$ sagt etwas darüber aus, inwieweit die Zitate der beiden Rater übereinstimmen.¹³⁹ Das Ergebnis des Krippendorffs $c\alpha$ dagegen ermöglicht Aussagen

¹³⁹ Vgl. „Atlas.Ti Windows User Manual. Analyse Anything, Faster“, S. 523.

darüber, inwieweit zwischen den Codes einer semantischen Domäne unterschieden werden kann.¹⁴⁰ In anderen Worten heißt das, dass mit dem Krippendorffs c_{α} geschaut werden kann, ob beispielsweise die Codes der semantischen Domäne *Reasoning* sich gut genug voneinander unterscheiden lassen. Für beide Krippendorffs Alphas gilt, dass die Reliabilität am höchsten ist, wenn dieser über einen Wert von 0,800 liegt.¹⁴¹ Die Reliabilität kann aber auch bei einem Wert über 0,667 noch als akzeptabel betrachtet werden.¹⁴² Die beiden Krippendorffs Alphas können mittels dem Programm Atlas.ti, das auch für das Codieren der Dialoge verwendet wurde, berechnet werden.

Bei dem ersten Versuch, die Krippendorffs Alphas zu berechnen, stellte sich heraus, dass den Zitaten dazu nur ein Code zugewiesen sein durfte. Bei den Zitaten der ersten Raterin war dies schon der Fall, aber bei der zweiten Raterin nicht überall. Entsprechend musste die zweite Raterin die Zitate und Codes nochmals überarbeiten. Bei der Überarbeitung wurden von der zweiten Raterin Notizen an den Zitaten, denen sie initial mehrere Codes zugeordnet hatte, hinzugefügt. Die erste Raterin hat sich diese Notizen angeschaut, um zu schauen, ob sie den Gedankengang der zweiten Raterin für das Zuordnen von mehreren Codes verstehen konnte. Dabei wurde festgestellt, dass dem Gedankengang in allen Fällen gefolgt werden konnte und somit keine Anpassungen mehr am Codebuch gemacht werden mussten. Meistens hat es sich um kleineren Versehen der ersten Raterin gehandelt; Es wurde zum Beispiel vergessen eine Frage, die als *Essential question* codiert werden musste, als Zitat zu markieren. Auf Grundlage der überarbeiteten Version der zweiten Raterin, diesmal mit jeweils nur einem Code einem Zitat entsprechend, wurden die in den folgenden Absätzen genannten Werte berechnet.

Zunächst wurde eine allgemeine prozentuale Übereinstimmung berechnet. Dieser Wert beträgt 78,3% und bedeutet, dass die zwei Raterinnen sich in 78,3% aller Fälle für den gleichen Code entschieden. Um die Validität der Interrater-Reliabilität genauer zu bestimmen, wurden auch die beiden Krippendorffs Alphas berechnet. Wie man in Tabelle 6 sehen kann, kann aus dem Krippendorffs $c_{\alpha_{\text{binary}}}$ geschlossen werden, dass die zweite Raterin an den Dialogen für bestimmte Codes (überall wo die Zahl nicht 1 entspricht) noch Zitate, zusätzlich zu den schon vorhandenen Zitaten, hinzugefügt hat. Die Werte für den Krippendorffs $c_{\alpha_{\text{binary}}}$ sind aber in allen Fällen gleich oder höher als 0,800, was bedeutet, dass die zwei Raterinnen sich oft genug über die Stellen der Zitate einig waren, um sie als zuverlässig einzustufen.

¹⁴⁰ Vgl. ebd., S. 524.

¹⁴¹ Vgl. Krippendorff, *Content Analysis*, S. 356.

¹⁴² Vgl. ebd.

Die Ergebnisse der Berechnung vom Krippendorffs c_{α} können auch der Tabelle 6 entnommen werden. Für die drei folgenden semantischen Domänen *Evaluative*, *Referential* und *Problem understanding* kann festgestellt werden, dass die beiden Raterinnen sich in allen Instanzen einig waren. Somit kann aus diesem Ergebnis geschlossen werden, dass die Codes, die diesen semantischen Domänen zugehören, ohne Probleme voneinander zu unterscheiden sind. Auch die semantische Domäne *Reasoning* ist mit einem Wert von $\alpha = 0,773$ ausreichend zuverlässig. Die einzige semantische Domäne, die laut der Berechnung vom Krippendorffs c_{α} nicht ausreichend vertrauenswürdig ist, ist mit einem Wert von $\alpha = 0,593$ die Domäne *Descriptive*. Das Problem dieser semantischen Domäne liegt bei dem Code *Descriptive: summary*, denn die beiden Raterinnen waren sich hier nur in vier Instanzen einig. Dafür gab es aber mehrere Zitate, die von der ersten Raterin den Code *Descriptive: default* und von der zweiten Raterin den Code *Descriptive: summary* zugewiesen bekamen und umgekehrt. Aus diesem Ergebnis kann dementsprechend geschlossen werden, dass für Rater nicht eindeutig festzustellen ist, ob ein Zitat eine Zusammenfassung oder eine Wiederholung von dem im Text Geschriebenen ist.

	Krippendorffs $c_{\alpha_{\text{binary}}}$	Krippendorffs c_{α}
Descriptive	0,886	0,593
Evaluative	0,875	1
Reasoning	0,800	0,773
Referential	0,872	1
Transformative reading	1	-
Essential question	0,827	-
Literary words	1	-
Problem understanding	1	1
Off topic	1	-
Total	0,941	0,777

Tabelle 6: Krippendorffs $c_{\alpha_{\text{binary}}}$ und c_{α} pro Domäne beziehungsweise allein stehenden Code.

5. Ergebnisse des Codierens

Das Codieren der schriftlichen Dialoge hat eine große Menge an Daten, die im Folgenden analysiert werden sollen, ergeben. Um gezielt an die Daten heranzugehen, ist es daher an dieser Stelle wichtig, sich die Forschungsfragen dieser Arbeit noch einmal anzuschauen. Die erste Frage – Wie interpretieren Schüler literarische Kurzgeschichten im Fremdsprachenunterricht vor und nach einer Unterrichtsintervention, die sich auf *Self-Questioning* richtet? – fokussiert darauf, was Schüler konkret machen, wenn sie einen Text interpretieren und über ihre Erfahrungen einen schriftlichen Dialog schreiben. Dabei ist es interessant, sich anzusehen, was

die Schüler ohnehin schon gemacht haben, was dem Pretest entnommen werden kann, und ob es einen Unterschied zwischen Pre- und Posttest gibt, was auf einen Effekt der Unterrichtsintervention deuten kann. Der Vergleich zwischen Pre- und Posttest führt auch zur Beantwortung der zweiten Forschungsfrage – Inwiefern hat die *Self-Questioning*-Unterrichtsmethode beim Lesen von Kurzgeschichten einen Effekt auf die Interpretationen der Schüler? –, weil man beim Herauslesen der Daten und insbesondere der Menge an Codes im Bereiche *Reasoning* sehen können müsste, ob dabei ein signifikanter Unterschied wahrzunehmen ist.

	Pretest D = 39	Posttest D = 39	Total D = 78
Descriptive: default	111	100	211
Descriptive: summary	22	6	28
Essential question	44	37	81
Evaluative: character	9	7	16
Evaluative: text	60	40	100
Literary words	38	25	63
Off topic	11	9	20
Problem understanding: self assessed	20	22	42
Problem understanding: without noticing	12	19	31
Reasoning: evaluative character	27	20	47
Reasoning: evaluative text	26	22	48
Reasoning: hypothetical	19	11	30
Reasoning: moral - beyond story	5	4	9
Reasoning: moral - in story	15	13	28
Reasoning: understanding character	25	38	63
Reasoning: understanding story events	16	39	55
Referential: real world	25	17	42
Referential: self	22	17	39
Transformative reading	30	21	51
Total	537	467	1004

Tabelle 7: Absolute Menge an Codes in Pre- und Posttest in 78 schriftlichen Dialogen.

In Tabelle 7 werden die absoluten Zahlen der Codemenge pro Code in jeweils Pre- und Posttest, sowie nochmal als Gesamtzahl, wiedergegeben. Da die Datenmenge relativ gering ist, wurde entschieden, beim Vergleich zwischen Pre- und Posttest keine 5% als Richtlinie für Signifikanz einzuhalten, sondern sich anzuschauen, an welchen Stellen der Unterschied zwischen Pre- und Posttest mehr als zehn Codes beträgt. Ein Einhalten der 5%-Grenze als Signifikanzkriterium beim Vergleich würde nämlich dazu führen, dass fast alle Codes einen relevanten Unterschied zwischen Pre- und Posttest aufweisen, was aufgrund der geringen Datenmenge nicht der Fall ist.

In Tabelle 7 fällt auf, dass sich die Gesamtzahl der Codes zwischen Pre- (537) und Posttest (467) deutlich unterscheidet. Ganz eindeutig lässt sich dieser Unterschied nicht erklären, aber eine mögliche Erklärung dieses Unterschieds wird in Kapitel 6 gegeben. Allgemein fällt auf, folgend auf die Verringerung der Gesamtzahl der Codes im Posttest, dass fast alle einzelnen Codes deswegen auch eine meistens kleine Verringerung im Posttest aufweisen.

Größere und deswegen auch auffällige Verringerungen der einzelnen Codes im Vergleich zwischen Pre- und Posttest treten bei den Codes *Descriptive: summary* (22 gegenüber 6), *Evaluative: text* (60 gegenüber 40) und *Literary words* (38 gegenüber 25) auf, wie in Tabelle 7 sichtbar ist. Genau genommen wäre laut den Signifikanzkriterien dieser Untersuchung von mindestens zehn Codes Unterschied zwischen Pre- und Posttest auch die Verringerung beim Code *Descriptive: default* auffällig. Da der Code *Descriptive: default* aber eine größere Codemenge (111 Codes im Pretest gegenüber 100 im Posttest) beinhaltet, ist hier der Unterschied von 11 Codes deutlich weniger signifikant als bei den anderen hier oben genannten Codes. Mögliche Erklärungen zur erheblichen Verringerung der Codes *Descriptive: summary*, *Evaluative: text* und *Literary words* im Posttest werden in Kapitel 6 diskutiert.

Die einzigen Codes, die beim Codieren der Dialoge im Posttest öfter und zudem signifikant häufiger zugewiesen wurden, sind die beiden *Problem understanding*-Codes und *Reasoning: understanding character*, sowie *Reasoning: understanding story events*. Zunächst soll an dieser Stelle auf die Zunahme der beiden *Reasoning* Codes eingegangen werden, weil das die Codes sind, die wirklich die Kerninhalte des Textes aufgreifen; Schüler tätigen Aussagen zu relevanten Inhalten des Textes, wie beispielsweise zu *Helm auf*: „Was könnte mit seinem Vater passiert sein?“ (Siehe das Codebuch im Anhang). Obwohl Tabelle 7 zeigt, dass es einen Anstieg der Codes *Reasoning: understanding character* und *Reasoning: understanding story events* gibt, ist es schwer, einzuschätzen, wie groß dieser Anstieg tatsächlich ist, weil nur absolute Zahlen gezeigt werden und die Gesamtanzahl der Codes sich im Posttest im Vergleich zum Pretest verringert hat. In Tabelle 8 werden daher die absoluten und die relativen Zahlen der semantischen Domänen der Codes genannt. Der Beschluss zur separaten Aufführung der semantischen Domänen wurde gefasst, weil der gesamte Bereich *Reasoning* als Interpretation(sansätze) von der Seite der Schüler betrachtet werden kann. Es geht nämlich in der ganzen Domäne *Reasoning* darum, dass Schüler sich Fragen zu dem Text stellen und diese auch versuchen, zu beantworten. In Tabelle 9 wird die semantische Domäne *Reasoning* daraufhin wieder in die einzelnen Codes aufgeteilt, um auch dort nochmal den

	Pretest		Posttest	
	Absolut	Prozentuale Verteilung	Absolut	Prozentuale Verteilung
Descriptive	133	24.77%	106	22.70%
Essential question	44	8.19%	37	7.92%
Evaluative	69	12.85%	47	10.06%
Literary words	38	7.08%	25	5.35%
Off topic	11	2.05%	9	1.93%
Problem understanding	32	5.96%	41	8.78%
Reasoning	133	24.77%	147	31.48%
Referential	47	8.75%	34	7.28%
Transformative reading	30	5.59%	21	4.50%
Total	537	100.00%	467	100.00%

Tabelle 8: Absolute und relative Menge an Codes pro semantische Domäne in Pre- und Posttest.

	Pretest		Posttest	
	Absolut	Prozentuale Verteilung	Absolut	Prozentuale Verteilung
Reasoning: evaluative character	27	20.30%	20	13.61%
Reasoning: evaluative text	26	19.55%	22	14.97%
Reasoning: hypothetical	19	14.29%	11	7.48%
Reasoning: moral - beyond story	5	3.76%	4	2.72%
Reasoning: moral - in story	15	11.28%	13	8.84%
Reasoning: understanding character	25	18.80%	38	25.85%
Reasoning: understanding story events	16	12.03%	39	26.53%
Total	133	100.00%	147	100.00%

Tabelle 9: Absolute und relative Menge an Codes der semantischen Domäne Reasoning in Pre- und Posttest.

relativen Anteil der jeweiligen Codes zu verdeutlichen. Was man in den beiden Tabellen 8 und 9 genau sehen kann, ist, dass der Anteil von *Reasoning*-Codes im Posttest angestiegen ist. Im Pretest ist der Anteil nur 24,77%, während der Anteil im Posttest auf 31,48% angestiegen ist. Schüler machen also im Posttest relativ mehr Äußerungen, die als *Reasoning* codiert wurden, als im Pretest. Wenn man sich anschließend die Tabelle 9 anschaut, sieht man, dass dieser Anstieg nur den Codes *Reasoning: understanding character* und *Reasoning: understanding story events* zu verdanken ist, weil das die einzigen Codes sind, die eine Zunahme zu verzeichnen haben, sowohl absolut als auch relativ. Anhand dieser Ergebnisse und unter Berücksichtigung der Tatsache, dass diese beiden Codes zum Kern des Textes gehen, kann man schlussfolgern, dass Schüler nach der Unterrichtsintervention im Posttest mehr interpretieren. Damit kann die zweite Forschungsfrage beantwortet werden, denn mit einem relativen Anstieg der Codes *Reasoning: understanding character* und *Reasoning: understanding story events* kann festgestellt werden, dass die Unterrichtsintervention einen positiven Effekt auf die Interpretationen von Schülern ergeben hat. Die Unterrichtsintervention gerichtet auf *Self-Questioning* hat sich also im Sinne von Interpretationen der Schüler ausgezahlt. Es ist dabei interessant zu vermerken, dass vor allem der Anstieg des Codes *Reasoning: understanding story events* groß ist. Wenn Schüler Aussagen machen, die mit diesem Code codiert werden, spielen sie das ‚Worum-Spiel‘, wie Culler es nennt (siehe Kapitel 2.2). Dieser Code wird daher von Schrijvers et al. als ‚klassische Interpretation‘ bezeichnet.¹⁴³ Es gibt im Vergleich zwischen Pre- und Posttest eine Zunahme von diesem Code um 120,53% ($= \frac{26,53-12,03}{12,03} \cdot 100$). Die Unterrichtsintervention hat also zwar einen positiven Effekt auf beide Codes, der positive Effekt auf den Code *Reasoning: understanding story events* ist jedoch deutlich ausgeprägter als auf den Code *Reasoning: understanding character*. Daraus kann geschlussfolgert werden, dass Schüler nach der Unterrichtsintervention vor allem öfter Äußerungen machen, die als klassische Interpretation bezeichnet werden können.

Interessanterweise stellen Schüler, sowohl im Pre- als auch im Posttest, deutlich mehr essenziellen Fragen bei der Kurzgeschichte *Helm auf* als bei *Stinkefisch*, was in Tabelle 10 sichtbar ist und noch einmal deutlicher in Tabelle 11, in der die absoluten Zahlen für die Teilnehmeranzahl korrigiert werden, wird. Dieses Ergebnis wird in Kapitel 6 diskutiert.

¹⁴³ Vgl. Codebuch von Schrijvers et al., das verwendet wurde für die Studie hinter dem Artikel: „Transformative Dialogic Literature Teaching“.

	Helm auf			Stinkefisch		
	Pretest D = 24	Posttest D = 15	Total	Pretest D = 15	Posttest D = 24	Total
Descriptive	82	34	116	51	72	123
Essential question	36	21	57	8	16	24
Evaluative	42	20	62	27	27	54
Literary words	19	17	36	19	8	27
Off topic	6	1	7	5	8	13
Problem understanding	11	8	19	21	33	54
Reasoning	81	78	159	52	69	121
Referential	42	19	61	5	15	20
Transformative reading	14	9	23	16	12	28
Total	591	366	957	360	476	836

Tabelle 10: Absolute Menge an Codes pro semantische Domäne sortiert nach Kurzgeschichte in Pre- und Posttest.

	Helm auf		Stinkefisch	
	Pretest	Posttest	Pretest	Posttest
\bar{x} essential questions	1,5	1,4	0,53	0,67

Tabelle 11: Durchschnitt (\bar{x}) essenzieller Fragen pro schriftlichen Dialog sortiert nach Kurzgeschichte in Pre- und Posttest.

		Helm auf		Stinkefisch	
		Pretest D = 24	Posttest D = 15	Pretest D = 15	Posttest D = 24
Problem understanding: self assessed	H	3	4	17	18
	\bar{x}	0,125	0,267	1,13	0,750
Problem understanding: without noticing	H	8	4	4	15
	\bar{x}	0,333	0,267	0,267	0,625

Tabelle 12: Absolute Menge an Codes (H) und Durchschnitt der Codes (\bar{x}) der semantischen Domäne *Problem understanding* sortiert nach Kurzgeschichte in Pre- und Posttest.

Weiter oben wurde schon festgestellt, dass auch eine Zunahme der beiden *Problem understanding*-Codes im Vergleich zwischen Pre- und Posttest sichtbar ist. Diese Zunahme ist interessant und soll näher betrachtet werden. Wenn man sich Tabelle 12 anschaut, kann festgestellt werden, dass der Anstieg der beiden *Problem understanding*-Codes in relativem Vergleich der Kurzgeschichten zwischen Pre- und Posttest zwar nicht sehr ausgeprägt ist, dafür aber der Unterschied zwischen den beiden Kurzgeschichten in der Menge Codes, die zugewiesen wurden, erheblich ist. Die beiden *Problem understanding*-Codes wurden nämlich jeweils im Pre- und Posttest häufiger in Dialogen, die von *Stinkefisch* handelten (54 Codes insgesamt), als in Dialogen, die von *Helm auf* handelten (19 Codes insgesamt), codiert. Vor allem der Code *Problem understanding: self assessed* wurde den *Stinkefisch*-Dialogen viel öfter als den *Helm auf*-Dialogen zugewiesen. Es scheint also so zu sein, dass die Schüler die Kurzgeschichte *Stinkefisch* als schwieriger erfahren. Der leichte Anstieg der beiden *Problem understanding*-Codes im Posttest könnte auf jeden Fall anhand dieser Erkenntnis erklärt werden, denn im Posttest war die Teilnehmergruppe, die die Kurzgeschichte *Stinkefisch* las, deutlich größer, nämlich vierundzwanzig Schüler, gegenüber nur fünfzehn, die die Kurzgeschichte *Helm auf* lasen. Ob die Kurzgeschichte *Stinkefisch* wirklich schwieriger als *Helm auf* sein könnte oder ob das nur die Erfahrung der Schüler ist, wird im folgenden Kapitel diskutiert.

6. Diskussion zu den Ergebnissen des Codierens

Anhand der Ergebnisse aus dem vorherigen Kapitel konnten die beiden ersten Forschungsfragen beantwortet werden. Die erste Frage, wie Schüler literarische Kurzgeschichten im Fremdsprachenunterricht vor und nach einer Unterrichtsintervention, die sich auf *Self-Questioning* richtet, interpretieren, wurde durch Auswertung der Resultate des Codierens in Kapitel 5 (siehe unter anderem Tabelle 7, 8 und 9) beantwortet. Um die zweite Frage, inwiefern die *Self-Questioning*-Unterrichtsmethode beim Lesen von Kurzgeschichten einen Effekt auf die Interpretationen der Schüler hat, zu beantworten, wurden die Mengen der Codes im Pre- und Posttest verglichen. Dabei stellte sich heraus, dass Schüler im Posttest deutlich öfter als im Pretest interpretierende Aussagen treffen. Vor allem der Anstieg der relativen Menge des Codes *Reasoning: understanding story events*, der als klassische Interpretation angesehen werden kann, war erheblich. Im Folgenden werden die Ergebnisse dieser beiden Forschungsfragen diskutiert.

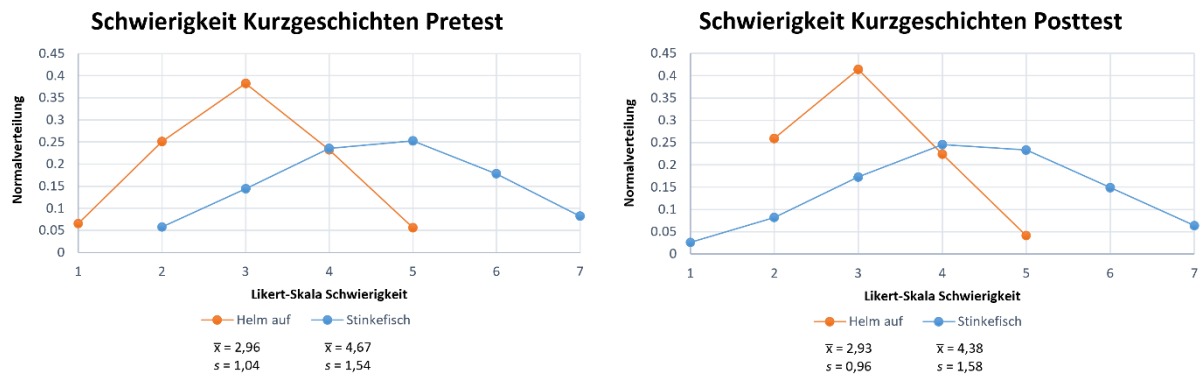
Wie vorhin schon besprochen, ist das erste, das an den Codierdaten auffällt, dass die Gesamtzahl der Codes im Pretest deutlich höher als im Posttest ist. Eine mögliche Erklärung dafür könnte sein, dass Schüler im Posttest ausführlicher zu weniger Aspekten geschrieben haben. Die Zeit, die die Schüler für die Aufgabe des schriftlichen Dialogs bekommen hatten, war in Pre- und Posttest gleich. Zudem war die Vorgabe, wie viele Seiten schriftlichen Dialog sie schreiben sollten, auch die Gleiche. Denkbar wäre also, dass Schüler längere Textstücken, denen nur ein Code zugeteilt werden konnte, geschrieben haben. Allerdings kann aus den Daten des Codierens leider nicht abgelesen werden, ob die Fragmente, die einen Code bekommen haben, möglicherweise einen Unterschied in ihrer durchschnittlichen Länge zwischen Pre- und Posttest aufweisen. Dementsprechend können zur Ursache dieser Abnahme der Gesamtzahl der Codes im Posttest im Vergleich zum Pretest keine endgültigen Aussagen getroffen werden.

Aufgrund der allgemeinen Verringerung der Anzahl der Codes im Posttest, ist es nicht verwunderlich, dass auch bei den einzelnen Codes vorwiegend eine (kleine) Reduzierung im Vergleich zwischen Pre- und Posttest wahrzunehmen ist. Größer (als zehn Codes) ist die Reduzierung der Codes *Descriptive: summary*, *Evaluative: text* und *Literary words* im Posttest, verglichen mit der Anzahl dieser Codes im Pretest. Die erhebliche Abnahme dieser drei Codes ist schwierig zu erklären. Vor allem für den Schwund des Codes *Literary words* ist es schwierig, eine Ursache zu ermitteln. Bei den Codes *Descriptive: summary* und *Evaluative: text* kann man sich hingegen vorstellen, dass Schüler im Pretest vielleicht mehr Aussagen dieser Art gemacht haben, weil sie eine bestimmte Menge an Text in den Dialog schreiben sollten und weil sie die Unterrichtsintervention noch nicht gehabt hatten und deswegen möglicherweise noch weniger gut wussten, was sie in dem schriftlichen Dialog überhaupt schreiben konnten. Das alles bleibt aber Spekulation und es könnte auch so sein, dass die größere Verringerung dieser drei Codes ein Zufall ist. Für weitere Forschung wäre es eventuell interessant, ob bei einer gleichartigen Studie vergleichbare Ergebnisse vorzufinden wären. Vor allem die Abnahme des Codes *Literary words* wäre dann besonders interessant, weil Schüler im Posttest schon deutlich öfter Aussagen machen, die als Interpretation angemerkt werden kann, und weil Schüler laut Anforderungen des Ministeriums beide Sachen können müssen, also sowohl interpretieren als auch sich mit literarischen Worten ausdrücken. Vielleicht ist diese spezifische Aufgabe des Schreibens eines schriftlichen Dialogs aber nicht geeignet, um beide Anforderungen gleichzeitig zu prüfen. Andererseits könnte das Problem eventuell auch damit umgangen werden, dass Schüler im Vornhinein schon wissen, was genau geprüft wird. Anzunehmen wäre, dass Schüler sich dann im schriftlichen Dialog mit mehr literarischen Worten ausdrücken, einfach weil sie wissen, dass das Teil der Prüfungsleistung ist.

Interessant war außerdem das Ergebnis, dass bei den essenziellen Fragen eine kleine Abnahme im Posttest, im Vergleich zum Pretest, wahrzunehmen ist, obwohl eine Zunahme dieses spezifischen Codes aufgrund der Unterrichtsintervention gerichtet auf *Self-Questioning* naheliegend wäre. Somit kann festgestellt werden, dass essenzielle Fragen nicht unbedingt mit *Reasoning: understanding character* und *Reasoning: understanding story events* einhergehen. Mit anderen Worten wurden die Interpretationsfähigkeiten der Schüler durch die Unterrichtsintervention gesteigert, aber diese Resultate waren unabhängig von der absoluten Anzahl essenzieller Fragen, die von den Schülern gestellt wurden. Dieses Ergebnis könnte man eventuell damit erklären, dass die Schüler nicht unbedingt mehr essenzielle Fragen infolge der Unterrichtsintervention stellen, aber dass sie sich wegen der Intervention ihrer Fragen bewusster sind. Zudem haben die Schüler während der Intervention gelernt, dass ihre eigenen Fragen legitim sind und sie es Wert sind, ihnen nachzugehen. Dadurch könnte der Anstieg von *Reasoning: understanding character* und *Reasoning: understanding story events*, während die essenziellen Fragen ungefähr gleich bleiben, erklärt werden. Eine andere Erklärung dafür, dass keine Steigerung der Menge an essenziellen Fragen wahrgenommen werden kann, könnte auch in der Art des Codierens liegen; Beim Codieren wurde der Begriff ‚Frage‘ sehr lose interpretiert. Das heißt, dass zum Beispiel auch Äußerungen wie „ich bin gespannt, ob ...“ als Frage codiert wurden. Es könnte sein, dass Schüler im Pretest weniger explizite Fragen als im Posttest gestellt haben, weil sie während der Unterrichtsintervention gelernt haben, ihre Fragen zu explizieren. Wenn der Begriff ‚Frage‘ enger definiert worden wäre – wenn beispielsweise ein Fragezeichen für das Codieren einer Frage unbedingt erforderlich gewesen wäre – dann könnten die Ergebnisse möglicherweise völlig anders aussehen. Beim Erstellen des Codebuchs wurde der Entschluss gefasst, dass der Inhalt wichtiger als die Form, beispielsweise die Interpunktion, angesehen wird. Dementsprechend kann jetzt im Nachhinein nicht mehr untersucht werden, ob die Ergebnisse mit einer engeren Definition von ‚Frage‘ anders aussähen.

Bei den essenziellen Fragen war zudem noch auffällig, dass Schüler deutlich mehr essenzielle Fragen zu der Kurzgeschichte *Helm auf* als zu *Stinkefisch* stellten. Das könnte mit dem Schwierigkeitsgrad der Texte zusammenhängen, denn es wäre durchaus vorstellbar, dass Schüler weniger essenzielle Fragen an einen Text stellen, wenn sie den Inhalt des Textes weniger gut verstehen. Die Vermutung lautet also, dass *Stinkefisch* für die Schüler schwieriger als *Helm auf* war. Dieses Ergebnis ist gewissermaßen auch in den beiden *Problem understanding*-Codes vorzufinden, denn die Codes *Problem understanding: without noticing* und vor allem *Problem understanding: self assessed* wurden bei den Dialogen zu *Stinkefisch*

vielmals öfter als bei denen zu *Helm auf* codiert. Die Schüler empfinden die Kurzgeschichte *Stinkefisch* also offenbar als schwieriger als die Kurzgeschichte *Helm auf*. Dieses Ergebnis kommt auch aus dem Fragebogen, den die Schüler nach dem Lesen ausgefüllt haben, hervor (siehe Graphik 3).



Graphik 3: Von Schülern laut Fragebogen empfundene Schwierigkeit (in dem: 1 = einfach, 7 = schwierig) der Kurzgeschichten in Pre- und Posttest.

Eine Frage gibt es dabei allerdings noch zu klären: sind die beiden Kurzgeschichten überhaupt unterschiedlich schwierig? Die Geschichten wurden schlussendlich darauf selektiert, dass der Wortschatz ungefähr auf vergleichbarem Niveau ist. Der Wortschatz ist aber nicht die einzige Ebene, auf welche eine Geschichte schwierig sein kann. Die narrative Komplexität dürfte vor allem bei literarischen Texten auch eine Rolle spielen. Zwar gibt es keine eindeutige Definition für den Begriff ‚narrative Komplexität‘¹⁴⁴, allerdings ist es naheliegend, dass bestimmte Eigenschaften eines Textes ihn einfacher oder schwieriger zu lesen machen. Die Eigenschaften, die in den beiden Kurzgeschichten deutlich differieren, sind in Tabelle 13 ersichtlich. Man könnte argumentieren, dass die Eigenschaften von *Stinkefisch* in diesem Zusammenhang den Text für die Schüler komplexer machen; Der Ich-Erzähler ist aus einer egozentrischen Perspektive normaler als der personale Erzähler, Zeitsprünge fordern allgemein von den Lesern mehr Aufmerksamkeit auf die zeitliche Komponente eines Textes und ein durchgehender innerer Monolog liest sich kontinuierlicher als eine Geschichte mit Unterbrechungen der

¹⁴⁴ Es gibt einige Wissenschaftler, die der Meinung sind, dass die Komplexität eines Narrativs eine Eigenschaft ist, die nur einige Narrativen besitzen, auf der anderen Seite gibt es aber auch Vertreter der Meinung, dass Komplexität eine Eigenschaft ist, die einem Narrativ inhärent ist. (Siehe: Grishakova und Poulaki, *Narrative Complexity*; Steiner, *Komplexes Erzählen - Literatur auf 2+n-ter Stufe*.) Zudem ist wissenschaftlich nicht eindeutig geklärt, was der Begriff ‚Komplexität‘ im Zusammenhang mit Narrativ beinhalten soll. (Siehe: Grishakova und Poulaki, *Narrative Complexity*; Rothmund, *Komplexe Welten: narrative Strategien in US-amerikanischen Fernsehserien*; Steiner, *Komplexes Erzählen - Literatur auf 2+n-ter Stufe: Zu einer Theorie literarischer Komplexität*.) In dieser Arbeit wird davon ausgegangen, dass Komplexität einem narrativen Text inhärent ist und sie nur unterschiedlich, auf einem Spektrum von kaum komplex bis sehr komplex sozusagen, ausgeprägt sein kann.

Erzählart durch Dialoge. Alles in allem ist es absolut vorstellbar, dass die Kurzgeschichte *Stinkefisch* für die Schüler aufgrund ihrer narrativen Komplexität schwieriger als *Helm auf* zu lesen war. Ob der Unterschied der essenziellen Fragen der Schüler zwischen den beiden Geschichten einen kausalen Zusammenhang mit der narrativen Komplexität eines Textes aufweist, ist durch diese Studie nicht festzustellen, wäre aber durchaus eine interessante Frage für weitere Forschung.

Helm auf	Stinkefisch
Ich-Erzähler	Größtenteils personaler Erzähler / Auktorialer Erzähler (ganz am Ende)
Kontinuierliche Zeit ¹⁴⁵	Zeitsprünge
Innerer Monolog	Größtenteils erlebte Rede inklusive Dialoge mit anderen Personen

Tabelle 13: Unterschiedliche Eigenschaften der Kurzgeschichten *Helm auf* und *Stinkefisch*.

¹⁴⁵ Der Junge in *Helm auf* erzählt als Erzählinstanz davon, dass er den Helm gekauft hat, wie andere darauf reagieren und wie er die Zukunft vor sich sieht. Die Erzählinstanz macht dabei selbst keine Zeitsprünge, sondern befindet sich immer am gleichen Zeitpunkt. Die Zeit der Erzählperspektive ist somit kontinuierlich.

7. Entwicklung eines Bewertungsschemas

In diesem Kapitel wird beschrieben, wie das Bewertungsschema mit Fokus auf Interpretation für die Aufgabe des Schreibens eines schriftlichen Dialogs zu einer Kurzgeschichte entstanden ist und welche Entscheidungen dabei aus welchen Gründen erfolgten. Das gesamte Bewertungsschema kann in Kapitel 7.7 eingesehen werden, aber in den einzelnen Kapiteln zu den unterschiedlichen Entscheidungen werden zur Verdeutlichung auch die für das Kapitel relevanten Teile des Bewertungsschema aufgeführt. Es wird zunächst auf die Kategorien auf der y-Achse (Vertikale), auf die bewertet werden soll, eingegangen. Danach werden die Kategorien und die Auswahl für ihre jeweiligen Abstufungen auf der x-Achse erklärt.

7.1 Auswahl der Kategorien

Literarischer Wortgebrauch
Textstruktur

Beschreiben	
Beurteilen	
Interpretieren	Hypothetisch
	Charaktere
	Geschichte
Referieren	

Tabelle 14: Kategorien (y-Achse) des Bewertungsschemas.

Die Auswahl der Kategorien, welche bewertet werden und die sich auf die y-Achse des Bewertungsschemas befinden, erfolgt anhand der Ergebnisse des Codierens und unter Berücksichtigung der wissenschaftlichen Literatur zu Interpretation und Bewertungsschemata.

Die Kategorien Textstruktur und literarischer Wortgebrauch wurden als Bewertungskriterien ausgewählt, sind jedoch unabhängig von den anderen Kategorien. Diese Wahl zur Trennung wurde getroffen, weil diese beiden

Kategorien weniger Abstufungen als die anderen Kategorien enthalten und sie zudem allgemeine Sprachfähigkeiten anstelle von Interpretationsfähigkeiten prüfen. Der literarische Wortgebrauch von Schülern wurde als wichtig genug erachtet, um in das Bewertungsschema aufzunehmen, weil literarischer Wortgebrauch in den *Eindtermen* vorgegeben wird. Zudem ist es von Belang, dass Schüler lernen, dass es Fachsprachen, die der Kommunikation im fachlichen Kontext dienen, gibt. Schüler lernen dadurch, sprachbewusst zu kommunizieren. Die Textstruktur wurde dagegen ausgewählt, auch wenn es dafür keinen Code im Codebuch gibt, weil eine gute Textstruktur von Seiten der Schüler die Lesbarkeit der Texte fördert. Eine bessere Textstruktur führt dazu, dass Argumente besser vermittelt werden und ist somit eine Kernkompetenz, die Schüler in der Schule erlernen sollen. Ein schöner Nebeneffekt einer besseren Textstruktur ist außerdem, dass die Arbeit des Dozenten erheblich erleichtert wird. Anders als man vielleicht erwarten könnte, werden die Kategorien als erstes aufgeführt und deswegen auch als erstes in dieser Arbeit erklärt, obwohl sie natürlich für das eigentliche

Thema des Bewertungsschemas, Interpretationsfähigkeiten, am uninteressantesten sind. Diese Wahl wurde getroffen, weil es in Bewertungsschemata üblich ist, dass Kategorien mit allgemeinen Voraussetzungen zuerst genannt werden.

Bei der Auswahl der vier beziehungsweise sechs anderen Kategorien war es wichtig, dass sie sich auf eine bestimmte Art auf die Codes, die verwendet wurden, um zu bestimmen, was Schüler in den schriftlichen Dialogen machten, beruhten und zudem, dass es auch einen wissenschaftlichen Hintergrund für die Kategorien gibt. Es wurde dabei die Wahl getroffen, die Kategorien mit einem Infinitiv zu bezeichnen, weil dadurch sofort klar wird, welches Verhalten die Schüler pro Kategorie global zeigen sollen; Sie sollen beschreiben, beurteilen, interpretieren und referieren. In den unterschiedlichen Abstufungen zu den Kategorien wird dann noch weiter erklärt, was die Schüler genau zeigen sollen.

Die Kategorie ‚Beschreiben‘ beruht auf den *Descriptive*-Codes und wurde gewählt, weil es beim Lesen in der Fremdsprache bereits eine Leistung ist, den Text wortwörtlich richtig zu verstehen. Schüler machen viele beschreibende Aussagen, wurde in den Ergebnissen des Codierens erkannt. Außerdem stellen McCarthy und Goldmann fest, dass Interpretieren nicht ein unabhängiger Prozess ist, sondern viel mehr auf Elementen der wörtlichen Story aufbaut.¹⁴⁶ Somit gibt es Gründe genug, ‚Beschreiben‘ als Kategorie in das Bewertungsschema aufzunehmen.

Die Kategorie ‚Beurteilen‘ basiert auf den *Evaluative*-Codes und wurde für den Prozess des Erlernens des Interpretierens als wichtig erachtet, weil einerseits Schüler von sich selbst aus bereits beurteilende Aussagen treffen, wie in den Ergebnissen des Codierens zu sehen ist. Andererseits wurde in Kapitel 2.2, ‚Interpretieren‘, zudem festgestellt, dass leistungsstarke Schüler sich öfter persönlich und subjektiv mit einem literarischen Text auseinandersetzen. Dementsprechend war es sinnvoll, die Kategorie ‚Beurteilen‘ in das Bewertungsschema aufzunehmen.

Die Kategorie ‚Interpretieren‘ stellt ein zentrales Element dar, denn das Bewertungsschema wird konzipiert, um die Interpretationsfähigkeiten von Schülern bewerten zu können. Es wurde beschlossen, die Kategorie in drei Subkategorien, nämlich ‚Hypothetisch‘, ‚Charaktere‘ und ‚Geschichte‘, den *Reasoning*-Codes *hypothetical*, *understanding character* und *understanding story events* entsprechend, zu unterteilen. Die Wahl zur Unterteilung in Subkategorien geschieht aus dem Grund, dass diese drei Codes sich

¹⁴⁶ Vgl. McCarthy und Goldman, „Comprehension of Short Stories: Effects of Task Instructions on Literary Interpretation“, S. 3.

doch stark voneinander unterscheiden und es deswegen auch schwierig wäre, sie in einer einzelnen Kategorie ‚Interpretieren‘ zusammenzufassen. Die Ergebnisse der Interrater-Reliabilität haben gezeigt, dass die *Reasoning*-Codes untereinander gut zu unterscheiden waren und deswegen kann darauf vertraut werden, dass auch Dozenten mit klaren Anweisungen im Bewertungsschema diesen Unterschied gut wahrnehmen können.

Die letzte Kategorie ‚Referieren‘ wurde ausgewählt, weil sie zum einen zu den *Referential*- und *Reasoning: moral*-Codes passt. Zum anderen wurde in Kapitel 2.2, ‚Interpretieren‘, festgestellt, dass Schüler sich beim Interpretieren oft auf ihre eigenen Erfahrungen mit der Welt berufen. In den Ergebnissen des Codierens kann gesehen werden, dass Schüler tatsächlich regelmäßig referierende Aussagen treffen. Aus diesen Gründen wurde also die Kategorie ‚Referieren‘ dem Bewertungsschema hinzugefügt.

Wie in Kapitel 2.4.3 beschrieben wurde, ist es wichtig, dass die verschiedenen Bewertungskriterien gut voneinander zu unterscheiden sind und dass die Arbeit des Bewertens für die Dozenten nicht zu viel Zeit beschlagnahmt. Deswegen empfiehlt CITO, maximal fünf Bewertungskriterien aufzustellen. Genaugenommen gibt es in dem in Kapitel 7.7 vorgestellten Bewertungsschema sechs Hauptbewertungskriterien und insgesamt acht Bewertungskriterien. Es wurde dabei bewusst von der Empfehlung von CITO abgewichen. Diese Wahl erfolgte aus dem Grund, dass das Bewertungsschema der Vielfalt der Aussagen von Schülern gerecht werden sollte; Das Bewertungsschema soll die Schüler in ihren Aussagen nicht einschränken.

7.2 Literarischer Wortgebrauch und Textstruktur

Literarischer Wortgebrauch	Der Schüler verwendet <u>keine</u> literarischen Terminologien.	Der Schüler verwendet einige literarische Terminologien.	
Textstruktur	Unzureichend.	Interpunktion ist in Ordnung.	Interpunktion ist in Ordnung + es werden konsequent Signalwörter verwendet, um Textstrukturen anzudeuten.



Tabelle 15: Abstufungen der Kategorien ‚literarischer Wortgebrauch‘ und ‚Textstruktur‘ des Bewertungsschemas.

Im Bewertungsschema sind die Kategorien ‚literarischer Wortgebrauch‘ und ‚Textstruktur‘ Sonderkategorien, weil sie sich nicht wirklich auf Interpretation beziehen und sie weniger Abstufungen als die anderen Kategorien enthalten. Die Abstufungen für diese Kategorien sind anders gestaltet worden, weil vier Abstufungen für diese beiden Kategorien keinen Sinn ergeben hätte. Für die Kategorie ‚literarischer Wortgebrauch‘ gibt es nur zwei Abstufungen; Entweder werden gar keine literarischen Wörter benutzt oder es werden einige verwendet. Die

Binarität dieser Kategorie ist beabsichtigt, weil die Schüler in *4havo* und *4vwo* das Ziel der *Eindtermen* der Verwendung von literarischem Wortschatz, das übrigens sowieso nur für die *vwo*-Abiturienten gilt, noch nicht erreicht haben sollen. Diese Kategorie ist nur als Signal da, dass die Verwendung von literarischen Worten wertgeschätzt wird. Schüler sollten nämlich lernen, dass es für Fachbereiche ein bestimmtes Vokabular, das von der Standardsprache abweicht, gibt. Mehrere Abstufungen hätten dazu führen können, dass Schüler irgendwelche schwierigen fachsprachlichen literarischen Worte, die gar nicht zu ihrem Wortgebrauch passen, gekünstelt einsetzen würden. Das wäre nicht wünschenswert. Beim Codieren der schriftlichen Dialoge (Code: *Literary words*) hat sich herausgestellt, dass Schüler literarische Worte von alleine bereits verwenden (siehe Bsp. 1: Titel, Bsp. 2: Thema, Bsp. 3: ik-personage).

Bsp. 1: Vind je de titel er goed bij passen?

Bsp. 2: Als je het thema in  n wordt moet omschrijven¹⁴⁷

Dan zou dat zijn geruchten

Bsp. 3: Nee, je krijgt ook bijna niets te weten over de ik-personage

Es scheint also fair, diese Schüler zu belohnen und den Schülern, die das vielleicht noch nicht von sich aus machen würden, einen Anreiz zu geben, auch einige literarische Worten zu verwenden.

Die Kategorie ‚Textstruktur‘ ist als Begriff noch ziemlich vage, wird aber in den Abstufungen erläutert und eingeschränkt. Wie man in Tabelle 15 lesen kann, werden lediglich die Interpunktion und die Verwendung von Signalwörtern bewertet. Die Entscheidung dazu wurde getroffen, weil beim Codieren auffiel, dass es ziemlich viele Schüler gab, die ihre Texte relativ schlecht strukturiert hatten, was dazu führte, dass es für die Raterin nicht immer klar war, was der Schüler überhaupt gemeint hat und die Raterin dementsprechend selbst eine Interpretation des Geschriebenen machen musste. In Bsp. 4 kann man beispielsweise vermuten, dass der Schüler eine Warum-Frage stellt, dafür wird aber das falsche Signalwort verwendet und die Frage wird inkorrekt mit einem Punkt signalisiert. Deswegen kann nicht eindeutig festgestellt werden, was der Schüler hier sagen möchte. Auch die fehlende, fehlerhafte und/oder inkonsequente Interpunktion bei einigen Dialogen hat das Codieren erschwert (siehe Bsp. 5, in dem die Sätze nicht immer mit einem Punkt abgeschlossen werden, und Bsp. 6, in dem die Fragezeichen konsequent fehlen).

Bsp. 4: Waar zou hij die helm kopen en nooit meer afdoen.

Bsp. 5: Nee, ik had hetzelfde verder was het verhaal niet zo spannend

Vond ik ook Ik snap het verhaal ook niet echt.

ik ook niet Ik weet alleen dat hij sigaretten wilde stelen met zijn vrienden.

¹⁴⁷ Die schwarzkarierten Fragezeichen in diesem und folgenden Beispielen sind da, weil das Codierungsprogramm nicht mit allen diakritischen Zeichen umgehen kann.

Bsp. 6: Vond je het een leuk verhaal
 nee
 Zou je het verhaal aanbevelen
 nee
 Waarom durfde niemand met die man te praten
 geen idee
 Was jij nieuwsgierig naar het einde

Die Aufnahme der Abstufungen mit Schwerpunkt Interpunktion und Signalwörter im Bewertungsschema soll also dazu führen, dass die Hantierbarkeit der schriftlichen Dialoge für die Dozenten deutlich gesteigert und somit die Arbeit des Prüfens erleichtert wird.

7.3 Beschreiben

Erzählt in eigenen Worten nach, was in der Geschichte passiert ist.	Erzählt in eigenen Worten nach, was in der Geschichte passiert ist + schildert dabei die Hauptlinien der Geschichte.	Erzählt in eigenen Worten nach, was in der Geschichte passiert ist + benutzt dies um Argumente zu unterstützen + identifiziert mindestens eine Lücke im Text.	Erzählt in eigenen Worten nach, was in der Geschichte passiert ist + benutzt dies um Argumente zu unterstützen + identifiziert mehrere Lücken im Text.
---	--	---	--

Tabelle 16: Abstufungen der Kategorie ‚Beschreiben‘ des Bewertungsschemas.

Die Kategorie ‚Beschreiben‘ ist wichtig für den Prozess des Erlernens des Interpretierens, weil ein gutes Verständnis (von den Hauptlinien) des Textes eine Voraussetzung für das Interpretieren ist. Um zu einer Interpretation zu gelangen, muss nämlich auf das im Text Geschriebene zurückgegriffen werden, um schließlich zu Inferenzen und Interpretationen zu gelangen. Die erste Spalte in Tabelle 16 entspricht dem niedrigsten Niveau vom Beschreiben, nämlich dass überhaupt irgendwelche Teile der Geschichte nacherzählt werden. In der zweiten Spalte sollen dann Schüler auch schon die Hauptlinien der Geschichte nacherzählen. Diese zwei Abstufungen wurden gewählt, weil beim Codieren aufgefallen ist, dass Schüler sich manchmal an einem Detail des Textes aufhängen und dann sowohl mit der Geschichte als auch beim Interpretieren nicht mehr weiterkommen. In Bsp. 7 kann man dies gut sehen, denn ein Schüler sagt, dass für ihn herausstach, dass Stinkefisch blind war, weil er das nicht erwartet hätte und geht daraufhin nur noch auf diesen kleinen relativ unwichtigen Teil der Geschichte ein. Der Rest des Textes wird einfach vergessen. Die Abstufungen, die einen Unterschied zwischen Nacherzählen und Nacherzählen der Hauptlinien machen, sollen einen Hinweis für die Schüler sein, dass nur irgendwelche Teile nachzuerzählen, weniger wünschenswert als das Schildern der Hauptlinien der Geschichte ist. Dabei ist ausdrücklich nicht gemeint, dass

Schüler die Geschichte zusammenfassen sollen. In den Ergebnissen wurde nämlich deutlich, dass nicht einfach zwischen Nacherzählen und Zusammenfassen unterschieden werden kann. Es geht bei dem Beschreiben von Hauptlinien vielmehr darum, dass Schüler ein paar wichtige Geschehnisse in ihrem Dialog aufgreifen. Das muss nicht direkt hintereinander passieren, sondern einfach nur über den Dialog verteilt.

Bsp. 7: dat die Stinkvis blind was.
hoezo dat?
dat verwacht niemand.
-
er zijn relatief maar weinig mensen blind.
wat zou dan de reden van die bril op hebben?
dat is een goeie vraag.
vertel....
misschien zich zo min mogelijk blind te [doen]
oké
heb jij een [ander] reden?
hij kan zich slimmer laten overkomen.
Nee, dat is onzin. blinde zijn bijna altijd kasplantjes.
hoho, niet zo snel!
waarom niet?
Ik ken een blinde die HBO heeft afgerond
[...]

Für die dritte und vierte Spalte, die höchste Abstufungen der Kategorie ‚Beschreiben‘, wurde der Code *Essential question* mit dem Ergebnis der Studie von Earthman (wie beschrieben in Kapitel 2.2), dass erfahrene Leser sich mehr als weniger erfahrene Leser auf schwieriger zu füllende Lücken im Text fokussierten, verbunden. Da es sich bei der Studie von Earthman um Studenten gehandelt hat, ist dieses Ergebnis nicht ohne Anpassung auf das Bewertungsschema übertragbar. Dementsprechend wird im Bewertungsschema kein Unterschied zwischen schwieriger und einfacher zu füllenden Lücken gemacht, sondern nur festgelegt, dass es wichtig ist, dass Schüler überhaupt erkennen, dass es Lücken im Text gibt. Ohne diese Erkenntnis ist es für die Schüler sowieso unmöglich, festzustellen, dass die Texte für die Lücken eine Interpretation erfordern. Das alles hängt mit dem Code *Essential question* zusammen, weil Schüler, um eine essenzielle Frage stellen zu können, auch erstmal eine Lücke identifizieren müssen.

Der Unterschied zwischen der dritten und vierten Spalte wird dadurch gemacht, ob Schüler nur eine Lücke oder bereits mehrere wahrnehmen. Wenn mehrere Lücken wahrgenommen werden, gibt es auch mehr Anhaltspunkte für Interpretationen. Das Wahrnehmen von Lücken wird noch zum Beschreiben hinzugerechnet, weil es hier wirklich erstmal nur um eine Wahrnehmung handelt. Die ist an sich schon wertvoll in dem Prozess hin zur Interpretation, auch wenn die Schüler noch nichts mit dieser Wahrnehmung machen.

Inwieweit sie versuchen, die Lücke(n) zu erklären und dabei zu interpretieren, wird in den Subkategorien der Kategorie ‚Interpretation‘ bewertet.

Ein Element der dritten und vierten Abstufung der Kategorie ‚Beschreiben‘ wurde bisher noch nicht beleuchtet, soll aber noch erklärt werden, gerade weil es in den ersten beiden Abstufungen noch nicht vorzufinden ist. Es handelt sich dabei um die Anforderung, dass Schüler Nacherzählungen der Geschichte in ihren Argumenten als Unterstützung für eben diese Argumente aufgreifen. Die Frage, die man sich stellen könnte, ist: wozu sollen die Schüler nacherzählen? Schlussendlich geht es darum, dass die Schüler ihre Meinungen und Interpretationen untermauern können und diese Untermauerung soll wenigstens zum Teil auf ausgewählten Textabschnitten beruhen, will sie eine bestimmte Qualität haben. Schüler machen diesen Versuch bereits vielfältig, wie man an den Mengen der *Descriptive-Codes* erkennen kann (siehe Tabelle 7) und wie man in Bsp. 8 und 9 lesen kann. Deswegen wurde die Einschätzung gemacht, dass die zweite auf die dritte Abstufung direkt mit zwei Kriterien (Beschreibung in Argumentation & Wahrnehmung von einer Lücke) aufgebaut werden kann.

- Bsp. 8:* Precies, de jongen zegt de hele tijd dat hij de helm niet zal afzetten, maar als zijn vader komt wel.
 Ja, misschien is er wel iets met de vader gebeurt.
 Zou kunnen, ik vraag me af wat.
 Ja, en wat maakt dat de jongen de helm draagt?
 Misschien heeft het wel met elkaar te maken
 [...]
 Maar ik denk dat de jongen er wel een reden voor heeft.
 Ja, wat dan?
 Misschien heeft de vader wel een ongeluk gehad. en voelt de jongen zich veiliger met helm
- Bsp. 9:* Het was wel opmerkelijk dat er een onderscheid werd gemaakt tussen de wijken.
 Zou de schrijver daarmee een punt willen maken?
 Ik denk wel dat de schrijver een verschil maakt tussen verschillende nationaliteiten en arm en rijk.
 De schrijver laat hier eigenlijk zien dat buitenlanders sneller overlast veroorzaken.

7.4 Beurteilen

Beurteilt den Text und/oder Charaktere + begründet dies <u>nicht</u> .	Beurteilt den Text + begründet dies.	Beurteilt Charaktere + begründet dies.	Beurteilt Charaktere + begründet dies + bezieht sich in der Begründung darauf, inwieweit er/sie sich mit Charakteren identifizieren kann.
--	--------------------------------------	--	---

Tabelle 17: Abstufungen der Kategorie ‚Beurteilen‘ des Bewertungsschemas.

Beurteilen als Kategorie legitimiert die Tatsache, dass Schüler eine Meinung über das Gelesene haben. Es gibt aber verschiedene Möglichkeiten das Gelesene zu beurteilen, denn zum einen kann der Text als Ganzes beurteilt werden und zum anderen kann ein oder mehrere Charakter(e) beurteilt werden. Dabei ist während des Codierens der schriftlichen Dialoge

aufgefallen, dass es Schüler gibt, die ihr Urteil begründen (siehe die beiden *Reasoning: evaluative*-Codes und Bsp. 10 und 11), und dass es auch Schüler gibt, die dies eben nicht tun (siehe die beiden *Evaluative*-Codes und Bsp. 12 und 13). Beurteilen und nicht begründen landet im Bewertungsschema in der ersten Spalte, da Schüler sich dafür kaum Gedanken über das Gelesene machen müssen. Diese Art des Beurteilens wird daher am wenigsten belohnt.

Bsp. 10: Ik vond het stukje van het kopen van een slot voor de helm wel apart
kun je mij vertellen waarom dat stukje zo apart was?

ja, hoor. Ik vond het een beetje overdreven dat hij een slot ging kopen.

Bsp. 11: maar waarom vond jij het precies zo interessant?

Doordat ik het heel spannend vond wat de blinde man zou zeggen, hoe hij erop zou reageren.

Bsp. 12: Wat vond je van Stinkfisch dan?

Een beetje raar

Waarom?

Ik vond het hele verhaal gewoon raar

Bsp. 13: ik vond het ook raar dat hij een ketting koopt

An zweiter und dritter Stelle befinden sich die beurteilenden Aussagen, bei dem die Schüler ihre Aussagen auch begründen. Das Begründen eines Urteils über das Gelesene wird als wertvoll für den Prozess des Interpretierens eingeschätzt, weil Schüler sich dabei fragen müssen, was genau an dem Text dran ist, weswegen sie ein bestimmtes Urteil fällen. Wenn sie auf diese Art und Weise über den Text nachdenken, ist der Weg frei, um festzustellen, dass ihnen zum Beispiel noch Informationen fehlen, für die es wünschenswert wäre, wenn sie einen Versuch unternehmen würden, zu interpretieren.

In den beiden Abstufungen des Beurteilens inklusive Argumentation (Spalte zwei und drei) wird ein Unterschied zwischen der Beurteilung des Textes in seiner Gänze und der Beurteilung eines oder mehrerer Charaktere gemacht. Dieser Entschluss wurde mit dem Gedankengang gefasst, dass man sich tiefer mit einem Text auseinandersetzen muss, um einen Charakter beurteilen zu können. Man muss sich fragen, was einen Charakter treibt, weshalb der Charakter bestimmte Entscheidungen trifft und der Weg von dahin bis zur Interpretation ist gar nicht so weit. Das Beurteilen eines oder mehrerer Charaktere wird demzufolge höher als das Beurteilen des Textes eingestuft.

Unter Berücksichtigung der Tatsache, dass leistungsstärkere Schüler sich häufiger emotional mit einem Text auseinandersetzen (siehe Kapitel 2.2), und der Einstufung der empathischen Identifikation mit Charakteren in Niveau 4F des niederländischen Referenzrahmen Sprache (siehe Kapitel 2.4.2), scheint es sinnvoll, die Identifizierung mit Charakteren in der höchst bewerteten Abstufung des Beurteilungsschema aufzunehmen. Beschlossen wurde, dass dieses Kriterium der Identifizierung gut zur Kategorie ‚Beurteilen‘ passt, weil denkbar ist, dass die (nicht) vorhandene Identifizierung mit einem Charakter auch zu einem Urteil über diesen Charakter führen kann. Die mögliche Identifizierung passt also gut

in der Argumentationsstruktur zu einem oder mehreren Charakteren hinein. Bei der Einführung dieses Kriteriums in das Bewertungsschema war, anders als im Niveau 4F, wichtig, dass den Schülern dabei freigelassen wird, ob sie argumentieren weswegen sie sich schon oder eben nicht mit einem Charakter identifizieren können. Es müssen alle Gefühle des Schülers wertgeschätzt werden, damit sie nicht für eine Prüfung in eine Richtung, in der sie sich nicht wohl fühlen, gedrückt werden. Dieses Kriterium entspricht natürlich der höchsten Abstufung der Kategorie ‚Beurteilen‘ und muss somit von den Schülern noch nicht erreicht werden. Leistungsstärkere Schüler könnte diese Abstufung aber auf die Idee bringen, dass begründete Aussagen zur möglichen Identifizierung mit einem Charakter auch wertvoll sind und dementsprechend in der Prüfung wertgeschätzt werden.

7.5 Interpretieren

7.5.1 Hypothetisch

Spekuliert über das Ende / alternative Handlungsverläufe oder macht Verbesserungsvorschläge an den Text.	Macht Verbesserungsvorschläge an den Text und basiert sich dabei auf eigene Wünsche / Leseerfahrungen.	Spekuliert über das Ende / alternative Handlungsverläufe oder macht Verbesserungsvorschläge an den Text + argumentiert dabei anhand des Textes.	Spekuliert über das Ende / alternative Handlungsverläufe oder macht Verbesserungsvorschläge an den Text + argumentiert dabei anhand Textes + zeigt dabei Verständnis für die Wahl des Autors und versucht diese Wahl auch zu erklären.
--	--	--	--

Tabelle 18: Abstufungen der Kategorie ‚Interpretieren: hypothetisch‘ des Bewertungsschemas.

Die Kategorie ‚Interpretieren: hypothetisch‘ beruht auf dem Code *Reasoning: Hypothetical*. Für die Entscheidung, welche Abstufungen für diese Kategorie gewählt werden sollten, wurde vor allem in den *Hypothetical*-Codes geschaut, was die Schüler geschrieben haben und welche Aussagen über den Text wertvoller als andere sind und warum dies der Fall ist. Erst für die vierte Spalte gibt es keine Beispiele aus den schriftlichen Dialogen und stattdessen wurde überlegt, was noch wünschenswert und eventuell für die leistungsstärkeren Schüler auch erreichbar wäre.

Bsp. 14: Komt de vader denk je terug?
Ik denk het niet jij?
Ik zou het niet weten, geen idee.

- Bsp. 15:* Zou je het vervolg willen lezen
 Ja, ben benieuwd.
 Ik ook, hoe denk jij dat het eindigt?
 Ik denk dat de politie de stenengooier straf geeft.
 Dat denk ik ook.
 Denk jij dat hij er van gaat leren?
 Ik denk dat hij het niet nog een keer zou doen.
- Bsp. 16:* Toen de jongens di foto zagen hoop ik dat ze erg schrokken
 Ja ik ook, dat ze er iets van leren
- Bsp. 17:* ik zou ook graag nog een stuk bij het verhaal willen lezen over het moment dat hij de helm af zet
 Ja, dat zou het verhaal wel leuker maken
 En ook iets duidelijk over waarom hij de helm afzet
 Ja, en ook over hoe de mensen om de ik-persoon heen reageren als de helm af is
- Bsp. 18:* Ik ben ook wel benieuwd hoe het af zou lopen.
 Ja, of Tom nog met jongens om gaat.
 En of de man nog leeft.
 Ik denk van wel, en dat ze hun excuses hebben aangeboden.
 Ik denk dat alleen Tom dat gedaan heeft.
 Waarom denk je dat?
Omdat Tom als eerste in zag dat de man helemaal geen moordenaar was.

Wenn man die Beispiele 14 und 15 mit den Beispielen 16 und 17 vergleicht, fällt auf, dass in den beiden ersten Fällen einfach spekuliert wird, während in Bsp. 16 und 17 zudem noch einen Wunsch ausgesprochen wird. Der Wunsch ist persönlich und wird nicht weiter erklärt, aber er stellt, im Gegensatz zu der reinen Spekulation in Bsp. 14 und 15, eine Argumentation dar. Aus diesem Grund wurde der Entschluss gefasst, dass diese Beispiele unterschiedlich bewertet werden sollten, nämlich die Spekulation inklusive Argumentation besser als die ohne. Schlussendlich ist es für jede Art von Beschäftigung mit einem Text, ob es jetzt um das Äußern einer Meinung oder um eine Interpretation des Textes handelt, wichtig, dass die Schüler Argumente liefern können. Sonst halten ihre Ansichten, wie gut sie auch sein mögen, schlichtweg nicht Stand.

Noch besser ist es aber, wenn Schüler aus dem Text heraus argumentieren, wenn sie über den Text spekulieren. Das heißt nicht, dass sie keine Wünsche mehr äußern dürfen, sondern dass sie sich zusätzlich noch Anhaltspunkte im Text suchen, um ihre Spekulation zu erklären. In Beispiel 18 kann man lesen, dass ein Schüler über das Ende des Textes heraus spekuliert und dabei den Gedankengang an einer Gegebenheit aus dem Text festmacht. Solche Überlegungen, bei denen anhand des Textes argumentiert wird, sollen besser bewertet werden als Argumentation anhand persönlicher Wünsche, weswegen für diese Überlegungen die dritte Abstufung gewählt wurde.

Für die vierte Spalte der Kategorie ‚Interpretieren: hypothetisch‘ gibt es schließlich kein Beispiel aus den schriftlichen Dialogen dieser Studie. Das ist aber nicht unbedingt schlecht, da die vierte Spalte sowieso die am schwierigsten zu erreichende Spalte sein soll und nur für die leistungsstärksten Schüler erreichbar sein soll. Die Verknüpfung zwischen Spekulation über

den Text und die Entscheidungen des Autors liegt nahe. Wenn Schüler nämlich darüber nachdenken, was auch anders hätte laufen können oder wie die Geschichte weitergehen könnte, ist die Frage, weshalb das aber vom Autor nicht so gewählt wurde, eine logische Fortsetzung. In den schriftlichen Dialogen dieser Studie gibt es Schüler, die sich Fragen, weswegen der Autor bestimmte Entscheidungen getroffen hat (siehe Bsp. 19 und 20). Diese Aussagen wurden mit *Reasoning: understanding story events* codiert. Dabei fällt auf, dass Schüler meistens auch einen Versuch unternehmen, diese Entscheidung zu erklären. Für den Prozess des Interpretierens ist dies auch von Belang, weil auf diese Art und Weise der Weg für ein tieferes Verständnis des Textes eröffnet wird. Entsprechend der Schwierigkeit wurde das Kriterium, dass Schüler ein Verständnis für die Entscheidung des Autors äußern und diese Wahl auch versuchen zu erklären, als vierte Abstufung für die Kategorie ‚Interpretieren: hypothetisch‘ ausgewählt.

Bsp. 19: Het was wel opmerkelijk dat er een onderscheid werd gemaakt tussen de wijken.

Zou de schrijver daarmee een punt willen maken?

Ik denk wel dat de schrijver een verschil maakt tussen verschillende nationaliteiten en arm en rijk.

De schrijver laat hier eigenlijk zien dat buitenlanders sneller overlast veroorzaken.

De **rijkere** kinderen werden door hun ouders naar school gebracht, die worden altijd gecontroleerd.

Dus eigenlijk hebben ze minder vrijheid?

Ja, zo zou je het kunnen benoemen

Bsp. 20: Zouden ze die man nou alleen zo noemen door geruchten?

Ja, maar ook door de stank.

Terwijl die man eigenlijk gewoon oud en blind is aan **zijn** oog.

Ze noemen hem zo uit vooroordelen.

Echter, als Tom hem van dichtbij ziet is het heel anders

Wat zou de schrijver hiermee willen bereiken?

Er zijn twee mogelijkheden.

Vertel

Hij wil duidelijk maken dat buitenlanders sneller vooroordelen.

Of?

Hij wil duidelijk maken dat voor oordelen fout zijn.

7.5.2 Karaktere

Stelt keine oder wenig Fragen über (das Handeln von) Charaktere(n) + versucht <u>nicht</u> , diese Fragen zu beantworten.	Stellt Fragen über (das Handeln von) Charaktere(n) + versucht, diese Fragen zu beantworten.	Stellt Fragen über (das Handeln von) Charaktere(n) + versucht, diese Fragen zu beantworten + bietet dafür mehrere Perspektiven an.	Stellt Fragen über (das Handeln von) Charaktere(n) + versucht, diese Fragen zu beantworten + bietet dafür mehrere Perspektiven an + argumentiert für die unterschiedlichen Perspektiven.
---	---	--	--

Tabelle 19: Abstufungen der Kategorie ‚Interpretieren: Charaktere‘ des Bewertungsschemas.

Die Kategorie ‚Interpretieren: Karaktere‘ stimmt met dem Code *Reasoning: understanding character* überein. Um zu bestimmen, wie die Abstufungen dieser Kategorie aussehen sollten, wurde zunächst auf Beispiele in den Codes geschaut. Außerdem wurde zur Kenntnis genommen, dass erfahrenere Leser oft mehrere Perspektiven einnehmen, wie in Kapitel 2.2 beschrieben wurde. Während des Codierens war bereits aufgefallen, dass manche Schüler verschiedene Perspektiven einnahmen, während andere das nicht machten. Vor allem im Bereich *Reasoning: understanding character* wurden unterschiedliche Perspektiven innerhalb eines schriftlichen Dialogs eingenommen. Dieses Kriterium, dass Schüler mehrere Perspektive einnehmen, sollte also in der Konzipierung der Kategorie ‚Interpretieren: Karaktere‘ einbezogen werden, aber erst in den fortgeschrittenen Abstufungen.

Bsp. 20: Ja, ze dachten al die tijd verkeerd voer de man.

Ja, inderdaad. maar waarom zouden ze dat denken.

Goeie vraag, ze kenden de man niet eens persoonlijk.

Bsp. 21: je weet niet wat voor helm en waarvoor.

wat had je in gedachten bij die helm?

zo'n mooie gele

waarvan?

gesponsort door Jumbo

Bsp. 22: Maar ik denk dat de jongen er wel een reden voor heeft.

Ja, wat dan?

Misschien heeft de vader wel een ongeluk gehad. en voelt de jongen zich veiliger met helm zou kunnen.

Bsp. 23: Waarom hield Tom de jongens eigenlijk niet tegen toen ze stenen gooiden.

Misschien was hij bang dat ze naar hem zouden gooien.


Für die erste und zweite Spalte bleiben zum Entscheiden der Kriterien der Abstufungen nur die Beispiele ohne Einsicht in mehreren Perspektiven übrig. In den Beispielen konnten zwei Möglichkeiten vorgefunden werden. Einerseits gibt es Schüler, die wenig oder keine Fragen über (das Handeln von) Charaktere(n) stellen und auch keinen Versuch unternehmen, diese Fragen zu beantworten (siehe Bsp. 20 und 21). Andererseits gibt es Schüler, die diese Fragen stellen und schon versuchen, sie zu beantworten (siehe Bsp. 22 und 23). Die Schüler, die einen Versuch machen, einen Charakter zu verstehen, sollen dabei besser als die, die es nicht versuchen, bewertet werden.

Bsp. 24: Die jongen spoort gewoon niet

Dat is duidelijk

Of hij mist zijn vader echt

Bsp. 25: Waarom zou die jongen dat denk je doen voor zijn vader

weet t niet zeker, is denk ik omdat hij een statement wou maken hoe bedoel je?

nou om te laten zien zovan pas als pap terug komt zet ik hem af, als chantage middel

Ik heb een ander idee waarom hij het zou willen

oh, vertel!

Omdat zijn vader is omgekomen bij een ongeluk, en als de vader een helm op had gehad zou hij niet dood zijn.

Bsp. 26: Zou de man het eigenlijk overleven?

Hij lag in het ziekenhuis op het laatst. Maar volgens mij ging de man toch dood.

Zijn huis werd wel gesloopt en er werd een garage voor auto's gemaakt Dus misschien leeft hij niet meer.

Het zou ook nog kunnen dat de man nu ergens anders woont.

Klopt, ik kan me voorstellen dat hij niet meer bij die kinderen in de buurt wil wonen.

Für die dritte und vierte Spalte wurde in den schriftlichen Dialogen geschaut, in welchen Fällen Schüler mehrere Perspektiven einnehmen und ob all diese Fälle gleich bewertet werden sollen. Daraus hat sich ein Qualitätsunterschied zwischen den Beispielen ergeben, denn in Beispiel 24 werden zwar unterschiedliche Perspektiven eingenommen, aber Bsp. 25 und 26 sind deutlich besser. Der Unterschied liegt darin, dass in Beispiel 24 einfach zwei Möglichkeiten genannt werden und diese jedoch nicht weiter begründet werden. In Beispielen 25 und 26 dagegen wird für beide Perspektive erklärt, warum der Schüler sie als Möglichkeit betrachtet. Dass die Schüler aus Bsp. 25 und 26 versuchen zu erklären, welche Perspektiven weswegen möglich wären, zeigt, dass sie sich intensiver mit dem Text auseinandersetzen. Das soll also besser bewertet werden und ist dementsprechend in der vierten Spalte eingesetzt worden, während bloß die Benennung mehrerer Perspektiven als dritte Abstufung gewählt wurde.

Übrigens ist Beispiel 24 ein gutes Beispiel dafür, dass Schüleraussagen nicht immer hundertprozentig in eine Abstufung hineinpassen. Der Schüler in Beispiel 24 stellt nämlich im schriftlichen Dialog nicht explizit die Frage, warum der Junge den Helm trägt, allerdings kann man aus seiner Aussage in Beispiel 24 schon ableiten, dass der Schüler sich diese Frage gestellt hat. Ziel der Unterrichtsintervention war unter anderem, dass Schüler lernen, solche Fragen für sich selbst zu explizieren. Das hat dieser Schüler nicht gemacht. In dem Fall ist es dem Dozenten überlassen, in welche Abstufung solch eine Aussage eingeteilt wird; Das könnte eventuell auch abhängig sein vom Kontext und der allgemeinen Qualität des schriftlichen Dialogs dieses Schülers. Die Erwartung ist aber, dass solche Fälle nicht mehr auftauchen, wenn die Schüler genau wissen, welche Anforderungen das Bewertungsschema stellt. Anders als in der wissenschaftlichen Situation, bei der die Schüler nicht wirklich wussten, was genau die Anforderungen waren, sollten Schüler in einer Prüfungssituation, bei der sie das Bewertungsschema schon eingesehen haben, Bescheid wissen, dass sie Fragen explizit nennen sollen, um bestimmten Anforderungen zu entsprechen.

7.5.3 Geschichte

Stellt keine oder wenig Fragen über Intentionen / Bedeutungen / Inhalte des Textes + versucht <u>nicht</u> , diese Fragen zu beantworten.	Stellt Fragen über Intentionen / Bedeutungen / Inhalte des Textes + versucht, diese Fragen zu beantworten.	Stellt Fragen über Intentionen / Bedeutungen / Inhalte des Textes + versucht, diese Fragen zu beantworten + argumentiert dabei aus dem Text heraus.	Stellt Fragen über Intentionen / Bedeutungen / Inhalte des Textes + versucht, diese Fragen zu beantworten + argumentiert dabei aus dem Text + aus den Kenntnissen der Welt / der Literatur / des Genres heraus.
---	--	---	---

Tabelle 20: Abstufungen der Kategorie ‚Interpretieren: Geschichte‘ des Bewertungsschemas.

Für die Einteilung der Abstufungen der Kategorie ‚Interpretieren: Geschichte‘ wurde zunächst auf die Beschreibung des Codes *Reasoning: understanding story events* im Codebuch (siehe Anhang) geschaut. Der Code umfasst Aussagen/Argumentationen von Schülern über Intentionen des Autors, Bedeutungen von Elementen des Textes und Inhalte des Textes. Damit ist dieser Code ziemlich umfangreich. Es wurde die Entscheidung getroffen, diese Kategorie im Bewertungsschema genauso umfangreich zu gestalten, weil der Umfang der Kategorie dazu führt, dass Schüler eine relativ große Freiheit darin haben, was sie an dem Text interessiert beziehungsweise worüber sie sprechen wollen. Das Bewertungsschema steuert sie in diesem Sinne nicht, nur über Autorenintentionen oder bestimmte Elemente des Textes zu sprechen.

Die ersten beiden Abstufungen dieser Kategorie gleichen den ersten beiden Abstufungen der Kategorie ‚Interpretieren: Charaktere‘. Die Wahl der Abstufungen wurde aus den gleichen Gründen wie bei ‚Interpretieren: Charaktere‘ getroffen, nämlich weil es wichtig ist, dass Schüler überhaupt Fragen stellen und dann lernen, dass es auch wichtig ist, einen Versuch zu machen, diese Fragen zu beantworten. Dies haben die Schüler in der Unterrichtsintervention auch gelernt. In den schriftlichen Dialogen konnten auch für diese Kategorie für beide Abstufungen ausreichend Beispiele gefunden werden. In Beispielen 27 und 28 werden nur Fragen über den Text gestellt, wie in der ersten Spalte dieser Kategorie, während Schüler in den Beispielen 29, 30 und 31 auch versuchen, die von ihnen gestellte Frage zu beantworten. Die Wahl für die Angleichung dieser Abstufungen zwischen den Kategorien ‚Interpretieren: Charaktere‘ und ‚Interpretieren: Geschichte‘ erhöht außerdem die Übersichtlichkeit des Bewertungsschemas.

Bsp. 27: Ik vroeg me af waarom hij die steen gooide
Dat vroeg ik me ook af?

Bsp. 28: Wat heeft het eerste deel eigenlijk met het verhaal te maken?
Dat stuk over hoe anders het dorp was waar het zich afspeelde?
Ja die, die kon ik niet echt plaatsen.

Bsp. 29: Misschien is er wel iets met de vader gebeurt.
[...]
Misschien heeft de vader wel een ongeluk gehad.

Bsp. 30: En waarom heeft die man altijd vieze kleren aan maar wel schone schoenen.
misschien heeft hij geen idee hoe hij kleren moet wassen.

Bsp. 31: Denk je dat de schrijver ook een boodschap aan de lezers mee wilde geven?
Ja, ik denk dat hij duidelijk wilt maken dat je niet alles moet geloven en niet van vooroordelen uit moet gaan.

Beim Versuch, Fragen über den Text zu beantworten, wäre es natürlich wünschenswert, dass Schüler auch Informationen aus dem Text aufgreifen und für ihre Argumentation verwenden. Schüler machen dies auch, wie man in den Beispielen 32 und 33 lesen kann. Im Vergleich zu den Beispielen 29, 30 und 31 erkennt man einen deutlichen Qualitätsunterschied, denn sobald die Schüler versuchen, anhand des Textes zu argumentieren, kommen sie auch zu Ergebnissen, die im Kontext des Textes Sinn ergeben. Das heißt nicht, dass die Schüler, die nicht anhand des Textes argumentieren keine guten Antworten auf ihre Fragen ausdenken können, aber der Bezug zum Text fehlt in einer solchen Argumentation. Deswegen wurde die Anforderung, dass Schüler aus dem Text heraus argumentieren sollen, in der dritten Abstufung hinzugefügt.

Bsp. 32: Misschien is de vader wel overleden of voor altijd in een ander land.
Dat zou kunnen kloppen aangezien hij wel zijn toekomst met de helm vertelde
Hij [lied] merken dat hij de helm voor de rest van z~~ijn~~ leven op zou hebben.

Bsp. 33: Ze hebben informatie gekregen van de politie, zouden ze daardoor hebben ingezien dat wat ze deden fout was?
Ik denk het niet.
Waarom?
De informatie folders lagen op de grond.

Die vierte Abstufung wurde im Vergleich zur dritten um ein Kriterium ergänzt, nämlich dass Schüler auch noch aus ihren Kenntnissen der Welt oder der Literatur oder des Genres heraus ihre Argumentation für eine mögliche Beantwortung der Frage unterstützen sollen. Dies wäre natürlich wünschenswert, weil die Argumentation der Schüler dadurch noch besser wird und man dagegen weniger einwenden kann. In den schriftlichen Dialogen dieser Studie konnte kein Beispiel von einem Schüler, der argumentiert wie in der vierten Abstufung gewünscht, gefunden werden. Trotzdem wurde das Kriterium der vierten Abstufung hinzugefügt, weil das Bewertungsschema für die Schüler auch kennzeichnen soll, was sie noch lernen müssen, um (sehr) gut interpretieren zu können.

7.6 Referieren

Erkennt Zusammenhänge zwischen (Teilen) der Geschichte und sich selbst.	Erkennt Zusammenhänge zwischen (Teilen) der Geschichte und (sich selbst +) der eigenen Lebeumgebung.	Erkennt Zusammenhänge zwischen (Teilen) der Geschichte und der realen Welt (weitreichender als sich selbst und die eigene Lebeumgebung).	Erkennt Zusammenhänge zwischen (Teilen) der Geschichte und der realen Welt + trifft moralische Aussagen in Bezug auf die Geschichte, argumentiert aus der realen Welt heraus.
---	--	--	---

Tabelle 21: Abstufungen der Kategorie ‚Referieren‘ des Bewertungsschemas.

Es gibt zwei Codes, die offensichtlich zur Kategorie ‚Referieren‘ passen, nämlich die beiden *Referential*-Codes. Im Codebuch wird ein Unterschied zwischen Aussagen, die sich auf das Selbst des Schülers (*Referential: self*), und Aussagen, die die reale Welt betreffen (*Referential: real world*), gemacht. Aber auch der Code *Reasoning: moral – beyond story* bezieht sich auf die reale Welt, weshalb dieser Code ebenfalls in der Kategorie ‚Referieren‘ verarbeitet werden kann.

Der Grund für die Aufnahme dieser Kategorie ins Bewertungsschema, ist, dass Schüler beim Vorhersagen, was in einem Text als nächstes passieren wird, sich auf ihre Erfahrungen mit der Welt berufen, wie im theoretischen Rahmen besprochen wurde (Kapitel 2.2). Dabei liegt eine Vorhersage basierend auf dem Verständnis des eigenen Selbst natürlich näher als eine Vorhersage basierend auf einem allgemeinen Verständnis der Welt, schlichtweg weil man nun mal sich selbst immer besser als alles, was außerhalb sich selbst liegt, kennen kann. Referieren auf sich selbst soll also eine niedrigere Abstufung als das Referieren auf die reale Welt sein.

Wenn Schüler eine moralische Aussage treffen, berufen sie sich auf Kenntnisse im kulturellen Bereich der realen Welt; Welche Verhaltensweisen sind normal, welche nicht und was ist wirklich nicht akzeptiert? Es handelt sich um bestimmte Normen und Werte einer Gesellschaft und das fordert ein größeres Verständnis der Welt, als wenn sich Schüler ohne moralische Aussagen auf die reale Welt beziehen. Deswegen werden moralische Aussagen, bei denen sich Schüler auf ihr Verständnis der realen Welt berufen, höher als Aussagen, die sich nur auf die reale Welt beziehen, eingestuft.

Damit sind drei Abstufungen für die Kategorie ‚Referieren‘ gefunden, allerdings soll noch eine Abstufung, analog zu den anderen Kategorien, hinzukommen. Dazu wurden die genauen Beschreibungen der relevanten Codes angeschaut. Für den Code *Referential: self* gibt es mehrere Möglichkeiten und eine davon ist, dass Schüler mit Bezug auf den Text auf das

eigene Leben, weiterführender als sich selbst allein, verweisen. Dies erfordert mehr Einsicht in die reale Welt, als wenn die Schüler nur bei sich bleiben. Es könnte also eine gute Möglichkeit sein, ein Verweisen auf das eigene Leben, weiterführender als sich selbst, als Abstufung zwischen der Abstufung, bei der Schüler nur auf sich selbst verweisen, und der Abstufung, bei der Schüler sich auf die reale Welt beziehen sollen, aufzunehmen. Um die Anwendbarkeit dieser Idee zu prüfen, wurde in die schriftlichen Dialoge geschaut. Wie man im Beispiel 35 sehen kann, gab es Schüler, die weiterführender als sich selbst auf das Leben referiert haben. Der Unterschied zu Beispiel 34 besteht darin, dass der Schüler sich im Beispiel 35 in eine andere, ihm bekannte Person der realen Welt hineinzusetzen versucht. Dieser Versuch könnte zu anderen Einsichten in die Charaktere der Geschichte führen, als wenn Schüler nur auf sich selbst referieren. Im Beispiel 36 wird noch weiter generalisiert, indem der Schüler fragt, ob Leute es eigentlich für normal halten würden, wenn.... Dieses Beispiel entspricht folglich auch einer höheren, nämlich der dritten Abstufung. Beispiel 37 wird schließlich in die vierte Abstufung eingeteilt, weil der Schüler sich nicht nur fragt, ob ein bestimmtes Verhalten normal ist, sondern diese Frage in seiner moralischen Aussage beantwortet und daraufhin auch generelle Aussagen über Erziehung in der realen Welt macht.

Bsp. 34: Ik zou er niet mee kunnen leven ♦ jij?

Nee, super vervelend

Dag in dag uit die helm op..

Bsp. 35: Stel jij zou dit ook hebben, zouden jou ouders dat dan goed vinden?

Ik zou niet weten wat er erg is aan een helm.

Ik ook niet, maar het kan wel.

Ik hoop dat mijn ouders het goed vinden. Tenslotte moet ik het dragen en niet zij.

Ben ik het mee eens. Ik denk dat mijn ouders het wel accepteren. En misschien ook wel een beetje leuk vinden.

Bsp. 36: denk jij eigenlijk dat het normaal gevonden zou worden als iemand de hele tijd een helm op heeft?

Ja, waarom niet?

nou aangezien het best raar is als je iemand die bijv. zijn hond aan het uitlaten is met een helm op ziet lopen.

Nou, ieder zijn eigen ding.

en trouwens mag je überhaupt wel een helm op bij bijv. je pasfoto.

Bsp. 37: Kom op zeg! je laat je zoon toch niet 3 weken onbepaald een helm op hebben.

Ja, daar zit ook wel weer wat in. En als je het zo bekijkt zie je dat de meeste problemen bij opvoeding beginnen.

inderdaad! Daarom moet de overheid daar ook wat aan doen.

Maar daar zullen ze vast geen geld voor overhebben

Ja, of de belasting gaat omhoog.

7.7 Das gesamte Bewertungsschema

In Tabelle 16 (weiter unten in diesem Kapitel) kann das gesamte Bewertungsschema eingesehen werden. In diesem Kapitel sollen noch ein paar Fragen, die das gesamte Bewertungsschema betreffen, geklärt werden. Es gibt nämlich einige Abstufungen, die eine Abstufung einer anderen Kategorie gleichen und die somit noch weiteren Erklärungen

bedürfen. Außerdem wurde bisher noch nichts über den eigentlichen Einsatz des Bewertungsschemas geschrieben. Auch das soll in diesem Kapitel besprochen werden.

7.7.1 Vergleich der verschiedenen Kategorien und ihrer Abstufungen

Wenn man Interpretationen bewerten möchte, muss dabei berücksichtigt werden, dass Interpretationen irgendwoher kommen. Diese Erkenntnis ist insbesondere wichtig, wenn man Schüler, die das Interpretieren noch erlernen müssen, auf ihre Interpretationsfähigkeiten bewerten möchte. Das Bewertungsschema ist demgemäß so konzipiert worden, dass es auch Phasen der Verarbeitung des Textes, bei denen man vielleicht noch nicht ganz von Interpretation sprechen kann, enthält. Dies führt dazu, dass manche Anforderungen in verschiedenen Kategorien und Abstufungen auftauchen.

Zunächst kann man sehen, dass die Anforderung der Kategorie ‚Beschreiben‘, um in eigenen Worten nachzuerzählen, was in der Geschichte passiert ist und dies zu benutzen um Argumente zu unterstützen, den Anforderungen der Kategorien ‚Interpretieren: hypothetisch‘ und ‚Interpretieren: Geschichte‘, um anhand des Textes beziehungsweise aus dem Text heraus zu argumentieren, ähnelt. Die Anforderung befindet sich in allen Kategorien in Abstufungen drei und vier und ist nicht die einzige Anforderung dieser Abstufungen. Die Abstufungen der unterschiedlichen Kategorien enthalten alle noch zusätzliche Anforderungen, die den Unterschied zwischen den Kategorien darstellen. Allerdings ist es für alle drei Kategorien von Belang, dass Schüler ihre Argumentation anhand des Textes darlegen, damit nachvollzogen werden kann, über welchen Weg der Schüler zu bestimmten Einsichten bezüglich des Textes gelangt. Die Anforderung, anhand des Textes zu argumentieren, stellt somit eine Basisfähigkeit für die höheren Abstufungen aller drei dieser Kategorien dar. Obwohl diese Anforderung sich in den Fällen gleicht, ist der Anlass für die Argumentation ein anderer. Deswegen wurde beschlossen, dass es nicht problematisch ist, dass die Anforderungen in mehreren Kategorien des Bewertungsschema auftauchen.

Zudem kann man sehen, dass die Anforderung des Begründens oder Argumentierens sich nicht immer in der gleichen Abstufung im Bewertungsschema befindet. Das liegt daran, dass pro Kategorie überlegt wurde, welche Anforderungen am wichtigsten sind. Bei der Kategorie ‚Beurteilen‘ wurde beispielsweise beschlossen, dass eine Begründung für eine Beurteilung äußerst wichtig ist. Deswegen ist das Begründen der Beurteilung in allen Abstufungen außer der ersten eine Anforderung. Dagegen wurde für die Kategorie ‚Interpretieren: Charaktere‘ der Entschluss gefasst, dass das Einnehmen mehrerer Perspektiven an der Stelle für die Interpretation bedeutsamer als die Argumentation war, auch weil Argumentation in anderen

Kategorien schon bewertet wird. Die Argumentation wurde in diesem Fall also dem Einnehmen mehrerer Perspektiven untergeordnet und ist erst in der vierten Abstufung als Anforderung vorhanden.

Weiterhin könnte man erkennen, dass die Anforderung der Kategorie ‚Beschreiben‘, um Lücken erkennen zu können, auch Ähnlichkeiten zu den Anforderungen der Kategorien ‚Interpretieren: Charaktere‘ und ‚Interpretieren: Geschichte‘, dass Schüler Fragen über (das Handeln von) Charaktere(n) beziehungsweise Intentionen/Bedeutungen/Inhalte des Textes stellen sollen, aufweist. Diese Anforderungen sind alle auf dem Code *Essential question* zurückzuführen. Der Code ist gewissermaßen für das Bewertungsschema aufgeteilt worden, um zu bestimmten Kategorien zu passen. Der Schwerpunkt ist aber für die jeweiligen Kategorien ein anderer. In der Kategorie ‚Beschreiben‘ geht es erstmal darum, dass Schüler überhaupt Lücken im Text entdecken. Das müssen nicht unbedingt Fragen sein, sondern dürfen auch Feststellungen sein. In der Kategorie ‚Interpretieren: Charaktere‘ geht es dagegen darum, dass Schüler Fragen über (das Handeln von) Charaktere(n) stellen. Das werden wahrscheinlich zum größten Teil essenzielle Fragen sein, obwohl das nicht zwingend notwendig ist. Auch für die Kategorie ‚Interpretieren: Geschichte‘ ist es nicht zwingend so, dass die Fragen, die dort vorgesehen werden, immer essenzielle Fragen sind. Es wäre durchaus denkbar, dass Schüler auch Fragen über beispielsweise den Inhalt stellen, die aber nicht wirklich den Kernpunkt des Textes treffen und somit auch keine essenziellen Fragen sind. Ein zusätzlicher Vorteil der Art und Weise, wie der Code *Essential Question* für das Bewertungsschema aufgeteilt wurde, stellt der Fakt da, dass Dozenten sich nicht fragen müssen, welche Fragen jetzt wirklich essenziell sind. Das Bewertungsschema gibt einfach vor, um welche Art von Fragen (oder Identifizierung im Falle der Kategorie ‚Beschreiben‘) es sich in der jeweiligen Kategorie handelt.

Ferner kann man in der Kategorie ‚Beurteilen‘ sehen, dass in der vierten Abstufung von Schülern verlangt wird, dass sie begründen sollen, inwieweit sie sich mit einem Charakter identifizieren können. Sie müssen dabei auf sich selbst verweisen und deswegen liegt eine Überschneidung mit der ersten Abstufung der Kategorie ‚Referieren‘ vor. Wenn Schüler die vierte Abstufung der Kategorie ‚Beurteilen‘ erreichen, schaffen sie automatisch auch die erste Abstufung der Kategorie ‚Referieren‘. Es wurde überlegt, ob das ein Problem darstellen könnte, aber es wurde entschieden, dass dies nicht der Fall ist, weil der Schüler dann nur die erste Abstufung direkt schafft. Wie vorhin schon angesprochen wurde, ist es schwierig, Anforderungen komplett freistehend zu machen, weil es auf jeden Fall in den niedrigeren Abstufungen vor allem Vorstufen des Interpretierens betrifft.

7.7.2 Empfehlungen für den Einsatz des Bewertungsschemas

Das hier (Ende des Kapitels) vorliegende Bewertungsschema gibt in verschiedenen Abstufungen zu unterschiedlichen Kategorien wieder, welche Schritte von Schülern unternommen werden können, um zu Interpretationen zu gelangen. Sinn und Zweck dieses Kapitels soll nicht sein, dass vorgegeben wird, welche Note ein Schüler unter bestimmten Bedingungen bekommen soll. Stattdessen gilt die Empfehlung, das Bewertungsschema formativ statt summativ einzusetzen. Das Bewertungsschema kann als Instrument dazu dienen, den Schülern *Feed-Up* über den Lernprozess hin zu Interpretationen zu verabreichen und den Fortschritt der Schüler (durch sich selbst) zu überwachen. Das Bewertungsschema ist dementsprechend auch mit dem Ziel, dass es als *Feed-Up* eingesetzt werden kann, entworfen worden. Demnach soll das Bewertungsschema im Vornhinein, vor der (formativen) Prüfung, mit den Schülern geteilt werden. Auf diese Art und Weise wird für die Schüler ersichtlich, was der Begriff ‚Interpretation‘ im Kontext dieser Prüfung heißt beziehungsweise was sie in der Prüfung konkret zeigen sollen. Außerdem verleiht das Bewertungsschema einen Einblick in den Weg, den die Schüler auch über die Prüfung hinaus noch folgen können.

Das Bewertungsschema wurde anhand von Daten aus der vierten Klasse (= 10. Klasse in Deutschland) der Niveaus *havo* und *vwo* konzipiert. Das Bewertungsschema eignet sich deshalb grundsätzlich für diese Klasse. Allerdings enthält das Bewertungsschema natürlich verschiedene Abstufungen auf dem Weg hin zu Interpretationen, weswegen es auch denkbar wäre, das Bewertungsschema in niedrigeren und höheren Klassenstufen zu verwenden.

Bei der Datenerhebung im Pre- und Posttest wurde den Schülern vorgeschrieben, dass sie mindestens drei Seiten schriftlicher Dialog schreiben sollten. Dies geschah um zu gewährleisten, dass die Schüler genug Aussagen zum Text treffen und dass die Daten vergleichbar sind. Man könnte sich überlegen, ob solch eine Vorgabe in Prüfungssituationen auch notwendig ist. Wenn die Schüler das Bewertungsschema im Voraus schon als *Feed-Up* bekommen haben, wissen sie ja bereits was von ihnen erwartet wird. Eventuell wäre hier eine gute Möglichkeit, einen Unterschied zwischen *havo* und *vwo* zu machen, in dem man für die *havo*-Schüler eine Vorgabe zur Minimallänge eines schriftlichen Dialogs macht und für die *vwo*-Schüler nicht. Der Grund dafür ist, dass *havo*-Schüler Anhaltspunkte brauchen; auf diese Art und Weise werden sie dazu gezwungen, noch einmal mehr über den Text nachzudenken, wenn sie die Minimallänge noch nicht erreicht haben. Von *vwo*-Schülern dagegen darf erwartet werden, dass sie anhand des Bewertungsschemas selbst einschätzen können, wie viel Text sie schreiben müssen.

Außerdem wird empfohlen, dass Schüler das Bewertungsschema während der Prüfung einsehen können. Der Vorteil davon, dass Schüler das Bewertungsschema zur Prüfung hinzubekommen, liegt darin, dass sie einerseits sich nicht merken müssen, auf welche Kategorien sie bewertet werden, und andererseits kann das Bewertungsschema auch eine extra Anregung dafür sein, noch einmal tiefer über den Text nachzudenken, wenn ein Schüler darauf schaut, welche Abstufungen bereits seiner Ansicht nach geschafft sein könnten und welche möglicherweise noch erreichbar wären. Man könnte dagegen einbringen, dass dies nicht üblich ist und dass Schüler dann nur versuchen werden, die Anforderungen abzuarbeiten. Allerdings ist das Bewertungsschema auf eine solche Art und Weise konzipiert worden, dass Schüler den Anforderungen nur entsprechen können, wenn sie wirklich selbst die Interpretationsfähigkeiten besitzen. Das Bewertungsschema wäre also nur ein guter Leitfaden dazu, welche Gedankenschritte die Schüler machen könnten. Die Gedankenschritte aber müssen die Schüler trotzdem selbst machen und in dem schriftlichen Dialog zeigen.

Das Bewertungsschema ist auf eine Art und Weise konzipiert worden, dass Schüler eine Vielfalt von Äußerungen, die auch bewertet werden, zur Kurzgeschichte schreiben können. Um diese Freiheit nicht allzu sehr einzuschränken, muss man sich fragen, ob es notwendig ist, alle Kategorien zu bewerten. Sollte dies nämlich der Fall sein, werden Schüler krampfhaft versuchen, die Anforderungen aller Kategorien innerhalb der ihr gegebenen Zeit zu erfüllen. Das führt möglicherweise zu weniger tiefen Interpretationen. Deswegen wird an dieser Stelle vorgeschlagen, das Bewertungsschema auf vier Kategorien (exklusive literarischem Wortgebrauch und Textstruktur) zu prüfen. Somit wird den Schülern genug Freiraum gelassen, selbst zu entscheiden, über welche Sachen bezüglich des Textes sie schreiben möchten. Die Anzahl von vier Kategorien wird empfohlen, weil Schüler dann immer mindestens eine ‚Interpretieren‘-Kategorie aufgreifen müssen. Dementsprechend wird gewährleistet, dass das Bewertungsschema wirklich die Interpretationsfähigkeiten von Schülern prüft.

Abschließend bleibt noch die Frage zu klären, welches Niveau Schüler ungefähr erreichen sollten. Für die vierte Klasse *havo* und *vwo*, in der die Daten erhoben wurden, müssten die dritten Abstufungen erreichbar sein. Allerdings gibt es natürlich immer Schüler, die weniger oder mehr erreichen. Man muss berücksichtigen, dass die Schüler den Text in einer Fremdsprache gelesen haben. Deswegen wird es Schüler geben, die den Text nicht gut genug verstanden haben, um überall die dritte Abstufung zu erreichen. Basierend auf der Menge an Daten, die für diese Studie untersucht wurden (78 schriftliche Dialoge), ist die Empfehlung für eine formative Prüfung, als Anforderung zu stellen, dass die Schüler in mindestens drei von vier Kategorien die dritte Abstufung erreichen sollen.

Literarischer Wortgebrauch	<input type="checkbox"/> Es werden <u>keine</u> literarischen Terminologien verwendet.	<input type="checkbox"/> Es werden einige literarische Terminologien verwendet.
Textstruktur	<input type="checkbox"/> Unzureichend.	<input type="checkbox"/> Interpunktion ist in Ordnung. + es werden konsequent Signalwörter verwendet, um Textstrukturen anzudeuten.

Beschreiben	<input type="checkbox"/> Erzählt in eigenen Worten nach, was in der Geschichte passiert ist.	<input type="checkbox"/> Erzählt in eigenen Worten nach, was in der Geschichte passiert ist + schildert dabei die Hauptlinien der Geschichte.	<input type="checkbox"/> Erzählt in eigenen Worten nach, was in der Geschichte passiert ist + benutzt dies um Argumente zu unterstützen + identifiziert mindestens eine Lücke im Text.	<input type="checkbox"/> Erzählt in eigenen Worten nach, was in der Geschichte passiert ist + benutzt dies um Argumente zu unterstützen + identifiziert mehrere Lücken im Text.
Beurteilen	<input type="checkbox"/> Beurteilt den Text und/oder Charaktere + begründet dies <u>nicht</u> .	<input type="checkbox"/> Beurteilt den Text + begründet dies.	<input type="checkbox"/> Beurteilt Charaktere + begründet dies.	<input type="checkbox"/> Beurteilt Charaktere + begründet dies + bezieht sich in der Begründung darauf, inwieweit er/sie sich mit Charakteren identifizieren kann.
Interpretieren	<input type="checkbox"/> Hypothetisch Spekuliert über das Ende / alternative Handlungsverläufe oder macht Verbesserungsvorschläge an den Text.	<input type="checkbox"/> Macht Verbesserungsvorschläge an den Text und basiert sich dabei auf eigene Wünsche / Leseerfahrungen.	<input type="checkbox"/> Spekuliert über das Ende / alternative Handlungsverläufe oder macht Verbesserungsvorschläge an den Text + argumentiert dabei anhand des Textes.	<input type="checkbox"/> Spekuliert über das Ende / alternative Handlungsverläufe oder macht Verbesserungsvorschläge an den Text + argumentiert dabei anhand des Textes + zeigt dabei Verständnis für die Wahl des Autors und versucht diese Wahl auch zu erklären.

Interpretieren	Charaktere	<input type="checkbox"/> Stellt keine oder wenig Fragen über (das Handeln von) Charaktere(n) + versucht <u>nicht</u> , diese Fragen zu beantworten.	<input type="checkbox"/> Stellt Fragen über (das Handeln von) Charaktere(n) + versucht, diese Fragen zu beantworten.	<input type="checkbox"/> Stellt Fragen über (das Handeln von) Charaktere(n) + versucht, diese Fragen zu beantworten + bietet dafür mehrere Perspektiven an.	<input type="checkbox"/> Stellt Fragen über (das Handeln von) Charaktere(n) + versucht, diese Fragen zu beantworten + bietet dafür mehrere Perspektiven an + argumentiert für die unterschiedlichen Perspektiven.
	Geschichte	<input type="checkbox"/> Stellt keine oder wenig Fragen über Intentionen / Bedeutungen / Inhalte des Textes + versucht <u>nicht</u> , diese Fragen zu beantworten.	<input type="checkbox"/> Stellt Fragen über Intentionen / Bedeutungen / Inhalte des Textes + versucht, diese Fragen zu beantworten.	<input type="checkbox"/> Stellt Fragen über Intentionen / Bedeutungen / Inhalte des Textes + versucht, diese Fragen zu beantworten + argumentiert dabei aus dem Text heraus.	<input type="checkbox"/> Stellt Fragen über Intentionen / Bedeutungen / Inhalte des Textes + versucht, diese Fragen zu beantworten + argumentiert dabei aus dem Text + aus den Kenntnissen der Welt / der Literatur / des Genres heraus.
Referieren	<input type="checkbox"/> Erkennt Zusammenhänge zwischen (Teilen) der Geschichte und sich selbst.	<input type="checkbox"/> Erkennt Zusammenhänge zwischen (Teilen) der Geschichte und (sich selbst +) der eigenen Lebeumgebung.	<input type="checkbox"/> Erkennt Zusammenhänge zwischen (Teilen) der Geschichte und der realen Welt (weitreichender als sich selbst und die eigene Lebeumgebung).	<input type="checkbox"/> Erkennt Zusammenhänge zwischen (Teilen) der Geschichte und der realen Welt + trifft moralische Aussagen in Bezug auf die Geschichte, argumentiert aus der realen Welt heraus.	

Tabelle 22: Bewertungsschema mit Fokus auf Interpretation für die Aufgabe des Schreibens eines Dialogs zu einer Kurzgeschichte.

8. Diskussion

8.1 Erste Forschungsfrage

Die erste Forschungsfrage, mit der sich diese Arbeit beschäftigt hat, lautete: Wie interpretieren Schüler literarische Kurzgeschichten im Fremdsprachenunterricht vor und nach einer Unterrichtsintervention, die sich auf *Self-Questioning* richtet? Eine ausführliche Antwort auf diese Frage wird in Kapitel 5 gegeben. Um diese Forschungsfrage zu beantworten, wurden insgesamt 78 schriftlichen Dialoge von Schülern der vierten Klasse *havo* und *vwo* codiert und anschließend analysiert. Dazu wurde das Codebuch, das Schrijvers et al.¹⁴⁸ für ihre Forschung verwendet haben, anhand von zehn Pilotdialogen für die Zwecke dieser Studie angepasst.

Eine der Veränderungen am Codebuch war, dass der Code *Descriptive* in zwei aufgeteilt wurde, nämlich in *Descriptive: default* und *Descriptive: summary*. Dies geschah aufgrund der Einsicht, dass ‚Nacherzählen‘ und ‚Zusammenfassen‘ laut der Taxonomie von Kwakernaak zwei unterschiedliche Stufen der Beherrschung des Materials darstellt¹⁴⁹; Schüler müssen mehr können und verstehen, um eine Zusammenfassung zu geben, als wenn sie nur nacherzählen. Später, bei der Analyse der Interrater-Reliabilität, stellte sich heraus, dass die semantische Domäne *Descriptive* die einzige Domäne ist, bei der die Interrater-Reliabilität so niedrig ist, dass man daraus schließen kann, dass die beiden *Descriptive*-Codes nicht ausreichend voneinander zu unterscheiden sind. Mit anderen Worten waren sich die beiden Rater häufig nicht darüber einig, ob Schüler nun nacherzählen oder schon zusammenfassen. Demnach kann man sich fragen, ob die unterschiedlichen Einstufungen vom ‚Nacherzählen‘ und ‚Zusammenfassen‘ in Kwakernaaks Taxonomie Sinn ergeben. Es wäre interessant, sich in weiterer Forschung mit der Frage, ob ‚Nacherzählen‘ und ‚Zusammenfassen‘ überhaupt wirklich unterschieden werden können, zu befassen.

Obwohl sorgfältig beachtet wurde, dass alle für Interpretationen auch nur ansatzweise relevante Daten anhand des Codebuches codiert werden konnten, kann es sein, dass die Codes eine einzelne Aussage eines Schülers nicht gerecht worden sind. Das Risiko geht mit der Einteilung in Codes einher. Für die Analyse der schriftlichen Dialoge war dieser Schritt des Codierens notwendig, allerdings führt er auch dazu, dass aus den qualitativen Daten

¹⁴⁸ Schrijvers et al., „Transformative Dialogic Literature Teaching Fosters Adolescents' Insight into Human Nature and Motivation“.

¹⁴⁹ Vgl. Kwakernaak, „Een taxonomie voor het vreemdetalenonderwijs“, S. 14.

quantitative Ergebnisse hervorgehen. Dies hat auch Auswirkungen auf die Analyse der Ergebnisse für die zweite Forschungsfrage.

8.2 Zweite Forschungsfrage

Im Rahmen der zweiten Forschungsfrage, inwieweit die *Self-Questioning*-Unterrichtsmethode beim Lesen von Kurzgeschichten einen Effekt auf die Interpretationen der Schüler hat, wurde spezifisch auf die Codes *Reasoning: understanding story events* und *Reasoning: understanding character* geschaut. Für diese beiden Codes gab es im Posttest, im Vergleich zum Pretest, einen deutlichen Anstieg. Das bedeutet, dass mehr Segmente einen dieser beiden Codes zugewiesen bekamen. Somit konnte festgestellt werden, dass die Schüler im Posttest mehr interpretierten und die Unterrichtsintervention also einen positiven Effekt auf die Interpretationsfähigkeiten von Schülern hatte. Wie gerade eben aber schon angesprochen wurde, sind die Codes quantitativ analysiert worden. Das heißt im Kontext der zweiten Forschungsfrage, dass zwar geschlussfolgert werden kann, dass es nach der Unterrichtsintervention einen Anstieg der beiden *Reasoning: understanding*-Codes gab, aber das aus diesen Ergebnissen keine verlässlichen Schlüsse über die Qualität der Daten gezogen werden können, obwohl die Aussagen von Schülern natürlich schon einem bestimmten Wert entsprechen sollen, bevor sie als *Reasoning: understanding* codiert werden. Jedoch sind innerhalb der *Reasoning: understanding*-Codes selbstverständlich trotzdem qualitative Unterschiede möglich. Theoretisch wäre es also möglich, dass die Quantität im Posttest anstieg, während die Qualität sich verschlechterte. Dies ist zwar nicht zu erwarten, weil es dafür keinen Anlass geben dürfte, aber auch nicht vollständig auszuschließen. Dieses Problem wurde bei der Erstellung des Bewertungsschemas berücksichtigt, indem Beispiele aus den Dialogen, also qualitative Daten, verwendet wurden, um die verschiedenen Abstufungen zu bestimmen.

8.3 Dritte Forschungsfrage

Die Erstellung des Bewertungsschemas erfolgte aus der dritten Forschungsfrage, wie ein Bewertungsschema für Interpretationsfähigkeiten von Schülern, zu einer Aufgabe, bei der von Schülern einen schriftlichen Dialog zu einer Kurzgeschichte geschrieben wurde, aussehen könnte, heraus. Das Bewertungsschema ist somit gewissermaßen eine Konklusion dieser Forschungsfrage. Für die Untermauerung der Einteilung des Bewertungsschemas wurden die Codes, die Ergebnisse der Analyse der Codes im Pre- und Posttest und Beispiele aus verschiedenen Dialogen verwendet. Zudem wurden verschiedene Erkenntnisse aus der wissenschaftlichen Literatur, die im theoretischen Rahmen aufgeführt werden, berücksichtigt.

Dabei gibt es noch einige ergänzende Details zu den Daten und anlässlich der wissenschaftlichen Literatur, die an dieser Stelle angesprochen werden sollen.

Erstens sind die Daten natürlich im Rahmen einer wissenschaftlichen Studie gesammelt worden. Das hat dazu geführt, dass es für die Schüler, anders als in einer Prüfungssituation, keine direkten Konsequenzen gab, wenn sie die Dialoge nicht seriös geschrieben haben. In den Daten konnte man dies teilweise erkennen, weil Schüler beispielsweise relativ schnell zu der Farbe des Helmes von Max Verstappen abgeschweift sind. Es könnte demnach also durchaus so sein, dass die Daten ein etwas negativeres Bild, als wenn die schriftlichen Dialoge in einer Prüfungssituation geschrieben worden wären, aufzeigten. Diese Tatsache hatte auf die Erstellung des Bewertungsschemas keinen wirklichen Einfluss, weil vor allem anhand von positiven Beispielen die Abstufungen bestimmt wurden. Dementsprechend war der Effekt von Negativbeispielen auf das Bewertungsschema eher gering.

Zweitens könnte man sich fragen, ob die vorgegebene Anfangsfrage der schriftlichen Dialoge, nämlich „Oké, welk stukje van het verhaal sprong er voor jou echt uit?“¹⁵⁰, die richtige ist, um den Schülern Interpretationen zu entlocken. Im theoretischen Rahmen wurde schon kurz angedeutet, dass die Fragestellung auch einen Einfluss auf die Endergebnisse hat. McCarthy und Goldman haben nach Fragestellungen, die zu Interpretationen führen sollten, geforscht und haben herausgefunden, dass konkrete offene Fragestellungen, bei denen nach dem Thema des Textes gefragt wird oder bei denen die Schüler direkt nach einer Argumentation gefragt wird, zu mehr interpretativen Äußerungen führen, als wenn die Fragestellung nach der Handlung oder dem Inhalt des Textes fragt.¹⁵¹ Demnach könnte noch ein Fortschritt im Bereich der interpretativen Äußerungen erzielt werden, wenn eine andere Frage am Anfang der schriftlichen Dialoge gestellt wird. Eine Frage wie: ‚Was glaubst du ist das Thema dieser Kurzgeschichte und kannst du das anhand des Textes erklären?‘ müsste laut der Studie von McCarthy und Goldman also zu mehr interpretativen Ergebnissen führen. Es wäre interessant zu untersuchen, ob die Fragestellung im Falle der schriftlichen Dialoge auch eine solche wahrnehmbare Auswirkung auf die Menge an interpretativen Äußerungen der Schüler haben könnte.

Für die Erstellung des Bewertungsschemas wurden die Ergebnisse des Codierens beachtet. Die Ergebnisse des Codierens haben Daten zu der Menge an Codes im Pre- und Posttest und der durchschnittlichen Menge der einzelnen Codes pro Schüler ergeben. Im

¹⁵⁰ Übersetzung: Okay, welcher Teil der Geschichte stach für dich echt heraus?

¹⁵¹ Vgl. McCarthy und Goldman, „Comprehension of Short Stories: Effects of Task Instructions on Literary Interpretation“, S. 18.

Nachhinein, bei der Konzipierung des Bewertungsschemas, stellte sich heraus, dass es auch interessant gewesen wäre, zu wissen wie die Codes über einzelne Dialoge verteilt sind. Es ist nämlich vorstellbar, dass einige Schüler Aussagen mit einem bestimmten Code treffen, während andere Codes in ihren Dialogen kaum auftreten. Es wäre also interessant gewesen, auch die Standardabweichung der unterschiedlichen Codes zu berechnen. Zudem wären auch die Daten zur durchschnittlichen Menge an Codes pro schriftlichen Dialog und die Standardabweichung dazu interessant gewesen. Leider konnten diese Daten aber nicht anhand des Codierprogramms Atlas.ti berechnet werden. Das Fehlen dieser Daten hat Auswirkungen auf das Bewertungsschema und die Empfehlungen zu dem Bewertungsschema. Es kann nicht festgestellt werden, wie unterschiedlich die Aussagen eines Schülers innerhalb eines schriftlichen Dialogs sind. Einerseits konnte dementsprechend die Menge an unterschiedlichen Codes bei der Erstellung des Bewertungsschemas in der Überlegung, wie viele Kategorien es enthalten sollte, nicht berücksichtigt werden. Andererseits führt es dazu, dass die Empfehlung, wie viele Kategorien des Bewertungsschemas bewertet werden sollen, nicht aus den Ergebnissen des Codierens heraus begründet werden kann. Noch im theoretischen Rahmen wurde festgestellt, dass leistungsschwächere Schüler Aussagen zu weniger Kategorien als leistungsstarke Schüler äußern. Im Rahmen dieser Arbeit wurde nicht festgestellt, wie viele unterschiedliche Codes durchschnittlich pro schriftlichen Dialog codiert wurden. Die Empfehlung, vier Kategorien des Bewertungsschemas zu bewerten, beruht also auf der Erfahrung der Autorin dieser Arbeit mit den schriftlichen Dialogen, jedoch kann aus der Analyse der Daten nicht abgeleitet werden, für wie viele Schüler diese Anforderung schaffbar ist.

Weiterhin konnte das Bewertungsschema nicht so aufgebaut werden, dass gar keine anderen Faktoren als Interpretationen bewertet werden. Interpretationen basieren sich nun einmal auf Kenntnisse, die auch über den Text hinausragen. Im Bewertungsschema schlägt sich das unter anderem darin nieder, dass Kenntnisse der Welt einbezogen werden. Schüler, die den Text oder Teile davon auf die reale Welt beziehen können, erreichen höhere Abstufungen als diejenige, die das nicht können. Es könnte dementsprechend hilfreich sein, im Vornhinein mit Schülern darüber zu sprechen, in welchem Sinne literarische Texte und die reale Welt miteinander zusammenhängen und wie sich das zu den Fähigkeiten des Interpretierens verhält. Sowieso wäre es eine gute Idee, mit den Schülern explizit zu besprechen, was Interpretation heißt und wozu sich Schüler alles im Rahmen von Interpretationen äußern können. Denn wie im theoretischen Rahmen besprochen, führen explizite spezifische Hinweise zu besseren und

tieferen Interpretationen. Das Bewertungsschema kann ein guter Anlass und guter Leitfaden für ein Klassengespräch zu diesem Thema sein.

Ferner muss beim Einsatz des Bewertungsschemas berücksichtigt werden, dass die Schüler den Text, zu dem sie einen schriftlichen Dialog schreiben, in einer Fremdsprache lesen. Das bedeutet, dass das Sprachniveau der Schüler Auswirkungen auf ihre Interpretationsfähigkeiten hat. Schüler können ihre Kenntnisse über das Interpretieren aus der Muttersprache erst in die Fremdsprache übertragen, wenn sie ein bestimmtes Sprachniveau erreicht haben, wie die *Threshold-Theory* besagt. Außerdem muss man damit rechnen, dass Schüler, die die Sprache relativ schlecht verstehen, im Vergleich zu Kommilitonen, die die Sprache besser verstehen, weniger variierte Aussagen zum Text treffen und sich stattdessen vor allem auf narrative Aspekte fokussieren werden. Wenn Schüler im Bewertungsschema relativ schlecht punkten, darf demnach nicht ohne weiteres angenommen werden, dass diese Schüler nicht ausreichend interpretieren können. Stattdessen sollte man herausfinden, ob nicht die Sprachfähigkeiten ein Problem darstellen und dazu führen, dass der Schüler den *Threshold*, um interpretieren zu können, nicht schafft. Interessanterweise ergab sich in der Forschung von Urlaub, dass die *Self-Questioning*-Methode für geringere Kenntnisse der Fremdsprache kompensieren könnte.¹⁵² Es wäre interessant zu untersuchen, ob dieses Ergebnis reproduzierbar ist und *Self-Questioning* eventuell wirklich eine kompensierende Wirkung bezüglich Sprachfähigkeiten erzeugt. Das wäre noch ein Grund mehr, um die *Self-Questioning*-Methode generell im Fremdsprachenunterricht einzuführen.

Im theoretischen Rahmen wurde schon kurz angesprochen, dass Dozenten einen Text mehrmals lesen und dadurch auf dem Gebiet von Interpretationen häufig Fragen, die die Schüler, die den Text nur einmal gelesen haben, nicht beantworten können, stellen. Es geht dabei um Muster im Text. Beispielsweise fiel der Autorin dieser Arbeit erst beim dritten Lesen des Textes *Helm auf* auf, dass der Protagonist zu einer Beerdigung war und dies darauf hindeuten könnte, dass der Vater gestorben ist. Man kann demnach von Schülern, die den Text in einer Prüfungssituation zum ersten Mal sehen, nicht die gleiche Interpretationsfähigkeiten bezüglich dieses Textes erwarten, als wenn sie den Text schon mehrmals gelesen hätten. Deswegen muss man sich fragen, ob es nicht sinnvoll wäre, Interpretationsfähigkeiten anhand eines Textes, den die Schüler vor der Prüfungssituation schon mindestens einmal sorgfältig gelesen haben, zu prüfen. Übrigens könnte das zweifach Lesen eines Textes auch ein guter

¹⁵² Vgl. Urlaub, „Reading Strategies and Literature Instruction: Teaching Learners to Generate Questions to Foster Literary Reading in the Second Language“, S. 302.

Weg sein, um möglichen Sprachproblemen vorzubeugen (alternativ erlaubt man die Verwendung von Wörterbüchern). Basierend auf dem Datenbestand, bei dem die Schüler den Text aber nur einmal gelesen hatten, wurde das vorliegende Bewertungsschema mit dem Gedanken entworfen, dass die Schüler den Text nur einmal gelesen haben. Dementsprechend müssten, wenn die Schüler den Text mehrmals lesen, auf jeden Fall die Anforderungen an die Schüler erneut durchdacht werden.

Sollte das Bewertungsschema, gegen die Empfehlung, summativ für eine Prüfung eingesetzt werden, stellt sich die Frage, ob eine Nachprüfung anhand des Bewertungsschemas möglich ist. Grundsätzlich ist die Antwort darauf ‚ja‘, aber die Chance ist groß, dass die Resultate nicht wirklich vergleichbar mit der ersten Prüfung sind – angenommen, es wird für die Nachprüfung eine neue Kurzgeschichte ausgewählt. Dies liegt daran, dass die Komplexität, sowohl auf Sprachniveau als auch auf Geschichtsebene, einen Einfluss auf die Interpretationsfähigkeiten von Schülern hat, wie in den Ergebnissen und in der Diskussion zu den Ergebnissen im Vergleich zwischen *Helm auf* und *Stinkefisch* beschrieben wurde. Zudem eignen manche Kurzgeschichten sich für bestimmte Abstufungen mehr als andere; Zum Beispiel könnte eine moralische Aussage zu *Stinkefisch* eventuell mehr auf der Hand liegen als zu *Helm auf*. Demnach müssten die Anforderungen, die an die Schüler gestellt werden, an die Geschichte angepasst werden. Als Dozent hat man aber mit sehr wenigen Schülern, die eine Nachprüfung ablegen, nicht genug Daten um die Anforderungen gerecht anzupassen, damit die Note der ersten Prüfung und Nachprüfung vergleichbar werden. Es wäre also nicht wünschenswert, eine Nachprüfung mit einem anderen Text abzunehmen.

Im Übrigen könnte man sich fragen, ob das Bewertungsschema auch in anderen Fremdsprachen eingesetzt werden könnte. Grundsätzlich müsste das bei einer gleichen Aufgabenstellung möglich sein. Allerdings muss man dabei beachten, dass das Bewertungsschema für den niederländischen Fremdsprachenunterricht Deutsch erstellt wurde. Man müsste sich also überlegen, ob der Kontext der Gleiche ist. Wenn man das Bewertungsschema beispielsweise für das Fach Französisch in den Niederlanden einsetzen wollen würde, muss man sich fragen, ob das Sprachniveau der Schüler vergleichbar ist oder wenigstens gut genug, um sich interpretativ mit einem literarischen Text in der Fremdsprache zu beschäftigen. Auch wenn man das Bewertungsschema schon für den DaF-Unterricht einsetzen möchte, aber in einem anderen Land als die Niederlande, wird sich der Kontext ändern. In diesem Fall ist wiederum der Kontext von Belang; Sind die Anforderungen an die Schüler zum Beispiel die Gleichen? Es gäbe in jedem Fall, wenn man das Bewertungsschema

in einem anderen Kontext als dem in dieser Arbeit erforschten, noch einige Fragen zu klären, bevor es dazu eingesetzt werden kann.

Abschließend soll noch angesprochen werden, dass es im Rahmen dieser Arbeit nicht mehr gelungen ist, das Bewertungsschema unter Dozenten zu testen, so wie es eigentlich vorgesehen war. Dies würde nämlich den Umfang dieser Arbeit sprengen. Geplant war, dass mehrere Dozenten das Bewertungsschema inklusive eines oder zweier Dialoge zur Bewertung vorgelegt bekommen. Auf diese Art und Weise sollte die Zuverlässigkeit und Anwendbarkeit des Bewertungsschemas geprüft werden. Es wäre wichtig, dass die Zuverlässigkeit in Zukunft noch geprüft wird, damit klar wird, ob Dozenten die gleichen Dialoge auch gleichmäßig bewerten.

Literaturverzeichnis

- Actualisatie Examenprogramma's Moderne Vreemde Talen | SLO. „Actualisatie examenprogramma's moderne vreemde talen“. Zugegriffen 27. Juni 2023. <https://www.actualisatiemvt.nl/>.
- Actualisatiemvt.nl. „Conceptraamwerk, selectie van concepteindtermen en conceptkarakteristieken Moderne Vreemde Talen. Tussenproducten voor Adviesperiode 2“. SLO, 28. März 2023.
- Altinay-Sezen, Deniz. „Literarische Kompetenz – ein sprachübergreifendes Phänomen? Eine empirische Untersuchung über die potenzielle Übertragung der muttersprachlichen literarischen Kompetenz in die L2“. Masterthesis, Universiteit Utrecht, 2014.
- Andringa, Els. „Ontwikkelingen in het omgaan met literaire complexiteit“. *Tijdschrift voor onderwijsresearch* 20, Nr. 3 (1995): 271–84.
- „Atlas.Ti Windows User Manual. Analyse Anything, Faster“. Atlas.ti, 5. März 2023. https://doc.atlasti.com/ManualWin/ATLAS.ti_ManualWin.23.pdf.
- Björnsson, Carl-Hugo. *Läsbarhet*. Stockholm: Bokförlaget Liber AB, 1968.
- Bloemert, Jasmijn, Sean de Goede, und Maron Goedhard. „Via een meervoudige benadering naar verrijkt mvt-literatuuronderwijs“. *Levende Talen Magazine* 104, Nr. 3 (2017): 10–15.
- Bloemert, Jasmijn, Ellen Jansen, und Wim van de Grift. „Exploring EFL Literature Approaches in Dutch Secondary Education“. *Language, Culture and Curriculum* 29, Nr. 2 (3. Mai 2016): 169–88. <https://doi.org/10.1080/07908318.2015.1136324>.
- CITOLab. „6 tips voor het maken van een beoordelingsschaal of -schema“. CITO, Juni 2023. https://cito.nl/media/wgdfw455/6-tips-voor-het-maken-van-een-beoordelingsschaal_def.pdf.
- College voor Toetsen en Examens. „Examenprogramma moderne vreemde talen en literatuur havo/vwo“. Examenblad. Zugegriffen 9. Mai 2023. https://www.examenblad.nl/examenstof/moderne-vreemde-talen-vwo/2023/f=/mvt_havovwo.pdf.
- Cornelissen, Gertrudis Elisabeth Caecilia. *'Maar als je erover nadenkt ... ': een jaar literatuuronderwijs in groepen 7 en 8 van de basisschool*. Amsterdam: Stichting Lezen, 2016.
- Culler, Jonathan. *Literaturtheorie. Eine kurze Einführung*. Übersetzt von Andreas Mahler. 2. Aufl. Stuttgart: Reclam, 2013.
- De Leescoalitie. „Oproep tot een ambitieus leesoffensief“, Oktober 2020. <https://tijdvooreenleesoffensief.nl/wp-content/uploads/2021/02/Manifest-Leesoffensief.pdf>.
- Dönszelmann, Sebastiaan, Catherine van Beuningen, Anna Kaal, und Rick de Graaff, Hrsg. *Handboek vreemdetalendidactiek: vertrekpunten - vaardigheden - vakinhoud*. Bussum: Uitgeverij Coutinho, 2020.
- Earthman, Elise Ann. „Creating the Virtual Work: Readers' Processes in Understanding Literary Texts“. *Research in the Teaching of English* 26, Nr. 4 (1992): 351–84.
- Europarat, Hrsg. *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Straßburg: Cambridge University Press, 2001. <https://rm.coe.int/1680459f97>.
- , Hrsg. *Common European Framework of Reference for Languages: Learning, Teaching, Assessment - Companion Volume*. Straßburg: Council of Europe, 2020. <https://rm.coe.int/common-european-framework-of-reference-for-languages-learning-teaching/16809ea0d4>.

- Fecteau, Monique L. „First-and Second-Language Reading Comprehension of Literary Texts“. *The Modern Language Journal* 83, Nr. 4 (Dezember 1999): 475–93. <https://doi.org/10.1111/0026-7902.00036>.
- Grishakova, Marina, und Maria Poulaki, Hrsg. *Narrative Complexity*. Frontiers of Narrative. University of Nebraska Press, 2019.
- Hall, Geoff. *Literature in Language Education*. London: Palgrave Macmillan UK, 2005. <https://doi.org/10.1057/9780230502727>.
- Hattie, John, und Helen Timperley. „The Power of Feedback“. *Review of Educational Research* 77, Nr. 1 (März 2007): 81–112. <https://doi.org/10.3102/003465430298487>.
- Heinrich, Finn-Ole. *Helm auf*, 2013. <https://www.finnoleheinrich.de/buch/>.
- Inspectie van het onderwijs. „De Staat van het Onderwijs 2023“. Ministerie van Onderwijs, Cultuur en Wetenschap, Mai 2023. <https://www.onderwijsinspectie.nl/documenten/rapporten/2023/05/10/rapport-de-staat-van-het-onderwijs-2023>.
- . „Peil.Leesvaardigheid einde (speciaal) basisonderwijs 2020-2021“. Ministerie van Onderwijs, Cultuur en Wetenschap, Dezember 2022. <https://www.onderwijsinspectie.nl/documenten/themarapporten/2022/12/13/peil-leesvaardigheid-einde-basisonderwijs-en-speciaal-basisonderwijs-2020-2021>.
- Iser, Wolfgang. *The Act of Reading. A Theory of Aesthetic Response*. Baltimore: Johns Hopkins University Press, 1978.
- Janssen, Tanja. „Instruction in Self-Questioning as a Literary Reading Strategy: An Exploration of Empirical Research“. *L1-Educational Studies in Language and Literature* 2, Nr. 2 (2002): 95–120. <https://doi.org/10.1023/A:1020855401075>.
- Janssen, Tanja, Martine Braaksma, und Michel Couzijn. „Self-Questioning in the Literature Classroom: Effects on Students’ Interpretation and Appreciation of Short Stories“. *L1 Educational Studies in Language and Literature* 9, Nr. 1 (März 2009): 91–116. <https://doi.org/10.17239/L1ESLL-2009.09.01.05>.
- Janssen, Tanja, Martine Braaksma, und Gert Rijlaarsdam. „Literary Reading Activities of Good and Weak Students: A Think Aloud Study“. *European Journal of Psychology of Education* 21, Nr. 1 (März 2006): 35–52. <https://doi.org/10.1007/BF03173568>.
- Kadir, Khairul Husna Abdul, Tengku Nor Rizan Tengku Mohamad Maasum, und Ravichandran Vengadasamy. „Conceptualising Literary Responses of L2 Readers: An Exploratory Study“. *International Journal of Social Science and Humanity* 2, Nr. 6 (2012): 565–68. <https://doi.org/10.7763/IJSSH.2012.V2.173>.
- Knaap, Ewout van der. *Literatuur en film in het vreemdetalenonderwijs*. Bussum: Uitgeverij Coutinho, 2019.
- Krippendorff, Klaus. *Content Analysis: An Introduction to Its Methodology*. Fourth Edition. Los Angeles: SAGE, 2018.
- Kwakernaak, Erik. „Een taxonomie voor het vreemdetalenonderwijs“. *Levende Talen Magazine* 99, Nr. 6 (2012): 10–15.
- Lazar, Gillian. *Literature and Language Teaching: A Guide for Teachers and Trainers*. Cambridge Teacher Training and Development. Cambridge [England]; New York: Cambridge University Press, 1993.
- Lee, Jeong-Won, und Diane Lemonnier Schallert. „The Relative Contribution of L2 Language Proficiency and L1 Reading Ability to L2 Reading Performance: A Test of the Threshold Hypothesis in an EFL Context“. *TESOL Quarterly* 31, Nr. 4 (1997): 713. <https://doi.org/10.2307/3587757>.
- Leeuw, Bart van der, und Clary Ravesloot. *Boekverslagen analyseren met het referentiekader*. Enschede: SLO, 2012. <https://www.slo.nl/publicaties/@4187/boekverslagen/>.

- Lenhard, Wolfgang, und Alexandra Lenhard. „Berechnung des Lesbarkeitsindex LIX nach Björnson“. *Psychometrica*, 2022 2014. <https://doi.org/10.13140/RG.2.1.1512.3447>.
- „Literatur im Unterricht. Unterrichtsmaterial für den Einsatz von deutschsprachiger Literatur am Dag van de Duitse taal“. Duitsland Instituut Amsterdam, 25. März 2020. <https://machmit.nl/assets/upload/Lesmateriaal/Idee%C3%ABnboekje2020.pdf>.
- Maassen, Nathalie, und Dorien den Otter. „Huh, moesten we dat ook leren? Over toetskwaliteit in het onderwijs“. *Van twaalf tot achttien*, Januar 2015, 36–37.
- Maassen, Nathalie, Dorien den Otter, Saskia Wools, Bas Hemker, Gerard Straetmans, und Theo Eggen. „Kwaliteit van toetsen binnen handbereik. Een reviewstudie van onderzoek en onderzoeksresultaten naar de kwaliteit van toetsen“. RCEC, CITO, 2014.
- Marshall, James D., Peter Smagorinsky, und Michael W. Smith. *The Language of Interpretation: Patterns of Discourse in Discussions of Literature*. NCTE Research Report, no. 27. Urbana, Ill: National Council of Teachers of English, 1995.
- Mason, Lucia, Fabio Scirica, und Laura Salvi. „Effects of Beliefs about Meaning Construction and Task Instructions on Interpretation of Narrative Text“. *Contemporary Educational Psychology* 31, Nr. 4 (2006): 411–37. <https://doi.org/10.1016/j.cedpsych.2005.12.001>.
- McCarthy, Kathryn S., und Susan R. Goldman. „Comprehension of Short Stories: Effects of Task Instructions on Literary Interpretation“. *Discourse Processes* 52, Nr. 7 (3. Oktober 2015): 585–608. <https://doi.org/10.1080/0163853X.2014.967610>.
- Meijer, Dick, und Daniela Fasoglio. „Handreiking schoolexamen moderne vreemde talen havo/vwo“. SLO, März 2007.
- Naji, Jeneen, Ganakumaran Subramaniam, und Goodith White. *New Approaches to Literature for Language Learning*. Cham: Springer International Publishing, 2019. <https://doi.org/10.1007/978-3-030-15256-7>.
- Nood, Lisanne de, und Marjolein van Herten. „‘Een open einde kan leuk zijn, maar dit is wel heel open’: Waarover leerlingen praten, als zij over verhalen praten“. *Levende Talen Tijdschrift* 21, Nr. 4 (30. Dezember 2020): 27–38.
- Paran, Amos. „The Role of Literature in Instructed Foreign Language Learning and Teaching: An Evidence-Based Survey“. *Language Teaching* 41, Nr. 4 (Oktober 2008): 465–96. <https://doi.org/10.1017/S026144480800520X>.
- Ramachers, Stefanie. „Wanderlust, dialogisch leren met korte verhalen binnen het schoolvak Duits“. Unveröffentlicht, o. J.
- Ramachers, Stefanie, und Roel Willems. „Lezen in de vreemde taal: Net zo plezierig als in de moedertaal?“ Unveröffentlicht, o. J.
- . „Reading Short Stories in a Foreign Language: Is It as Pleasurable as in the L1? Absorption and Appreciation of Short Stories in Dutch Secondary School Students Learning German as a Foreign Language“. Unveröffentlicht, o. J.
- Rosenblatt, Louise M. *The Reader, the Text, the Poem: The Transactional Theory of the Literary Work*. Paperback ed. Carbondale: Southern Illinois University Press, 1994.
- Rosenshine, Barak, Carla Meister, und Saul Chapman. „Teaching Students to Generate Questions: A Review of the Intervention Studies“. *Review of Educational Research* 66, Nr. 2 (1996): 181–221.
- Rothmund, Kathrin. *Komplexe Welten: narrative Strategien in US-amerikanischen Fernsehserien*. 2., korr. Aufl. Deep Focus 15. Berlin: Bertz + Fischer, 2013.
- Sanders, Piet, Erik Roelofs, Jacqueline Visser, Saskia Wools, Erna Meijer, José Noijons, Tom Erkens, Hans Kuhlemeier, Huub Verstralen, und Bas Hemker. *Toetsen op school*. Herziene versie 2017. Arnhem: Stichting Cito Instituut voor Toetsontwikkeling, 2011.

- Schaper, J.C., A.M. Wennekers, and J. de Haan. „Trends in Media:Tijd“. Den Haag: Sociaal en Cultureel Planbureau, Dezember 2019. <https://digitaal.scp.nl/trends-in-mediatiid/lezen/>.
- Schiedermaier, Simone. „Gegenwartsliteratur im Unterricht Deutsch als Fremdsprache. Ein Vorschlag zur Arbeit mit Textnetzen“. In: *Was zu beginnen nicht aufhört. Facetten von Gegenwartsliteratur in der internationalen Germanistik und im Fach Deutsch als Fremdsprache*, herausgegeben von Almut Hille und Oliver Niels Völkel, 95–113. Literatur, Kultur und Medien in Deutsch als Fremd- und Zweitsprache. München, 2020.
- Schirach, Ferdinand von. „Stinkefisch“. In: *Strafe: Stories*. München: btb, 2018.
- Schoonen, Rob, Jan Hulstijn, und Bart Bossers. „Metacognitive and Language-Specific Knowledge in Native and Foreign Language Reading Comprehension: An Empirical Study Among Dutch Students in Grades 6, 8 and 10“. *Language Learning* 48, Nr. 1 (März 1998): 71–106. <https://doi.org/10.1111/1467-9922.00033>.
- Schrijvers, Marloes, Tanja Janssen, Olivia Fialho, Sven De Maeyer, und Gert Rijlaarsdam. „Transformative Dialogic Literature Teaching Fosters Adolescents’ Insight into Human Nature and Motivation“. *Learning and Instruction* 63 (Oktober 2019): 101–216. <https://doi.org/10.1016/j.learninstruc.2019.101216>.
- Schweikle, Günther, und Irmgard Schweikle. *Metzler Lexikon Literatur. Begriffe und Definitionen*. Herausgegeben von Dieter Burdorf, Christoph Fasbender, und Burkhard Moennighoff. 3. Aufl. Stuttgart, Weimar: J.B. Metzler, 2007.
- Singer, Harry, und Dan Donlan. „Active Comprehension: Problem-Solving Schema with Question Generation for Comprehension of Complex Short Stories“. *Reading Research Quarterly* 17, Nr. 2 (1982): 166. <https://doi.org/10.2307/747482>.
- SLO. „Referentiekader taal“. Internetseite. Zugegriffen 29. Juni 2023. <https://www.slo.nl/thema/vakspecifieke-thema/nederlands/referentiekader-taal/>.
- Slujsmans, Dominique, Annemieke Peeters, Linda Jakobs, und Saskia Weijzen. „De kwaliteit van toetsen onder de loep“. *Onderwijsinnovatie*, Dezember 2012, 17–25.
- Steiner, André. *Komplexes Erzählen - Literatur auf 2+n-ter Stufe: Zu einer Theorie literarischer Komplexität*. 1. Aufl. Bd. 3. Literaturtheorie. Bielefeld: Transcript Verlag, 2021. <https://doi.org/10.14361/9783839458495>.
- Urlaub, Per. „Reading Strategies and Literature Instruction: Teaching Learners to Generate Questions to Foster Literary Reading in the Second Language“. *System* 40, Nr. 2 (Juni 2012): 296–304. <https://doi.org/10.1016/j.system.2012.05.002>.
- „Voorstellen ontwikkelteam Engels Moderne vreemde talen. Voorstel voor de basis van de herziening van de kerndoelen en eindtermen van de leraren en schoolleiders uit het ontwikkelteam Engels/Moderne Vreemde Talen“. curriculum.nu, Oktober 2019.
- Witte, Theodorus Cornelis Hendrikus. *Het oog van de meester: een onderzoek naar de literaire ontwikkeling van havo-en vwo-leerlingen in de tweede fase van het voortgezet onderwijs*. Stichting lezen reeks 12. Delft: Uitg. Eburon, 2008.

Anhang

Ausgewählte Dialoge inkl. Mengenangabe der Codes zur Berechnung der Interrater-Reliabilität

	Helm auf oder Stinkefisch	Descriptive: default	Descriptive: summary	Essential question	Evaluative: character	Evaluative: text	Literary words	Off topic	Problem understanding: self assessed	Problem understanding: without noticing	Reasoning: evaluative character	Reasoning: evaluative text	Reasoning: hypothetical	Reasoning: moral - beyond story	Reasoning: moral - in story	Reasoning: understanding character	Reasoning: understanding story events	Referential: real world	Referential: self	Transformative reading	Total	
Posttest	Sf	5	1	1		3			2			1		2	1		2		2			20
Pretest	Sf	4		2		1	4				1	2		2	2	2			1	3		24
Posttest	Sf	1			1	6		2			1									1		12
Pretest	Ha	2		2	1						1					1		1	1	1		10
Pretest	Ha	10	2	1		1				1	1		1					1				18
Posttest	Sf	1				3				4			2			1			1			12
Pretest	Ha	4		3		2	1			1			2			2						15
Pretest	Ha	3		4		3	1		1		2					2	1	1	1	2		21
Posttest	Sf	2		3		1			1	1			1			1	2		1			13
Pretest	Sf	4	2	1		1	3	1	1	1	1	3					3	2				23
Posttest	Ha					2	1	1	1													5
Pretest	Sf					1		1	3			1										6
Total		36	5	17	2	24	10	5	9	8	7	7	6	4	3	9	8	5	7	7		179

Codebuch

Dit codeboek is gebaseerd op het codeboek dat ontwikkeld werd voor het onderzoek van Schrijvers et al.¹⁵³

In dit onderzoek wordt elke dialoog als eenheid onderzocht. De eerste vraag, die voor iedereen hetzelfde is, wordt niet gecodeerd. Eveneens worden gespreksafsluitingen – bijv. ‘doei’, ‘fijne dag nog’, etc. – niet gecodeerd.

Niet alle uitspraken van leerlingen in de dialogen worden gecodeerd. Soms blijven woorden, vragen of hele zinnen ongecodeerd, bijvoorbeeld omdat leerlingen dingen zeggen om het gesprek op gang te houden, die echter inhoudelijk niet relevant zijn voor dit onderzoek. Het doel is namelijk niet, om alle uitspraken van leerlingen te coderen. Uiteindelijk gaat het om uitspraken die een interpretatief karakter hebben of een voorstadium daarvan of voorwaarde voor interpretatie zijn én het gaat om uitspraken, waarin leerlingen laten zien dat ze de tekst niet begrepen hebben, omdat begripsproblemen van invloed zijn op de interpretatievaardigheden van leerlingen.

Soms kom je stukken tekst tegen, waar een code in een code gecodeerd moet worden.

Voorbeeld:

Ik denk dat de jongen zijn helm uiteindelijk af gaat doen als hij uit het rouwproces komt.

Denk ik ook

Dan gaat hij beseffen dat zijn vader niet terug k~~o~~n komen.

Ja, als net iemand is overleden ga je rare dingen doen

En uiteindelijk komt alles goed

Het hele bovenstaande stuk hoort bij de code *Reasoning: hypothetical*, maar tegelijkertijd is de uitspraak ‘als net iemand is overleden ga je rare dingen doen’ ook een veralgemenisering die met *Referential: real world* gecodeerd dient te worden. Let er dus op dat er ook codes in codes toegewezen kunnen en moeten worden.

¹⁵³ Schrijvers et al., „Transformative Dialogic Literature Teaching Fosters Adolescents’ Insight into Human Nature and Motivation“.

Algemeen geldt, als leerlingen hier meer dan één mogelijkheid voor een situatie benoemen, die apart coderen. Bijv.:

Wat zou de schrijver hiermee willen bereiken?

Er zijn twee mogelijkheden.

Vertel

Hij wil duidelijk maken dat buitenlanders sneller vooroordelen. → Reasoning: understanding story events

Of?

Hij wil duidelijk maken dat voor oordelen fout zijn. → Reasoning: understanding story events

Had jij eigenlijk medelijden met de jongen?

Nee. Hij kiest er zelf voor om een helm te dragen.

Huh? licht toe!

Nou als hij dan gepest wordt of jeuk heeft is het zijn eigen schuld. Want hij heeft er zelf voor gekozen om de helm te dragen

Dat vind ik gemeen van je. Deels heb je gelijk, maar die jongen is wel zijn vader verloren !!!! en hij zit nog op school dus hij is nog aardig jong

Let op: soms loopt eenzelfde redenering door over meerdere regels. Als het allemaal bij dezelfde redentatie hoort, dan maar één code toekennen!

Specifieke opmerking betreffende de codes met *Reasoning*:


- Reasoning hoeft niet per se causaal te zijn. Het valt op, dat leerlingen niet altijd duidelijke signaal woorden gebruiken voor hun redentaties en om die reden kan er bij het volgende fragment over getwist worden, of de leerling de tekst interessant vond om te lezen, omdat ze wilde weten wat er op het einde zou gebeuren of dat deze twee uitspraken door middel van correlatie samenhangen. Ook voor zulke niet duidelijk causale, maar wel samenhangende uitspraken wordt de code *Reasoning* gebruikt. In onderstaande geval zou dat zijn: *Reasoning: evaluative text*.

Ik vond het wel interessant om te lezen.

Ja, mee eens. ik wilde weten wat er op het einde ging gebeuren.

Descriptive: default

Deze code gebruiken als er feiten / gebeurtenissen uit het verhaal naverteld worden, zonder diepere interpretatie. Leerlingen vertalen of parafaseren wat er letterlijk in de tekst staat.

Terwijl die man eigenlijk gewoon oud en blind is aan n oog.

het stukje waar de jongen zegt dat hij de helm af doet als de vader er weer is.

Descriptive: summary

Deze code wordt gebruikt voor uitspraken die weliswaar descriptief zijn, maar waarin leerlingen wel laten zien, dat ze verschillende onderdelen van de tekst samenvatten om tot deze uitspraak te komen. Deze code is gebaseerd op niveau R4, samenvatten, van de taxonomie van

Kwakernaak.¹⁵⁴ Een voorbeeld van een uitspraak die deze code krijgt is te vinden helemaal achter in het codeboek onder het kopje ‘voorbeelddialoog’ en wordt hieronder uitgebreider uitgelegd.

Op de eerste vraag: ‘Oké, welk stukje van het verhaal sprong er voor jou echt uit?’ antwoordt de leerling: ‘dat hij zoveel moeite deed voor een helm.’ Dit antwoord omvat in feite meerdere stukken uit de tekst, namelijk dat de jongen in *Helm auf* geld bij elkaar geschraapt heeft om de helm te kopen, dat hij volhardt in het dragen van de helm ondanks het feit dat het hem veel lichamelijk ongemak brengt, dat hij een ketting koopt zodat niemand de helm van zijn hoofd af kan halen, etc. Er kan dus gesteld worden dat de leerling die deze uitspraak doet, verschillende gebeurtenissen uit het verhaal samenvat. Daarom wordt aan deze uitspraak de code *Descriptive: summary* toegewezen.

Essential question

Deze code gebruiken voor essentiële vragen, ook wel hamvragen genoemd. Janssen et al. definiëren dit soort vragen als volgt: “[Essential questions are] questions that appear to go to the heart of the story, that do not have one obvious ‘right’ answer, and keep lingering in your mind after having read the story.”¹⁵⁵ Deze code wordt dus gebruikt als leerlingen vragen stellen die 1. aan de kern van het verhaal raken, 2. niet duidelijk beantwoord worden in de tekst en 3. in je achterhoofd blijven sluimeren als het verhaal uitgelezen is. Soms worden vragen niet als een klassieke vraag gesteld, maar zeggen leerlingen bijvoorbeeld dat ze benieuwd zijn naar ... → Dit kan als vraag opgevat worden en moet dus gecodeerd worden als *Essential question* als aan de overige criteria voldaan wordt.

Deze code ook gebruiken, als leerlingen vragen stellen over waarom een auteur een bepaalde keuze zou kunnen hebben gemaakt.

Let op!: soms stellen leerlingen een *Essential question* zonder dat ze proberen er een antwoord op te vinden. De vraag wordt ook in dat geval als *Essential question* gecodeerd.

Waarom zou die jongen dat denk je doen voor zijn vader

*Ja, misschien is er wel iets met de vader gebeurt.
Zou kunnen, ik vraag me af wat.*

¹⁵⁴ Vgl. Kwakernaak, „Een taxonomie voor het vreemdetalenonderwijs“, S. 14.

¹⁵⁵ Janssen, Braaksmā, und Couzijn, „Self-Questioning in the Literature Classroom: Effects on Students’ Interpretation and Appreciation of Short Stories“, S. 100.

Evaluative: character

Deze code gebruiken, als leerlingen zeggen wat zij van een karakter vinden, zonder dat er een argumentatie plaatsvindt over het waarom. Algemene uitspraken waarin leerlingen ‘alles’ of ‘gewoon’ en dergelijke als argumentatie aanvoeren horen ook thuis bij de code *Evaluative: character*, omdat de redenering te zwak en te algemeen is om onder *Reasoning: evaluative character* te vallen.

*Wat vond je van Stinkfisch dan?
Een beetje raar*

Evaluative: text

Deze code gebruiken, als leerlingen zeggen wat zij van (een gedeelte van) de tekst / het verhaal vinden, zonder dat er een argumentatie plaatsvindt over het waarom. Algemene uitspraken waarin leerlingen ‘alles’ of ‘de hele tekst’ ‘Duits’ en dergelijke als argumentatie aanvoeren (zoals in het eerste voorbeeld hieronder) horen ook thuis bij de code *Evaluative: text*, omdat de redenering te zwak en te algemeen of niet *textspezijsch* genoeg is om onder *Reasoning: evaluative text* te vallen.

*Ik vond het verhaal nogal vreemd
Wat vond je vreemd dan?
alles eigenlijk*

het was wel een best goed verhaal.

Literary words

Deze code gebruiken als leerlingen een zekere vaktaal met betrekking tot literatuur gebruiken. Voorbeelden hiervan zijn: ik-persoon, personage, schrijver / auteur, open einde, symbool staan voor, etc. Per literair woord maar één keer per dialoog coderen. De woorden ‘verhaal’ en ‘tekst’ worden niet gecodeerd.

Omdat de ik-verteller je helemaal meenam in zijn gedachtes

Hoe vind je het thema van het verhaal

het is een best open einde

Problem understanding: self assessed

Deze code gebruiken als een leerling iets zegt over in hoeverre hij/zij de tekst begrepen heeft.

*aan het begin vond ik het wel moeilijk te begrijpen.
Ja, maar een stukje verder snapte ik het wel*

*voordat we gaan nadenken over het verhaal, had jij moeilijk woorden
Ja, [brizenscheider], geen idee wat dat is
die vond ik ook moeilijk, heb jij het gevonden
nee, stap 1 lukte niet
hmm dat had ik ook bij stap 2+3
Ja, vervelend woord is het
en stap 4?
Nee, ik weet nog steeds niet wat het betekend
ik ook niet dan maar zonder dat woord
Straks even op internet opzoeken!*

Problem understanding: without noticing

Deze code gebruiken als een leerling (een onderdeel van) de tekst duidelijk niet begrepen heeft en dat zelf niet door heeft. Daarbij gaat het niet om de overdraagbare zin of interpretatie, maar juist om de dingen die letterlijk in de tekst staan en die door een leerling dus niet begrepen zijn als gevolg van taalproblemen; Bijvoorbeeld gebruiken als er descriptieve fouten gemaakt worden.

Als een leerling duidelijk aangeeft, dat hij de hele tekst moeilijk vond om te begrijpen, dan moet dat als *Self assessed* gecodeerd worden. Dikwijls wordt daarna toch geprobeerd over de tekst te spreken. De dingen die dan daar gezegd worden en niet kloppen: coderen met *without noticing*.

*Waar ging de tekst volgens jou over
over een man die veel dode vissen kocht en waar zijn kinderen bij betrokken waren
ooh oke wie vond jij de hoofdpersoon?
de vader want die kwam het vaakst aan bod
de kinderen niet?
ik vond van niet zij waren niet het belangrijkste
waarom niet
kwamen later pas echt voor
ja dat snap ik.
fijn ik ook
waar dacht je dat het afspeelde
ergens in een kelder.*

Reasoning: evaluative character

Deze code gebruiken als een leerling een evaluatieve uitspraak over een personage doet en ook een argumentatie voor die evaluatie laat zien. Ook hierbij is soms wel enigszins interpreterend vermogen van de beoordelaar nodig, om de argumentatie van de leerling te kunnen zien (zoals uitgelegd bij *Reasoning: evaluative text*). Als er twijfel bestaat over wel of geen argumentatie, dan wordt het voordeel van de twijfel verleent en gekozen voor deze code, codeer sowieso het stukje van de redentatie mee.

Het valt op, dat leerlingen vaker descriptief redeneren, zeker in het geval van *Helm auf*. Daar zou je een soort argumentatie achter kunnen vinden, die verwijst naar de bekende ‘echte’ wereld van de leerling; een uitspraak in de trend van: ‘ik vond het raar, dat hij die helm altijd op heeft’ zou dus argumenterend opgevat kunnen worden in de richting van ‘want het is niet normaal dat je de hele tijd een helm draagt’. Zulke uitspraken dus ook coderen met *Reasoning*.

*Ik vond de jongen wel dapper en zeker van zijn zaak.
Hij zou er alles voor over hebben voor die helm.*

Reasoning: evaluative text

Deze code gebruiken als een leerling een evaluatieve uitspraak over de tekst / het verhaal doet en ook een argumentatie voor die evaluatie laat zien. Hierbij is soms wel enigszins interpreterend vermogen van de beoordelaar nodig, om de argumentatie van de leerling te kunnen zien:

*het stukje waar de jongen zegt dat hij de helm af doet als de vader er weer is.
Ja, dat was inderdaad best apart.
Precies, *want* de jongen zegt de hele tijd dat hij de helm niet zal afzetten, maar als zijn vader komt wel *en dat snap ik niet helemaal / want dat is niet normaal**

In dit voorbeeld zijn de sterretjes en de woorden daartussen ter verduidelijking toegevoegd, die staan dus niet in de dialoog van deze leerling. Hier valt op, dat de leerling geen signaalwoorden gebruikt die een argumentatiestructuur laten zien. Eigenlijk doet de leerling alleen maar een descriptieve uitspraak, waarop een evaluatieve uitspraak volgt en sluit weer af met een descriptieve uitspraak, die in feite een herhaling is van de voorgaande descriptieve uitspraak. Tussen sterretjes is een signaalwoord en duidelijkere redenering toegevoegd, die aangeven hoe je deze uitspraak als argumentatie zou kunnen interpreteren. Over het algemeen geldt voor het coderen in zulke gevallen, dat als er twijfel is tussen *Evaluative: text* en *Reasoning: evaluative text*, dat de code *Reasoning: evaluative text* van toepassing is. Het bovenstaande voorbeeld krijgt dus het voordeel van de twijfel en wordt gecodeerd met *Reasoning: evaluative text*.

*Wat vond je van het verhaal?
Ik vond het best apart, vanwege die helm.*

*Ik vond het wel interessant om te lezen.
Ja, mee eens. ik wilde weten wat er op het einde ging gebeuren.*

Reasoning: hypothetical

Deze code gebruiken wanneer een leerling:

- Speculeert over wat er na afloop van het verhaal nog zou kunnen gebeuren.
- Verbetersuggesties ten opzichte van het verhaal doet.
- Speculeert over wat er ook had kunnen gebeuren / wat anders had kunnen gaan.

*Ik denk dat de jongen zijn helm uiteindelijk af gaat doen als hij uit het rouwproces komt.
Denk ik ook
Dan gaat hij beseffen dat zijn vader niet terug k~~o~~n komen.
Ja, als net iemand is overleden ga je rare dingen doen
En uiteindelijk komt alles goed*

Reasoning: moral – in story

Deze code gebruiken voor morele redeneringen die betrekking hebben op de verhaalsituatie.

Onder morele redeneringen vallen:

- Uitspraken over of iets in het verhaal rechtvaardig, goed, terecht, verdiend, etc. is.
- Uitspraken over schuld en onschuld.
- Uitspraken waarin een moreel oordeel wordt gegeven over het gedrag van personages: dat kan echt niet, is echt niet normaal, zou je niet moeten doen, etc.

Let goed op of uitspraken terugverwijzen naar (een personage in) het verhaal of worden uitgebreid naar ‘de wereld’ / ‘mensen’ → kies dan voor *Reasoning: moral – beyond story*.

*Ik snap niet waarom die jongen hem nou ging uitschelden
nee, ik ook niet hij had niet naar zijn vrienden moeten luisteren*

Maar ik vind dat er wel iemand voor hem op moet nemen.

Reasoning: moral – beyond story

Deze code gebruiken voor morele redeneringen die refereren aan situaties buiten het verhaal. Het gaat om soortgelijke uitspraken als bij *Reasoning: moral – in story* (rechtvaardig, verdiend, terecht, schuldig, etc.), maar in vorm lijken ze op de code *Referential: real world* (zie de code direct hieronder); Er wordt vaak gerefereerd naar de echte wereld buiten het verhaal door te spreken over ‘mensen’, ‘iedereen’, ‘niemand’, etc.

Vooroordelen zijn ook niet goed maar ze bestaan wel

Reasoning: understanding character

Deze code gebruiken als de leerling beredeneert waarom een personage iets doet, denkt, voelt, zich op een bepaalde manier gedraagt. De leerling probeert het personage te ‘begrijpen’ of te karakteriseren: wat is het voor iemand, waarom is diegene zo, waaruit blijkt dat? Gebruiken voor:

- Uitspraken waarin beredeneerd wordt waarom personage X, Y, Z doet/zegt/denkt/voelt.
- Uitspraken waarin een leerling aangeeft te denken dat er X, Y, Z met een personage aan de hand is.
- Uitspraken waarin beredeneerd wordt waarom een personage is zoals hij is.

*Zouden ze die man nou alleen zo noemen door geruchten?
Ja, maar ook door de stank.
Terwijl die man eigenlijk gewoon oud en blind is aan n oog.
Ze noemen hem zo uit vooroordelen.*

*Maar ik denk dat de jongen er wel een reden voor heeft.
Ja, wat dan?
Misschien heeft de vader wel een ongeluk gehad. en voelt de jongen zich veiliger met helm
zou kunnen.*

Reasoning: understanding story events

Deze code gebruiken als een leerling beredeneert wat bepaalde gebeurtenissen in het verhaal zouden kunnen betekenen. Dit oogt als ‘klassieke interpretatie’: waarom denk ik dat dit-en-dat gebeurt in het verhaal? Wat denk ik dat het betekent? Gebruiken voor:

- Uitspraken waarin de leerling een gebeurtenis of ander element in het verhaal niet begrijpt en een betekenis / verklaring zoekt.
- Uitspraken waarin beredeneerd wordt waarom de auteur bepaalde keuzes heeft gemaakt.

*Dat klopt. Heb je enig idee waarom die geruchten eigenlijk waren ontstaan
Ja ik denk omdat hij teruggetrokken is in de maatschappij*

*Wat zou de schrijver hiermee willen bereiken?
Er zijn twee mogelijkheden.
Vertel
Hij wil duidelijk maken dat buitenlanders sneller vooroordelen.
Of?
Hij wil duidelijk maken dat voor oordelen fout zijn.*

misschien is er wel iets met de vader gebeurt.

Referential: real world

Deze code gebruiken als een leerling in zijn uitspraken verwijst naar de echte wereld buiten het verhaal. Deze code heeft soms overlapping met *Reasoning: moral –beyond story* → als een uitspraak een duidelijke moreel bevat, dan kiezen voor *Reasoning: moral – beyond story* ipv. *Referential: real world*. Specifiek moet deze code dus gebruikt worden, als uitspraken van leerlingen algemener zijn dan alleen betrekking hebbend op het verhaal. Signaalwoorden hiervoor zijn: mensen, iedereen, personen, in het echt, etc.

Het laat zien dat iemands omgeving heel anders kan zijn dan de persoon zelf.

*En als je het zo bekijkt zie je dat de meeste problemen bij opvoeding beginnen.
inderdaad! Daarom moet de overheid daar ook wat aan doen.
Maar daar zullen ze vast geen geld voor overhebben
Ja, of de belasting gaat omhoog.
Maar dan gaan ze weer veel gezeik krijgen*

Referential: self

Deze code gebruiken als een leerling in relatie tot het verhaal verwijst naar zichzelf. Specifiek wordt deze code gebruikt als leerlingen uitspraken doen:

- Waarbij ze zich afvragen: Wat zou jij doen? Hoe zou jij reageren? Waar zou jij voor kiezen? (Leerlingen verwijzen in dit geval meestal met ‘jij’ naar zichzelf, omdat ze in deze opdracht een dialoog met zichzelf voeren.)
- Waarbij ze hun eigen gedrag vergelijken met wat er in het verhaal gebeurt: ik zou dat ook doen, ik zou dat niet doen, etc.
- Waarbij ze verwijzen naar hun persoonlijke leven – bijv. ouders, vrienden, etc. – zolang er een directe relatie met het verhaal is.

*Ik zou zelf nooit zoiets doen.
Altijd een helm dragen bedoel je?
Ja, dat lijkt me echt verschrikkelijk.*

Transformative reading

Deze code gebruiken als:

- Een leerling iets uit het verhaal herkende van zichzelf / iets herkenbaar vindt.
- Een leerling zegt een gebeurtenis of persoon uit het verhaal voor zich te kunnen zien.
- Een leerling zich kan verplaatsen in het verhaal of een personage. (inplaatsen, inbeelden, alsof je erbij bent, begrijpen hoe iemand zich voelt.)
- Een leerling sympathie aan een personage betuigt / meeleeft met een personage. Het gaat om gevoelens van compassie, medeleven, medelijden, etc.

*Had jij eigenlijk medelijden met de jongen?
Nee. Hij kiest er zelf voor om een helm te dragen.
Huh? licht toe!
Nou als hij dan gepest wordt of jeuk heeft is het zijn eigen schuld. Want hij heeft er zelf voor gekozen om de helm te dragen
Dat vind ik gemeen van je. Deels heb je gelijk, maar die jongen is wel zijn vader verloren !!!! en hij zit nog op school dus hij is nog aardig jong*

*Nee klopt, maar ik heb wel een idee van hoe hij eruit ziet.
Ja, dat is wel zo.*

Off topic

Deze code gebruiken als leerlingen praten over dingen, die niet langer tekst gerelateerd zijn.

Nee, dat is onzin. blinde zijn bijna altijd kasplantjes.

hoho, niet zo snel!

waarom niet?

Ik ken een blinde die HBO heeft afgerond

dat is knap.

Jazeker

maar dan zijn veel mensen die blind zijn een kasplantje.

nee, want ze kunnen ook iets nuttigs doen voor de maatschappij.

echt niet

echt wel

hoezo wel?

Met braille en hun laptop kunnen ze vanalles opschrijven wat anderen kan helpen.

waarmee, ze moeten eerst die kennis opdoen en ik heb nog nooit een schoolboek met braille gezien.

die vragen ze apart aan.

oh, dat wist ik niet.

Voorbeelddialoog

Schriftelijke dialoog van een leerling - Helm auf	Codes
Ok, welk stukje van het verhaal sprong er voor jou echt uit?	Niet coderen – eerste vraag is voor iedereen hetzelfde
dat hij zoveel moeite deed voor een helm	Descriptive: summary
Ja, vond ik ook	
wat vond jij dat eruit sprong?	
ik vond het [...] dat hij die helm niet af wilde doen	Descriptive
Ja, alsof het hem dierbaar was	Reasoning: understanding character
voordat we gaan nadenken over het verhaal, had jij moeilijk woorden	Problem understanding: self assessed
Ja, [brizenscheider], geen idee wat dat is	
die vond ik ook moeilijk, heb jij het gevonden	
nee, stap 1 lukte niet	
hmm dat had ik ook bij stap 2+3	
Ja, vervelend woord is het	
en stap 4?	
Nee, ik weet nog steeds niet wat het betekend	
ik ook niet dan maar zonder dat woord	
Straks even op internet opzoeken!	
Terug naar het verhaal	
ja, welke vragen riep het bij jou op?	
waarom droeg hij überhaupt die helm?	Essential question**
ja, dat was ook het eerste dat bij mij opkwam	
ik kon het ook niet opmaken uit het verhaal	
ik denk gewoon dat de helm hem dierbaar is, misschien is iemand overleden?	Reasoning: understanding character Reasoning: understanding story events
Ja, dat zou best kunnen en dat zou best de vader kunnen zijn	
Ja, eens!	
hij zou alleen zijn helm afdoen voor zijn vader	Descriptive
daarom, hij heeft het erover dat hij hem terug wil of iets, dat hij gewoon weer thuis is	Descriptive
Ja, als hij dood zou zijn zou hij er echt niet meer zijn	
Ja, en is di jongen zo verdrietig dat hij de helm die miss. van zijn vader is niet af wilt doen.	Reasoning: understanding character
Ja, er kunnen natuurlijk ook meer redenen zijn waarom dat di vader er niet is.	Reasoning: understanding story events
Ja, tuurlijk! bijv. hij is weggegaan of hij heeft een nieuwe vriendin	
daarom!	
ja, ik denk persoonlijk dat hij niet dood is	
oh, mij juist wel het lijkt mij juist heel logisch dat hij dood is	Reasoning: understanding story events
oh nou ok. dat kan.	
hij is er wel heel erg van overtuigd dat hij de helm niet meer af gaat doen	

ja, en dat is super knap!	
Knap? ik vind het ergens ook wel overdreven. <u>Je kan niet voor altijd een helm ophouden</u>	Reasoning: evaluative character Referential: real world
vind ik niet, hij moet fijn doen waar hij zin in heeft en dat is goed.	
bedankt voor het gesprek. was leuk! Ja, vond ik ook, dankjewel!	Niet coderen – afsluitingen nooit coderen

*Er was niet leesbaar wat de leerling hier geschreven had. Hoewel het begin van deze uitspraak, ‘ik vind het’, erop lijkt alsof hier een evaluatieve uitspraak gedaan wordt, kan deze geen code toegewezen krijgen, omdat niet na te gaan is, wat deze uitspraak van de leerling precies betekent. Dat de jongen de helm niet af wilde doen, in het vervolg van de uitspraak van de leerling, is wel duidelijk een descriptieve uitspraak.

**Hoewel in het verhaal duidelijk wordt, dat de jongen de helm op de een of andere manier vanwege zijn vader draagt, wordt de vraag niet bevredigend beantwoord door de tekst. Er is dus strikt genomen wel één antwoord op de vraag in de tekst, maar deze vraag wordt vanwege de onbevredigende aard van het antwoord in de tekst desalniettemin gecodeerd als *Essential Question*.

Format schriftlicher Dialoge

<p>Oké, welk stukje van het verhaal sprong er voor jou echt uit?</p> 