



## UvA-DARE (Digital Academic Repository)

### The story, the self, the other

*Developing insight into human nature in the literature classroom*

Schrijvers, M.S.T.

#### Publication date

2019

#### Document Version

Final published version

#### License

Other

[Link to publication](#)

#### Citation for published version (APA):

Schrijvers, M. S. T. (2019). *The story, the self, the other: Developing insight into human nature in the literature classroom*. [Thesis, fully internal, Universiteit van Amsterdam].

#### General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

#### Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

MARLOES SCHRIJVERS

THE STORY  
THE SELF  
THE OTHER

DEVELOPING INSIGHT INTO HUMAN NATURE  
IN THE LITERATURE CLASSROOM

THE STORY, THE SELF, THE OTHER  
DEVELOPING INSIGHT INTO HUMAN NATURE  
IN THE LITERATURE CLASSROOM



UNIVERSITY OF AMSTERDAM  
Research Institute of Child Development  
and Education

Cover design: Thomas Bijen

Printed by: Gildeprint

ISBN: 978-94-6323-589-1

This publication was supported by a grant from NWO, the Dutch Organization for Scientific Research, grant number 360 30 240.



THE STORY, THE SELF, THE OTHER  
DEVELOPING INSIGHT INTO HUMAN NATURE  
IN THE LITERATURE CLASSROOM

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor  
aan de Universiteit van Amsterdam  
op gezag van de Rector Magnificus  
prof. dr. ir. K.I.J. Maex  
ten overstaan van een door het College voor Promoties  
ingestelde commissie, in het openbaar te verdedigen  
in de Agnietenkapel

op vrijdag 3 mei 2019, te 14.00 uur

door

Marloes Sophie Thomas Schrijvers

Geboren te Baarle-Nassau

## **Promotiecommissie**

Promotor: prof. dr. G.C.W. Rijlaarsdam  
Universiteit van Amsterdam

Co-promotores: dr. T.M. Janssen  
Universiteit van Amsterdam

dr. O. da Costa Fialho  
Universiteit Utrecht

Overige leden: prof. dr. M.L.L. Volman  
Universiteit van Amsterdam

prof. dr. C.A.M. van Boxtel  
Universiteit van Amsterdam

prof. dr. I. Pieper  
Universität Hildesheim

prof. dr. W.L.H. van Lierop-Debrauwer  
Tilburg University

prof. dr. E.M.P. van Gemert  
Universiteit van Amsterdam

Faculteit der Maatschappij- en Gedragwetenschappen

## TABLE OF CONTENTS

CHAPTER 1	Introduction	1
CHAPTER 2	The impact of literature education on students' perceptions of self and others: Exploring personal and social learning experiences in relation to teacher approach	7
CHAPTER 3	Gaining insight into human nature: A review of literature classroom intervention studies	41
CHAPTER 4	Designing a literature classroom intervention to foster 10 <sup>th</sup> grade students' insight into human nature	87
CHAPTER 5	Effects of dialogic literary instruction on 10 <sup>th</sup> grade students' insight into human nature	125
CHAPTER 6	Transformative Dialogic Literature Teaching fosters adolescents' insight into human nature	155
CHAPTER 7	General discussion	183
REFERENCES		217
AUTHOR INDEX		231

SUMMARY	237
SAMENVATTING	247
APPENDICES	259
CONTRIBUTIONS OF AUTHORS	297
RELATED PUBLICATIONS AND PRESENTATIONS	299
CURRICULUM VITAE	301
DANKWOORD	303
ICO DISSERTATION SERIES	309

# CHAPTER 1

## INTRODUCTION

*Books challenge and interrogate;  
they give us windows into the lives of others  
and mirrors so that we can better see ourselves.*

– John Green

### 1 AIM AND SCOPE

Ever since ancient times, mankind faces questions about self and others: who are we, and how do we relate to other people? Where are the boundaries between “us” and “them”? Peeking through a window into the lives of others or taking a look in the mirror to face ourselves may add to our understanding of what it means to be human. Reflecting on how we position ourselves in relation to others in the world may be particularly important in view of fundamental developments of the past few decades, including globalization, migration, and responses thereto, ranging from tolerance to bigotry. Education has often been considered a place where people may learn to reflect on their own nature as well as the nature of others (Biesta, 2007; Nussbaum, 1997; 2010), as is illustrated by particular approaches to teaching and learning, such as social and emotional education (Elias et al., 1997), citizenship education (Derricott, 2014), and moral and character education (Nucci, Krettenauer, & Narváez, 2014).

The literature classroom pre-eminently offers a space of opportunity for adolescents to develop insight into ourselves and others. As John Green (2016) suggested, reading fictional and literary texts – novels, stories, poems – may offer us such insights. When we read, we are temporarily drawn into a simulated social world, in which we can safely experience what it would be like to be in situations that may either resemble or be very different from situations in our own lives (Mar & Oatley, 2008). This experience is considered to be “transformative”: it may change our insight into ourselves and into ourselves in relation to

others (Fialho, 2012). In this dissertation, we define this insight in a broad sense: we understand it as insight into *human nature*, which may include insight into our own previously unrecognized qualities or shortcomings, insight into self-other relations, understandings of and altered attitudes toward individual others and groups of people, and considerations of difficulties or moral dilemmas that people may face. Reading fictional or literary texts may result in gaining such insights, for adults, adolescents and younger children alike, (e.g., Fialho, 2018; Fialho, Hakemulder, & Bal, 2016; Lysaker, Tonge, Gauson, & Miller, 2011; Richardson & Eccles, 2007).

We focus in this dissertation on literature classrooms in upper secondary education in the Netherlands. The Dutch literature curriculum is not subject to any nation-wide regulations: schools and teachers have much freedom in deciding on reading lists, instructional approaches, and examinations. Formal literature education starts at the beginning of upper secondary education, in grade 10, when students are about 15 years old. Up to 9<sup>th</sup> grade, students usually are stimulated to read children's and young adult literature, but from 10<sup>th</sup> grade onwards their teachers mostly expect them to read increasingly complex literary texts intended for adult readers. Students work toward three global objectives that are addressed in the final examinations: acquiring literary-historical knowledge, applying structural-analytical skills, and reflecting on their own literary reading experiences and development (Dutch Institute for Curriculum Development, 2012). Although the third objective can be considered a form of self-development, this dissertation goes beyond the concept of the self as a literary reader: it considers the self as a whole, in continued social interaction with others. As such, developing insight into human nature is not explicitly included in the global objectives for literature education in upper secondary education in the Netherlands. Moreover, in most schools, literature education does not have the status of a separate school subject. Usually, it is considered to be a subdomain of Dutch language teaching, like writing, grammar, and vocabulary. In upper secondary education (grade 10-12), a relatively small amount of time is devoted to literary instruction: of about 155 minutes of Dutch language teaching per week, students receive on average 35 to 40 minutes of literary instruction (Oberon, 2016).

Despite its implicitness in the global objectives for literature teaching and the time constraints teachers face, insight into human nature seems to be valued as a potential outcome of literary instruction by teachers as well as policy makers in the Netherlands. For example, teachers reported in previous studies that they consider personal development – including identification, empathy for others, and learning about yourself and the world – to be an important ob-

jective of literature teaching (Janssen, 1998; Oberon, 2016). Likewise, after the Dutch Ministry of Education announced a curricular reform, to be implemented in 2021, teachers in the curriculum development team for Dutch language and literature stated that literary reading familiarizes students with other worlds, contributes to moral development, and helps them to think about “why people make particular choices, and about themselves, the other and the world” (Curriculum.nu, 2018a, p. 4, translation *MS*). Along the same lines, the Dutch Institute for Curriculum Development suggested that “[...] literature education has the capacity to expand students’ social and cultural horizons and to stimulate their empathetic capabilities” (2015, p. 25, translation *MS*). However, these claims remain at a rhetoric level: in the Dutch context, there is no empirical support for the assumption that literary instruction may foster students’ insight into human nature. With this dissertation, we aim to fill this void. Thereby, we also intend to contribute to the development of valid intervention studies in the literature classroom, which enable us to draw conclusions about the effectiveness of a particular instructional approach. The following overarching question guides this research project:

*Which instructional approach to literature teaching in Dutch upper secondary education is appropriate for fostering students’ insight into human nature?*

## 2 STRUCTURE OF THIS DISSERTATION

This dissertation consists of five studies, as shown in Figure 1.1. The research project as a whole can be characterized as a form of educational design research. The design of an instructional approach lies at the heart of the project. Study 1 and 2 provided an empirically grounded framework for designing two subsequent versions of an intervention, in Study 3. The effects of the two versions were tested in two intervention studies (Study 4 and 5). Here, we introduce the five studies included in this dissertation in closer detail. Each study is set up as a separate journal article. As a consequence, when presenting these five studies in this dissertation, some overlap in theoretical frameworks and discussion sections is inevitable. On the other hand, the advantage for the reader is that each chapter can be read on its own.

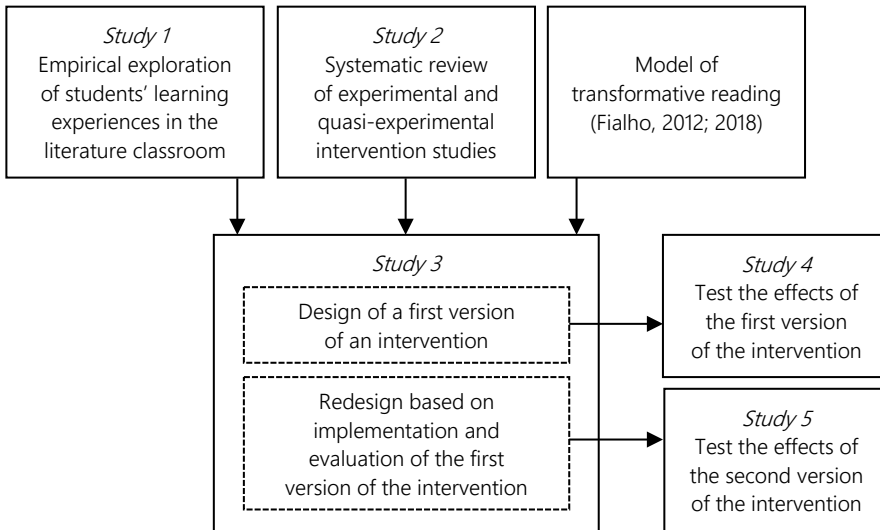


Figure 1.1. Overview of the research project.\*

## 2.1 Coming to Terms

As a first step, we needed to explore the context in which the research project were to take place. Therefore, in Study 1 (Chapter 2), we conducted an empirical exploration of current practices and learning experiences in Dutch upper secondary school literature classrooms. We investigated whether students reported to gain any insight into themselves and others as a result of regular literature education, and whether those experiences varied among classrooms, depending on the instructional approach taken by the students' teachers. The findings – relations between learning experiences and characteristics of literary instruction – provided first indications of parameters that might play a role in the design of an instructional approach.

In addition to the outcomes of empirical explorations of the field, the framework for educational design research consisted of instructional design principles. In Study 2 (Chapter 3), we identified these principles via a systematic review of previous experimental and quasi-experimental literature classroom intervention studies that aimed to foster students' insight into human nature. In this review, thirteen studies were included. Nine of these provided empirical support for fostering students' insight into human nature – that is, insight into themselves, fictional others, and/or real-world others. The outcomes of Study 1 and 2 informed the design of an instructional approach for literature education in 10<sup>th</sup> grade classrooms in the Netherlands.

## 2.2 *From Design Principles to an Intervention Design*

In Study 3 (Chapter 4), we developed an instructional approach in an iterative design process. As Figure 1.1. shows, the design was informed not only by the outcomes of Studies 1 and 2, but also by a theoretical-empirical model of transformative reading (Fialho, 2012; 2018). The model of transformative reading includes self-other insights as a component, and indicates that several other components underlie these insights – reading experiences such as vividly picturing the setting and characters in a story, and feeling sympathy or compassion for characters. These and other transformative reading components are given particular attention in the design of the instructional approach. In the design process, we collaborated with teachers of Dutch language and literature who, for example, suggested titles of literary texts to be used, and pointed out time constraints they face in their everyday teaching practices. We designed two subsequent versions of an intervention, in which the second was a redesign of the first, by focusing in particular on the validity and practicality of both versions.

## 2.3 *Testing the Interventions*

Study 4 and 5 are mixed-methods intervention studies in which we tested the effects of the two versions of the intervention on students' transformative reading experiences, including – most importantly – their insight into human nature. In these studies, we relied on questionnaire data as well as writing task data. These two studies added to Fialho's (2018) work with adult readers, as we developed instruments and coding schemes to detect adolescent students' transformative reading experiences. For example, the newly developed Transformative Reading Experiences Questionnaire (TREQ) for adolescents was used in both intervention studies and yielded consistent validity indices for both samples of students, across measurement moments.

More specifically, Study 4 (Chapter 5) focused on assessing students' perceptions of their own learning as a result of the first version of the intervention, as well as on effects of the intervention on students' insight into human nature. Study 5 (Chapter 6) investigated the effects of the redesigned version of the intervention on students' insight into human nature as well as their support of eudaimonic motivations for reading, that is, reading for meaningfulness and insight into human conditions. In addition, we aimed to alleviate prominent challenges that students face in the literature classroom: their limited capability to deal with difficulties that may emerge during reading literary texts, and their

rather low motivation for literature education. Therefore, we assessed whether the intervention enhanced students' reported use of strategies to deal with difficulties during literary reading, and, as indicators of their motivation, the extent to which they experienced autonomy, competence and relatedness in the literature classroom.

## 2.4 *General Discussion*

In the General Discussion (Chapter 7), we bring together our main findings by providing answers to the research questions addressed in Studies 1 to 5. Subsequently, we discuss four key concepts that guided these studies: insight into human nature, transformative reading, literariness, and dialogic learning in literature classrooms. Next, we address potential validity issues in our studies regarding the intervention-as-designed and the intervention-as-implemented, the instruments and the research designs we applied, followed by discussing the external validity of our studies and the risk of a potential researcher bias. Finally, we discuss directions for future studies as well as the implications of our research for educational practice.

---

\* *This PhD research has been conducted as part of the project *Uses of Literary Narrative Fiction in Social Contexts*, which was funded by the Dutch Organization for Scientific Research (NWO; grant number 360 30 240). The grant application was prepared by Prof. Dr. F. Hakemulde and Dr. O. Fialho (Utrecht University), Dr. T.M. Janssen and Prof. Dr. G.C.W. Rijlaarsdam (University of Amsterdam), and Dr. M. Bal (Free University, Amsterdam). The grant was awarded to Hakemulder and Rijlaarsdam. Aim of the project was to develop and validate a theoretical-empirical model of transformative reading (Fialho, 2018) and to investigate how this model could inform learning about self and others in two social contexts: the workplace (Brokerhof, 2018), and the literature classroom (this dissertation).*

## CHAPTER 2

### THE IMPACT OF LITERATURE EDUCATION ON STUDENTS' PERCEPTIONS OF SELF AND OTHERS

#### Exploring personal and social learning experiences in relation to teacher approach\*

The Dutch Institute for Curriculum Development argues that literature education is important for broadening students' personal, social and cultural horizons. Indeed, reading literary fiction may alter readers' self- and social perceptions, but little is known about whether adolescents gain such personal and social insights through reading in the secondary literature classroom, nor about how these perceived learning outcomes are related to their teachers' approaches to various aspects of literature teaching. Thus, the aims of this study were to examine the impact of literature education on students' self- and social perceptions and to explore relationships between students' learning experiences and their teachers' classroom practices. Dutch students ( $N = 297$ , grades 10-12) wrote a learner report about what they learned about themselves and other people through literature education, and completed a measure on familiarity with fiction. Teachers ( $N = 13$ ) completed the Teachers' Approaches to Literature Education Questionnaire, which indicated more analytical-interpretative or personal-experiential approaches to three aspects of teaching. Students of teachers with distinct approaches to these aspects were grouped to compare their learning experiences. Findings showed that nearly all students (99%) reported to have learned about themselves and others, mainly personal characterizations of oneself and others, learning about oneself and others as literary readers, descriptions and evaluations of people's behaviors, and lessons for life. In addition, teachers' reports of more classroom interaction and student autonomy were related to students' more frequent reports of personal and social insights, but this may also partly be explained by students being more familiar with fiction and having a more positive attitude toward literary reading. Implications for personal and social learning in the literature classroom are discussed.

---

\* Schrijvers, M., Janssen, T., Fialho, O. & Rijlaarsdam, G. (2016). *The impact of literature education on students' perceptions of self and others: Exploring personal and social learning experiences in relation to teacher approach*. L1 – Educational Studies in Language and Literature, 17, 1-37. doi:10.17239/L1ESLL-2016.16.04.01

## 1 INTRODUCTION

Recently, the Dutch Institute for Curriculum Development contended that literature education “has an important value for developing citizenship, [for instance by] broadening social and cultural horizons and developing empathic capabilities” (2015, p. 15). This aim is not so far-fetched: not only has reading literary fiction often been considered an inherently cultural and social activity (e.g., Bloome & Green, 2002; Galda & Beach, 2001) and not a monolithic form of experience (Miall & Kuiken, 1998; 1995), literary scholars have also connected reading literary fiction to readers’ abilities to imagine other people’s situations and to make inferences about their thoughts and emotions (e.g., Keen, 2006; 2007; Palmer, 2004; Zunshine, 2006; 2015). Moreover, the experience of reading literary fiction has been considered a life experience that can be self-modifying and thus may have impact on readers’ self-development (Fialho, 2012; Miall & Kuiken, 2002).

The Dutch Institute for Curriculum Development, then, appeared to value what reading literary fiction may bring about: personal and social development. The Dutch secondary literature classroom may therefore precisely be the place to attend to personal and social aspects of literary reading. However, as Fialho (2012) noted from a global perspective, “there still is no consensus (and perhaps may never be) about the aims of literary education, and little know”ledge of how literary reading is processed, as social and cultural factors are involved (p. 3). The same holds true for the Netherlands. There is no prescribed curriculum for literature teaching in secondary schools: teachers are allowed much freedom and may use different approaches to literature teaching (Janssen & Rijlaarsdam, 2007).

In this study, we examine whether a representative sample of Dutch secondary school students experiences any personal and social learning in their literature classrooms, and we explore how such experiences may be related to aspects of their literature teachers’ practice. First, we outline the national context in which this study takes place. We then present the theoretical-empirical framework in which we situate this research.

### 1.1 *Literature Education in the Netherlands*

In line with the history of the institutionalization of literature as described by Graff (2007), literature education in the Netherlands originated at the end of the 19<sup>th</sup> century. In those days, it focused on historical-biographical knowledge. In the early 1970’s, structuralist approaches in literary studies emerged (Witte, Rijlaarsdam, & Schram, 2012). Close reading and structural analysis became

important. From 1980, there was a transition from text-centered to more reader-centered approaches. Dutch literary educators were influenced by reader response critics from the United States, such as Rosenblatt, Bleich and Holland, and by German reception aesthetics scholars, such as Iser and Jauss. Their views were increasingly acknowledged and have influenced the practices of at least part of the Dutch literature teachers (Janssen & Rijlaarsdam, 2007; Witte et al., 2012).

Attention for close reading and structural analysis, however, has not disappeared. In 1998, as part of a larger Dutch educational reform, a new examination program for literature education has been introduced, which still applies today and intertwines three core elements: literary-historical knowledge, structural-analytical skills, and individual literary development. This program is confined to the upper grades of secondary education, which is the focus of this study. To obtain a satisfactory grade for literature education at the end of secondary school, a student must be able to give a substantiated report of his or her reading experiences of a number of self-selected literary works; to recognize and distinguish between literary text types and be able to use literary terms for interpreting literary texts; and to give an overview of the outlines of literary history and place literary works in a historical perspective (Dutch Institute for Curriculum Development, 2012, p. 25).

Higher general education students\* must read a minimum of eight literary works, whereas pre-university students must read twelve, three of which must be published before 1880. All works must originally be written in Dutch; therefore, books by Flemish, Antillean and Surinam authors are allowed as well. In addition, teachers often require students to read literature intended for adults. Generally, students are required to keep a reading portfolio. This usually contains a reading autobiography and several "book reports", which may include book summaries, literary analyses, comparisons with films, and reviews. The portfolio allows for assessing students' individual literary development (Dirksen, 2007) and thus provides some space for a variety of reading preferences and choices among students. It is common, but not compulsory, that students round off literature education at secondary school with an oral exam about the books they have read and reported about in their portfolio.

All in all, the domain of literature education in the Netherlands allows for much freedom: "[t]eachers themselves decide which texts to work on, and de-

---

\* *The Dutch secondary educational system distinguishes between higher general education (havo, five years), which prepares for higher vocational education, and pre-university education (vwo, six years), which prepares for university.*

cide which objectives to emphasize and how much time to devote to literature" (Witte et al., 2012, p. 2). Such curricular freedom, however, may not necessarily contribute to students' motivation for literature education. Students not always become engaged with school-assigned texts. In many Dutch schools, they choose from a teacher-selected list of literary works. Conceivably, students may feel obliged to read texts they would not have chosen themselves, which might cause resistance to reading (Bintz, 1993). This could potentially impede transportation into a story, "a convergent process, where all mental systems and capacities become focused on events occurring in the narrative" (Green & Brock, 2000, p. 701) which "may be a mechanism for narrative-based belief change" (p. 703). If there is little freedom of choice, students' individual reading preferences may not be fully acknowledged, whereas research suggests that attending to their preferences, offering them a certain freedom of choice and supporting them in choosing the book that fits them best at a particular moment may be crucial for their engagement in the literature classroom (Beach, Appleman, Hynds, & Wilhelm, 2011; Lenters, 2006; Smith & Wilhelm, 2002; Witte, 2008).

### *1.2 Reading Literary Fiction: Changes in Self and Social Perceptions*

Reading literary narrative fiction has the potential to change readers' sense of self (Fialho, 2012; Sikora, Kuiken, & Miall, 2010) and their perceptions of others (Hakemulder, 2000). Current pressing questions concern the processes involved in a mode of reading that impacts self- and social perceptions (Fialho, 2012) and the outcomes of this mode of reading (Koopman & Hakemulder, 2015).

It has been suggested that Theory of Mind (Zunshine, 2006) and narrative empathy (Keen, 2013) are process components of reading literary fiction (e.g., Bal & Veltkamp, 2013; Kidd & Castano, 2013). Zunshine (2006) defined Theory of Mind as "our ability to explain people's behavior in terms of their thoughts, feelings, beliefs, and desires" (p. 6). Keen (2013) defined narrative empathy as "the sharing of feeling and perspective-taking induced by reading [...] narratives of another's situation and condition" (n.p., see also Keen, 2007). Without being exhaustive, we outline some of the growing empirical support for these claims, based on both quantitative and qualitative studies (see Hakemulder, Fialho, & Bal, 2016; Mar & Oatley, 2008, for overviews).

Experimental studies have shed light on the role of Theory of Mind and empathy in reading literary fiction. In five online experiments among adults of around 34 years old, Kidd and Castano (2013) found that reading literary fiction

enhanced readers' affective and cognitive Theory of Mind, which they defined, from a neuropsychological rather than a literary perspective, as "the ability to detect and understand others' emotions" and "[the] inference and representation of others' beliefs and intentions" (p. 377). They contended that the (temporary) effects on Theory of Mind were specific to what they selected as literary fiction, and not to popular fiction. This distinction, which may seem somewhat artificial, was based on work of Barthes, Bakhtin and Bruner (Kidd & Castano, 2013, p. 377-378) and was operationalized by selecting novels awarded literary prizes, against texts that did not receive awards. Despite the fact that the effects found by Kidd and Castano were not all confirmed in replication studies (Dijkstra, Verkoeijen, Van Kuijk, Chow, Bakker, & Zwaan, 2015; Liu & Want, 2015), their study opened up the possibilities for further investigations on the role of Theory of Mind in literary fiction reading.

In a study among adult readers of about 25 years old, Bal and Velkamp (2013) focused on empathy. Although, unlike Keen (2013), they defined the concept from a psychological perspective, considering it "the cognitive and intellectual ability to recognize the emotions of other persons and to emotionally respond to other persons" (p. 2), they expected that this broader notion of empathy could be related to the literary reading process. They found that empathy increased a week after fiction reading, but only in case of high transportation (cf. Green & Brock, 2000).

Through interviews with 16-year-old readers, Charlton, Pette and Burbaum (2004) found that reading (literary) fiction made them compare their own lives to story situations and experience empathic engagements with characters' feelings. Moreover, adolescents regarded reading fiction as a way of understanding others' experiences, through which they might feel connected to others or see new possibilities for their own lives (Rothbauer, 2011). Finally, in a survey study among Dutch elementary and secondary school students, aged 9 to 17, Van der Bolt (2000) found that more than half of 3025 participants reported having experienced sympathy (52%) and empathy (57%) when reading fiction, and that these affective reading responses seemed to occur more often among avid readers.

More specifically, Theory of Mind and narrative empathy may be important components of reading experiences that can be characterized as self-modifying. First, in 19-year-old psychology students' think-aloud responses to a literary story, Fialho (2012) identified changes in their positioning towards the story and the main character, thus changing the way they perceived themselves as the reading unfolded. Findings revealed two types of self-modifying

reading experiences: the first characterized by empathic engagement with the story setting and blurred boundaries between oneself and the narrator or characters, suggesting personal identification, and the second characterized by sympathetic engagement with characters and blurred boundaries between oneself and others, suggesting more general identification. Second, Sikora, Kuiken and Miall (2010) showed through questionnaire responses of 24-year-old literature students that self-modifying feelings and a deepened self-perception were evoked if readers, who lost a loved one, experienced aesthetic emotions when encountering stylistically striking passages in a poem. Third, Richardson and Eccles (2007) found in their interview study among adolescents that voluntary reading sometimes made them explore their possible selves (Markus & Nurius, 1986): it made them think about who they are, who they would like to be and who they do not want to become.

Theory of Mind and narrative empathy may also play important roles in modifying readers' social perceptions through literary reading. For example, Hakemulder (2000) found that identification with a story character who represented an outgroup positively affected readers' beliefs about this particular outgroup. These results were supported in a second study, in which Hakemulder found similar outcomes after readers were purposefully instructed to actively take the role of the character, by means of an empathy-building instruction. Similarly, Johnson (2013) found that adult readers of fiction who were more transported into a story reported less negative outgroup perceptions. Based on these findings as well as other studies, Koopman and Hakemulder (2015) proposed that such during-reading empathy and role-taking may result in real-world empathy as an after effect.

Clearly, there is support for the notion that reading literary fiction may "change the reader for the better" (Hakemulder et al., 2016, p. 19): it may enhance self-examination or self-reflection as well as social understandings. This may have important implications for the potential of using literature in specific environments, for example, teachers' professional development (Kooy, 2006), people's professional behavior in general (Bal, Butterman, & Bakker, 2011), and domain-specific education in literary studies (e.g., Fialho, Zyngier & Miall, 2011; Fialho, Miall & Zyngier, 2012) as well as the secondary literature classroom that forms the context for the present study.

### 1.3 *Adolescents in the Literature Classroom*

Studies about the effects of literary reading on self-insights and social insights have rarely focused specifically on adolescents (e.g., Bal & Veltkamp, 2013;

Fialho, 2012; Johnson, 2013; Kidd & Castano, 2013; Sikora et al., 2010). Yet, research has shown that reading processes of novice adolescent readers differ from reading processes of more experienced adult readers. Expert readers have, for example, a large variety of reading strategies at their disposal and are capable of regarding literary texts from multiple perspectives, as well as of analyzing them on various levels. Novice readers of literature, on the other hand, often mainly focus on the events in a story and regard a story from a single perspective (e.g., Andringa, 1990, 1995; Earthman, 1992; Peskin, 1998; Zeitz, 1994; for overviews also see Goldman, McCarthy, & Burkett, 2015; Hanauer, 1999). These differences in literary reading processes suggest that the effects of literary reading on readers' selves and their social perceptions could differ as well: literary reading may affect expert adult readers and novice adolescent readers in different ways. The question arises, then, what is known about the impact of literary reading on *adolescents'* personal and social insights in the context of the *literature classroom*. Two terms are purposefully italicized here.

First, little is known about the extent to which adolescents gain personal and social insights from literary reading. Based on his interpretation of exemplary responses, Appleyard (1991) contended that adolescents may draw connections between stories, themselves and the social world. They may "experience involvement with the story and identification with the character" (p. 100), but often with more than one character, which fits their growing ability to take various social perspectives. A character's identity may not resemble their own identity, but rather represent the kind of person they would like to become (cf. "the desired self"; Markus & Nurius, 1986). Furthermore, Appleyard found that adolescents "talk about the realism of the story" (p. 100), for instance about how accurately a story reflects their own experiences or how easily similar situations can be imagined. Finally, according to Appleyard, adolescents stated that "a good story makes them think" (p. 100): they may reflect on characters' motives and emotions and compare these to their own, or they may think about the meaning of a story.

Second, like Appleyard's (1991) research, most studies about adolescents' personal and social insights as a result of reading fiction are conducted in the context of voluntary leisure reading, like the studies mentioned earlier (Charlton et al., 2004; Richardson & Eccles, 2007; Rothbauer, 2011). Research conducted in the context of the literature classroom among adolescents, on the other hand, has often been confined to reading engagement and/or analytical skills in terms of interpreting literary texts (Eva-Wood, 2004; Janssen, Braaksma, & Couzijn, 2009; Peskin, 1998; 2007; Pieper & Wieser, 2012; Tengberg, Olin-

Scheller, & Lindholm, 2015). Although many of these studies incorporated the perspective of what readers bring to the text – in particular Eva-Wood (2004), who developed a think-and-feel-aloud pedagogy – the perspective of what literary fiction might mean to adolescent readers and what they can take away from it for their (social) lives, remained largely unexplored.

Possibly, the literature classroom might hinder the gain of personal and social insights, if adolescents feel resistance toward literary fiction reading. In the Netherlands, this seems to be the case for at least part of the students. Van Schooten (2005) found that Dutch students' attitude toward literary reading became more negative in higher grades of secondary education: as the years of literature education increased, these adolescents seemed to enjoy literary reading less and less. Witte (2008) drew similar conclusions. Potentially, such negative attitudes impede with gaining personal and social insights from literary reading in the classroom.

On the other hand, the literature classroom might foster personal and social insights if the social nature of literary reading (Bloome & Green, 1984/2002) is acknowledged and valued, for instance if students are encouraged to bring personal experiences to texts and to share their reading responses (Applebee, Langer, Nystrand, & Gamoran, 2003; Beach et al., 2011). Sharing their responses may broaden their minds (Holland & Schwartz, 1975), which may, by extension, broaden their perceptions of themselves and others. In conclusion, then, the teacher's perspective or approach taken in the literature classroom may, in part, determine whether students take away personal and social insights from literary reading and accompanying learning activities.

#### 1.4 *Teachers' Approaches to Literature Education*

Which perspective is taken and which learning activities are emphasized in literature classrooms, is largely based on what teachers value. A conversational inquiry amongst Australian and Dutch literature teachers showed that teachers had their own opinions about what literature education should look like and that they developed their praxis accordingly (Van de Ven & Doecke, 2011). Approaches to literature teaching, thus, may vary strongly among literature teachers (cf. Applebee, 1994) and are therefore not easy to describe or define. In the Netherlands, this may perhaps be even more pronounced than in other countries, given the curricular freedom in this context (Janssen & Rijlaarsdam, 2007; Witte et al., 2012).

In the Dutch context, Janssen (1998) noted that teachers often taught eclectically, although they tended to emphasize one of four approaches (p. 311):

- An author-oriented, literary history approach (cultural development);
- A text-oriented, structural analysis approach (aesthetic awareness);
- A context-oriented, sociological approach (social awareness);
- A reader-oriented, text-experiencing approach (personal development).

Based on survey data of former students about the goals and text use in their literature education, Verboord and Van Rees (2003) brought these four approaches back to a subject matter- or culture-oriented approach (combining the author-oriented and text-oriented approach) and a student-oriented approach (combining the context-oriented and reader-oriented approach).

Other distinctions have been made as well, such as an interpretative versus an experiential approach (Fialho et al., 2011; 2012), while other scholars have argued for bridging the gap by attending both to meaning of literary texts as well as bringing personal experiences to these texts (Wilhelm, 2007), for example by creating "a social community that supports learning literature" (Beach et al., 2011, p. 8). Similarly, Van de Ven and Doecke (2011) noted that the teachers in their study connected the interpretation of literary texts to "the need to negotiate the social relationships that comprise any classroom" (p. 219).

Characterizing Dutch teachers' approaches to literature education, thus, may be challenging. Yet, attempting to do so is relevant, because previous studies into Dutch literature education have shown that different teacher approaches generate different learning outcomes in students (Janssen, 1998; Janssen & Rijlaarsdam, 1996). In the context of the current study, teachers' emphasis on particular goals or pedagogical activities may be related to whether their students perceive any impact of literature education on their personal and social insights.

### 1.5 *Aims and Research Questions*

The aim of the present study is to find out whether and to what extent a sample of Dutch upper secondary school students reports to experience gains in personal and social insights through literature education, and to explore whether this is in any way related to their teacher's classroom practice. Our research questions are:

1. Do students report learning experiences concerning their self-perceptions and social perceptions attributed to literature education, and if so, which kinds of learning experiences are these?
2. Are the (kinds of) learning experiences of students within a class related to their teacher's approach to literature education?

## 2 METHOD

### 2.1 *Participants*

We approached upper secondary school literature teachers in our network by e-mail and published a call for participation on a Facebook page for Dutch language teachers. 21 teachers were willing to participate. Of those, we selected 13 teachers (seven females), based on gender and location. Four teachers were from schools located in smaller towns, six in larger towns and three in major cities, distributed over the mid-western and southern part of the Netherlands. Their age ranged from 23 to 63 years ( $M = 42.2$ ,  $SD = 11.6$ ). Their experience as literature teachers in upper secondary education varied considerably ( $M = 11.6$  years,  $SD = 11.1$ , range 2-40), as did the percentage of time they allocated to the domain of literature within their Dutch language lessons ( $M = 29.3\%$ ,  $SD = 15.1\%$ , range 10-60).

Each teacher selected, in consultation with the first author, one class to participate in the study. We strived for variation in school levels and grades and included therefore grades 10 and 11 at higher general education level, and grades 10-12 at pre-university level. In addition, teachers only selected those classes which time schedules allowed for participation. We asked parents for consent for their child's participation in the study. None of them withheld their consent. In total, 297 students of 13 classes participated, of which 49% were females. Participants' age ranged from 14 to 20 years ( $M = 16.4$ ;  $SD = 1.1$ ). The number of students in a class ranged from 18 to 26 ( $M = 22.9$ ;  $SD = 2.7$ ).

### 2.2 *Instruments*

*Learner report.* To collect responses about what students thought they learned about themselves and other people through literature education, we asked them to complete a learner report: a semi-open reflective writing assignment, originally developed by De Groot (1980a; 1980b). A learner report allows for the explication of learning experiences that remain implicit in other measures and was found to be a valid and reliable instrument in previous research (Janssen, 1998; Janssen & Rijlaarsdam, 1996; Van der Kamp, 1980; Van Kesteren, 1993). It has been designed to collect fundamental learning outcomes that cannot be assessed by tests or demonstrations, but that may have impact on students and therefore can be reported.

A learner report contains open questions, to be answered in free writing. Introductory phrases are provided, for instance, "I learned that (I)..." and "I now know that it is not true that (I)...". In this study, these prompts were only in-

tended to support the participants to word their experiences: using them was not compulsory and rephrasing was allowed. Apart from these introductory phrases, we provided examples of learning experiences from another school subject (History) and both oral and written instructions before starting the task.

A first version of the learner report was tested in a pilot ( $N = 93$ ) and revised based on students' comments. The final learner report contained four sections, as Table 2.1 shows. Each section started on a new page to provide enough space for writing. Students were randomly assigned to one of four different orders to avoid test effects as threats to internal validity. The learner report was preceded by a page with background questions, and two questions to trigger students' memories of specific reading experiences: "What was the last book you read for school?" and "Which book, of all books you've ever read, do you remember best?" At the final page of the booklet, we asked students to indicate on a Likert scale how difficult it had been to complete the learner report, ranging from 1 (*very easy*) to 5 (*very difficult*).

*Author Recognition Test (ART)*. To determine students' familiarity with fiction, we used an adapted version of the Author Recognition Test (Stanovich & West, 1989). An ART consists of a list of author's names and foils, on which participants indicate real authors' names. The number of correctly recognized names is an indicator of one's familiarity with fiction: the ART was shown to have predictive validity for real-world reading, while avoiding socially desirable answering to questions about reading frequency and motivation (Mar, Oatley, Hirsh, dela Paz, & Peterson, 2006; Rain & Mar, 2014). Because one may be familiar with other authors than those included in the test, scores are relative: if Philip scores 15 whereas Emma scores 4, Philip is assumed to be *more* familiar with fiction than Emma.

We adjusted the ART for Dutch adolescents. It included eighty names: forty authors and forty foils. We strived for variation in author's names based on their original language (either Dutch or Flemish, or translated from other languages), main intended audience (either adults or youth), gender and canonicity. The instruction read: 'Encircle those names which you know for sure are authors' names. Some of these people are not authors, so do not guess.' Test scores ranged from minus 40 to plus 40.

Table 2.1. Overview of learner report sections with questions, general writing instruction and introductory phrases

Section	1	2	3	4
Prompt	Try to recall the literature lessons you attended. What did you learn about others during the literature lessons?	Think about the books you read for Dutch class. Try to remember what they were about. What did you learn about others by reading them?	Try to recall the literature lessons you attended. What did you learn about yourself during the literature lessons?	Think about the books you read for Dutch class. Try to remember what they were about. What did you learn about yourself by reading them?
Instruction	Write as many sentences as you can. Use and/or change the introductory phrases, if it helps you.			
Introductory phrases	By literature lessons, I learned ... In literature class, I discovered ... Because of what we do in literature class, I noticed ... Because of literature lessons, I know it is not true that ... What I know now by the lessons, is ...	By reading ... I learned that ... When reading stories, I noticed ... Because of the book ... I know it is not always true that ... By reading, I experienced it is not true that ... What I know now I've read ... is ...	Similar to section 1.	Similar to section 2.
	In literature class, I experienced ...	By reading ... I discovered that ...		

We piloted the ART along with the learner report. We substituted author's names that were not recognized at all and created four final versions, in which only the order of names varied. The ART was provided halfway the task booklet, so that deeper, reflective thinking about learning experiences was alternated with a cognitively less demanding recognition task.

*Teachers' Approaches to Literature Education Questionnaire (TALE-Q).* To gain insight into their approaches to literature teaching, we asked teachers to complete the newly constructed Teachers' Approaches to Literature Education Questionnaire (TALE-Q) online. Since various distinctions have been made in approaches to literature teaching (Fialho et al., 2011; 2012; Janssen, 1998; Verboord & Van Rees, 2003, see Section 1.4, p. 15), the TALE-Q emphasized what these studies have in common, thereby relying on a continuum with a so-called analytical-interpretative approach at one end, and a personal-experiential approach at the other end.

We considered the analytical-interpretative approach as more text-oriented, focused mainly on literary analysis and interpretation of texts substantiated by literary elements, and regarding works of literature as cultural, aesthetical, canonical objects. We considered the personal-experiential approach as more reader-oriented, mainly focused on personal experiences of literary texts, on drawing connections to the real, outer-textual world, and on sharing those literary experiences and resulting insights with others. Emphasis is given to discussing what a text means to its readers and to exploring in which ways readers may connect themselves to texts. Importantly, we did not regard the two approaches as being in dichotomy: teachers were not expected not employ "a" personal-experiential approach or "an" analytical-interpretative approach. Rather, for various aspects of their teaching, they might lean toward each of the extremes of the continuum, or they might be somewhere in between both approaches.

The TALE-Q consisted of six scales. Scale 1 (goals) indicated what teachers aim for in literature lessons. The five other scales represented a selection of pedagogical practices to pursue these goals, namely (2) text use, (3) focus on literary analysis or literary reading experiences, (4) degree of classroom interaction, (5) allowance for student autonomy, and (6) emphasis on ambiguity of literature. Each bipolar item consisted of two statements, with a Likert scale in between, which ranged from 1 (the most analytical-interpretative option) to 5 (the most personal-experiential option). Table 2.2 shows examples of these statements. In half of the items, statements were mirrored. After answering several background questions, teachers indicated for each item which of both

statements best fitted their literature lessons to the class that participated in this study.

Pilot participants ( $N = 17$ ) found a first version of the questionnaire, containing 51 items, too long. Twelve items were removed if this did not negatively affect the reliability of the then existing scales. Once the actual participants ( $N = 13$ ) completed the TALE-Q, two scales had a low reliability. From scale 2 (text use;  $\alpha = .43$ ), three items were removed, which resulted in  $\alpha = .74$ . From scale 7 ("literature as object vs. as tool for yourself and the world";  $\alpha = .29$ ), two items were moved to scale 1, resulting in  $\alpha = .77$ , and one item to scale 3, resulting in  $\alpha = .86$ . Scale 7 then consisted of only two items and was deleted from the data, so that six scales remained (see Table 2.2).

To determine whether TALE-Q scales might represent possible underlying factors, principal components analysis with Varimax rotation was used. Due to the small sample size, the Kaiser-Meyer-Olkin value was rather low (.33), yet the result for Bartlett's Test of Sphericity was significant ( $p < .001$ ). Analysis revealed three components which together accounted for 81% of the total variance. First, Attitude toward literary reading (eigenvalue 1.68, 28% of variance) included scale 3 and 6 and was considered to represent the extent to which a teacher promotes literary reading as a personal, ambiguous experience. Second, Students' roles in classroom processes (eigenvalue 1.62, 27% of variance) included scale 4 and 5 and was considered to represent teachers' self-reported practices in terms of student interaction and student autonomy in their literature classroom. Third, Intended teaching content (eigenvalue 1.55, 26% of variance) included scale 1 and 2 and was considered to represent what teachers intend to achieve in literature education and which types of literary texts they apply to achieve this (for factor loadings, see Table 2.3).

Table 2.2. *TALE-Q scales, item examples per scale, and Cronbach's Alpha*

Scale	Approach characterizations			Example of item: two statements	Items [deleted]	$\alpha$
	Analytical-interpretative	Personal-experiential	Analytical-interpretative			
Goals	More subject matter-oriented, analytical-interpretative goals; building textual and contextual knowledge.	More student-oriented, personal-experiential goals; developing literary taste, experiencing personal value of literature.	"My goal is that students learn to place literary texts in a literary-historical context."	"My goal is that students learn to relate literary texts to their own experiences."	7 [-]	.77
Text use	Use of texts more removed from students' world; more traditional and canonical texts.	Use of texts that concur more with students' worlds; texts with recognizable elements.	"I mainly pay attention to historical literary texts."	"I mainly pay attention to contemporary texts."	4 [3]	.74
Analysis vs. experience	A more cognitive, analytical focus on structural and literary characteristics; importance of analysis, understanding, interpretation.	A more personal focus on experiencing literary texts; importance of feeling, experience, recognition, imagination.	"In class, I mainly discuss the content and structural characteristics of a literary text."	"In class, I mainly discuss how a literary text affects me and what the text means to me."	7 [-]	.86
Classroom interaction	Less interaction: more use of frontal teaching and explanations by teacher; mainly individual learning, e.g. through writing.	More interaction: classroom conversations and student talk; mainly shared learning, e.g. through group discussions.	"Students mainly work individually."	"Students mainly work together in duos or small groups."	6 [-]	.89
Student autonomy	More control and decision making by the teacher; less freedom of choice for students.	Less control and decision making by teacher; students have a say in choice of topics or literary works.	"As the teacher, I usually choose which literary texts we use in class."	"Students usually have a say in which literary texts we use in class."	5 [-]	.79
Ambiguity of literature	Regarding and propagating literature as having an assumed definite meaning; importance of finding the author's message.	Regarding and propagating literature as ambiguous; accepting multiple interpretations; indeterminacy of what the author means.	"By reading a literary work, the reader can discover what the author wants to say with it."	"Every reader has an own interpretation of what a literary work could mean."	5[-]	.91

Table 2.3. Results of principal components analysis of TALE-Q scales

Scale	Factor loadings		
	Aspect 1	Aspect 2	Aspect 3
	Attitude toward literary reading	Students' roles in classroom processes	Intended teaching content
Goals	.35	.32	.75
Text use			.83
Analysis vs. experience	.90		
Classroom interaction		.74	-.43
Student autonomy		.88	
Ambiguity of literature	.83	-.31	

Note. Factor loadings  $\leq .30$  and  $\leq -.30$  are not displayed.

### 2.3 Procedures

Teachers completed the TALE-Q before students wrote their learner reports. Three trained research assistants visited the classes during one 50- or 60-minute lesson to collect the ART and learner reports. Apart from incidental student talk, the assistants reported no irregularities. The procedure consisted of three steps:

1. Guided by a PowerPoint presentation, the assistants discussed with the students their literature education so far for about five minutes, to focus their attention on the topic.
2. The assistants explained the learner report and ART task, guided by a PowerPoint slide, and announced that a gift card would be put up for raffle among the participants.
3. The task booklet with instructions and examples, background questions, learner report questions and the ART was distributed in class. Students were asked to work individually. After reading the instruction and examples, about 30-35 minutes were left to complete the task.

### 2.4 Data Analysis

*Segmentation.* Learner reports were typed out verbatim and imported into Atlas.ti. Because a sentence in a learner report could contain more than one learning experience (e.g., "I learned that I like to read novels *but also that I do not like to analyze them*", emphasis MS), we set rules for identifying single

learning experiences. To assess reliability, an independent rater segmented a randomly selected set of learner reports ( $n = 3$ , containing 33 learning experiences). Agreement was substantial (90%, Cohen's  $\kappa = .79$ ). In Atlas.ti, the first author then distinguished and numbered all individual learning experiences. In case of doubt, she discussed sentences with the second researcher.

*Coding system.* We developed a coding system using LEX-NAP procedures (Lexical Basis for Numerically Aided Phenomenology; Fialho, 2012) for bottom-up qualitative analysis. LEX-NAP is grounded in phenomenology (analysis of experiences) and linguistics. Aiming for intersubjectivity among individuals, LEX-NAP analytical procedures involve seeking for 'the knowledge shared by a community of experiencing subjects' (Fialho, 2012, p. 103), thus formulating the essence of those experiences, as well for intersubjectivity among the researchers who study those experiences, measured by inter-rater reliability tests.

We applied LEX-NAP procedures to a randomly selected sample of 65 learner reports (five of each class). The learning experiences in these reports were subject to comparative analysis. Once we found two or more experiences similar in both content and form, we formulated a paraphrase to capture their essence (see Table 2.4), which functioned as a code.

Table 2.4. Paraphrase code based on intersubjective responses by students

Code	< <u>through / by X</u> I <b>learned that</b> (modified) <u>books / literature / texts</u> (can) emotionally affect <b>me</b> >
Responses	By <u>reading stories</u> , I <b>noticed that</b> <u>some of them</u> can be quite touching. In 'Tonio', <u>the story in particular</u> affected <b>me</b> a lot. <u>These books</u> stir up many emotions <b>in me</b> . By <u>reading stories</u> , I [ <b>unreadable</b> ] that <u>they</u> can quite affect <b>you</b> . I <b>also noticed that</b> <u>books</u> can affect <b>me</b> rather fast on the emotional level.

Note. Underlining, bold and italics represent similarities in form.

Comparative analysis of the sample resulted in 168 paraphrase codes. To address the reliability of the coding system, a second rater coded 10% of the learner reports ( $n = 28$ , containing 283 learning experiences). Agreement was sufficient (75%, Cohen's  $\kappa = .75$ ).

When coding the remaining learner reports, we found experiences that were not present in the first sample of 65, which we labeled as 'other responses'. In a second round, we checked for additional shared experiences, which we

then assigned an own paraphrase code. We also merged paraphrase codes that highly resembled each other. Eventually, 114 paraphrase codes were left. Within these, three kinds of learning experiences emerged, on which we elaborate below:

1. *Learning experiences about oneself and others: content and evaluations.* Ten categories of content learning experiences (i.e., referring to what is learned about oneself or others, including both real and fictional people) were formed here, which are presented in Table 2.5, with examples of students' responses. These categories included the most relevant answers to the questions asked in the learner report and represented almost one third (64%) of all responses. Additionally, a number of responses about oneself and others were evaluations, which referred to one's own and others' (assumed) attitudes toward encounters with literature. We distinguished five categories (see Table 2.6).
2. *Learning experiences about literature and its context.* In addition to responses concerning the self and others, we distinguished six categories about literature and its context (see Table 2.7).
3. *Irrelevant or incomprehensible learning experiences.* Some learning experiences were irrelevant or incomprehensible, for example: 'I learned that you should not knock on some people's doors, because someone might chase after you with a hatchet', which seemed a generalization of a book passage to the real world, and 'We also went far back in time. Thus in the culture of that time', without further specification of what this student learned.

Table 2.5. Categories and examples of learning experiences about self and others

1. Learning about oneself and others as language learners and comprehensive readers.	
- No subcategories	"I learned that I am good at formulating and spelling." "I learned to determine / improve my reading level."
2. Learning about oneself and others as literary readers.	
- One's own literary skills and literary reading habits	"When reading stories, I noticed that I find it very hard to understand a story at once."
- Assumed knowledge of how others evaluate and think about literature	"In literature class, I discovered that opinions about books highly differ per individual."
- Affective literary responses (comparisons, recognition, identification, narrative empathy and sympathy)	"Through the literature, I also often compare the story to my own life." "By reading books, I learned that I recognize little or nothing from my own life." "Yet, these books ( <i>Tirza</i> , <i>Het diner</i> ) also often evoked a bit of compassion in me." "When reading, you feel like being the main character. I experienced her experiences."
3. Learning about oneself and others as persons: personal characterizations.	
- Insights in own personality, personal development and character traits	"I learned that my personality is rather a bit unique. I see things differently than others."
- Comparisons of other people or literary characters to each other (differences and similarities)	"I also discovered that the variety in world views and norms and values of different people is much larger than I initially thought."
- Notions of sympathy for and empathy with others, and understanding their emotions	"I read many books about war. These books evoke your sympathy and compassion for something that is not that long ago."
- Understandings of how others can be, what they can go through, that they can change, and why they make certain choices	"Because of <i>Maar buiten is het feest</i> , I know that it is not all roses out there. That people can have many problems, even though you can't see it from the outside."
- Understandings of others as impressionable by other people, culture, historical time, religion or what they go through in life (passively <i>being</i> influenced; compare to 6)	"By reading <i>De helaasheid der dingen</i> , I learned that your childhood has much influence on the person you become, so on how you think, how you react and how social you are."
4. Learning about oneself and others as thinkers.	
- The experience that literature makes the student think	"A book can make you think differently about something, like your view on the world."
- Understandings of how people think and form or change opinions	"I also learned something about how other people can think about the world."

---

5. Learning about others in former times and other cultures.

- Understandings of how people thought and behaved in former times, and differences to today "In the literature lessons, I learned that people in former times didn't have such a good life. They had to work hard for little money."
  - Understandings of what other cultures are like and differences to the student's own culture "I learned about people in other cultures / everyday surroundings, because in books, suddenly you come very close to different people with other philosophies of life."
- 

6. Learning about others as agents: behavioral characterizations.

- Understandings of how people can behave or react (general / specific) "By reading *Sonny boy*, I learned people can be very unfriendly, only because you have a different skin color."
  - Awareness of consequences, strangeness and unpredictability of behavior "Because of *Bezonen rood*, I know that characters not always react in a predictable way."
  - Awareness of influential behavior (actively *having* influence; compare to 3) "What I know because of literature lessons, is that people in difficult situations can influence each other very much. This can have positive or negative consequences."
  - Evaluations of behavior "In literature lessons, I learned that I don't like it when people are being selfish."
- 

7. Learning about one's future (un)desired selves: responses on what one does (not) want to be, do, become or achieve.

- "Through books, I know what kind of man and father I would like to be in the future."  
 "By reading *Gelukkige slaven*, I discovered I never want to be blinded by money."
- 

8. Learning lessons for life: responses on awareness of the importance, relativity or complexity of (social) phenomena, often formulated using the inclusive "you" and "we". Experiences are close to the self, but are (implicitly) extended to "everyone".

- "In literature lessons, I experienced that we should be very glad we're so well off."  
 "I learned through literature lessons that every culture should be equal."  
 "When reading stories, I noticed you should not judge people too quickly."  
 "By reading *Grip*, I learned you actually don't have a grip on anything (life, time)."  
 "I noticed that by reading books, I realize better that life is not as easy as it seems."
- 

9. One's negations of learning: statements of having learned little or nothing about self or others.

- "I did not learn anything about myself in literature lessons."  
 "Further, I didn't learn much about others."
- 

10. Other responses about the self and other people, which do not fit in the categories above.

- "I noticed that I do believe there's something, but I don't think it is God."  
 "We did a bit of literary history, Charles novels. If I had to learn something, it is that actually everyone wants to be a lifesaver."
-

*Table 2.6. Categories and examples of evaluative responses*

Type of evaluation	Example
1. Positive evaluations of reading and (specific types or works of) literature	"During literature lessons, I learned that I find literature quite interesting."
2. Negative evaluations of reading and (specific types or works of) literature	"By reading <i>Erik of het kleine insectenboek</i> , I discovered that I find it nonsense."
3. Positive evaluations of (specific activities during) literature lessons	"During literature lessons (and just before), I am always very cheerful and I always try to help people who don't understand it."
4. Negative evaluations of (specific activities during) literature lessons	"Many of the lessons were not very useful for me, in particular the lessons about how you can best read and understand a text."
5. Unclear or other evaluations (e.g., of school subjects, movies)	"I think it's nicer to watch a movie because you quickly miss out on details and therefore you have to watch the movie again."

*Table 2.7. Categories and examples of learning experiences about literature and context*

Topic	Example
1. Literary history	"By reading <i>Mariken van Nimwegen</i> [a medieval text], I learned that literature in former times was very different."
2. Depth and profundity of literature	"Because of the lessons, I know it's not true that a book is written just like that, but that there's often much more to it."
3. Range and scope of literature	"Through literature lessons, I learned that there are many different types of literature."
4. Function or deploy of literature	"People don't write a book just to entertain other people, but also to make something clear or to spread a message."
5. Authors	"I learned in the literature lessons that many authors insult other authors because of a different literary vision."

### 2.5 *Data Description of Student Data*

In total, students reported 2,997 learning experiences ( $M = 10.1$ ,  $SD = 4.4$ ), ranging from 2 to 29 experiences in a learner report. On a 5-point scale, students indicated that completing the learner report was moderately difficult ( $M = 3.4$ ,  $SD = 1.0$ ). There was no significant correlation between number of learning experiences and perceived difficulty, which indicated that students who found the task more difficult did not write down fewer learning experiences, and vice versa.

On the Author Recognition Test, students scored on average 8.9 ( $SD = 5.3$ , range -2 to 26). Pre-university students scored significantly higher ( $M = 10.7$ ,  $SD = 5.4$ ) than higher general education students ( $M = 6.0$ ,  $SD = 3.5$ ):  $t(295) = 8.34$ ,  $p < .001$ . However, pre-university students were on average older and more experienced with literature education than higher general education students. Yet, even when we only took 10<sup>th</sup> and 11<sup>th</sup> grade into account, pre-university students ( $n = 113$ ) scored on average 10.4 ( $SD = 5.8$ ) and were still more familiar with fiction than higher general education students:  $t(226) = 6.85$ ,  $p < .001$ . School levels thus differed in familiarity with fiction, which indicated the validity of the ART.

### 2.6 *Description of Teacher Data*

We used TALE-Q factor scores to map similarities and differences among teachers on three aspects of their teaching: 1) their attitude toward literary reading; 2) their self-reported practice in terms of students' roles in classroom processes; and 3) their intentions with regard to teaching content. We compared teachers' mean scores on each aspect of teaching to the overall mean of that aspect. In Table 2.8, an A represents a more analytical-interpretative approach to an aspect of teaching, indicated by scores of at least one standard deviation below the mean, whereas a P indicates the opposite: the teacher scored at least one standard deviation above the mean, suggesting a more personal-experiential approach to that aspect of teaching. In this way, Table 2.8 demonstrates the overlap as well as the variety among teachers: some of them centered around the mean on each of the three aspects (i.e., Eva, Peter, and Tess), whereas the others showed unique patterns of A's, P's, and neutral aspects of teaching.

Table 2.8. Indications of approach to three aspects of teaching, based on teachers' TALE-Q scores

Teacher	Attitude toward literary reading		Students' roles in classroom processes		Intended teaching content	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
	3.19	.63	2.74	.71	3.21	.48
Alice				A		
Anna	A				A	
Daniel					A	
Jeff	A		A			
Eva						
Peter						
Tess						
Karl				P	P	
Margaret					P	
Martin				P		
Milly				P		
Olaf	P					
Rebecca	P					

Note. A = analytical-interpretative, P = personal-experiential approach to aspect; empty cell = neutral approach to aspect. Teachers' names are pseudonyms.

### 3 RESULTS

#### 3.1 *Frequencies of Learning Experiences*

Table 2.9 demonstrates which types and categories of learning experiences students reported most frequently. Calculation of the percentages in the middle column of Table 2.9 was based on the total number of reported experiences. The middle column shows that responses about oneself and others were most frequent. Most of these learning experiences were formulated in terms of personal and social content: students reported *what* they learned about themselves or others. Responses about oneself and others as persons were most frequent, covering 19.6% of all responses. For example, one student reported: 'I discovered that the variety in world views and norms and values of different people is much larger than I initially thought'.

Table 2.9. Frequencies of categories of learning experiences

Category of learning experiences	% of all learning experiences	% of students responding at least once
1. Oneself and others	86.6	100.0
Content learning experiences: learning about...	63.8	99.3
Oneself and others as persons	19.6	80.1
Oneself and others as literary readers	13.3	64.6
Others as agents	7.0	48.8
One's lessons for life	5.2	33.3
Other responses on roles of selves	4.3	32.3
Negations of learning about self or others	4.3	26.3
Others in former times and other cultures	4.1	29.2
Oneself and others as thinkers	2.6	19.2
Oneself and others as learners or comprehending readers	1.8	14.5
One's future (un)desired selves	1.6	12.5
Evaluations	23.0	81.8
Positive evaluations of reading and literature	9.2	51.2
Negative evaluations of reading and literature	6.2	38.4
Other evaluative responses	3.2	25.9
Negative evaluations of literature lessons	2.6	19.9
Positive evaluations of literature lessons	1.8	15.2
2. Literature and context	11.0	55.2
Authors	2.7	18.9
Other responses on literature and context	2.3	20.2
Literary history	2.3	16.5
Function / deploy of literature	1.5	11.4
Profundity of literature	1.2	10.8
Scope of literature	1.1	8.8
3. Other responses	2.2	19.5

*Note.* Middle column = percentages of total number of learning experiences ( $N = 2997$ ; categories ordered from highest to lowest frequency); right column = percentages of number of students ( $N = 297$ ) responding at least once.

Second, responses were frequently about oneself and others as literary readers (13.3%), such as "When reading, you feel like being the main character. I experienced her experiences". This category was followed by responses about others' behavior (7.0%), for example, "By reading *Sonny boy*, I learned that people can be very unfriendly, only because you have a different skin color", and lessons for life (5.2%), with responses like "I noticed that by reading books, I realize better that life is not as easy as it seems", and "I learned through literature

lessons that every culture should be equal." Learning about others in former times and other cultures, about oneself and others as thinkers and about one's future (un)desired selves proved to be rather small categories. Almost a quarter of all responses were evaluations.

Percentages of the total number of responses could be influenced by a few students who repeatedly reported similar responses, either to emphasize them or because they could not think of other experiences. We therefore calculated for each category the percentage of students who reported at least one learning experience in that category. The right column of Table 2.9 shows that nearly all students reported at least one content learning experience about themselves or others (99%), most of them about oneself and others as persons, oneself and others as literary readers, and others as agents. Although the percentages in the middle and right column differ, the order of proportions is similar. Thus, the effect of students repeating themselves in their learner reports seemed negligible, and therefore we deemed it justifiable to perform further analyses on percentages of all learning experiences.

For only three categories of content learning experiences, there was a weak linear association between students' familiarity with fiction and the number of learning experiences they reported: for learning about oneself and others as persons ( $r = .21, p < .001$ ), for learning about others as agents ( $r = .18, p = .002$ ) and for learning about one's future (un)desired selves ( $r = .21, p < .001$ ). These few weak correlations suggested that reporting more learning experiences about oneself and others not necessarily went hand in hand with a higher familiarity with fiction.

### 3.2 *Relations between Aspects of Teaching and Learning Experiences*

Our second research question concerned whether students' learning experiences are related to their teacher's approaches to literature education. The total number of learning experiences about self and others differed significantly among classes ( $F(12, 284) = 5.15, p < .001$ ). To further explore whether the differences among classes could be explained by the approaches taken by teachers, we first calculated correlations between teaching aspects and students' content learning experiences (i.e., the mean frequency for each class on each category). For Students' roles in classroom processes, we found rather strong significant correlations with two categories of learning experiences: learning about oneself and others as literary readers ( $n = 13, r = .62, p = .025$ ) and learning about others as agents ( $n = 13, r = .56, p = .045$ ). Students of teachers who said to stimulate interaction and classroom autonomy thus reported more

learning experiences in these two categories. For Teachers' attitude toward literary reading and Intended teaching content we found no significant correlations with categories of learning experiences.

Table 2.8 showed, for each teaching aspect, which teachers reported a more pronounced analytical-interpretative practice (A's) or a more pronounced personal-experiential practice (P's). This enabled us to examine whether differences in students' content learning experiences could be found in relation to approaches to teaching aspects. For each teaching aspect, we grouped students of A-teachers and students of P-teachers. We assessed differences in frequencies of content learning experiences through independent sample *t*-tests. Results are presented below, accompanied by comparisons of ART scores and evaluations of literary reading and literature lessons, because these student variables may offer additional explanations for potential differences in learning experiences.

*Attitude toward literary reading.* Four teachers deviated at least one standard deviation from the mean: Jeff and Anna reported a distinct analytical-interpretative approach to this teaching aspect, whereas Olaf and Rebecca reported a distinct personal-experiential approach. We found that students taught by Jeff and Anna reported fewer learning experiences about themselves and others as literary readers ( $n = 36$ ,  $M = 1.0$ ,  $SD = 1.0$ ) than students taught by Olaf and Rebecca ( $n = 45$ ,  $M = 1.8$ ,  $SD = 1.7$ ). Levene's test indicated unequal variances ( $F = 13.37$ ,  $p < .001$ ), so degrees of freedom were adjusted from 79 to 75. The difference was statistically significant:  $t(75) = 2.58$ ,  $p = .012$ .

This result could not be attributed to Olaf's and Rebecca's students being more familiar with fiction than Jeff's and Anna's students: we found no significant difference on ART scores between both groups. However, Olaf's and Rebecca's students reported more positive evaluations of reading literature ( $M = 0.8$ ,  $SD = 1.0$ ) than Anna's and Jeff's students ( $M = 0.4$ ,  $SD = 0.6$ ). Due to unequal variances ( $F = 8.96$ ,  $p = .004$ ), degrees of freedom were adjusted from 79 to 75. The difference was statistically significant ( $t(75) = 2.48$ ,  $p = .015$ ). For students' negative evaluations of literary reading and their evaluations of literature lessons, we found no significant differences between both groups.

*Students' roles in classroom processes.* Five teachers deviated at least one standard deviation from the mean: Alice and Jeff reported more teacher-led classroom processes, whereas Karl, Martin, and Milly reported more student interaction and student autonomy in their literature classrooms (see Table 2.8). For six categories, we found significant differences in frequencies of learning experiences, as presented in Table 2.10. Results showed that the students of

Karl, Martin, and Milly, who favor interaction and student autonomy in their lessons, reported more learning experiences in these six categories than Alice's and Jeff's students.

However, these differences may additionally be explained by students' familiarity with fiction: Karl's, Martin's, and Milly's students ( $n = 74$ ) obtained higher ART scores ( $M = 8.6$ ,  $SD = 4.9$ ) than Alice's and Jeff's students ( $n = 39$ ,  $M = 5.2$ ,  $SD = 3.1$ ). Due to unequal variances ( $F = 6.47$ ,  $p = .012$ ), degrees of freedom were adjusted from 111 to 108. We found a statistically significant difference:  $t(108) = 4.51$ ,  $p < .001$ . Similarly, students of Karl, Martin, and Milly wrote more positive evaluations of literary reading ( $M = 1.4$ ,  $SD = 1.4$ ) than Alice's and Jeff's students ( $M = 0.6$ ,  $SD = 1.0$ ). Again, Levene's test indicated unequal variances ( $F = 5.28$ ,  $p = .023$ ), so degrees of freedom were adjusted from 111 to 103. The difference was statistically significant:  $t(103) = 3.28$ ,  $p = .001$ . We found no significant differences for negative evaluations of literary reading and for evaluations of literature lessons.

*Table 2.10. Differences in frequencies of learning experiences between students taught by Alice and Jeff and students taught by Karl, Martin, and Milly*

Category	Mean [SD] per group		<i>t</i> -test result
	Alice, Jeff ( $n = 39$ )	Karl, Martin, Milly ( $n = 74$ )	
Oneself and others as literary readers	1.1 [1.2]	2.0 [1.9]	$t(107) = 3.14$ , $p = .002$
Oneself and others as persons	1.3 [1.2]	2.0 [1.6]	$t(111) = 2.36$ , $p = .020$
Others in former times, other cultures	0.2 [0.5]	0.6 [1.1]	$t(108) = 2.56$ , $p = .012$
Others as agents	0.5 [0.7]	0.9 [1.1]	$t(109) = 2.55$ , $p = .012$
One's future (un)desired selves	0.1 [0.2]	0.2 [.5]	$t(111) = 2.15$ , $p = .034$
One's self-extended lessons for life	0.4 [0.9]	0.9[1.3]	$t(98) = 2.17$ , $p = .032$

*Note.* For nearly all tests, Levene's test indicated unequal variances; degrees of freedom adjusted from 111 for 'literary readers' ( $F = 11.85$ ,  $p = .001$ ;  $df$  to 107), 'former times' ( $F = 10.30$ ,  $p = .002$ ;  $df$  to 108), 'agents' ( $F = 5.18$ ,  $p = .025$ ;  $df$  to 108), 'future selves' ( $F = 14.49$ ,  $p < .001$ ;  $df$  remained 111), 'lessons for life' ( $F = 13.83$ ,  $p < .001$ ;  $df$  to 98).

*Intended teaching content.* As Table 2.8 shows, four teachers deviated at least one standard deviation from the mean: Daniel and Anna reported to strive for more analytical-interpretative goals and to choose texts that concur with these goals, whereas Karl and Margaret reported more personal-experiential aims and text choices. We compared the frequencies of their students' learning ex-

periences, but for none of the categories a significant difference between both groups was found.

#### 4 DISCUSSION

Does reading literary fiction teach us something about who we are, what other people can be like and how we can relate to them and their (inner) lives? According to previous research, this appears to be the case for (avid) adult readers, but these studies have told us little about how *adolescents* perceive personal and social learning in the institutional context of the literature classroom. We therefore asked students to reflect in writing on the learning experiences about themselves and others they gained from their literature education. The present study shows that students report learning experiences concerning self- and social perceptions which they attribute to (aspects of) their literature education: reading literature for school, and attending literature lessons.

Our second research question concerns relations between students' learning experiences and their teachers' approaches to literature education. Based on the results, we ought to be careful to conclude the existence of such relations. The extent to which teachers report to promote interaction and student autonomy in their literature classroom (the aspect Students' roles in classroom processes) is positively correlated with frequencies of students' learning experiences about themselves and others as literary readers, and about others as agents. For the two other aspects, Teachers' attitude toward literary reading and Intended teaching content, there are no linear relations with students' learning experiences. When focusing on teachers who most pronouncedly favor an analytical-interpretative approach or a personal-experiential approach to an aspect of teaching, significant differences in some categories of their students' learning experiences occur: for Teachers' attitude toward literary reading, there is a difference in one category and for Students' roles in classroom processes, there are differences in six categories. In all cases, students taught by teachers with a personal-experiential approach have reported more learning experiences that demonstrate insights into themselves and others than students of teachers with an analytical-interpretative approach to an aspect of teaching. However, students' familiarity with fiction and their evaluations of literary reading may also offer an explanation for more frequent reports of learning about themselves and others.

#### 4.1 *Perceived Personal and Social Learning Through Literature Education*

Most often, amongst the students in this study, the perceived impact of literature education on their self- and social perceptions takes the form of personal characterizations: as subcategories and examples of student responses show (see Table 2.5, category 3, p. 25), students report that literature education offers them insights into their own personality, understandings of how people are, and notions of empathy or sympathy with others. These latter statements are of particular interest: they incorporate learning to imagine what it is like to be in the shoes of real other human beings and to feel for them. For example, when asked what was learned about others through reading books for school, a 17 year-old girl writes: "I learned to look a bit further than my own surroundings. I can better imagine what it is like to be in other people's situations (at least try to do so)". It seems that literary stories or literature lessons may function as mediators between students and the social world around them. Even though Kidd and Castano's (2013) finding that reading literary fiction improves Theory of Mind was not replicated by other researchers, the present study suggests that students, through reading literary fiction, sometimes do imagine what it is like to experience someone else's situations and feelings. Whether literary reading may result in improvement or growth of Theory of Mind or related empathic skills remains a question to be answered, in particular for literary reading in a secondary school context.

Furthermore, perceived personal and social learning through literature education often refers to literary reading itself: students report to have learned about their own literary reading habits and about different views on literature, but also report experiences of narrative empathy and sympathy (Keen, 2007; Koopman & Hakemulder, 2015). These empathic and sympathetic engagements with literary characters and situations are in line with the results of previous studies of adolescents' reader responses (Appleyard, 1991; Charlton et al., 2004; Rothbauer, 2011; Van der Bolt, 2000).

Next, almost half of the participants reports perceived learning about people's behavior. Most of these learning experiences are formulated in a descriptive manner: students say to have learned *how* people can behave and sometimes evaluate other people's behavior. They note, for instance, that people can be dishonest, cruel or unpredictable, or that they can behave deceptively. These descriptions and evaluations of behavior seem rather basic learning experiences, compared to, for example, self-other comparisons, empathic or sympathetic engagements with characters or real people, or lessons for life,

which appear to reflect more profound insights than observing how people can act and expressing an opinion about that.

Self-extended lessons for life form another large category: they are reported by one third of the students. For instance, students report that they have come to appreciate their own lives, to have learned valuable lessons about social (in)equality or prejudices, or to have realized they cannot control everything in life. These lessons for life seem typical adolescent responses. As noted by Applebee (1978), readers of about sixteen years old consider literature as "one of many statements of how life might be understood" (p. 125) and reflect on whether it changed their own views on life. In other words, adolescent readers may transfer ideas and experiences from within a book to their own, outer-textual world. They reflect on what kind of impact a work of literature may have had on their own life, as this student statement demonstrates: "Because of reading *Die zomer*, I know that you don't always have to follow your friends, but that you should choose your own way".

In addition, some categories of learning experiences emerge relatively infrequently in students' learner reports, such as understandings of people in former times and other cultures. These learning experiences seem to suggest that literature education may sometimes evoke historical empathy in students, which refers to "the ability to see and judge the past in its own terms by trying to understand the mentality, frames of reference, beliefs, values, intentions and actions of historical agents [...]" (Yilmaz, 2007, p. 331; see also Davis, Yeager, & Foster, 2001) as well as change their attitudes toward multiculturalism and "the other" (Hakemulder, 2000; Johnson, 2013). Furthermore, students sometimes report to experience literature as a catalyst for thinking (cf. Appleyard, 1991). Whereas these responses are in this study rather infrequent, there is growing attention for the potential of literature education as a stimulus for critical thinking (e.g., Bean & Moni, 2003; Faust, 2000; Koek, Janssen, Hakemulder, & Rijlaarsdam, 2016). Finally, in line with findings by Richardson and Eccles (2007), we assumed that the participants in this study would also report to experience explorations of their possible future selves (Markus & Nurius, 1986). However, such experiences are reported by only 12.5% of the students, which is a rather small percentage compared to many other content learning experiences. An explanation may be that Richardson and Eccles asked for responses about "voluntary reading" (p. 342), while in our study students reflect on compulsory reading and literature lessons. It is conceivable that literature education that allows for (more) freedom in book choice might result in more students reflecting on what kind of person they would or would not like to become.

In addition to learning experiences concerning self-perceptions and social perceptions, part of the learning experiences concerns literature and its context, which concurs with responses found by Janssen (1998). She distinguished, for instance, responses on "literary works", and "literary-historical backgrounds", including biographical knowledge of authors. The same holds true for evaluative responses the students in this study report: Janssen found similar responses, and labeled them as positive and negative attitudes. For this type of learning, thus, the results of this study also concur with results of previous research.

#### 4.2 *Aspects of Teaching, Perceived Learning Outcomes and Student Variables*

We have shown that in their literature education, students perceive a variety of personal and social learning outcomes. It has been challenging to relate this perceived impact of literature education to the multidimensional array of teaching practices in the literature classroom. For the purposes of this study, we made use of a continuum, ranging from an analytical-interpretative to a personal-experiential extreme. We are aware that this opposition may seem rather obsolete in light of recent research (e.g., Beach et al., 2011; Van de Ven & Doecke, 2011; Wilhelm, 2007), but this does not mean that literature teachers are able to bridge the gap as well (Hillocks, 2011). Although teachers may be increasingly aware of the potential benefits of valuing their students' experiences, in practice a divergence between more engaging, experiential instruction and more formalist, knowledge-oriented instruction in the literature classroom appears to remain (Ives, 2012; Malo-Juvera, 2014), which may partly be due to the convenience of text analysis for testing and evaluation. Of both extremes described here, neither is "best": they represent *different* approaches, here seen as a continuum, which are likely to have different outcomes as well (cf. Janssen & Rijlaarsdam, 1996).

TALE-Q scores have shown that the participating teachers, to a certain extent, differ in which perspective they emphasize: some teachers report neutral scores, while others report to use either analytical-interpretative approaches or personal-experiential approaches. Yet, if they do so, they are not fully consistent: there is no teacher who has reported to distinctively enact either of both approaches across all three aspects of teaching.

When we compare perceived personal and social learning outcomes of students whose teachers report either an analytical-interpretative or a personal-experiential approach to an aspect of their teaching, some significant differences occur, in particular with respect to students' roles in classroom processes. In classes of teachers who report to attend more to classroom interaction

about literary experiences, to students' personal preferences and to (a certain degree of) freedom of choice, students report more personal and social learning experiences than students in classes of teachers who report to provide more teacher-led literature instruction with limited interaction and less freedom of choice.

This result concurs with previous studies that emphasize the importance of agency and freedom of choice in the literature classroom (e.g., Beach et al., 2011; Lenters, 2006; Smith & Wilhelm, 2002) as well as the potential benefits of interaction and dialogue (e.g., Applebee et al., 2003; Eeds & Wells, 1989; Galda & Beach, 2001; Nystrand, 1997). The current study expands these insights to perceived personal and social learning in the literature classroom, although we should bear in mind that no causal relations can be detected.

In addition, student variables seem to play a role. Results indicate that the students of teachers who report more interactive and autonomous student roles not only write down more personal and social learning experiences, but also evaluate literary reading more positively and that they are more familiar with fiction than students of teachers who report to apply more teacher-led instruction. There may be complex mediating or moderating relations among students' perceived learning outcomes, their evaluations of literary reading and literature lessons, their familiarity with fiction, and their teachers' approaches. While being beyond the scope of this paper, further study of such relations would certainly add to the existing body of research on literature education.

### 4.3 *Limitations*

This study is based on students' self-reports. Even though the learner report is a validated instrument, we should take into account that completing the task can be challenging. We are quite sure that not all internal learning experiences are explicated via this instrument: it may just be the tip of the iceberg. Yet, the small share of irrelevant or incomprehensible responses strengthens the validity of the learner report: in general, participants were inclined and able to report relevant learning experiences, and did not find the task very difficult.

A second limitation is that the conclusions of this study are mostly based on frequencies of students' learning experiences, which does not tell us anything about how an experience affects an *individual* student. By relying on frequencies, valuable and insightful learning experiences with much significance for individual students may remain under the radar. In future research, asking students to reflect on their learner report in a face-to-face conversation with the researcher may reveal which experiences have been most meaningful to them.

Furthermore, the representativeness of the sample can be questioned. The student sample size ( $N = 297$ ) is satisfactory and has enabled us to achieve variety in terms of school level and grades, which is reflected in the range of ART scores: we have not merely included avid readers. However, as a consequence of keeping the amount of qualitative student data manageable, the teacher sample is relatively small ( $N = 13$ ). Although there are some indications that teachers' approaches to students' roles in classroom processes are related to personal and social learning outcomes as perceived by students, results are by no means generalizable to all Dutch literature teachers.

Moreover, teacher data are also based on self-reports. For a small sample, the TALE-Q has shown to measure approaches to literature education on several teaching aspects in a reliable way. The fact that different learning outcomes are found for different approaches to an aspect of teaching, as reported by the teachers themselves, strengthens the validity of the TALE-Q. However, we have not confirmed through observations whether teachers actually enact what they claim to do, due to limited time and resources. Classroom observations would be of added value in future studies: now that approaches to teaching aspects appear to be related to students' gain of personal and social insights, it is recommendable to study what teachers and students actually do in literature classes.

With regard to the TALE-Q, data analysis is subject to a statistical limitation. We have applied multiple  $t$ -tests, which was appropriate in terms of the nature of the data set. However, multiple testing may increase the chance of finding significant results. Next to students' familiarity with fiction and their evaluations of literary reading as additional explanations, then, multiple testing is yet another reason to cautiously consider the results on relations between approaches to teaching aspects and students' learning experiences.

A final limitation is that we cannot draw any conclusions on growth or development in learning about oneself and others through literature education. We therefore suggest that small-sample longitudinal cohort studies (cf. Witte, 2008) can provide more insight into this kind of development.

#### *4.4 Pedagogical Implications*

The Dutch Institute for Curriculum Development (2015) contends that literature education has the capacity to expand students' social and cultural horizons and to stimulate their empathic capabilities. In this respect, our finding that Dutch adolescents report to attribute valuable lessons about themselves, other people and the world around them to literature education, is encouraging: the ob-

jectives mentioned by the Institute for Curriculum Development appear to be met, mainly in terms of the larger categories we have described above. However, some other categories occur less often, like reflecting on oneself and others as thinkers and on the future self, which suggests there is room for widening the scope of personally and socially relevant literature education. If literature education has the capacity to impact adolescents' sense of self and their social understandings, we might want to stimulate these insights, either by focusing even more on experiences students may already be familiar with or on experiences which seem rather infrequent.

This study supports the notion that, for adolescents, literary fiction can be a vehicle for gaining insights in themselves and others, even when reading takes place in secondary schools. The weak correlations between students' reported content learning experiences about themselves and others and their familiarity with fiction suggests that adolescents do not necessarily need to be 'bookworms' to learn about themselves and others: students who are less familiar with fiction have also reported valuable learning experiences. Reading and discussing a relevant and thought-provoking story may have impact on students' personal and social insights, even for students with a low reading motivation. Moreover, if students are allowed to make their own choices, this may facilitate their learning even more (Lenters, 2006; Smith & Wilhelm, 2002), as is exemplified by this response of an 18 year-old girl: 'I learned not much from the books I read for school, neither about myself nor about others, partly because I often don't like those books'. Responses like these also concur with the role of transportation into a story (cf. Green & Brock, 2000). If students do not feel transported into the texts they read and discuss, learning about themselves and others may be less likely to occur.

The main contribution of the present study is the description of the kinds of personal and social insights students report to take away from literature education. To our knowledge, this study is the first to systematically analyze a large number of student responses on this particular topic. The study further shows that these learning experiences occur in a complex, multidimensional context. Relations between approaches to certain aspects of teaching and perceived personal and social learning in the literature classroom cannot be pinpointed easily, and apply only to the Dutch educational context. Yet, there are some indications that teachers' practices with regard to the role of the student in the literature classroom are related to what students report to have learned about themselves and others. This implies that teachers and educational designers *can* engage in educational approaches that aim at enhancing these kinds of personal and social insights.

## CHAPTER 3

### GAINING INSIGHT INTO HUMAN NATURE

#### A REVIEW OF LITERATURE CLASSROOM INTERVENTION STUDIES \*

In this review, we explore whether and how literature education may foster adolescent students' insight into human nature. A systematic search of five databases was complemented with citation tracking, hand searches, and expert consultation. We included 13 experimental and quasi-experimental intervention studies. Methodological quality and quality-of-intervention descriptions were assessed. Analysis of empirical support for expected intervention effects indicated that, under certain conditions, literature education may foster students' insight into human nature. One intervention affected students' insight into themselves, two affected their understanding of fictional others, and six affected their understanding of, views on, or intended behavior toward real-world others. Subsequent analysis of interventions with full or partial empirical support yielded instructional design principles on (1) text selection; (2) activating, annotating, and reflecting on personal life and reading experiences in writing activities; and (3) verbally sharing these experiences with others in exploratory dialogues. Limitations and implications for future studies are discussed.

#### 1 INTRODUCTION

Contemporary society finds itself in turbulent times. In an era of globalization, migration, and polarization, there appears to be a need for people to be able to reflect on their own nature as well as on the nature of others. This may include their own position in the world, their views of themselves, and their perceptions of and relationships with other people. Numerous approaches to teaching and learning indicate that education may play a pivotal role in helping young people to gain insight into human nature, such as social and emotional

---

\* Schrijvers, M., Janssen, T., Fialho, O. & Rijlaarsdam, G. (2018). *Gaining insight into human nature: A review of literature classroom intervention studies*. *Review of Educational Research*. Advance online publication. doi:10.3102/0034654318812914

learning (Elias et al., 1997), citizenship education (Derricott, 2014), moral and character education (Nucci, Krettenauer, & Narváez, 2014), and values education (Halstead & Taylor, 1996). Despite having their own backgrounds, frames of reference, and terminology, these approaches all indicate that teachers of any subject may attend to “human nature”.

In this paper, we address how students’ insight into human nature may be fostered in the context of reading and responding to fictional or literary texts, more specifically in secondary school literature classrooms. We focus in particular on adolescent students in the upper grades, between 15 to 18 years old. There seems to be considerable interest for the role literature teaching may play in fostering students’ insight into human nature. In Belgium, for instance, one of the examination requirements for the domain of literature is that students are “able to put their reading experiences in a societal context” (Curriculum, 2017, n.p.). In the Netherlands, the Dutch Institute for Curriculum Development stated that literature education has “an important value for developing citizenship, [for instance by] broadening social and cultural horizons and developing empathic capabilities” (2015, p. 15). In the U.S., the U.K., and Australia, language teachers’ associations appear to value learning about human nature, stating that “students who read literature learn that literary texts are often relevant to their own lives [and offer] perspectives which may contrast and conflict with their own experiences” (National Council of Teachers of English, 2012, p. 21) and suggesting that students “need opportunities to explore how their own perspectives, values and assumptions compare with those in the texts they encounter” (UKLA, 2016, p. 5). Such statements, however, remain at the rhetoric level. Little is known about whether students’ insight into human nature may indeed be fostered via encounters with fictional and literary texts, and, if so, which instructional approaches may be particularly suitable for achieving this objective.

Therefore, we report on a synthesis of empirical intervention studies that have investigated whether and how insight into human nature – one’s own nature and the nature of both fictional and real-world others – may be fostered in literature education. This paper is positioned on the crossroads of educational studies and research into reading fiction and literature. First, we explain how we understand and relate the key concepts used in this paper: fictional texts, literary texts, and insight into human nature. We then discuss which characteristics of instructional approaches to literature teaching seem theoretically promising for fostering students’ insight into human nature, which will lead up to the research questions we aim to answer in this study.

### 1.1 *Fictional and Literary Texts*

The literature classroom includes all sorts of written fictional texts, for instance, stories, novels, poetry, drama, song texts, and so forth. The term *fictional text* refers to texts in which characters are “not presented as existing in the real world” (Koopman, 2016, p. 106). The world these characters inhabit does not exist in reality but may function as a safe abstraction and simulation of the real world (Mar & Oatley, 2008). The borders between fictional and *nonfictional* texts are blurred, for example, when fictionalized stories are based on true events or refer to real-world places. In addition, although fictional texts and stories are often used synonymously, a fictional text is not necessarily a narrative. Poems are considered nonnarrative; however, they are oftentimes fictional texts.

Defining *literary texts* is more challenging, even though the closely related term *literature* is frequently used by language teachers and in curriculum documents. Notions of literariness are partly based on readers’ perceptions and influenced by social conventions (e.g., Bourdieu, 1996; Ellis, 1974). Researchers have also argued that text features may distinguish literary texts from nonliterary ones, such as more complex characters (Mar & Oatley, 2008) or language use that deviates from conventional language use and is thus perceived as “striking” (Miall & Kuiken, 1999; Mukařovský, 1976). Literariness, then, is not a fixed, universal concept. What is considered complex and unconventional depends on the reader and the (historical) context. When we refer in this Introductory section to other studies, we adopt the terms originally used by the authors (e.g., fictional, nonfictional, literary texts), even though these terms may not always be well defined.

### 1.2 *Insight into Human Nature*

In this section, we relate reading fictional and literary texts to gaining insight into human nature and attempt to characterize this kind of learning in the literature classroom. Research in developmental psychology indicates that learning about human nature is pivotal during adolescence – a stage of life during which humans develop their sense of self and their social and moral identity (Hart, Atkins, & Ford, 1998). Adolescents become increasingly aware of their inner self and realize that others have an inner self as well, thereby acknowledging the relativity of their own perceptions, thoughts, and feelings (Kohlberg & Gilligan, 1971). Thus, adolescents come to understand “the mutuality of perspectives [which] includes a view of both self and other as complex psychologi-

cal systems of values, beliefs, attitudes etcetera" (Selman, 1975, p. 40). As such, adolescents' insights, attitudes, beliefs, responses, and behavior related to themselves and others are constantly evolving. Reading fiction and literature, it seems, may play a role in fostering adolescents' insight into human nature.

*The role of fictional and literary reading.* Insight into human nature may come about during and after reading fictional and literary texts, as Koopman and Hakemulder (2015) indicate in a synthesis of research. They distinguish between empathy (i.e., insight into the nature of others) and reflection (i.e., insight into one's own nature). The various definitions of empathy to which Koopman and Hakemulder refer all relate to the metaphor of putting oneself in the shoes of others, either cognitively or emotionally. Research included in Koopman and Hakemulder's synthesis indicated, for example, that reading fictional texts was found to enhance adult readers' scores on various empathy measures and the accuracy of their perceptions of social interactions (e.g., Bal & Veltkamp, 2013; Mar, Oatley, Hirsh, dela Paz, & Peterson, 2006). Moreover, reading fictional and literary texts may positively affect readers' outgroup perceptions (i.e., people's attitudes toward groups of human beings other than the group with which they identify) and may also be closely related to feeling empathy for others (Hakemulder, 2000; Johnson, 2013; Kaufman & Libby, 2012).

Reflection is defined by Koopman and Hakemulder (2015) as thinking about "oneself, often in relation to others and/or society" (p. 82). This definition resonates with Nussbaum (1995), who suggests that literary reading may help us to examine ourselves and to think about how we relate to others, ethical issues, and life in general. Like empathy, reflection may be the result of reading. For example, a qualitative study by Richardson and Eccles (2007) indicated that voluntary reading of both fictional and nonfictional texts made adolescent readers consider their future selves; that is, it made them reflect on what kind of human being they would or would not like to become. Two other studies found that adolescents who talked about reading fictional texts perceived connections with their own lives and how they understood others. German adolescents indicated that reading fictional texts made them compare their own lives to story situations and thus experience empathetic engagements with characters' feelings (Charlton, Pette, & Burbaum, 2004). Canadian teenagers regarded reading fictional texts as a way of understanding others' experiences, which made them feel connected to others and offered new possibilities for their own lives (Rothbauer, 2011).

Other studies have investigated which concepts and processes may underlie these effects on insight into human nature and which relationships may exist among them (Fialho, 2012; Fialho, Zyngier, & Burke, 2016; Hakemulder, Fialho,

& Bal, 2016). These studies suggest, for instance, that experiencing changes in notions of self and others evolve in particular when readers respond to passages that are highly metaphoric or stylistically deviant from conventional language use (Miall & Kuiken, 1994). They further indicate close relationships between readers' perceptions of themselves and their perceptions of others, which supports the view that there is, conceptually, no other without the self (Zahavi, 2014).

Although underlying processes and relationships must be further elucidated, ample research indicates that reading fictional and literary texts may result in gaining insight into human nature. Therefore, literature teaching may be a suitable domain to foster this kind of learning. A descriptive study in the Netherlands supports this claim by showing that students in upper secondary education reported learning experiences ascribed to their literature education that concerned their own nature and the nature of others, such as learning about their own and others' personalities, feeling empathy for others, and considering their future selves (Schrijvers, Janssen, Fialho, & Rijlaarsdam, 2016).

Therefore, we deliberately position the concept of "insight into human nature" in the context of the literature classroom. In the concept of human nature, self and others are inextricably linked (Zahavi, 2014). Furthermore, we assume that gaining insight into human nature may take place in the transactional space of meaning-making, which is created and inhabited by the reader and the text (Rosenblatt, 1938/1968). Reading a fictional or literary text may evoke, for example, readers' feelings, memories, or associations with other human beings. Finally, we expect that gaining insight into human nature may also take place in the space beyond the text, in which textual elements, such as fictional characters or events, function as representations of the real world (Mar & Oatley, 2008). For example, readers may compare characters, events, concepts, and themes in stories to real-life situations.

### 1.3 *Instructional Approaches to Literature Teaching*

Because we investigate not only whether literature education may foster students' insight into human nature but also *how* this may come about, we address instructional approaches to teaching literature. Relevant aspects are text selection, tasks, the role of the teacher, and stances taken toward texts.

*Text selection.* In literature curricula, numerous choices must be made in terms of what kind of texts are read in the classroom (genre), what these texts are

about (theme), and to which extent they can be characterized as fictional or literary texts (literariness).

In terms of genre, studies involving adult readers have included poetry (e.g., Sikora, Kuiken, & Miall, 2011) and prose (e.g., Fialho, 2012); therefore, we expect that both genres may be used in literature classroom interventions that focus on fostering students' insight into human nature. In addition, genres that are assumed to appeal to adolescents might also be used in interventions, such as young adult literature, song texts, or graphic novels.

Text themes are given little attention in overviews of previous studies with adult readers. Rather than theme, fictionality and literariness appeared to be determinative text characteristics for researchers to select texts for their studies (e.g., as shown in the synthesis of research by Koopman & Hakemulder, 2015). Yet, in some studies, researchers chose texts that were thematically related to the aim of the study. For example, Hakemulder (2000) investigated the effects of a story about a woman who opposes traditional gender roles in a fundamentalist Islamic country on participants' perceptions of such women. However, in other studies, multiple texts with various themes are used (e.g., Kidd & Castano, 2013). Thus, across previous studies, considerations of theme were scarce and inconsistent. Therefore, we can only speculate on the text themes that may be used in literature classroom interventions.

Finally, literariness has been identified as an explanatory factor to glean the impact of fiction on readers' sense of self and social perceptions (Hakemulder et al., 2016; Kuiken & Miall, 1994). To our knowledge, however, no studies featuring adolescent participants have been conducted that compare the effects of reading fictional texts to reading literary texts on adolescents' insight into human nature. We assume that reading fictional texts is the default reading activity in this context, that these fictional texts may be literary to a greater or lesser extent, and that the perception of their literariness may be mutually divergent among students, teachers, and researchers. Therefore, we will analyze which definition of literariness is used in intervention studies, if any.

*Tasks.* Literature teachers have many different types of tasks at their disposal to ask their students to respond to fictional and literary texts. These may include, for example, dialogues and discussions, formal and creative writing, performing drama, and creating visual and audiovisual arts (e.g., Beach, Appleman, Hynds, & Wilhelm, 2011; Galda & Beach, 2001; Soter, Wilkinson, Connors, Murphy, & Shen, 2010; Wilhelm, 2016). Such tasks are intended to prompt students' learning activities (Rijlaarsdam, Janssen, Rietdijk, & Van Weijen, 2017). Learning activities are cognitive or affective activities in relation to, in the case of this paper, fictional or literary texts. For example, a particular kind of writing task may

prompt various learning activities, such as evaluating a story and explaining the evaluation – did it appeal to students, why or why not?

A single task can prompt a variety of learning activities, which may not always be explicated in curriculum descriptions, lesson plans, or intervention studies. Moreover, tasks operate at an intentional level, but it is not always clear whether an intended learning activity actually had an effect on students. To avoid speculation about cognitive and affective processes that may or may not have taken place in students' minds, we focus our analysis of intervention studies on the tasks as described.

*Teachers' roles.* The role of literature teachers goes beyond selecting appropriate texts and designing tasks; they are also a determinative factor in the classroom discourse. We understand "discourse" here as any response to fictional and literary texts that may be expressed, regardless the type of task: by talking, writing, performing drama, creating art, and so forth. We would expect that if literature teaching focuses on fostering students' insight into human nature, teachers would aim or at least allow for *dialogic* discourse (Nystrand, 1997). This discourse opposes monologic interactions in classrooms (i.e., when the teacher controls what is being expressed), which are often preceded by an "initiation–response–evaluation" pattern: The teacher poses a question, a student responds (e.g., by speaking or writing), and the teacher evaluates this response (Applebee, Langer, Nystrand, & Gamoran, 2003). It implies that the teacher knows the "correct" answer to a question or the "true" interpretation of a fictional or literary text. If monologic discourse prevails, there is little opportunity for sharing individual experiences that are evoked by a text, including experiences and insights related to self and others.

Dialogic discourse in the literature classroom, on the other hand, allows students and the teacher to explore and share ideas amongst each other, which is what we would expect if the intention were to foster students' insight into human nature. The focus would be on stimulating students' divergent thinking and on developing and deepening ideas and experiences. The authenticity of questions and responses, expressed by both students and the teacher, is key for learning in dialogic literature classrooms. Students must be seen as capable partners in open conversations, response writing, and creative performances, which may be achieved by working in small groups. The teacher's task is to guide and to support students in their responses. They may do so, for example, by offering prompts for exploratory talk (Mercer & Dawes, 2008), thinking aloud during reading to model their own authentic responses to texts, or making explicit their own difficulties in response writing (Wilhelm,

2016). Specifically for facilitating small-group talk in response to texts, Wei, Murphy, and Firetto (2018) identified a taxonomy of subtle discourse moves that teachers may use, such as backchanneling (indicating that they are listening to their students), clarifying (inviting a student to provide a clearer response), prompting (helping students to construct an elaborate response, by asking for reasons and evidence from a text), and summarizing (giving an overview of what has been said during the talk). In our analysis, we will examine to which extent teachers' roles are addressed and explicated – as Wei et al. (2018) note, information about how teachers interact with their students may not always be given, or may remain implicit.

*Stance toward texts.* Rosenblatt (1938/1968; 1978/1994) proposed the concept 'stance toward texts' In her transactional theory of reading, she distinguishes between an efferent stance, where the reader primarily attends to information to be acquired, solutions to problems, or actions to be carried out after reading; and an aesthetic stance, where the reader focuses on what he or she is living through while reading a particular text.

Murphy and colleagues redefined the two categories and added a third one, resulting in three stances: an efferent, expressive, and critical-analytical stance (Murphy, Wilkinson, Soter, Hennessey, & Alexander, 2009; Soter et al., 2008). Similar to Rosenblatt's original definition, Murphy et al. define an *efferent stance* as text-focused responses, where reading is primarily meant to acquire and retrieve particular information. The authors refine the aesthetic stance to an *expressive stance*, in which the focus is on affective responses to the text or on the reader's spontaneous, emotional connection to all aspects of the experience with the text. The later-developed *critical-analytical stance* is meant to lend prominence to interrogating or querying the text while searching for underlying arguments, assumptions, world-views, or beliefs.

We expect that taking an expressive and a critical-analytical stance may be apt for fostering students' insight into human nature in the literature classroom. Research with adult participants suggests that insight into oneself and others is preceded by various kinds of spontaneous responses and emotional connections to a literary text, such as imagery of setting and characters, identification with characters, and feeling sympathy for characters (Fialho, 2012; 2018) – experiences that may be addressed in particular if an expressive stance toward texts is taken. Furthermore, insight into human nature may include students' understanding of complex social situations. Because fictional and literary texts may function as simulations of the real social world (Mar & Oatley, 2008), they may evoke numerous relevant questions, for example: Why do characters in this text think or behave in a particular way? What in the text may explain

their thoughts and behavior? Can we classify their behavior as being “right” or “wrong”? Would people in real life behave similarly? Addressing such issues requires students to investigate and reason about worldviews, beliefs, assumptions, and so forth, that are represented in a fictional or literary text. Therefore, we assume that students’ insight into human nature may be fostered by taking a critical-analytical stance in the literature classroom.

#### 1.4 *Aims and Research Questions*

Literature education is a promising domain for fostering students’ insight into human nature. However, an overview of research that investigates whether literature teaching – in particular, what kind of instructional approach – is effective in doing so is not yet available. Therefore, we systematically reviewed empirical intervention studies in the literature classroom that deliberately focus on fostering adolescents’ insight into human nature.

Our purpose is to identify a set of instructional design principles that can be used in future educational design studies. Design principles can be considered parameters that increase the likelihood of a particular objective being achieved – in this case, developing insight into human nature. Therefore, design principles are often formulated as heuristic “if/then” statements (Van den Akker, 1999; Reigeluth, 1999), such as: *If we want to increase the probability of achieving purpose X, then we are best advised to give a curriculum or instructional approach the characteristics A, B, and C.* As such, they are of a prescriptive nature and are design-oriented rather than learning-oriented: they “relate to creating learning environments and products rather than describing how learners acquire knowledge and skills from these environments and products” (Merrill, 2002, p. 44). In contrast with making ad hoc and random decisions, using design principles is likely to result in better-informed teaching and learning – particularly if we understand why previous interventions were effective.

To ultimately arrive at a set of instructional design principles, we developed the following research questions:

1. What effects did researchers expect to achieve by implementing the interventions included in this review?
2. To what extent was empirical support provided for these expected effects?
3. What instructional approaches were implemented in interventions with empirical support, in terms of (a) text genres, themes, and literariness; (b) tasks that were applied; (c) teachers’ roles; and (d) stances toward texts?

## 2 METHOD

We conducted a review of experimental and quasi-experimental intervention studies. First, we determined search terms and inclusion and exclusion criteria. Second, we searched five databases, complemented by citation tracking, hand searches, and expert consultation. Next, we set criteria for quality assessment and assessed the included studies accordingly.

### 2.1 *Search Terms*

We combined four clusters of search terms. The first cluster contained terms regarding the intended student population, that is, adolescents in upper secondary education. Examples of search terms included *secondary education*, *high school*, *secondary school students*, and *grade (9 to 12)*. The second cluster focused on texts used in literature teaching. We understood the literature classroom as a first-language context in secondary schools. We broadly defined the term *texts* to include fictional texts, such as novels, short stories, plays, and poems written by published authors. Search terms included *literature*, *novels*, *poetry*, *fiction*, *literature education*, *literature lessons*, and *literature class*. The third and fourth clusters included terms related to human nature. We distinguished between insight into one's own nature and into the nature of others and social relationships. Examples of search terms in the third cluster included *personality*, *self-understanding*, *possible selves*, *emotional experience*, *identity*, and *self-concept*. In the fourth cluster, search terms included *cultural awareness*, *empathy*, *social attitudes*, *social experience*, *prejudice*, *equality*, and *outgroup*. Appendix A contains all terms in search syntaxes and shows how we applied variations of search terms and adjacent terms.

### 2.2 *Inclusion and Exclusion Criteria*

Exploratory searches indicated that the search would return a large number of results but that few would meet the inclusion and exclusion criteria outlined below. We therefore did not limit the search to a particular time period, but we did confine it to peer-reviewed journals and edited book chapters. This restriction would exclude records that were not likely to meet the inclusion criteria, such as teacher association documents and government reports, and would provide an initial quality threshold via the academic system of peer-review and editing. We acknowledge that this decision may have biased the search because intervention studies with nonsignificant statistical results may not always pass peer-review procedures.

We further selected intervention studies based on five inclusion and exclusion criteria regarding type of intervention, hypotheses and measures, classroom context, research design, and publication language (see Table 3.1).

*Table 3.1. Inclusion and exclusion criteria*

	Included	Excluded
1. Intervention type	Interventions in the literature classroom context, e.g., literature projects, reading instructions, in-class reading, assigned literature homework	Interventions outside the literature classroom context, e.g., voluntary leisure reading, bibliotherapy, book clubs, after-school programs
2. Hypotheses and measures	Interventions in which some form of gaining insight into human nature is expected and measured, e.g., affecting readers' insight into themselves or perceptions of others	Interventions with other foci, e.g., expected effects on reading comprehension, literary analysis, interpretative skills
3. Classroom context	Studies conducted in regular, first-language classrooms at secondary education level	Studies conducted in other classrooms, e.g., special needs, foreign language, and primary and higher education classrooms
4. Research design	Intervention studies with an experimental, quasi-experimental or posttest-only with comparison condition research design	Other types of studies without comparison conditions, e.g., action research, case studies, longitudinal studies, cross-cultural studies
5. Language	Studies published in English	Studies published in other languages

For the first criterion, we broadly defined "intervention types." We also included intervention studies in which activities were initiated in the literature classroom, but ultimately performed outside of it (e.g., school-assigned book readings at home), and studies that focused on reading various types of texts with particular reading instructions given in the classroom. If researchers used self-written stories or manipulated text features (e.g., Andringa, 1996), the study was excluded because we sought to identify design principles based on published texts.

Regarding the second criterion, we focused on intervention studies' central hypotheses and the measures used to assess them. If researchers expected that an intervention would in some way affect readers' insight into human nature (e.g., their perceptions of self, attitudes toward others, understandings of oth-

ers, ways of handling particular social situations, etc.), their studies were included – provided that these expected effects were measured. When relevant variables were presented as side effects to other variables (e.g., studies that examine both text comprehension and empathy), we included the study but focused on the variables relevant for this review.

For the third criterion, we deliberately only included intervention studies conducted in first-language classrooms because a foreign language teaching context introduces comprehension challenges that potentially interfere with gaining insight into human nature. We were particularly interested in intervention studies with adolescent participants in the upper grades of secondary education. However, we decided to also include studies conducted in lower grades of secondary school. If they occurred in the records, studies in grade 7 or 8, for example, can perhaps inform interventions in upper grades. Thus, we decided to include intervention studies with adolescent participants ranging in age from 13 to 18 years old. We excluded intervention studies conducted at the college or university level because these often involve older students who voluntarily enroll in a particular program, such as Sociology, medical school, or literary studies, which hampers the generalizability of outcomes to secondary school students.

For the fourth criterion, we only included intervention studies in which the effects of the experimental condition were compared to the outcomes of students in a comparison condition, such as another approach in the literature classroom or a nontreated control condition. Without a comparison condition, it cannot be determined whether potential effects can actually be ascribed to the intervention (Shadish, Cook, & Campbell, 2002).

Finally, we decided that only intervention studies published in English were eligible for inclusion. This was a practical decision in terms of analyzing the studies; for example, one abstract referred to an article in Russian, a language not mastered by any of the authors.

### 2.3 Search Procedure

In November 2017\*, we searched for relevant intervention studies in five educational databases: Education Resources Information Center (ERIC), Web of Science, PsycINFO, Linguistics and Language Behavior Abstracts (LLBA), and Scopus. The search yielded 7,933 results, of which 6,554 remained after deduplication.

---

\* *Date refers to the final update of the search; previous searches were conducted in July 2015 and March 2016.*

*Screening database records.* The first and second author screened titles and abstracts. The second author screened results from Scopus, and the first author screened records from the other databases. Figure 3.1 shows an overview of the selection procedure. The large majority of records was excluded in this first round of screening, mostly because titles and abstracts indicated they were not experimental or quasi-experimental intervention studies, but mainly theoretical essays about the value of literature teaching, ethnographic descriptions of literature classroom practices, and literary analyses of how social themes such as disability or discrimination are represented in books for children or young adults (e.g., Cummins, 2013; Curwood, 2013). Other recurring reasons for immediate exclusion were the context of higher education (e.g., Blackie & Wear, 2015; Weber, 2010) and intervention studies in foreign language classrooms (e.g., Benavides Buitrago, 2017; Fredricks, 2012).

If titles and abstracts did not expressively provide the necessary information to determine inclusion or exclusion, the full text was screened. During this round, 194 records were screened full-text to determine their relevance. The first author primarily conducted this phase. Most intervention studies that were screened full-text did not include a control or comparison group and were therefore excluded (125 of 194 studies). For example, Banks (2009) addressed literacy, sexuality, and the values of LGBTQ young adult literature, but screening the full article revealed that he “[drew] on personal experience to show the importance of reading LGBT[Q] young adult literature empathetically and critically” (p. 33) without including a comparison condition. Similarly, Bender-Slack (2002) described a humanist approach to teaching literature in the aftermath of 9/11 but did not compare the effects of the lesson series to another condition.

In cases of doubt about inclusion after screening full-text articles, all four authors screened those intervention studies and discussed together whether or not to include them. Discussions mainly focused on whether or not to adapt the third inclusion criterion, which specified including intervention studies conducted at secondary education level. Even though the search syntax was specified for secondary school students, it returned studies in primary or higher education that seemed highly relevant. For example, Fialho, Zyngier, and Miall (2011) investigated the effects of experiential versus interpretative literature teaching on empathy and related variables, but their participants were first-year university students. We concluded that our review would lose focus if we attempted to bridge the differences between primary, secondary, and higher education. Eventually, the database search yielded only seven intervention studies to be included.

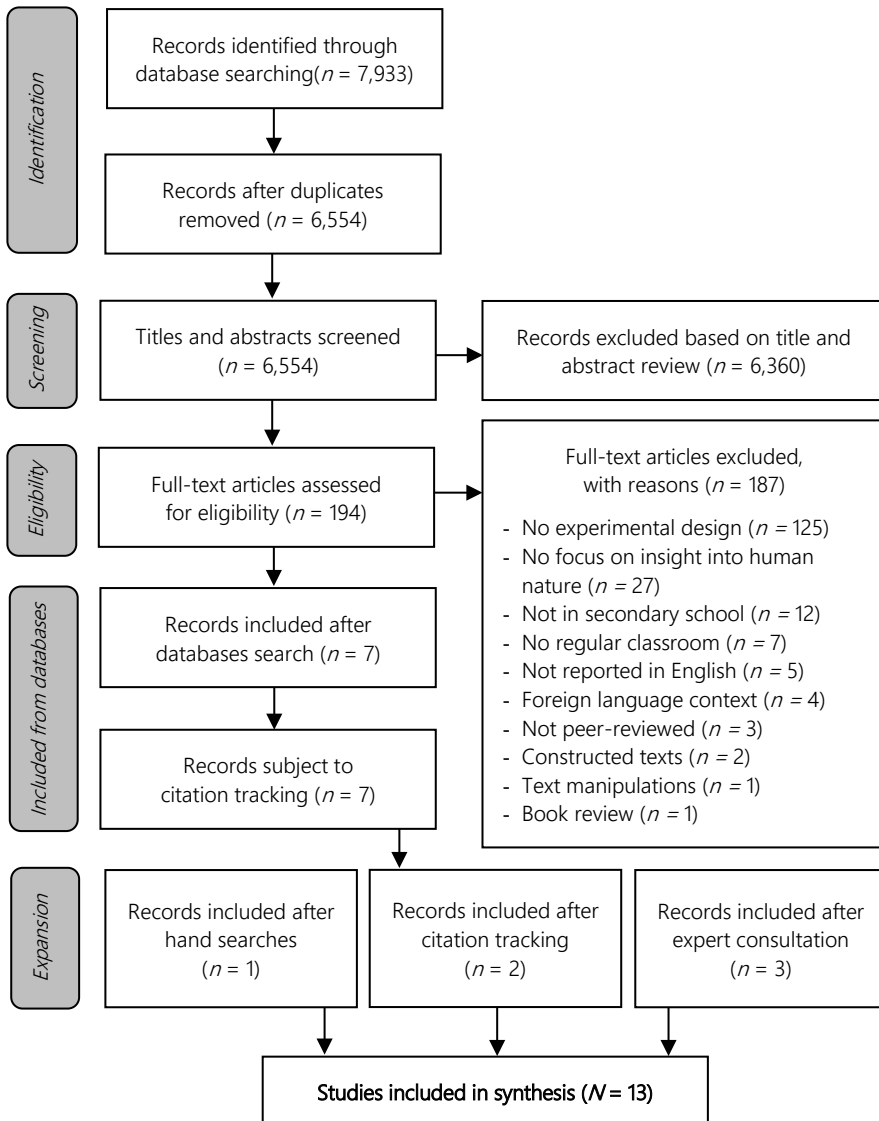


Figure 3.1. Flow chart of study selection procedure.

*Search expansion.* We expanded the search by applying citation tracking, by conducting hand searches, and by consulting experts in the field. First, for the seven intervention studies included from the database search, the first author screened whether they referred to relevant studies or whether these seven

studies themselves were cited in other relevant studies. We included two additional intervention studies (Darragh, 2015; Malo-Juvera, 2016), both of which cited a study from the database search (Malo-Juvera, 2014). For these two new intervention studies, we also performed citation tracking, which yielded no new studies to be included.

Furthermore, we hand-searched two journals that were not in the databases. First, we screened abstracts from *Pedagogy: Critical Approaches to Teaching Literature, Language, Composition, and Culture* from 2011 to 2017. Because no empirical intervention studies were published in this period, we waived further screening. Second, we screened *Study and Scrutiny: Research on Young Adult Literature* from its start in 2015. Apart from one study (Malo-Juvera, 2016), which was already included via citation tracking, this journal yielded no other studies to be included. We further consulted the *Annotated Bibliography of Research in the Teaching of English* (National Council of Teachers of English, 2003–2014), one online bibliography (Runge, 2012), and seven reviews and meta-analyses (De Leon, 2017; Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011; Galda & Beach, 2001; Koopman & Hakemulder, 2015; Murphy et al., 2009; Nystrand, 2006; Sigvardsson, 2016). From Koopman and Hakemulder (2015), we included an intervention study by Adler and Foster (1997). Citation tracking of this study yielded no other studies to be included.

Finally, we asked three academic experts in the field of literature education for suggestions. This yielded three additional intervention studies (Halász, 1991; Stevahn, Johnson, Johnson, & Real, 1996; Stevahn, Johnson, Johnson, Green, & Laginski, 1997). Citation tracking of these studies did not yield additional studies. Not all suggestions led to inclusion; for example, we excluded a study by Slone, Tarrasch, and Hallis (2000), which aimed to alter ethnic stereotyping in Israeli children, because the participants were in primary school.

As Figure 3.1 shows, the number of intervention studies included in the expansion phase was rather large compared to the number retrieved from the database search. This may be explained by the fact that most studies found in the expansion phase did not include keywords. Only Adler and Foster (1997) included one keyword, “bibliotherapy”. This particular keyword was not included in our search syntax; however, because the study tested a literature classroom intervention rather than small-group therapeutic sessions, we included it. Another reason for appearance in the expansion phase rather than the database search may be that studies were published in small electronic journals that were not included in large databases (Darragh, 2015; Malo-Juvera, 2016). Finally, our search syntax may not have complied fully with descriptions in ab-

stracts. For example, the relevance of the work by Stevahn et al. (1996; 1997) lies in its descriptions of “conflict resolution”, a term not included in the syntax.

Because our units of analysis were individual intervention studies rather than publications, we analyzed a single study from a publication reporting on two studies because one was conducted in higher education and another in secondary education (Hakemulder, 2008). Similarly, in addition to testing an intervention, Halász (1991) reported on another experiment that addressed differentiation between text genres, which we excluded. In total, we included 13 studies which all examined a different intervention.

## 2.4 *Quality Assessment*

Reviewing intervention studies for the purpose of informing both research and educational practice means that quality assessment must be performed both at the level of methodological characteristics of the study and at the level of intervention description. After all, a methodologically sound study may provide insufficient descriptions of the intervention, thereby hampering an answer to the question of “what works” (Rijlaarsdam et al., 2017).

### 2.4.1 *Methodological quality*

We assessed methodological quality to evaluate the validity of the conclusions of the included studies. We used indicators based on standards available from reviews of previous educational intervention studies (e.g., Engberg, 2004; Gersten et al., 2005; Hebert, Simpson, & Graham, 2013; O'Donnell, 2008; Pyle, Pyle, Lignugaris/Kraft, Duran, & Akers, 2017), as well as handbooks about research design (e.g., Shadish, Cook, & Campbell, 2002) and curriculum design (e.g., Van den Akker, Gravemeijer, McKenney, & Nieveen, 2006).

*Analysis.* Table 3.2 shows the coding scheme, which contains 15 indicators distributed over five categories: (a) description and rationale of comparison condition(s); (b) reliability of measures; (c) instructors and implementation; (d) data, results, and conclusions; and (e) attrition rate. Indicators that were not applicable (e.g., reliability of qualitative measurements if a study only included quantitative measures) were coded accordingly. If a study compared multiple experimental conditions, without using a control condition, indicators for the control condition were coded as “not applicable.” Of 195 scores (13 studies\*15 indicators), 144 were applicable. The two attrition rate indicators were nominally scored: 0 (*not reported*) or 1 (*reported*). The other 13 indicators were scored from 0 (*not at all* or *very poor*) to 4 (*completely* or *excellent*). We used scale

scores because intervention characteristics could be described in both more or less detail. For example, Henschel, Meier, and Roick (2016) specifically reported the duration of the comparison condition (135 minutes) and scored a 4 on this aspect, whereas Eva-Wood (2004) reported that the comparison group followed a four-week program, without specifying the duration of these lessons, resulting in a score of 2. Scale scores were also used to score reliability of measures because most researchers used multiple instruments. If the reliability of all measures was sufficient, a score of 4 was assigned. If the reliability of one or more measures was insufficient, a lower score was assigned. In case researchers reported only percentage agreement to report interrater reliability of coding schemes, without taking chance agreement into account, we assigned a score of 0 (Adler & Foster, 1997).

An independent second rater also scored the methodological indicators for all studies. For the two attrition rate indicators, agreement between both raters was 100%. To establish interrater reliability for scale scores, we calculated an intraclass correlation coefficient (ICC) for consistency of single measures using a two-way mixed model, which quantified the degree to which the raters provided consistent rank-ordering in their quality scores across studies and indicators (McGraw & Wong, 1996). The intraclass correlation was in the excellent range,  $ICC = 0.83$ ,  $p < .001$  (Cicchetti, 1994). The quality of intervention descriptions, thus, was rated reliably by both raters. Consequently, the scores of the first author were used for interpretation.

*Outcomes.* As Table 3.2 shows, some researchers adequately described what happened in the comparison conditions (Henschel et al., 2016; Vezzali, Stathi, & Giovannini, 2012; White, 1995). This was not the case in other studies. For example, Garrod (1989) only reported that “the comparison group curriculum had some features in common with the treatment,” followed by two examples of selected texts (p. 68). Other researchers merely mentioned that comparison group students read an unrelated text (Hakemulder, 2008) or were part of a wait-list (Malo-Juvera, 2016). Rationales for the choice of texts and tasks (A3, A5) were seldom reported. For instance, Adler and Foster (1997) only mentioned that texts in the comparison condition were part of the regular curriculum. Malo-Juvera (2014) reported that students in the comparison group read a text by Shakespeare – instead of the young adult novel that was related to the intervention theme, which was read by the experimental group – but in both groups, “similar instructional methods” were used (p. 416). However, the purpose of designing the comparison condition in this particular way was not discussed.

Table 3.2. Assigned quality scores (0-4) to methodological quality indicators, per category

Study	A. Description and rationale of comparison condition(s)				B. Reliability of measures				C. Instructors, implementation		D. Data, results and conclusions		E. Attrition rate		
	A1	A2	A3	A4	A5	B1	B2	B3	B4	C1	C2	D1	D2	E1	E2
Adler & Foster (1997)	2	2	1	2	0	-	-	0	-	0	2	2	4	1	0
Darragh (2015)	1	0	0	4	1	0	-	0	-	0	2	1	2	0	-
Eva-Wood (2004)	2	2	2	3	3	4	3	3	4	2	2	4	3	0	-
Garrod (1989)	1	1	1	1	1	-	-	0	-	0	0	4	2	0	-
Hakemulder (2008)	0	1	0	-	-	0	-	-	-	-	0	1	1	0	-
Halász (1991)	-	-	-	-	-	-	-	4	4	-	2	1	3	0	-
Henschel et al. (2016)	4	4	3	-	-	4	4	4	4	4	3	4	1	1	1
Malo-Juvera (2014)	3	4	0	1	0	4	4	-	-	4	0	4	2	1	0
Malo-Juvera (2016)	2	0	0	0	0	4	4	-	-	2	0	4	3	0	-
Stevahn et al. (1996)	-	-	-	-	-	-	-	4	4	4	1	4	3	0	-
Stevahn et al. (1997)	4	4	0	3	2	-	-	1	1	4	0	3	4	0	-
Vezzali et al. (2012)	4	3	3	2	2	4	4	-	-	2	4	4	4	1	1
White (1995)	1	4	4	4	4	-	-	0	-	3	2	1	4	0	-

*Note:* Indicators not applicable denoted with -. A1. duration of comparison condition(s); A2. selected texts; A3. rationale for texts; A4. tasks; A5. rationale for tasks; B1. reliability reported for quantitative measures; B2. reliability of B1 > .70; B3. reliability reported for qualitative measures; B4. reliability of B3 > .70; C1. possible instructor effects taken into account; C2. information about implementation fidelity reported; D1. statistical descriptives sufficiently reported (i.e., group sample sizes, means, standard deviations); D2. conclusions legitimately based on data and results; E1. attrition rate reported; E2. attrition rate < 10%.

Although applicable assigned scores in the categories for measurement reliability were generally high, they occurred less often for qualitative measures such as written responses (B3) than for quantitative measures such as standardized questionnaires (B1). Yet, if reliability was statistically reported, the values (e.g., Cronbach's Alpha or Cohen's Kappa) mostly met the widely accepted threshold of .70 (B2, B4).

Possible instructor or teacher effects (C1) and implementation fidelity (C2) that may have affected the validity of the results of interventions were often neglected. Some researchers avoided instructor effects by working with trained instructors in all conditions (Henschel et al., 2016) or with two teachers who both taught the experimental and the comparison condition (Malo-Juvera, 2014). In other studies, two different teachers taught the experimental and comparison groups (Adler & Foster, 1997; Darragh, 2015; Garrod, 1989). Very few researchers reported on implementation fidelity. In order to assess implementation, Vezzali et al. (2012) evaluated whether students had actually read their books by looking at their written summaries. In some studies, examples of students' writing or transcripts of classroom discussions provided information about the implementation of those tasks (e.g., Darragh, 2015; White, 1995). Other researchers did not report on implementation fidelity at all. Researchers did not use, for example, teacher logs, objective classroom observations, or other measures of implementation fidelity.

Some studies lacked sufficient descriptive statistics (e.g., group sample sizes, means and standard deviations or standard errors) or overestimated their conclusions in light of their own statistical results. For instance, an intervention was found to have a statistically significant effect, but this effect applied to a subgroup of students of which only five were in the experimental condition (Garrod, 1989). In another study, a "small disordinal interaction between the treatment condition and empathy" was reported (Henschel et al., 2016, p. 17). At the posttest, mean empathy scores did not differ between the experimental and the comparison condition (both  $M = 2.40$ ). At the pretest, the comparison group scored higher than the intervention group, but the researchers did not report whether this initial difference was statistically significant, nor how pretest scores were taken into account in the analysis. If a pretest score were included as a covariate, a statistically significant effect would not have occurred.

Finally, attrition rate was reported in only four of the 13 studies. In two of these, more than 10% of the students dropped out between the pretest and posttest (Adler & Foster, 1997; Malo-Juvera, 2014). If sample sizes differ at the pretest and posttest, internal validity may be affected, which may result in in-

correct statistical analyses. Such issues, however, were not considered in the included studies.

#### 2.4.2 Quality of intervention descriptions

In Table 3.3, the coding scheme and outcomes of the assessment of the quality of intervention descriptions are presented.

Table 3.3. Assigned quality scores (0-4) to characteristics of intervention descriptions

Study	Duration	Texts	Rationale for texts	Tasks	Rationale for tasks
Adler & Foster (1997)	2	4	4	1	0
Darragh (2015)	1	4	3	4	1
Eva-Wood (2004)	2	2	2	3	3
Garrod (1989)	1	2	1	2	2
Hakemulder (2008)	1	4	4	4	4
Halász (1991)	2	4	4	3	4
Henschel et al. (2016)	4	4	3	3	3
Malo-Juvera (2014)	4	4	4	4	3
Malo-Juvera (2016)	3	4	4	1	3
Stevahn et al. (1996)	4	4	0	2	2
Stevahn et al. (1997)	4	4	0	3	2
Vezzali et al. (2012)	4	3	3	2	2
White (1995)	1	4	3	4	4

*Analysis.* We assessed to which extent the interventions were described in detail, by focusing on five basic characteristics (see Table 3.3): information about (a) the duration of the intervention, (b) the selected texts, (c) the tasks students were asked to complete, (d) rationales for selecting these texts, and (e) rationales for designing these tasks. Rationales are important because they indicate why the selected texts and tasks would be suitable to achieve a particular purpose, thereby helping educational designers to make well-informed choices (Rijlaarsdam et al., 2017). The quality of descriptions for these five indicators was scored on a scale from 0 (*not at all or very poor*) to 4 (*completely or excellent*).

To establish interrater reliability, an independent second rater scored the five indicators for all 13 studies. The intraclass correlation between the first and second rater was excellent, ICC = 0.78,  $p < .001$  (Cicchetti, 1994). Consequently, the scores of the first author of this paper were used for interpretation.

*Outcomes.* First, Table 3.3 indicates a rather large variety across studies regarding information about the duration of interventions. For example, Malo-Juvera (2014) reported in detail that the instructional unit “lasted 5 weeks and consisted of 12 classes (each lasting 1 hour and 45 minutes)” (p. 415). In contrast, Garrod (1989) mentioned neither the exact number of lessons nor their duration; the only information given was that the intervention was part of a year-long curriculum (which received a score of 1). Second, most researchers reported which texts were used in interventions by providing titles, authors, and sometimes summaries (Adler & Foster, 1997). A low score on this criterion is exemplified by Eva-Wood (2004) who reported author’s names but not how many and which poems were selected. Third, most studies contained good or excellent descriptions of intervention tasks.

Finally, rationales for text selection and tasks received, overall, lower scores than their descriptive counterparts. There was a rather large variety across studies, ranging from 4 (e.g., Hakemulder, 2008, who provided an explanation of why a passage from a particular multicultural novel was selected and a theory-based rationale for implementing a reading task) to 0 (e.g., Stevahn et al., 1996; 1997, who provided no rationale for why particular novels were selected).

#### 2.4.3 *Quality assessment conclusions*

Sufficient intervention descriptions were provided by the majority of the included studies, although the rationales for these interventions were sometimes suboptimal. Some studies were subject to validity issues. In particular, some interventions were taught by a single teacher while the comparison condition was taught by another. In most cases, it remained unclear to which extent the implementation resembled the original intervention design. Yet, because selected texts and tasks were sufficiently described in most studies, we assumed that the selection would offer valuable input for further analysis.

#### 2.5 *Data Analysis*

First, we analyzed the expected outcomes of the included interventions. In some cases, insight into human nature was only expected as an ancillary effect, for instance, if the primary aim was to enhance text comprehension but an additional effect on empathy for characters was expected (Henschel et al., 2016). Thus, in the analysis, we opted to use the term *expected effects* rather than *aims*.

Next, we analyzed the outcomes of the interventions to determine whether empirical support was provided for the expected effects in comparison to another condition (e.g., different approach to teaching literature or business-as-usual). We determined whether studies indicated full empirical support (as indicated by all measures that were applied), partial support (as indicated by some of the measures), or no support for the intervention that was implemented. We did so via critical appraisal of empirical support presented in the studies, as indicated in the Methodological Quality section: we thoroughly analyzed whether researchers might have overestimated empirical support found in their studies (e.g., Garrod, 1989).

We subsequently analyzed instructional approaches of only those interventions with full and partial empirical support. We addressed genres, themes, and literariness of selected texts, tasks, the role of the teacher, and stances toward texts taken in these interventions. Details of all studies as reported in the original publications (including expected effects, research designs, instruments, and demonstrated outcomes) are presented in Appendix B.

### 3 RESULTS

#### 3.1 *Expected Effects and Empirical Support*

Few experimental or quasi-experimental intervention studies in literature classrooms have focused on students' insight into human nature. Only 13 studies fully met our inclusion criteria. In this section, we present an overview of the expected effects and the empirical support found in these intervention studies.

As Table 3.4 shows, we distinguished three categories of expected effects on students' insight into human nature. Researchers expected to affect (a) students' insight into themselves, (b) their understanding of fictional others, and/or (c) their understanding of, views on, or intended behavior toward real-world others. These categories emerged from researchers' use of measures that focused either on insight into oneself (e.g., a scale with items such as "Reading literature makes me sensitive to aspects of my life that I usually ignore"; Miall & Kuiken, 1995, p. 55; in Eva-Wood, 2004), into fictional others (e.g., "I can easily empathize with one of the characters from the text"; Henschel et al., 2016, p. 16), or into real-world others (e.g., "When girls wear low cut tops and short skirts they're just asking for trouble"; Malo-Juvera, 2014, p. 419). One researcher expected effects on both students' self-insight and their understanding of fictional characters (Eva-Wood, 2004).

Table 3.4 also indicates to which extent there was empirical support for the expected effects. There was empirical support for nine out of 13 interventions,

although two of these could only provide partial support (Adler & Foster, 1997; Eva-Wood, 2004). Effect sizes are included when they were reported in the study or could be calculated based on the data, and annotated if they were not applicable or could not be calculated.

*Insight into oneself.* Two researchers expected that their intervention would affect students' insight into themselves. One of these interventions provided full empirical support for this expected effect (Halász, 1991), whereas the other provided no empirical support (Eva-Wood, 2004).

Halász (1991) asked students to write down memories and associations that were evoked during reading a text or to do so in response to salient words from the text. First, he expected that students would rely more on personal experiences when responding to a literary text than to an expository text or an essay. Second, he expected the same result if, after reading, students' writings responded to high frequent, salient words from these texts. Third, he expected that responding to salient words after reading would result in more personal responses than responding to the same words without reading the texts. All three hypotheses were confirmed. In all cases, reading the literary text evoked more personal, affective, and detailed responses, which most often consisted of personal references and indications of emotion, than reading the other texts or reading no text. Halász also observed that students, in their personal responses, predominantly referred to secondary sources such as fiction, music, art, or experiences they garnered from others. This finding indicates that drawing upon such sources may help students to interact with a literary text with respect to their own lives. Because analyses were based upon merged categories (e.g., "personal references" consisted of four categories) and means and standard deviations were only reported for individual categories, effect sizes of merged categories could not be calculated.

Table 3.4. *Expected and demonstrated intervention effects*

Study	Insight into oneself	Understanding of fictional others	Understanding of, views on, or intended behavior toward real-world others	Empirical support
Halász (1991)	Insight into personal memories			Full
Malo-Juvera (2014)			Views on sexual harassment	Full
Malo-Juvera (2016)			Views on others' sexual orientation	Full
Stevahn et al. (1996)			Understanding of, intended behavior in and views on conflict situations	Full
Stevahn et al. (1997)			Views on and intended behavior toward immigrants	Full
White (1995)		Understanding characters and their behavior		Full
Adler & Foster (1997)			Views on caring for others	Partial <sup>a</sup>
Eva-Wood (2004)	Insight into own qualities and world *	Empathy for characters, understanding poems' speakers *		Partial <sup>b</sup>
Darragh (2015)			Views on and intended behavior toward disabled people	No
Garrod (1989)			Understanding people's moral dilemmas	No
Hakemulder (2008)			Views on immigrants	No
Henschel et al. (2016)		Empathy for characters *		No

*Note.* \* Ancillary effect rather than primary aim. <sup>a</sup> Empirical support from one out of three measures that were used. <sup>b</sup> Empirical support only provided for understanding poems' speakers.

Eva-Wood (2004) implemented a think-and-feel-aloud pedagogy in response to poetry. The pedagogy was primarily expected to enhance students' transactions with poems, in terms of higher levels of engagement in analysis and more sophisticated responses, but Eva-Wood also assessed whether it enhanced students' insight into previously unrecognized qualities in themselves and in their world. To assess this, she used the *Insights* scale of the Literary Response Questionnaire (Miall & Kuiken, 1995). On this scale, no differences were found between the experimental condition and a comparison condition that focused on structural analysis of the poems.

*Understanding of fictional others.* Three researchers expected intervention effects on students' understanding of fictional others. One study provided full empirical support for the expected effect (White, 1995), another provided partial empirical support (Eva-Wood, 2004), and the third could not provide empirical support (Henschel et al., 2016).

White (1995) expected that students who completed a thematically relevant autobiographical writing task prior to reading a short story would express more sophisticated understandings of fictional characters and their behavior than students who had not written before reading. The effect was confirmed by analyzing classroom discussions: Students who completed the writing task more often moved beyond literal understandings toward more abstract understandings of characters and their actions. If students had not written, they less often showed such understandings of characters and their behavior. Effect sizes could not be calculated because group sample sizes were not given.

In addition to their self-insight, Eva-Wood (2004) expected students' empathy for fictional characters to be affected (measured using the *Empathy* scale of the Literary Response Questionnaire; Miall & Kuiken, 1995) and expected students to personally engage and identify with poems' speakers as measured by contributions to classroom discussions. No differences on empathy were found between the experimental and the comparison condition. However, students in the experimental condition showed greater personal engagement and identification with poems' speakers, for example, by expressing their understanding of a speaker who had lost a loved one. Thus, Eva-Wood's intervention study showed partial empirical support for fostering students' understanding of fictional others; effect sizes for personal engagement and identification could not be calculated, as standard deviations were not reported.

Henschel et al. (2016) expected that completing reader-oriented tasks, which focused on personal emotional engagement and creative responses, would increase students' empathy for fictional characters – as opposed to

completing text-based tasks that stimulated text analysis via cognitive activities. The hypothesis was tested by using three items adapted from the *Fantasy empathy* scale of the Interpersonal Reactivity Index (Davis, 1983). As indicated in the Method section, posttest mean scores did not differ between conditions, and pretest scores were not taken into account in the statistical analysis. Based on the reported results, we could not infer that the intervention study yielded empirical support for an expected effect on students' empathy for characters.

*Understanding of, views on, or intended behavior toward real-world others.* In nine out of 13 studies, researchers expected to affect students' understanding of, views on, or intended behavior toward other human beings in the real world. The measures they used did not apply specifically to reading (as did the scales used by Eva-Wood, 2004, and Henschel et al., 2016) but to the world beyond students' experiences with the text. Of nine interventions in this category, five provided full empirical support (Malo-Juvera, 2014; 2016; Stevahn et al., 1996; 1997; Vezzali et al., 2012); one provided partial empirical support (Adler & Foster, 1997), and three provided no empirical support (Darragh, 2015; Hakemulder, 2008; Garrod, 1989).

Malo-Juvera (2014) expected that a dialogic approach to teaching a young adult novel about sexual harassment would result in reduced rape myth acceptance in students. Students were expected to reject the ideas that victims provoke rape or falsely claim it happened. In the comparison condition, students were taught a classic novel via a similar dialogic approach. The Adolescent Rape Myth Scale, a questionnaire based on previous studies, was administered as a pre- and posttest to assess the effect. A main effect of condition was found ( $d = 0.84$ ), indicating that students' rape myth acceptance scores at the posttest were lower if they received the intervention than if they were in the comparison group.

Malo-Juvera (2016) implemented the same approach to teaching a young adult novel about the coming-out of a male adolescent character who identifies as being homosexual. By using the researcher-developed Adolescent Homophobia Index as a pre- and posttest, a main effect of condition was found. Students in the experimental condition had lower posttest homophobia scores than students in an untreated control condition ( $d = 0.87$ ).

Stevahn et al. (1996; 1997) expected that conflict-resolution training, in which conflicts from fictional novels were used, would result in improved understandings of how to solve conflicts, better application of this knowledge in conflict scenarios (intended behavior), and more constructive, positive views on conflicts. In the comparison condition, students read the same novel but completed tasks that did not focus on conflicts. Measures included writing down

steps to solve a conflict (understanding), writing short essays about how a conflict scenario could be solved (intended behavior), and writing words associated with conflict (views on conflict) – coded as negative/destructive, neutral, or positive/constructive. Students in the experimental condition scored higher on understanding how to solve conflicts and on intended behavior in conflict situations than students in the control condition, both at the posttest and the delayed posttest ( $d$ 's > 1.00). Students in the experimental condition also listed more positive associations than students in the control condition. More specifically, Stevahn et al. (1996) compared two versions of the intervention: a cooperative and an individualistic version. Results indicated an interaction effect of condition and version. The cooperative version of the intervention most effectively fostered understandings of conflict resolution, intended behavior to solve conflicts, and positive views on conflict.

Finally, Vezzali et al. (2012) asked students to read a novel featuring immigrant characters and to complete a writing and evaluation task afterwards. The authors expected that the intervention would result in more positive views on immigrants and more positive intended behavior toward them. The authors compared the intervention to students reading a nonintercultural novel, and completing the same tasks, as well as to students not reading or completing tasks at all. This was tested by administering various measures, such as a Word Association Task to assess immigrant stereotypes, a Hypothetical Contact Scenario Test to assess behavioral intentions toward immigrants, and items that assessed students' desire for future contact with immigrants. The researchers used two planned contrasts to test their hypotheses. First, they compared the intercultural reading condition with the two control conditions to test the effects of indirect contact through book reading; second, they compared the intercultural reading condition to the nonintercultural reading condition, to ensure that effects would not be due to reading *any* book. Compared to the two other conditions taken together (contrast 1), students in the experimental group scored higher on all measures at the posttest ( $d$ 's ranging from 0.56 to 1.16). Compared to the nonintercultural reading condition (contrast 2), similar results were found ( $d$ 's ranging from 0.52 to 1.22).

Adler and Foster (1997) developed an intervention which they expected to increase students' support for the value "caring for others." Students read three novels in which this theme was prominent and participated in classroom discussions and exercises that were designed to reinforce the theme of the books. In the comparison condition, students read novels from the regular curriculum. All students completed three essays as pretests and the same three essays as

posttests, which included topics about caring for a family member, caring for strangers who lost their home to a fire, and about friends as stand-ins for family. In the essays about friends, more students in the experimental group showed positive change in valuing "caring for others" than in the control group, a difference that was statistically significant. For the other two pre- and posttest essays, no differences were found. Thus, there was partial empirical support for the expectation that the intervention would foster students' support for the value of caring for others. An effect size could not be calculated because group sample sizes were not given.

Three studies remained in which no empirical support was found for the expected effects. Darragh (2015) asked students to read a young adult novel featuring a disabled character and to respond to it in writing tasks. She expected that students would develop more positive views on and intended behavior toward disabled people, as compared to students who read novels without disabled characters. One measure assessed views on disabled people via a word association task; the other asked students whether they would undertake particular activities with a disabled person. No differences between conditions were found.

Hakemulder (2008) expected that reading an excerpt from an intercultural novel while following a role-taking reading instruction would evoke a positive view of immigrants. The intervention was compared to three other conditions: (1) reading the same text but focusing on its structure, (2) reading an essay about the same theme without instruction, and (3) reading an unrelated text without instruction. Five items were used to assess intervention effects. On two of these, a statistically significant difference between the experimental and a comparison condition was found; however, these effects were in favor of the comparison conditions. Thus, the intervention effect contradicted the hypothesis; however, it should be noted that this was determined using a single item rather than a validated attitude scale.

Finally, Garrod (1989) developed an intervention in which students engaged in Socratic discussions about particular moral dilemmas in literary texts. He expected that this would result in students developing their moral reasoning abilities compared to a condition in which Socratic discussions focused on other texts. Moral reasoning development was assessed using the written version of the Kohlberg Moral Judgment Interview. Of all students in both conditions ( $N = 44$ ), 17 students formed a subgroup with the lowest pretest scores. In this subgroup, the intervention had the expected effect: Students in the experimental condition ( $n = 5$ ) achieved more growth than those in the comparison condition ( $n = 12$ ), a difference that was statistically significant. However, the small

condition sample sizes in this subgroup may have affected the statistical results. In addition, there was no overall difference between the experimental and the comparison condition.

*Interventions with empirical support.* Nine interventions provided full or partial empirical evidence for fostering students' insight into human nature, in terms of insight into themselves (Halász, 1991), their understanding of fictional others (Eva-Wood, 2004; White, 1995), and their understanding of, views on, or intended behavior toward real-world others (Adler & Foster, 1997; Malo-Juvera, 2014; 2016; Stevahn et al., 1996; 1997; Vezzali et al., 2012).

### 3.2 *Instructional Approaches*

We analyzed the instructional approaches applied in the nine interventions for which full or partial empirical support was found. Because design principles should be based on interventions with empirical support (Merrill, 2002; Van den Akker, 1999), we left aside the four studies in which no empirical support was found.

#### 3.2.1 *Texts used*

In this section, we describe which texts were selected in the interventions with full or partial empirical support, thereby addressing genres, themes, and literariness. We observed that students mostly could not choose their own reading materials. Only Vezzali et al. (2012) allowed students to choose a book from a list with preselected titles.

*Genres.* Researchers used fictional texts – mostly novels and short stories. In one study wherein poetry was read, Eva-Wood (2004) reported that poems by Emily Dickinson, Robert Frost, e.e. Cummings, and Langston Hughes were used but did not specify any titles (p. 178). If novels were used, most researchers provided an additional genre specification: young adult literature (*Speak* by Anderson, in Malo-Juvera, 2014; *Geography Club* by Hartinger, in Malo-Juvera, 2016), historical fiction (*Days of terror* by Smucker, in Stevahn et al., 1996), a coming-of-age-novel (*Crabbe* by Bell, in Stevahn et al., 1997), or an intercultural novel (e.g., *Le nuvole da latte* by Frescura, in Vezzali et al., 2012). Only Adler and Foster (1997) did not specify the genre of their three selected novels (*Friends Are Like That* by Hermes, *Red Cap* by Wisler, and *The Clay Marble* by Ho). The short stories that were used were characterized as a complex meta-

phorical story (*The Vulture* by Kafka, in Halász, 1991) and two literary stories (*Indian Camp* and *The End of Something* by Hemingway, in White, 1995).

*Themes.* Most researchers considered text theme to be a relevant factor for selection with two exceptions: Eva-Wood (2004) and Halász (1991) did not reflect on the theme of the texts they used. In the seven remaining studies, clear lines of reasoning about text themes were provided. Malo-Juvera (2014; 2016) selected thematically relevant young adult literature because discussing such themes was expected to alter adolescents' social beliefs (e.g., Kaywell, 1993; in Malo-Juvera, 2014). Likewise, Vezzali et al. (2012) selected novels in which the targeted outgroup – immigrants – played a role. Adler and Foster (1997) similarly chose novels in which caring for others was an important theme, as indicated by protagonists who actively considered their behavior toward others (p. 277). Stevahn et al. (1996; 1997) indicated that the theme of "conflicts" in the two novels they selected would be suitable for learning conflict-resolution strategies, although they did not explicate why the particular novels by Schmucker and Bell were selected. Finally, to enhance students' understandings of characters and their behavior, White (1995) selected two stories that portrayed the difficult and painful aspects of social relationships (i.e., dating and parent-child relationships).

*Literariness.* Because there is little consensus about how to determine literariness, we analyzed whether researchers themselves made any references to the concept, and if so, whether they provided a rationale for labeling a text as such. Researchers who exclusively used the terms "books" and "novels" and not "literature" made no assumptions about possible literary features of these texts (Adler & Foster, 1997; Stevahn et al., 1996; 1997; Vezzali et al., 2012). In contrast, Malo-Juvera (2014; 2016) referred to "young adult literature." Similarly, Halász (1991) and White (1995) described the stories they used as "literature" or "literary." Halász put literary texts on a par with fictional texts, stating that *The Vulture* is "a literary [text], presenting fictitious events with fictitious characters" (p. 249). None of these three researchers further explained why the texts they used could be perceived as literary texts nor whether any textual features would point to this classification (e.g., see Mar & Oatley, 2008; Miall & Kuiken, 1999; Mukařovský, 1976).

Only Eva-Wood (2004) considered literariness from a theoretical perspective, stating that reading poems defamiliarizes readers when they encounter stylistic devices that are specific to literary texts, such as metaphors and similes. Literary texts thus deviate from the conventional understandings of words and the relationships among them (Eva-Wood, 2004, p. 175-176; Miall & Kuiken,

1994). All in all, literariness seemed of little concern in the studies included in this review.

### 3.2.2 Tasks

Two types of tasks were identified as the most salient: writing tasks and dialogues. We will characterize them below, followed by a short characterization of tasks that occurred less frequently in the interventions.

*Writing.* Writing tasks were found to have three aims and were implemented in corresponding moments: (1) to activate previous personal experiences relevant to a text theme prior to reading (Malo-Juvera, 2016; White, 1995); (2) to annotate, during the reading process, spontaneous responses evoked by the text (Eva-Wood, 2004; Halász, 1991); and (3) to reflect on and respond to issues addressed in the text and/or one's experiences with reading the text directly after finishing the full text or a distinctive excerpt, such as a scene or a chapter (Adler & Foster, 1997; Halász, 1991; Malo-Juvera, 2014; 2016; Vezzali et al., 2012).

*Writing to activate previous personal experiences.* Malo-Juvera (2014; 2016) asked students to respond to a young adult novel in dialogic sequences. These sequences consisted of three steps: students (1) completed an individual writing task, (2) shared their written responses in a small group, and (3) presented the group conclusions in a whole-class discussion. In the 2016 study, one of seven implemented sequences was completed before students started reading the novel. They were asked to write about bullying in their school – a task that activated their previous personal experiences with this theme.

Likewise in White's (1995) intervention, students were given the theme of a short story (i.e., parent-child or dating relationships) and were asked, prior to reading, to write about relevant background knowledge and personal experiences. Two task characteristics stood out. First, the task prompted students to write about experiences from their own lives, but the wording also allowed students to refer to events they had not experienced but had observed (e.g., "write about parents *you know*"; White, 1995, p. 184, emphasis in original). Thereby, experiences of both a primary and secondary nature were prompted. Second, the tasks explicitly encouraged students to explore multiple perspectives on themes. Students were asked to write about parents who are good as well as bad teachers and also about the characteristics of healthy and fun dating relationships and why such relationships might end.

*Writing to annotate spontaneous responses.* Both Eva-Wood (2004) and Halász (1991) implemented writing tasks to stimulate students to notice and annotate their responses during the reading process. Eva-Wood's intervention centered on a think-and-feel-aloud pedagogy, in which students were taught how to notice and verbalize the thoughts and emotions that a particular poem evoked in them. After observing their teacher model the reading strategy, students practiced the strategy in pairs: one student verbalized responses while reading and the other took notes, and after this procedure they switched roles. Writing down responses enabled students to analyze them after reading.

Similarly, one of the writing tasks in Halász's (1991) study took place while students read a literary text. During the reading process, students were asked to annotate the text in terms of personal experiences, memories, and associations they extrapolated from the story. They read with a pencil in hand and immediately wrote down their responses. These two studies showed that a writing task may function as an effective tool to help students express their initial reading experiences.

*Reflective response writing.* Reflective response writing occurred most often in the included interventions. This happened either after a full text had been read or after students had finished part of the text. Vezzali et al. (2012) asked students to identify key activities in the story by writing a summary. Halász (1991) asked students to respond in writing to salient words found in the text after they had read the full text. Similar to the annotation task, students were asked to respond in terms of personal experiences and memories that the salient words evoked. The writing task thus enabled students to formulate their responses in terms of personal references and to express emotions. In addition, students often referred to secondary experiences from fiction, movies, art, and hearsay from others around them. Halász argued that such experiences should also be regarded as legitimate responses in the literature classroom.

In dialogic sequences implemented by Malo-Juvera (2014; 2016), reflective response writing was frequently implemented. In the study addressing rape myth acceptance, students were asked to imagine the situation of the protagonist by writing her a letter in which they convinced her to seek help; to evaluate in writing whether they thought the protagonist was raped or not (directly after reading a particular scene); to write a conversation with the antagonist, convincing him that he was guilty of rape; and to write about circumstances in which they would or would not believe a girl who said she had been raped (after finishing the novel; Malo-Juvera, 2014).

In the study addressing sexual orientation and homophobia, students were first asked to write about three issues after finishing excerpts from the novel:

why gay teenagers more often attempt suicide than heterosexual teenagers; a conversation with a character about supporting another character who questioned his own sexual orientation; and their opinions about which sexual orientation the protagonist might choose if he had a choice. After finishing the novel, students completed three more writing tasks: identifying the three most important themes in the book; determining their favorite characters and explaining their choices; and considering whether their own school needed a "gay-straight alliance club" (Malo-Juvera, 2016, p. 11). Writing tasks thus required formulating opinions and evaluations as well as adopting the perspective of the novel's characters and relying on one's own imagination, for example, when writing conversations.

Finally, Adler and Foster (1997) implemented individual reflective writing tasks (e.g., writing journal entries about personal feelings and favorite quotes from the books) and organized students to write in groups. Students were asked to produce artifacts such as a group collage about the theme of a book and their feelings related to it or to create a mind-map-like organizer in their group to visualize a character's social relationships (e.g., friends, family).

*Dialogues.* Students were asked to engage in dialogues in all interventions except one (Vezzali et al., 2012). In some interventions, dialogues took place exclusively in small groups (Stevahn et al., 1996; 1997) or exclusively as whole-class activities (White, 1995). More often, however, combinations or sequences of small-group and whole-class dialogues were implemented (Adler & Foster, 1997; Eva-Wood, 2004; Malo-Juvera, 2014; 2016).

In several interventions, writing prepared students to engage in small-group or whole-class dialogues. Dialogues, then, added a layer of responses, perspectives, and interpretations to the kind that individual students explored in writing. After writing and reading, White (1995) engaged students in classroom dialogues about the characters, their actions, and the consequences of these actions. The teacher guided the dialogues by following a protocol. The protocol contained ten questions. Seven questions initiated describing characters or explaining their actions. These were followed by a prediction question about what might happen next, a question about which message for the real world might be embedded in the text, and one question that addressed why the author made particular choices to include or to describe a character. The study indicated that autobiographical writing prior to reading helped students to explore in these dialogues what fictional characters are like and why they behave in a certain way.

Likewise, writing to record thoughts and feelings in response to a poem enabled students to talk about these responses (Eva-Wood, 2004). In dialogues, students were asked to focus on evoked emotions, on specific words and phrases they responded to, and on interpretative questions and remarks. Similarly, in dialogic sequences (Malo-Juvera, 2014; 2016), students were asked to engage in small-group dialogues to share written responses, followed by an exchange of each groups' conclusions. These tasks appeared to invite students to express and to compare their experiences with the text and its theme, but little information was given about the exact instructions that guided these dialogues.

In other interventions, the dialogues did not follow on writing. Adler and Foster (1997) integrated writing and talking. Groups of students were asked to produce a collage and a mind map. In these tasks, a certain amount of writing was involved, but students were also required to talk about what they created. Adler and Foster applied classroom dialogues as well, but did not report any information on them. Stevahn et al. (1997) applied various forms of dialogue not combined with writing tasks. As part of learning how to solve conflicts, students were asked to talk about conflicts that are common for teenagers. Thus, relevant previous knowledge and experiences were activated. After students observed their teacher model a resolution strategy, students talked in pairs or triads about the conflicts that they identified in the novel and about what a character might say to solve these conflicts. In the less effective individual learning condition (Stevahn et al., 1996), students did not talk amongst each other; rather, during other activities, students explained to their teacher the conflicts they found in the novel and wrote a script to describe how they would solve them.

*Infrequently implemented tasks.* Tasks other than writing and dialogues occurred less frequently in interventions with empirical support: observation tasks, role-playing, and multiple-choice evaluation of reading experiences. Eva-Wood (2004) asked students to observe their teacher demonstrate the think-and-feel-aloud pedagogy. In the intervention by Stevahn et al. (1996; 1997), students observed their teacher and their peers who acted out the resolution of a conflict, by which role-playing was also implemented. Finally, Vezzali et al. (2012) asked students how much they liked the book they read, to which extent it was interesting and pleasant, and whether they had problems reading it; they answered these questions by circling their evaluations.

*Conclusions.* All in all, writing and dialogues were the most salient types of tasks in interventions with full or partial empirical support. Notably, some inter-

ventions featured minimal instructions, such as a single writing task (Halász, 1991) or a written summary and encircled evaluations (Vezzali et al., 2012). Some tasks functioned as intervention activities – in the sense that they were presented to students as response tasks – and research instruments simultaneously. For example, Vezzali et al. used students' evaluations to control for appreciation and difficulty in statistical analyses, Halász used students' written responses for analysis, and White (1995) analyzed students' responses in classroom dialogues.

### *3.2.3 The role of the teacher*

In interventions with full or partial empirical support, we expected to see descriptions of how teachers were asked to foster dialogic discourse, for example, by offering students exploratory prompts and open-ended questions to guide and to support their dialogues, by making subtle discourse moves that facilitated student talk, or by interacting with students about their written responses to texts. Only two researchers reflected on the role of the teacher (Malo-Juvera, 2014; Eva-Wood, 2004).

Malo-Juvera (2014) reported that the teacher only interfered minimally during small-group dialogues that followed on individual response writing, which reduced the authoritarian role of the teacher. The teacher did not collect and assess the responses that students had written. The small-group dialogues with minimal teacher interference allowed students to share responses in an authentic, genuine way, which included a variety of perspectives and opinions. In whole-class dialogues, the teacher urged students to ask each other questions in order to explore multiple perspectives and contradictions. This called for an atmosphere in which students felt free to express their responses. Teachers avoided "correcting" students' opinions and beliefs from their position of authority. Thus, students were able to explore, express, and compare authentic responses, both in writing as well as in small groups where they guided the dialogue themselves. Eva-Wood (2004) emphasized the role of teachers as models of the think-and-feel-aloud pedagogy. Additionally, she reported that teachers asked targeted questions during whole-class dialogues, such as what students saw and felt while reading or what surprised them. Finally, she indicated that teachers helped students to draw connections between their experience with a poem and its literary elements.

In short, Malo-Juvera (2014) and Eva-Wood (2004) appeared to envision somewhat different roles for teachers. Nonetheless, both roles allow for dialogic discourse in the classroom, in which the teacher acts as a facilitator rather

than an authoritative figure. No details were provided on specific, subtle discourse moves that teachers might make to facilitate students' small-group talk – even though 'modeling' was included in the Teacher Move Taxonomy by Wei et al. (2018) as well as in Eva-Wood's work, the latter did not address it as a discourse move, but as an instructional strategy for a particular mode of reading.

Other interventions did not shed light on teachers' roles, neither in terms of guiding students' talk, nor in terms of facilitating or providing feedback on their writing or role-playing. Researchers often dispensed the interventions (Halász, 1991; Vezzali et al., 2012) or provided no information on how teachers interacted with students during group talk, writing tasks or role-playing (Adler & Foster, 1997; Stevahn et al., 1996; 1997). White (1995) described the questions that teachers asked during classroom dialogues, but did not describe how teachers guided students' autobiographical writing (e.g., if students encountered difficulties or were hesitant to complete the task). All in all, the data in this review were not fit to draw firm conclusions about teachers' roles. Therefore, our analysis remains inconclusive with regard to this aspect of instructional approaches.

### 3.2.4 *Stance toward texts*

As a final aspect of instructional approaches, we analyzed which stance toward the text can be inferred from the interventions: an efferent, expressive, or critical-analytical stance (Murphy et al., 2009).

An efferent stance was taken in one intervention (Vezzali et al., 2012); students were asked to write a summary after reading an intercultural novel. Rather than querying worldviews or beliefs underlying the text or expressing their spontaneous responses, students were asked to identify specific information from the text (i.e., key events). An expressive stance toward the text prevailed in two other studies (Eva-Wood, 2004; Halász, 1991). In both interventions, students were invited to express their spontaneous, personal, affective responses to a text verbally and/or in writing.

A critical-analytical stance toward the texts was taken in four interventions (Malo-Juvera, 2014; 2016; Stevahn et al., 1996; 1997). Malo-Juvera asked students to critically interrogate texts in terms of ideas, assumptions, and worldviews regarding the sexual harassment and sexual orientation they presented. The tasks prompted students, for instance, to evaluate whether a protagonist was telling the truth and to compare their evaluations with other students. According to Malo-Juvera, students were invited to explore "moral rea-

soning about sensitive topics" (p. 421). Stevahn et al. asked students to search the text for conflicts and to reason about how these conflicts could be solved by characters; as such, students used the text rather instrumentally to deepen their understanding of conflicts and to enhance their skills at solving them.

In two interventions, an expressive and critical-analytical stance appeared to be combined. Even though the autobiographical writing task implemented by White (1995) was completed before the text was even introduced, students were asked to connect their own experiences to issues presented in the texts, which pointed toward an expressive stance toward the texts. Yet, the ultimately purpose appeared to be that students would better understand the characters and their behaviors, which suggests a critical-analytical stance. Adler and Foster (1997), although they provided little information about the intervention, asked students to express the feelings that the text evoked by writing individual journal entries and by constructing a group collage – this suggests an expressive stance. However, students were also invited to take a more analytical stance toward the theme of caring for others, by investigating social relationships among characters and by creating a mind map to visualize them.

All in all, the existing categorization suggests that taking both an expressive and a critical-analytical stance toward texts seems most promising for fostering students' insight into human nature. Indeed, taking an expressive stance may foster students' insight into themselves (Halász, 1991), fictional characters (Eva-Wood, 2004; White, 1995), and real-world others (Adler & Foster, 1997). The potential relationship between a critical-analytical stance and intervention effects was more evident. In one case, students' insight into fictional characters was fostered (White, 1995); in five other interventions that featured this stance, students' understanding of, views on, or intended behavior toward real-world others were affected (Adler & Foster, 1997; Malo-Juvera, 2014; 2016; Stevahn et al., 1996; 1997).

#### 4 DISCUSSION

We set out to examine whether and how literature education may foster adolescents' insight into human nature. Overall, we included 13 intervention studies in this review. Nine of these studies provided full or partial empirical support for the expected effects on students' insight into human nature, compared to four studies in which no empirical support for the expected effects was found. All in all, this review suggests that literature teaching, under certain conditions, may foster students' insight into human nature.

One intervention fostered students' insight into themselves in terms of their personal memories as evoked by a literary text (Halász, 1991), whereas two interventions affected students' understanding of fictional characters (Eva-Wood, 2004; White, 1995). Finally, six studies indicated that literature teaching could foster students' understanding of, views on, or intended behavior toward real-world others, in terms of caring for others (Adler & Foster, 1997), sexual harassment and sexual orientation (Malo-Juvera, 2014; 2016), conflict resolution (Stevahn et al., 1996; 1997), and immigration (Vezzali et al., 2012).

It should be noted that the three categories distinguished in this review merely served as a means to analyze what kind of intervention effects researchers expected to demonstrate. From a conceptual point of view, the overarching term *insight into human nature* clarifies that it is virtually impossible to make clear distinctions between the terms "self" and "other" (see Zahavi, 2014) or between fictional and real-world others (see Mar & Oatley, 2008). For example, Malo-Juvera (2016) asked students to write and to talk about which sexual orientation the protagonist might choose if he had a choice. Such a prompt asked students not only to consider the perspective of a fictional character but also (by extension) their own views on sexual orientation. In doing so, students might rely on their knowledge of real-world people and situations to shape their opinions; they might also see the protagonist as a representative of human beings in the real world.

In addition to analyzing what gaining insight into human nature may entail in the literature classroom, our review addressed which instructional approaches may particularly foster this insight in adolescent students. Based on instructional approaches for which empirical support was found, we will identify a set of instructional design principles, which may be used as guidelines for classroom practices and future interventions. Subsequently, we discuss the limitations of the current study and offer suggestions for future research.

#### 4.1 *From Instructional Approaches to Design Principles*

In this review study, we considered design principles to be parameters for future intervention design that increase the likelihood of a particular objective being achieved, which can be captured in an if/then-statement. Based on the analysis of instructional approaches, we formulate the following statement, containing three individual design principles:

*If we want to increase the likelihood that adolescent students gain insight into human nature in the literature classroom, we are best advised to:*

1. *Select fictional texts such as novels, short stories, passages, or poems, that are thematically relevant for the intended outcomes of the intervention;*
2. *Design writing tasks related to fictional texts and text themes that prompt students to (a) activate previous personal experiences before reading, (b) notice and annotate their experiences during the reading process, and/or (c) reflect on evoked experiences directly after reading; and*
3. *Design exploratory dialogic activities that stimulate students to verbally share their personal experiences related to fictional texts and text themes.*

In this section, we discuss why using each principle as a guideline may lead to designing literary instruction that increases the likelihood of fostering students' insight into human nature.

*The principle of text selection.* The first principle suggests that we may increase the probability of fostering students' insight into human nature by selecting fictional texts that are thematically relevant for the intended outcomes of the intervention. We found that researchers used an array of fictional texts, including young adult novels, short stories, and poetry. These texts were often thematically relevant for the intended outcomes of an intervention, for example, when a minority group was represented by characters (e.g., Malo-Juvera, 2016; Vezzali et al., 2012) or when a particular kind of behavior or social relationship played a prominent role (Adler & Foster, 1997; Malo-Juvera, 2014; White, 1995).

Two lines of reasoning explain why thematically relevant fiction may foster insight into human nature. First, from a psychological perspective, indirect contact theory (Turner et al., 2007; see Vezzali et al., 2012) states that reading fictional texts is a form of indirect, imagined contact which has similar positive effects on intergroup attitudes as direct contact, while also producing less anxiety (Turner, Hewstone, Voci, & Vonofakou, 2008). Therefore, reading fictional texts may positively alter people's personal views on and attitudes toward real-world others. Second, from the perspective of empirical literary studies, Mar and Oatley (2008) aptly state "the function of fiction is the abstraction and simulation of social experience" (p. 173), which "facilitates the communication and understanding of social information and makes it more compelling, achieving a form of learning through experience" (ibid.). Vividly experiencing a simulation of social life through reading fictional texts may thus help readers to better understand their own lives and the lives of others.

Notably, researchers did not attend to the possibility that social or moral themes in fiction may be sensitive issues in students' lives, such as family relationships (White, 1995), sexual harassment (Malo-Juvera, 2014), sexual orientation (Malo-Juvera, 2016), and immigration (Vezzali et al. 2012). Creating a safe learning environment to talk about such themes was not addressed in the reviewed intervention studies. Students' sense of safety may be considered a default, that is not always critically reflected upon (Boostrom, 1998).

Another aspect of text selection that researchers hardly considered, except for Eva-Wood (2004), was whether selected texts were "literary" texts as indicated, for example, by language use that deviates from conventional language use (Van Peer, Zyngier, & Hakemulder, 2007) or by "gaps" that the reader must fill (Iser, 1980). Yet, it may be worthwhile to select texts that are considered to be literary because the concept of literariness may explain the impact that fictional texts have on readers' sense of self and their social perceptions (e.g., Hakemulder, Fialho, & Bal, 2016).

*The principle of writing about personal experiences.* The second principle suggests that designing writing tasks related to fictional texts and text themes, that prompt students to (a) activate previous personal experiences before reading, (b) notice and annotate their experiences during the reading process, and/or (c) write reflective responses directly after reading, may increase the likelihood of fostering students' insight into human nature.

Most interventions included writing tasks, either as stand-alone activities or combined with dialogues. Reflective response writing after finishing a story or novel or after reading a well-delineated excerpt occurred most often (Adler & Foster, 1997; Malo-Juvera, 2014; 2016; Vezzali et al., 2012), but we also observed prereading writing tasks that activated previous personal experiences (White, 1995; Malo-Juvera, 2016) and tasks in which students annotated responses during the reading process (Eva-Wood, 2004; Halász; 1991). Such writing tasks prompted students to activate, notice, and reflect on personal experiences related to a story theme (e.g., thoughts, feelings, memories, questions, and associations). These experiences may stem from students' own lives as well as from secondary sources such as situations they have heard or read about or have seen in a movie (Halász, 1991; White, 1995). In the interventions with empirical support, writing tasks pointed to an efferent stance toward texts (Vezzali et al., 2012), an expressive stance (Eva-Wood, 2004; Halász, 1991), a critical-analytical stance (Malo-Juvera, 2014; 2016), or a combination of the latter two (Adler & Foster, 1997; White, 1995).

Research on learning processes has demonstrated the importance of establishing a meaningful basis of prior knowledge in which new information can be

embedded (Dochy, Segers, & Buehl, 1999; Merrill, 2002; Pressley et al., 1992). This basis of prior knowledge, we argue, may well include previous life experiences. From a more domain-specific perspective, writing prior to reading may enhance students' emotional involvement in a text (Janssen & Braaksma, 2016). In addition, writing tasks assigned during and directly after reading may help students to engage in internal dialogues with the texts. This line of reasoning traces back to Rosenblatt's transactional theory of reading (1938/1968), which outlines how activities like engaging, constructing, and imagining are part of people's reading experience. Beach (1993) built upon Rosenblatt's work and argued for experiential approaches to teaching literature that stimulate these aspects of the reading experience, which appear to be crucial for reading experiences to impact oneself and oneself in relation to others (Fialho, 2018; Fialho, Hakemulder, & Bal, 2016). All in all, completing writing tasks may function as "writing-to-learn" (Klein, Boscolo, Kirkpatrick, & Gelati, 2014). If it allows students to draw upon personal experiences in particular, writing-to-learn is considered a valuable activity in literature classrooms (Newell, 1996).

*The principle of verbally sharing personal experiences.* The third principle is based on the most salient type of tasks in interventions with empirical support, and suggests that the likelihood of fostering students' insight into human nature may be increased if we design exploratory dialogic activities that stimulate students to verbally share their personal experiences related to fictional texts and text themes.

In most interventions, dialogues followed after students read a particular text and completed individual writing tasks about personal experiences in relation to that text (Eva-Wood, 2004; Malo-Juvera, 2014; 2016; Stevahn et al., 1996; 1997; White, 1995). Most external dialogues appeared to be *exploratory* in nature. Students were asked to express and to compare their personal experiences in relation to the text and its theme (Adler & Foster, 1997; Eva-Wood, 2004; Malo-Juvera, 2014; 2016). By verbally sharing their experiences, students can form a connection between the internal dialogue they have with the text to external dialogues with others. These external dialogues may take place in small groups or as whole-class dialogues or in combination, where the latter follows the former (Malo-Juvera, 2014; 2016). Such a buildup, from the individual to peer groups to the classroom level, creates multiple layers of sharing responses, interpretations, and perspectives. External dialogues may imply taking an expressive or a critical-analytical stance toward the text, or combining both.

The effectiveness of sharing experiences in groups resonates with a constructivist perspective on teaching and learning (Vygotsky, 1978) and with dialogic learning theories (Barnes, 1976; Mercer & Dawes, 2008; Nystrand, 1997). In the context of the literature classroom, this implies that students benefit from responding to texts as authentically as possible, using their own language. Moreover, theories of reading that trace back to Rosenblatt (1938/1968) suggest that fictional and literary texts by nature allow for multiple interpretations to be constructed by readers. Because reading has been theorized to be an inherently social activity (Beach, 1993; Steen & Schram, 2001), readers may share ideas, experiences, and interpretations that relate to themselves and others in the social domain of the classroom.

If reading remains an individual activity, it is confined to a single reader's experiences of feelings evoked by a text, imaginations of what it would be like to be in the position of a character, questions that come up while reading, and so forth. Talking to peers about such experiences adds another layer; it offers students the opportunity to verbalize and thus to consider a wider array of thoughts, questions, feelings, ideas, and perspectives. By implication, if literature teaching is expected to foster students' insight into human nature, social aspects of learning and reading can hardly be ignored.

#### *4.2 Interventions Without Empirical Support*

The design principles are based on the analysis of instructional approaches in interventions with full or partial empirical support. However, when reviewing the interventions without empirical support, we concluded that their instructional approaches were also in line with the design principles. For example, thematically relevant fictional texts were used (Darragh, 2015; Hakemulder, 2008), writing tasks were applied (Darragh, 2015; Hakemulder, 2008; Henschel et al., 2016), and students were asked to engage in dialogues (Garrod, 1989).

One might argue that all interventions, either with or without empirical support, were rather alike in their instructional approaches. Therefore, it may seem invalid to consider these instructional approaches as being informative for designing future literature classroom interventions. However, numerous methodological or contextual reasons may explain why four out of 13 studies found no empirical support for their interventions. For example, the instruments used may not have been apt for capturing these effects, an intervention may not have been implemented as originally intended, or the contrast between the experimental and comparison condition may not have been large enough. Such reasons cannot be determined with certainty in this review be-

cause sufficient information about instruments, coding schemes, implementation fidelity, and comparison conditions was not always provided in the studies without empirical support. These shortcomings repeatedly resulted in rather low scores on methodological quality indicators.

Moreover, the similarities in instructional approaches of interventions with and without empirical support suggest that theoretical underpinnings were rather similar across all 13 studies. Due to methodological and contextual factors, empirical support for interventions based on these theoretical notions cannot be guaranteed, but the overlap in instructional approaches does seem to strengthen the selection of included studies from a theoretical point of view. In our view, these similarities are no cause for concern; rather, they suggest there were solid theoretical grounds for the design principles inferred from those interventions that did provide full or partial empirical support.

#### 4.3 *Limitations*

Both this review study and the studies included in it are subject to limitations. First, as with all review studies, comprehensiveness may have been at stake. Although we systematically searched databases and used citation tracking, hand searches, and consulted experts, relevant publications may have escaped our attention. Because we limited our search to peer-reviewed results, our selection may have been subject to publication bias. However, the search results returned several relevant studies with statistically nonsignificant outcomes that were published in peer-reviewed journals.

The included intervention studies were not without limitations either. These came to light via quality assessment procedures, which also informed the admissibility of empirical support presented in the studies. As such, thorough methodological quality appraisal functioned as a gatekeeper for the overall validity of this review study.

Implementation fidelity was a major issue in the included intervention studies. Although it is crucial to know whether interventions are implemented as intended (O'Donnell, 2008), few studies sufficiently accounted for it. In addition, our review remains inconclusive regarding the role of the teacher. In the majority of the included studies, descriptions of teachers' roles were insufficient. The two researchers who addressed teachers' roles (Malo-Juvera, 2014; Eva-Wood, 2004) both seemed to suggest that the teacher should allow dialogic discourse in the classroom, but operationalized this in different ways. Moreover, in neither of these studies, teachers' subtle discourse moves were addressed, even though such moves may be 'influential in promoting or hin-

dering students' learning outcomes' (Wei et al., 2018, p. 579). All in all, our review yielded too little information to formulate a design principle about teacher–student interactions.

Finally, one might argue that a limitation of our study lies in the fact that nearly half of the included intervention studies were identified in the search expansion phase rather than via database searches. However, we would suggest that this study demonstrates the importance of conducting a search via a variety of sources: rather than settling for search results from databases, future researchers are advised to include citation tracking procedures, hand searches and expert consultation. In conclusion, reviewing previous intervention studies may be troublesome in various ways. Nonetheless, it remains an important step in gaining insight into evidence-based educational practices.

#### 4.4 *Future Research*

The search and screening procedures of this review indicate that few studies on gaining insight into human nature in the literature classroom used experimental or quasi-experimental designs. This points to a need to expand this type of research in order to shed further light on whether and how literature education may foster insight into human nature. In addition, this review remained inconclusive about the roles of teachers. As such, future studies should further develop ways to describe teacher–student interactions and investigate their effects on students' insight into human nature. Finally, future intervention studies should further explore whether or not only fictional but also literary texts impact students' insight into human nature. Such studies may, for example, use available indexes of foregrounding and literariness (Miall & Kuiken, 1994; Shen, 2008) to compare literary reading to nonliterary reading conditions.

#### 4.5 *Conclusion*

One of the potential values of literature education is its capacity to foster young people's reflections on how they position themselves in the world with respect to others. Our review critically investigated whether and how literature education may foster adolescents' insight into human nature. Analysis suggests that this insight may be developed by reading and responding to fictional texts in the literature classroom. Moreover, our study sheds light on design principles based on empirically supported instructional approaches. Students' insight into human nature may be fostered if they read thematically relevant fictional texts and participate in writing activities that focus on activating, annotating, and reflecting on personal experiences in relation to fictional texts and themes.

Doing so may prepare students for exploratory small-group and/or whole-class dialogues, in which experiences are verbally shared. By identifying these principles, we hope that this study functions as a stepping stone for those who wish to design literary instruction to foster students' insight into human nature.



## CHAPTER 4

### DESIGNING A LITERATURE CLASSROOM INTERVENTION TO FOSTER 10<sup>TH</sup> GRADE STUDENTS' INSIGHT INTO HUMAN NATURE

We describe the design of a literature classroom intervention for 15-year-old students in the Netherlands, which aimed to foster their insight into human nature – insight into themselves, fictional others, and real-world others. The design was informed by a theoretical-empirical model of transformative reading, an explorative study in Dutch literature classrooms, and three initial design principles identified in a review of previous intervention studies. We investigated the effects of an iterative design process on two quality indicators – validity and practicality – and on the initial design principles underlying the intervention, to contribute to theory and classroom practice. A first intervention was developed in collaboration with teachers, tested in trial studies, and taught by 13 teachers to 22 classes. In four lessons, students focused on internal and external dialogues with and about short stories addressing various social-moral themes. We assessed validity and practicality by using implementation and evaluation data from teacher logs, time on task observations, students' evaluations, and teachers interviews. Suggestions for improvement were derived from the data (e.g., make learning objectives more explicit, select stories centering around a single social-moral theme). A redesigned intervention was taught by 6 teachers to 6 classes. Teacher and student data supported the validity and practicality of the redesigned intervention and yielded sub-principles for operationalization of the initial design principles. All in all, this study suggested that an iterative design process, grounded in previous empirical research and utilizing both teacher and student data, may result in the design of valid and practical domain-specific interventions.

#### 1 INTRODUCTION

Intervention studies in literature classrooms empirically evaluate whether a particular instructional approach helps students to achieve predetermined objectives, such as improving their interpretative skills (Janssen, Braaksma, & Couzijn, 2009; Levine & Horton, 2013) or rethinking certain social-moral attitudes (Malo-Juvera, 2014; 2016). Researchers usually develop such an instructional approach in an educational design research project, ideally in close collaboration with teachers. Educational design research has been conceptualized as consisting of three phases (Plomp, 2013): the preliminary research phase, in which relevant

literature is reviewed and a theoretical framework is built; the development phase, in which an intervention is developed, improved and refined; and the assessment phase, in which its implementation and effectiveness are evaluated, compared to predetermined specifications. Following these phases increases the probability of designing high-quality interventions.

Researchers have established several quality indicators for interventions: validity, practicality, effectiveness, sustainability and replicability. First, Nieveen (1999) distinguished between content validity, which dictates that the components of an intervention should be based on state-of-the-art knowledge and should be relevant to those using the intervention, and construct validity, which means that all components should be consistently linked. Second, for the intervention to be practical, teachers must consider it to be usable and use it in a way that is generally compatible with the designers' intentions (see O'Donnell, 2008). Third, Nieveen indicated that high-quality interventions should result in the desired outcomes: the intervention should be effective. Rietdijk, Janssen, Van Weijen, Van den Bergh and Rijlaarsdam (2017) pointed to continued use of the developed intervention after the research project (sustainability). Finally, Rijlaarsdam, Janssen, Rietdijk, and Van Weijen (2017) called for more comprehensive descriptions of interventions that strengthen their replicability. Metaphorically speaking, interventions all too frequently remain "black boxes": it is unclear what happened in the classroom, why it happened, and how what happened was developed. This not only threatens the validity and replicability of interventions, but also hampers detailed insights into domain-specific instructional activities that are designed to achieve particular aims (see also Schrijvers, Janssen, Fialho, & Rijlaarsdam, 2018).

With this paper, we aim to respond to the call by Rijlaarsdam et al. (2017) for more comprehensive and replicable intervention descriptions, by describing in detail the iterative design process of an intervention for 10<sup>th</sup> grade literature classrooms that aims to foster students' insight into human nature – insight into themselves, fictional others, and real-world others. We focus on validity and practicality; effectiveness and sustainability are beyond our current scope, as their evaluation requires different types of studies (e.g., quasi-experimental or longitudinal). As Figure 4.1 shows, the design project included a preliminary research phase, two development phases and two subsequent assessment phases (Plomp, 2013). Both development phases were informed by the preliminary research phase, and development phase 2 was additionally informed by the results of assessment phase 1.

We operated on the micro level of curriculum design (Van den Akker, 2013): the level of the classroom and the instructional materials and strategies used in

it. The intervention was designed for 10<sup>th</sup> grade of the higher general secondary education track in the Netherlands, which is the second highest track in Dutch secondary education and prepares for higher vocational education but not for university.

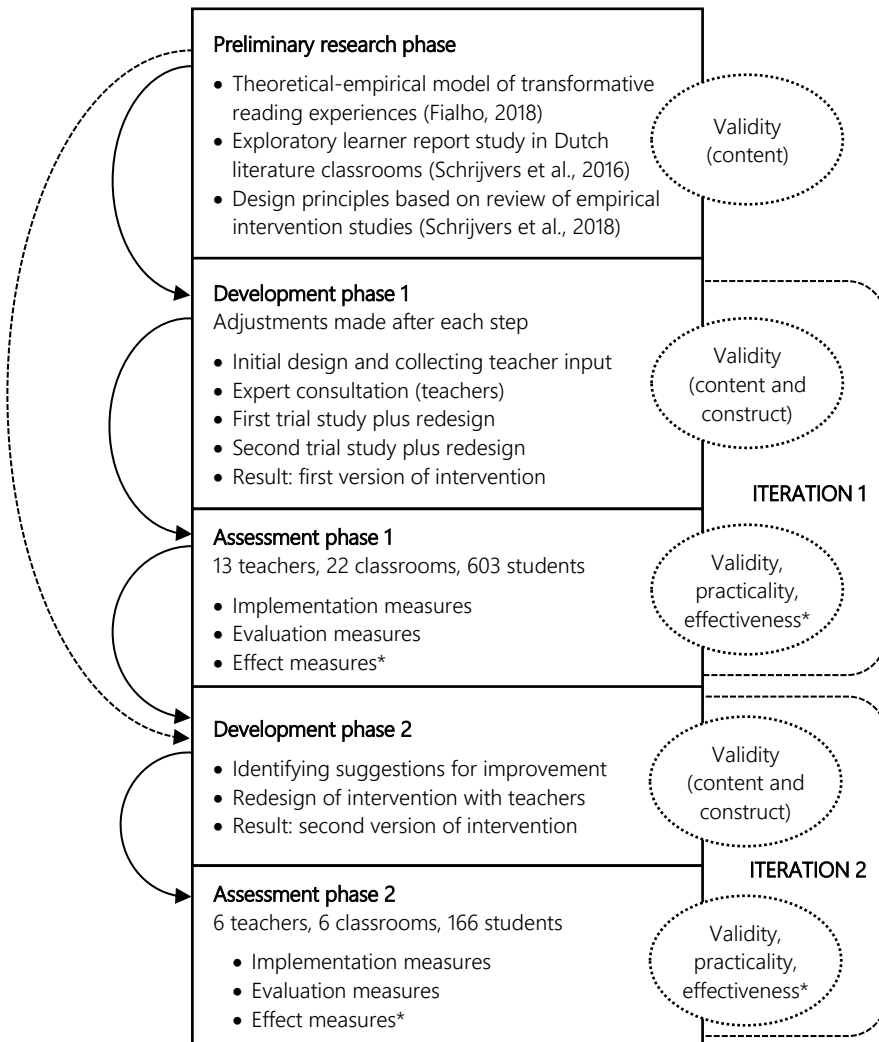


Figure 4.1. Overview of the design process (\* see Chapters 5 and 6).

In Dutch schools, literature education usually does not have the status of a separate subject. It is a sub-domain within Dutch language classes, like writing and rhetoric. In lieu of regulations, standardized tests and nationwide exams, teachers have much freedom in selecting literary texts and designing tasks and instructions. Students are merely required to work toward three intertwined objectives: acquiring literary-historical knowledge, utilizing structural-analytical skills, and reflecting on their literary reading experiences and development (Dutch Institute for Curriculum Development, 2012). The current design project is related to the third objective, as its focus is on particular literary experiences in relation to “human nature”, for example, relating a story theme to the own life, or considering how thoughts, feelings and behaviors of characters in fictional situations represent human responses to similar real-life situations.

Figure 4.1 summarizes the design process. We first outline preliminary research results, including three initial design principles. In the Method section, we describe the design process and instruments used to assess the implementation and evaluation of two subsequent versions of the intervention, which yielded information about their validity and practicality. We then present the two versions of the intervention and their implementation and evaluation outcomes, and indicate how the design process affected the initial design principles.

### 1.1 *Outcomes of Preliminary Research Phase*

The preliminary research phase consisted of the development of a theoretical-empirical model of transformative reading (Fialho, 2012; 2018), an exploratory learner report study in Dutch literature classrooms (Schrijvers, Janssen, Fialho, & Rijlaarsdam, 2016), and a review study that resulted in a set of design principles upon which the intervention construct was based (Schrijvers et al., 2018).

*Theoretical-empirical model of transformative reading.* One of the acclaimed merits of reading fictional and literary texts is that, via processes of empathy and reflection, it offers readers insights into who they are, how they position themselves in the world, and how they see themselves in relation to other human beings (for overviews, see Koopman & Hakemulder, 2015; Hakemulder, Fialho & Bal, 2016). In short, it may offer readers “insight into human nature”. This mode of reading has been conceptualized as a ‘transformative experience’, because it may alter readers’ perceptions of themselves and themselves in relation to others (Fialho, 2012).

In a theoretical-empirical model, Fialho (2018) distinguished two outcomes of transformative reading – insights into oneself and into others – and identi-

fied six underlying components. Adult readers who talked in Fialho's phenomenological interview studies about reading experiences that had a transformative impact on them, indicated that they vividly imagined the setting and characters in a story (*imagery*), recognized something of themselves or others in characters (*identification*), enacted and embodied the experiences of a character (*experience-taking*), evaluated characters positively or negatively (*character evaluation*), felt sympathy and compassion for characters (*sympathy*), and noticed which words, phrases or sentences were particularly striking to them (*aesthetic awareness*). For adult readers, these particular experiences preceded new or deeper insights into themselves and others (*self-other insights*).

Whereas the transformative reading model (Fialho, 2012; 2018) has not yet been validated for adolescent readers, studies suggested that similar experiences may occur in them, even though researchers used other terms than "transformative reading" or "insights into self and others", or worked with non-literary texts. For example, adolescents were found to consider their possible future selves as a result of fiction reading: they reflected on who they would (not) like to become (Richardson & Eccles, 2007). In addition, they were found to compare their own lives to story situations and to experience empathetic engagements with characters' feelings (Charlton, Pette, & Burbaum, 2004), as well as to regard fiction reading as a way of understanding others' experiences, which made them feel connected to others or offered them new options for their own lives (Rothbauer, 2011). Therefore, it seems likely that adolescents may as well engage in a "transformative" mode of reading.

*Exploratory learner report study.* In an exploratory, descriptive study in which we asked students to complete a written learner report, we found that learning about self and others was among the learning outcomes in Dutch upper secondary literature classrooms (Schrijvers et al., 2016). In their learner reports, students reported that literature education, for instance, made them learn about their own and other people's personalities, relations and behavior, consider their future selves, and identify life lessons in literary texts. Such experiences occurred more frequently if students' teachers reported that they allowed for more student autonomy and interaction in the classroom.

Literature education may thus foster students' insight into human nature, departing from themes and issues raised in texts. This potential learning outcome appears to be valued by curricular organizations and teachers in the Netherlands. For example, a team of Dutch language and literature teachers, who are working on an intended curriculum reform, suggested that literary reading may familiarize students with other worlds, contributes to moral devel-

opment, and helps them to think about people's choices, about themselves, others, and the world (Curriculum.nu, 2018a). Moreover, teachers reported to consider fostering students' personal growth or personal development as an important aim of literature teaching (Janssen, 1998; Oberon, 2016). However, little is known about *how* literature education might foster these learning outcomes.

*State-of-the-art: design principles.* A review of intervention studies shed light on instructional approaches that may foster students' insight into human nature in the literature classroom (Schrijvers et al., 2018). Studies were included in this review if they aimed at fostering some form of insight into self, fictional others, or real-world others, were conducted in regular, first-language secondary education classrooms, and used (quasi-)experimental research designs. Thirteen studies were included. Nine studies provided empirical support for fostering students' insight into human nature (e.g., Adler & Foster, 1997; Malo-Juvera, 2014; 2016; Vezzali, Stathi, & Giovannini, 2012; White, 1995), from which three design principles were derived.

The first principle suggested that fictional texts should be read that are thematically relevant for the aim of an intervention. If fictional texts address relevant social situations, readers may consider how they would position themselves in those situations and how they would impact themselves and others (see Mar & Oatley, 2008). As researchers seldomly considered whether the fictional texts used would be "literary" texts, the review remained inconclusive as for how to conceptualize "literature" and whether it might successfully foster students' insight into human nature.

The second principle suggested to design writing tasks, related to texts and themes, that stimulate students to activate previous personal experiences before reading, to notice and annotate responses during reading, and/or to write down responses or reflections directly after reading. These tasks stimulate students to engage in an *internal dialogue* with the text (see Janssen et al., 2009). Such dialogues may generate transactional processes of meaning-making (Rosenblatt, 1938/1983), because students may become aware of the responses evoked by texts and of how these responses are related to the outer-textual world.

The third principle suggested to design exploratory dialogic activities that stimulate students to verbally share personal experiences related to texts and text themes. Responses and reflections noticed in internal dialogues, thus, may be shared in *external dialogues*, which allows for exploring multiple perspectives on a text and the issues it addresses. This may alter students' opinions,

views and perceptions of themselves and others, or offer them new ones. Such dialogues may take place in pairs, small groups, or the classroom.

In short, the preliminary research phase yielded a central premise for the intervention design: to optimize the probability of fostering their insight into human nature, students should be invited to engage in internal and external dialogues with and about texts and to focus on transformative reading experiences such as imagery, identification and sympathy.

### 1.2 *Aims and Research Question*

Designing an intervention can be compared to designing a questionnaire. Both processes start by reviewing relevant literature, conducting exploratory work, developing a theoretical model, and, in case of an intervention, identifying design principles. In a first iteration, an initial construct is designed: an intervention prototype, or a set of questionnaire items. In practice tests, data are collected to assess validity and practicality (of interventions) or reliability (of questionnaires). If these are unsatisfactory, a second iteration will follow: the construct is adjusted and tested again. Iterations continue until a valid, practical intervention or reliable questionnaire is constructed. Just like questionnaire design may contribute to theory about the measured variable, intervention design may add to the state-of-the-art by further specifying the initial design principles in terms of their operationalization in practice.

The primary aim of this study was, therefore, to design a literature classroom intervention for 10<sup>th</sup> grade students, in which the model of transformative reading, the text selection principle, the internal dialogue principle, and the external dialogue principle were integrated. Because validity – both at the content and construct level – and practicality are considered to be amongst the indicators of the quality of an intervention (Nieveen, 1999; O'Donnell, 2008), the study was guided by the following research question:

To what extent is the intervention that is designed a valid and practical instructional approach for upper secondary literature classrooms, according to both students and teachers?

An additional aim of the study was to contribute to the state-of-the-art by reflecting on how the initial design principles are affected by their operationalization in the classroom.

## 2 METHOD

### 2.1 *Design Process and Participants*

The two iterations in the design process both consisted of several steps (see Figure 4.1), which we describe below. For both assessment phases, we also describe the sample of teachers and students involved, as we collected implementation and evaluation data from them to assess validity and practicality.

#### 2.1.1 *Development phase 1*

*Initial design.* As the research team, we collected teachers' input via email, asking for practical suggestions, for instance, concerning the number of units, as well as for examples of suitable short stories and tasks. This led to the decision to use short stories that could be read in about 10-15 minutes, to ensure that reading could be accompanied by both internal and external dialogue activities within a single unit. Teachers also indicated that devoting four lessons (of 50 minutes, the conventional length at Dutch secondary schools) to the project would be preferable, in view of their regular teaching program. We therefore decided that each unit should fit within a single lesson. We further discussed how transformative reading experiences could be integrated in the units and designed a prototype of the intervention.

*Expert consultation.* The prototype was discussed in a meeting with four experienced teachers. This resulted in two major changes, concerning text choice and external dialogues. First, we originally considered offering students freedom of choice in reading materials, because this might increase their engagement (e.g., Lenters, 2006). However, the research literature is ambiguous at this point: in only one out of the nine intervention studies on which the text choice principle was based, students chose their reading material (from a preselected list; Vezzali et al., 2012). Moreover, the consulted teachers suggested that if students would read the same stories, this would enable "deeper" talk about story themes and related insights into themselves and others. Second, the teachers suggested to end the intervention by a teacher-led classroom dialogue in which students would discuss what they had learned, to establish closure and to talk about how students could apply what they learned in future literature lessons. We adjusted the units according to their suggestions.

*First trial study.* One female teacher taught each of the four units separately to four different classes. The first author observed these lessons. Overall, the teacher felt that teaching the units was feasible and that students were en-

gaged in the tasks. To further operationalize the external dialogue principle, she suggested to include, in unit 1, a task with “questioning cards” to guide students in using follow-up questions. As the first author also observed that students struggled with follow-up questions, a task was designed accordingly. Related to the text choice principle, the teacher questioned one story: she felt its theme was not familiar enough for students. We decided to assess the appropriateness of this story again in a second trial study. Some other minor adjustments were made, mainly to materials and organization (e.g., bundling tasks in a booklet; ways to compose small groups of students).

The teacher also commented on a first version of the teacher guideline she received. She felt it was clear and manageable, but suggested to include information about the background, goals and structure of the units, as well as concrete suggestions for teachers’ roles and their interaction with students. We cross-checked her suggestions with the second trial teacher.

*Second trial study.* Another female teacher, at another school, taught the units to a single class, in four lessons over the course of two weeks. Again, the first author observed. The trial resulted in two adjustments following up on the previous ones. The story we specifically attended to was replaced, as the second teacher also felt it was unclear for students. Further, we designed an observation task in unit 1, in which teachers would demonstrate using follow-up questions during a dialogue about reading experiences, as the teacher felt that using the questioning cards did not suffice. Additionally, she confirmed the suggestions for the teacher guideline, which was revised accordingly.

All in all, consulting experts and conducting trials led to adjusting the initial design in operationalizing the text choice and external dialogue principle. The trial teachers were satisfied with the operationalization of the internal dialogue principle.

### *2.1.2 Assessment phase 1*

Assessment focused on implementation and evaluation of the intervention, which we call Transformative Dialogic Literature Teaching, version 1 (for short: TDLT-1). We assessed content validity by searching evaluation data for endorsements and skepticism regarding the intervention and its relevance. Construct validity was assessed by focusing on data concerning the coherence and clarity of the units. Finally, we assessed practicality by analyzing whether units were implemented as intended, and by focusing on data that concerned feasibility. We collected data of teachers and students. Whereas validity and practi-

quality of interventions has previously been conceptualized mainly from teachers' perspectives – for example, when assessment is conceptualized as '[evaluating] whether target users can work with the intervention and are willing to apply it in their teaching' (Plomp, 2013, p. 30; italics added) – we regarded students as equally important stakeholders, who can provide valuable information for a redesign.

The phase started with individual preparatory meetings with thirteen teachers from six schools across the Netherlands, for a walk-through of the teacher guideline and the units. They then taught TDLT-1 to either one or two classes, in two to four weeks. The teachers had 13.3 years of teaching experience; two of them were male. Student participants ( $N = 603$ , from 22 classes) were on average 15.9 years old; 52.5% was female. Their parents received a passive consent letter and could object to their child's participation; none of them withheld their consent.

### *2.1.3 Development phase 2*

By analyzing data from assessment phase 1, we identified suggestions for improvement. These were discussed during a meeting of the first author and three teachers who had been involved in the first assessment phase, resulting in adjustments of the operationalization of the initial design principles. For example, a suggestion for improvement was that students needed more time to learn to engage in internal and external dialogues; therefore, it was decided that the redesigned intervention, called TDLT-2, would consist of six units. The first author adjusted the teaching and learning materials and returned them to the teachers for final comments, which only led to some improvements in formulation.

### *2.1.4 Assessment phase 2*

Six teachers taught TDLT-2 to a single class: those involved in development phase 2, as well as three new teachers. They had on average 18.7 years of teaching experience; all were females. The new teachers participated in a workshop led by the first author, which consisted of: (a) information about the theoretical background of the intervention, (b) a walk-through of the material, (c) an exercise to put themselves in the role of students when writing down first responses to a story, (d) practicing to give feedback on students' dialogues, and (e) time for questions. One of the teachers involved in the redesign was present and discussed, for example, student talk she had observed and challenges she faced in giving feedback on dialogues.

Again, we collected implementation and evaluation data from teachers and students. TDLT-2 was taught to 166 students of six classes in four schools. They were on average 15.5 years old; 49.2% was female. Parents were again asked for passive consent. After the second assessment phase, we reevaluated validity and practicality, after which a third design iteration was not deemed necessary.

## 2.2 *Instruments and Data-Analysis*

In design research, triangulation of data sources and data collection methods, as well as empirical testing of the practicality of interventions is important for data interpretation (McKenney, Nieveen, & Van den Akker, 2006). Therefore, we used five instruments to collect data of teachers and students: (a) teacher logs, (b) time on task observations, (c) teacher interviews, (d) student evaluation forms, and (e) student evaluation tasks. Table 4.1 shows the indicators of validity and practicality (appreciation and relevance, structure and coherence, clarity and feasibility) in which the instruments provided insight, whether they were used to assess TDLT-1, TDLT-2, or both, and how many responses and observations were collected. Below, we provide further details of instruments and analysis.

*Teacher logs.* Teachers completed one online log per unit, which consisted of several phases; for each phase, teachers indicated whether it was fully, partly or not completed. If they did not complete a phase, they indicated the reason (e.g., "not enough time", "forgot about it") and were asked to add elaborations. We also asked teachers to evaluate fully or partly completed phases. They indicated on 5-point scales (letters A-D refer to items in Table 4.1):

1. How interested and engaged students seemed to be (A);
2. How clear the phase seemed for students (B);
3. How attainable it was to teach the phase (C);
4. How much was order and discipline there was in the classroom (D).

Here, too, we asked them to elaborate. We analyzed how many phases were fully, partly, and not completed and how units were evaluated. Furthermore, teachers' elaborations informed the interview guidelines.

Table 4.1. Overview of instruments and derived quality indicators

Instrument	Quality indicators	TDLT	Response rate or <i>N</i> observations
Teacher logs	<i>Practicality:</i> - Percentages fully, partly, not completed phases - Evaluation of feasibility of teaching (C) - Evaluation of maintaining order in class (D) <i>Content validity:</i> appreciation and relevance for students (A) <i>Construct validity:</i> clarity for students (B)	1	96.6% for 21 phases from 4 units, in 22 classes
		2	94.5% for 31 phases from 6 units, in 6 classes
Time on task observations	<i>Practicality:</i> percentages time on task (overall, and for intended activities)	1	<i>N</i> = 1690 in 22 visits; 1 unit in each class
		2	<i>N</i> = 877 in 12 visits; 2 units in each class
Teacher interviews	<i>Practicality, content and construct validity:</i> indications of appreciation and relevance, structure and coherence, clarity and feasibility	1	100%
Student evaluation form	<i>Content validity:</i> - Overall appreciation and relevance - Appreciation and relevance of external dialogues - Appreciation and relevance of internal dialogues - Appreciation of stories read - Sense of safety as aspect of appreciation <i>Construct validity:</i> - Overall clarity and comprehensibility - Clarity and comprehensibility of dialogue tasks <i>Additional support for indicators:</i> strengths, suggestions for improvement	1	85.4%
Student evaluation task	<i>Content validity:</i> relevance of intervention elements, story appreciation <i>Construct validity:</i> coherence between intervention elements and objectives <i>Practicality:</i> story difficulty	2	90.4% (story items) 70.5% (intervention elements)

*Time on task observations.* To assess the proportion of available learning time that students were engaged in intervention tasks, we conducted time on task observations (Karweit, 1984). Behavior was coded "on task" if students worked on the given task and did what was asked of them, for example, listening to the teacher or a peer, talking about a task, reading a story, writing in the workbook, talking to the teacher, or asking questions. We defined "off task" behavior as obvious non-learning behavior, when students were not working on the given task but were, for instance, being disruptive, talking about something irrelevant, or looking at cell phones.

We formulated two additional coding instructions: if students seemed disengaged or were waiting for a next task without being disruptive, we coded this as "on task", because they might still be listening or thinking. If we could not get a clear view on a student (e.g., if another student moved into the line of sight) we used "unclear". In addition to task behavior, we coded for each observation moment the learning activity the teacher intended at that moment: teacher-led activity (explanation, instruction); individual student task; reading and/or listening to a (read-aloud) story; dialogic pair or small-group activity; whole-class activity (presenting, discussing); or unclear.

During a classroom visit, we randomly selected six students – or, in case of group work, one student from each group – and observed them in multiple rounds. We observed the first student for twenty seconds, coded task behavior and intended activity, observed the same student for another twenty seconds, again coded task behavior and intended activity, and observed the next student. After observing all six students twice for twenty seconds, there was a one-minute break, after which the second round started, which continued until the end of the unit. If a student left the classroom, we chose a substitute. All units were observed at least once. We analyzed which proportion of observed time students were on and off task, exploring differences between units and activities.

*Teacher interviews.* The first author held semi-structured interviews with the teachers involved in teaching TDLT-1. To stimulate recall, they were encouraged to browse through materials (e.g., workbook, teacher guideline) whenever needed. Interviews were analyzed directly from notes and audio recordings. We focused in particular on experiences shared by multiple teachers. The interview consisted of three parts:

1. General evaluation of main intervention components: Teachers were asked to tell about a moment in TDLT-1 that went particularly well and one that went not as well. Next, stories, internal and external dialogue tasks, the

teacher's own role, achievement of intervention goals, teacher guidelines and the preparatory meeting were discussed.

2. Teacher-specific questions based on teacher logs. Questions were asked about the teacher's elaborations in the logs. For example, one teacher noted: "I need a lot of words to ask a student: 'Do you mean to say...?' Then he or she replies 'Yes, exactly,' and I switch to another student. I will take that more into account in other lessons." In the interview, she was asked to elaborate: why did she feel this was important, and how was she taking it into account in other units?
3. Tips for new teachers. Teachers were asked about tips and tricks for new teachers who would teach TDLT-1. Finally, they were asked if there was anything else left to discuss and they received a gift card as a token of appreciation.

*Student evaluation form.* Students evaluated TDLT-1 by filling in a form. All items were evaluated on 5-point agreement scales, unless indicated otherwise. First, students scored evaluation words that followed the phrase "I found the units...", for instance "fun", "useful", "clear" and "confusing" (10 items). Negative items were recoded. Principal components analysis with Varimax rotation ( $KMO = .87$ ; Bartlett's test  $p < .001$ ) revealed two components which accounted for 57% of the total variance:

- Overall appreciation and relevance (eigenvalue 4.27, 42.7% of variance), including words like 'useful', 'fun' and 'boring';
- Overall clarity and comprehensibility (eigenvalue 1.45, 14.5% of variance), including 'clear', 'difficult' and 'confusing'.

Next, students completed 12 items on how meaningful, difficult (items recoded) and enjoyable they found internal and external dialogues. Principal components analysis with Varimax rotation ( $KMO = .74$ ; Bartlett's test  $p < .001$ ) revealed three components, which together accounted for 55% of the variance:

- Appreciation and relevance of external dialogues (eigenvalue 3.48, 29.0% of variance), indicating how meaningful and enjoyable small-group and whole-class dialogues were, such as "Talking in small groups about stories was meaningful".
- Clarity and comprehensibility of dialogue tasks (eigenvalue 2.08, 17.3% of variance) containing the recoded difficulty items, such as "Talking in small groups about stories was difficult".
- Appreciation and relevance of internal dialogues (eigenvalue 1.09, 9.1% of variance; content validity indicator), indicating how meaningful and enjoyable internal dialogue tasks were.

Students also indicated their appreciation of the stories they read (3 items), by assigning each story a grade (1-10). When asked to indicate which story they read in unit 3 and to evaluate it, only 396 of the 515 students who completed the form did so (i.e., almost 25% left blank which story they read). Apparently, many students could not remember which one they read in the third unit.

Further, we assessed students' sense of safety as an aspect of appreciation. Four items formed an internally consistent scale (Cronbach's  $\alpha = .71$ ). Finally, we asked students to list strengths and suggestions for improvement. They listed 683 strengths and 528 suggestions. Of the latter, 36 (6.8%) concerned the research in which TDLT-1 was embedded, such as comments about effect measures (e.g., "boring", "repetitive"). Ten responses (1.9%) referred to how students themselves could improve (e.g., "Pay more attention in class"). We decided to leave out these responses, after which 482 suggestions for improvement remained. We analyzed responses inductively to see if they would support other validity and practicality indicators.

*Student evaluation task.* In assessment phase 2, to assess story appreciation and difficulty, we asked students to evaluate these aspects for the story they read for the final task of TDLT-2, which they had chosen from four options. As two of these stories were also used in TDLT-1, this was deemed to give a valid impression of story appreciation and difficulty.

In addition, students evaluated the relevance of various intervention elements and the coherence between these elements and four main intervention objectives. They indicated whether or not elements (e.g., dialogue guidelines, teachers' explanations) were helpful for achieving progress in learning to 1) notice responses during reading, 2) deal with incomprehension during reading, 3) gain insights into reading experiences, and 4) actively contribute to dialogues about stories and reading experiences. We analyzed the frequency of binary responses (helpful or not) across elements and objectives. The 117 students who responded could all have indicated that each element was helpful for each objective, thus assigning "helpful" 32 times. In practice, they were more selective and indicated 1,784 times that an element had been helpful (on average 15 indications per student).

## 3 RESULTS

3.1 *First Development Phase*

This phase resulted in TDLT-1, a four-unit intervention. Its overall objective was, as described in the teacher guideline, “to help students identify connections between short literary stories and themselves (i.e., their personalities, the way they are, the way they think) and their view on the social world (i.e., how other people are, behave and think)”. TDLT-1 consisted of one preparatory unit, aimed at understanding and applying strategies for engaging in and deepening external dialogues about stories. The subsequent units were “reading-and-dialogue” units. The goal of each unit was presented in students’ workbooks. The workbook further contained all instructions, tasks, and space to write answers and notes during group dialogues. Stories were bundled in a separate booklet. We also designed PowerPoint slides with instructions. The units of TDLT-1 are described in closer detail in Appendix C; here, we indicate how we operationalized the initial design principles.

3.1.1 *Operationalization of text choice principle*

In line with this principle – thematically relevant fictional texts should be selected – we used texts addressing peculiar, presumptuous or painful social interactions between characters (see Table 4.2 for descriptions). In response to the teacher input collected during the initial design, we selected both canonical short stories often used in 10<sup>th</sup> grade (e.g., *A plate with spaghetti* and *Blood*) and more recently published stories (e.g., *She was everywhere* and *Flight behavior*). Their thematic appropriateness was confirmed via expert consultation and trial studies. Students read a story in units 2, 3, and 4. For unit 3, they selected one story: in unit 1, they read several sentences and a description of each story and indicated which two they preferred. The teacher decided which of these a student would read, to group students for the external dialogue task in unit 3.

3.1.2 *Operationalization of the internal dialogue principle*

To operationalize the second principle – design writing tasks, related to texts and themes, that stimulate students to activate previous personal experiences before reading, notice their experiences during reading, and/or reflect on evoked experiences directly after reading – we used pre-reading tasks, reading instructions and individual reflection tasks that focused on story themes and

transformative reading experiences. As an example of a pre-reading task, students wrote in unit 4 about their ideas of an afterlife, to prepare for reading about someone in a plane crash, who has, to his own disbelief, a religious experience. Personal views were thus activated in a safe and time-efficient way. An example of a reading instruction was: "Try to pay close attention to your own responses while reading the story: which thoughts, ideas and feelings does it evoke in you? What in the story stands out to you?" Finally, as an example of reflection directly after reading, students were asked to indicate which part of the story stood out most to them (aesthetic awareness) and the extent to which the story evoked, for example, imagery and sympathy.

Table 4.2. Stories used in TDLT-1

Story	Author	Description	Unit
1. <i>Ze was overal</i> ( <i>She was everywhere</i> )	Ed van Eeden	Confused and suspicious man keeps thinking about ex-girlfriend; he ends up on the roof of a library, while people in the street below are staring at him.	2
2. <i>Een bord met spaghetti</i> ( <i>A plate with spaghetti</i> )	Adriaan van Dis	Man in restaurant thinks a black man stole his plate; he passive-aggressively confronts him, but finds out he misjudged.	3*
3. <i>De biefstuk van het zoete water</i> ( <i>The freshwater steak</i> )	Hans Dorresteyn	Boy is nervous during a fishing trip, because his father always physically punishes him several days after he has misbehaved.	3*
4. <i>Het recht</i> ( <i>The right</i> )	Annelies Verbeke	Man distrusts his black cleaning lady; he tries to trick her into stealing, but she gets the better of him.	3*
5. <i>Bloed</i> ( <i>Blood</i> )	Gerard Reve	Child is physically abused by guardian and takes revenge by causing him to take a deathly fall.	3*
6. <i>Merkwaardig verhaal</i> ( <i>Curious story</i> )	Elke Geurts	Girl is phoned by her grandmother, who instructs her to act as if she has died; awkward conversation about the cause of grandmother's made-up death.	3*
7. <i>Vluchtgedrag</i> ( <i>Flight behaviour</i> )	Bertram Koeleman	Man experiences a plane crash and, in the final moments of his life, is confronted with how religion may play a role in such circumstances.	4

Note. \* In unit 3, students read one of these five stories.

### 3.1.3 *Operationalization of the external dialogue principle*

Expert consultation and trial studies already indicated that the third principle – design exploratory dialogic activities that stimulate students to verbally share personal experiences related to texts and text themes – might be difficult to operationalize: teachers felt that students needed guidance for talking in small groups or pairs, not led by the teacher. We indicate below how the external dialogue principle was operationalized in the preparatory unit dedicated to dialogue guidelines, as well as in subsequent reading-and-dialogue units.

*Preparatory unit.* Students first considered what characterizes good and less good dialogues and were provided with five guidelines: Listen carefully, Ask follow-up questions, Postpone a first judgment, Distribute speaking time equally, and Deepen the content of the talk (in Dutch, these formed an acronym that translates as “fluent”; students were encouraged to engage in a “fluent dialogue”). After observing their teacher modeling how to ask follow-up questions, they talked in small groups about a reading-related topic (i.e., what is important when choosing a book to read), using cards with follow-up questions, for instance, “Can you give an example?” or “Who has a different opinion?”. Small-group experiences were then shared in class and students reflected individually, in their workbooks, on strengths and points for improvement of their dialogues.

*Reading-and-dialogue units.* Whenever students started an external dialogue task, teachers reminded them of the guidelines. In unit 2, students had indicated which reading experience was most prominent and were grouped accordingly: for example, students who strongly pictured the story in their minds (imagery) formed a group. Each group completed a dialogue task to deepen their prominent reading experience. For instance, the imagery group talked about what the images in their minds looked like, explored the atmosphere the images evoked, discussed and selected photographs that represented the atmosphere, and talked about which five words from the story fit the chosen photographs best. In unit 3, the dialogue started with deciding on the most important story moments, by drawing a story board. Students then talked about which life lesson they derived from the story. In unit 4, the dialogic task was a speed date: students talked in the first round about how they would feel and react if they were the story characters; in the second, about which associations the story and its characters evoked; in the third, about what happened at the end and how they thought it connected to the theme of religion, which was prominent in the story.

External dialogues were also operationalized at the classroom level at the end of each unit. For instance, in unit 2, group representatives explained what their group had talked about, while other students indicated in their workbooks whether what they heard made them reconsider their own small-group task. In unit 4, students first individually wrote down what they learned from the project, shared this in pairs, and exchanged it in a teacher-led classroom dialogue.

### 3.2 *First Assessment Phase*

In examining the validity and practicality of TDLT-1, we first address the intervention as a whole, followed by selected stories, internal dialogue tasks, and external dialogue tasks, in line with the design principles. Quantitative results from teacher logs, time on task observations, and evaluation forms are substantiated with results from teacher interviews (see Table 4.1, p. 98) and strengths and suggestions for improvement listed by students (the latter are summarized in Figure 4.2 and 4.3).

#### 3.2.1 *Intervention as a whole*

*Content validity.* In their logs, teachers generally agreed students were interested and engaged in TDLT-1 ( $M = 3.8$ ,  $SD = .8$ ). Students themselves neither highly valued nor highly disliked it (Overall appreciation and relevance:  $M = 3.0$ ;  $SD = .8$ ): 47.6% responded negatively ( $M < 3.0$ ), and 45.4% responded positively ( $M > 3.0$ ; for 7%,  $M = 3.0$ ). Students agreed to feel safe during the intervention ( $M = 3.9$ ,  $SD = .7$ ).

Students' mixed responses were reflected in strengths and suggestions they listed. Of the strengths, 23.7% referred to valuable, relevant learning outcomes. 7.5% concerned general strengths (TDLT-1 was "important" or "fun"). A small share (2.3%) concerned a safe social atmosphere in class. Of the suggestions for improvement, 21.7% concerned TDLT-1 not being fun or engaging. Its purpose and relevance were sometimes questioned as well (6%). Teachers agreed in interviews that students needed clearer goals and more insight into steps to work toward those goals. One teacher suggested that a rubric for "noticing reading experiences and talking about it" might be helpful.

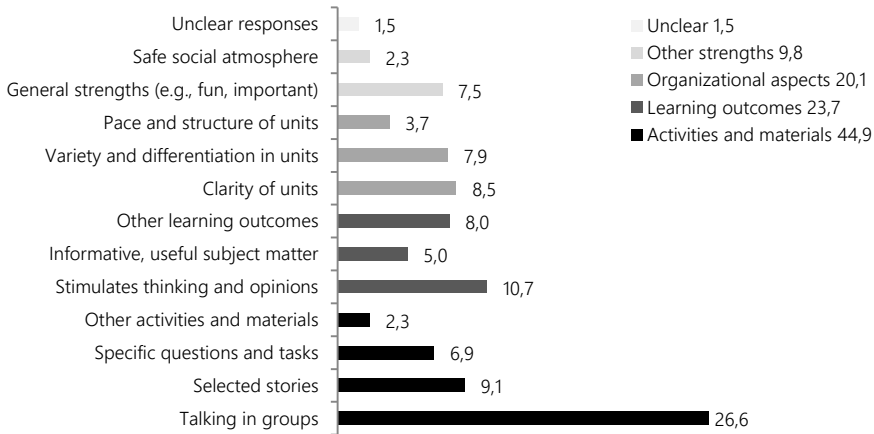


Figure 4.2. Categories of strengths (% of 683 strengths).

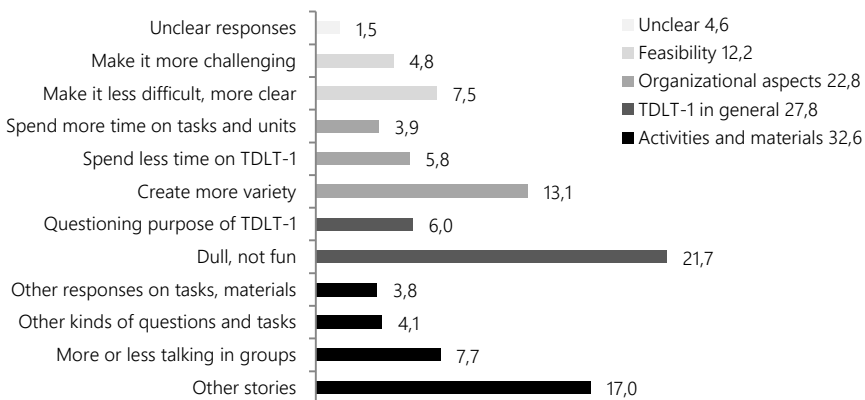


Figure 4.3. Categories of suggestions for improvement (% of 482 suggestions).

*Construct validity.* In their logs, teachers indicated the units were clear for students ( $M = 4.1$ ,  $SD = .5$ ). Students also were neutral to positive about Overall clarity and comprehensibility ( $M = 3.6$ ,  $SD = .8$ ). In addition, 8.5% of the strengths they listed concerned the clarity of the units, and 3.7% their pace and structure. As 13.1% of the suggestions for improvement concerned calls for more variety, we hypothesized that TDLT-1 might even have been too coherently structured.

*Practicality.* Teacher logs showed that 88% of the phases was fully completed as intended; 9.4% was partly completed, and 2.7% was not completed. For

partly completed phases, teachers mostly noted they had spent somewhat less time on them than planned. In general, organizing various dialogic activities around a short story in a single 50-minute unit seemed manageable. Teachers agreed that teaching the phases was attainable ( $M = 4.2$ ,  $SD = .5$ ) and that there was order and discipline in the classroom ( $M = 4.0$ ,  $SD = .6$ ).

Time on task findings, however, showed different results. Students were on task in 72.5% of the observed time and off task in 24.9% of the time (in 2.5%, task behavior was "unclear"). The on task percentage was below the standard of 80% that has been suggested in studies on effective teaching (Kauchak & Eggen, 1993; Muijs & Reynolds, 2010). In addition, it varied significantly among units ( $\chi^2(6) = 15.09$ ,  $p = .020$ ), which was due to unit 3 (67.7% versus 73.1%, 73.9%, 74.2%). The same was true for variance among teachers ( $\chi^2(24) = 124.21$ ,  $p < .001$ ): only two of them scored above the 80% norm (range 54.1% – 84.8%). Similarly, on task percentages varied significantly among activities ( $\chi^2(8) = 56.54$ ,  $p < .001$ ), as will be illustrated in relation to the design principles.

For teachers, feasibility seemed related to the intended active pedagogy of TDLT-1 (e.g., short phases in a high pace, group work, organizing short presentations). In interviews, teachers either said to appreciate it ("it's something else", "it challenges students", "you can organize so much more in a lesson than I expected"), or suggested it was "too much fuss" or "strange for students, who are too passive for this". Not for all teachers, thus, an intervention that required such a strict organization was feasible. Although they all appreciated the introductory meeting, some suggested to organize an active workshop to become familiar with the intervention.

Students considered feasibility a point for improvement (12.2% of all suggestions). They felt that units were either not feasible, for example, "It was hard to complete everything in detail in the time we were given", or not challenging enough, for instance, "I'd like to go deeper into the discussions but the tasks didn't really allow for it". In addition, in 3.9% of the cases they suggested to devote more time to the project.

### 3.2.2 Selected stories

*Content validity.* Students rated story appreciation on a scale of 1 to 10. Two stories were read by all students were rated on average 6.3 (*She was everywhere*,  $SD = 1.3$ ) and 6.7 (*Flight behavior*,  $SD = 1.3$ ). Stories in unit 3 were evaluated by fewer students; mean scores varied between 6.1 and 6.9, with standard deviations similar to the ones above. Overall, although standard deviations indicated considerable variety, students evaluated the stories neutrally to posi-

tively. However, 17% of the suggestions for improvement concerned stories ("Select better stories, these were vague"), against 9.1% of the strengths ("Nice stories").

Teachers evaluated the stories positively, in particular *Flight behavior*. For example, one teacher said the story was "really imaginable for [students] and concrete enough to talk about, they can really put themselves in the [plane crash] situation". Some teachers had reservations about *She was everywhere*: on the one hand, it was well-chosen because of its recognizable setting and comprehensible style, but it was also characterized as "intangible" and "too open". Yet, when asked if they would replace a story, none of the teachers indicated that they would.

*Construct validity.* In student data, we found no responses about the coherence of the stories. Three teachers commented on it during interviews. Even though they appreciated most stories, they suggested to select stories centering around a single theme. One teacher said: "I think that stories centered around a theme help students to identify relations between stories, themselves and how others think. It prompts them to consider a theme in multiple units. Something like 'injustice', maybe."

*Practicality.* In the interviews, most teachers indicated the stories were practical in terms of reading time: only occasionally, students could not finish their story in time. Some teachers read stories aloud, so that students finished and started the next task simultaneously; others had their students read individually, to fully focus on the responses stories evoked. Several teachers mentioned that, in practical terms, they preferred short stories over excerpts from novels, to offer students "the sense of a beginning and end", as one teacher said. The on task percentage for story reading was high (91%).

### 3.2.3 Internal dialogues

*Content validity.* For Appreciation and relevance of internal dialogues, students' mean score was 2.7 ( $SD = .7$ ). On average, thus, they did not value internal dialogue tasks highly. In strengths and suggestions for improvement, they did not refer to internal dialogues. We suspect that internal dialogues were not prominent enough for students to reflect upon them and, potentially, value them. Several teachers endorsed this: they said that the purpose of internal dialogues could be made more explicit, as these were "strange" for students. For instance, one teacher said: "What students find strange is that these writing tasks are very open. For them it feels [...] as if anything goes."

Apart from this suggestion, teachers evaluated pre-reading activation tasks, tasks to notice responses, and reflection tasks after reading positively. For example, one of them said: "Thinking about a theme prior to reading is valuable and safe. We should do it more often." Teachers also recalled that, despite the purpose perhaps not being entirely clear, students were generally engaged in the tasks, as this response illustrates: "My students completed the tasks rather seriously and felt they were heard and taken seriously."

*Construct validity.* Students evaluated Clarity and comprehensibility of dialogue tasks negatively to neutrally ( $M = 2.8$ ;  $SD = 1.0$ ; applies to internal and external dialogues). Most teachers felt the tasks were clear and well-structured. For instance, one of them said: "The tasks had a clear buildup. I think that helped my students to focus." In terms of overall coherence, some teachers noticed that internal dialogue tasks created starting points for external dialogues, which were then more profound than they usually observed. However, others said this happened "not as much as you would expect".

Some teachers observed that students struggled to notice responses during reading and suggested they might benefit from teacher think-alouds, to model that all sorts of responses may be evoked during reading. Another suggestion was to teach strategies for when students encounter difficulties in stories. Several teachers noticed that students tended to simply ignore difficulties and continue reading, which potentially hampered their thinking about story themes in relation to themselves and others. Teachers felt that a "step-by-step strategy to deal with difficulties" could be coherently connected to attention for reading responses. As one teacher said: "Things like 'Huh?' or 'I don't get it' are genuine responses that occur frequently among 10<sup>th</sup> graders, but they often just don't know what to do when it happens".

*Practicality.* Teachers mostly felt that the internal dialogue tasks were well-organized and were practical to work with. For example, a teacher said: "This went really well. The students understood the steps and had enough time to complete the tasks." Teachers further indicated that annotating or highlighting the text when noticing a response and guided reflection after reading (e.g., scoring statements about imagery and sympathy) worked well for students, "because it narrows down what students are asked to reflect upon", as another teacher said. The on task percentage for individual student activities was 77%.

### 3.2.4 External dialogues

*Content validity.* Like the internal dialogues, students did not value external dialogues highly (Appreciation and relevance of external dialogues:  $M = 2.8$ ;  $SD = .8$ ), even though “talking in groups” was most frequently mentioned as a strength of the project (26.6% of all strengths). Of the suggestions for improvement, only 7.7% concerned group talk activities.

Teachers appreciated that external dialogue tasks invited for talking about stories, which they considered a valuable activity in itself. They were most enthusiastic about the speed dates in unit 4. Evaluations of other tasks were mixed. Some teachers were positively surprised, as this response illustrates: “There was some real improvement in my class. I didn’t expect it, but students were actually engaged in practicing to talk about reading experiences. Even now, after the project, they sometimes refer to personal experiences.” However, the relevance of external dialogues was not always obvious for students, as another teacher indicated: “It was quite difficult to get them talking. You know, they did do it, they came up with new ideas, but they just find it weird to share personal experiences.” Other teachers observed a lack of follow-up questions and general disinterest in literature. “The dialogues were... well, mediocre,” one teacher said. “Students are too easily satisfied. They don’t ask follow-up questions, it just doesn’t interest them. The talk remains artificial.” Teachers also indicated that students were “quickly done talking” or “completed tasks superficially”, even though there were always groups that engaged in more extensive, serious dialogues. Teachers suggested to offer students a set of follow-up questions they could use anytime and to clarify why talking about literature can be beneficial.

External dialogues also took place at the classroom level. Teachers noticed these dialogues were not always necessary for students, who felt “they had already discussed a topic,” as one teacher said. Students might as well reflect individually on dialogues, “by scoring themselves and their group”, as another teacher proposed.

*Construct validity.* As indicated, students evaluated Clarity and comprehensibility of dialogue tasks negatively to neutrally ( $M = 2.8$ ;  $SD = 1.0$ ). In terms of coherence, teachers mostly felt that the guidelines, introduced in unit 1, helped students to engage in dialogues, although “it doesn’t come naturally, I constantly reminded them”, as one teacher said.

*Practicality.* Concerns were raised about the practicality of external dialogue tasks. Although only 2.7% of all unit phases was not completed, this mostly

occurred toward the end of a lesson, when a classroom dialogue would take place. Mostly, teachers indicated there was not enough time left. Three of them suggested that whole-class dialogues might be shifted to a next lesson. Units, thus, could potentially cross the boundaries of 50-minute lessons. This suggested that teachers felt the coherence among units could be released somewhat. During whole-class activities, students were on task in 74.4% of the time.

In contrast, students were on task in only 65.5% of the time devoted to dialogues in pairs or small groups. Several teachers said that students needed more time to get used to talking about reading experiences, which they expected to result in less superficial talk and more on task behavior during small-group dialogues. Finally, several teachers explained how they prompted students to deepen a dialogue whenever they fell silent or went off task, for example: "I asked something like: 'Okay, I hear your conclusion. Could you now talk about whether others could have alternative opinions?' Then they explored other perspectives as well." They suggested to include a list of such questions and prompts in the teacher guideline, to offer new teachers more guidance in scaffolding students' dialogues.

### 3.3 *Second Development Phase*

The second development phase resulted in a six-unit intervention (TDLT-2). Like TDLT-1, it started with a preparatory unit in which the main objective was to understand and apply strategies for engaging in and deepening dialogues about stories and reading experiences. The five subsequent units were reading-and-dialogue units. To relevantly connect TDLT-2 to the regular curriculum, we designed a final individual writing task in which students were asked to apply what they learned. Teachers provided feedback and included the task in students' curricular literature portfolios. TDLT-2 is described in closer detail in Appendix D; here, we present specific reoperationalizations of the initial design principles.

#### 3.3.1 *Reoperationalization of text choice principle*

Although all stories in TDLT-1 were considered to be thematically relevant for its aim, teachers suggested that more thematic coherence would help students to identify relations between stories and to consider more deeply how a story theme might be connected to themselves and their perceptions of others. In following the suggestion to select stories centering around "justice and injustice", we maintained some stories in TDLT-2 but substituted others (see Table

4.3). The story from TDLT-1 that teachers and students appreciated most, *Flight behavior*, did not fit this theme, but was given another role: we recorded two short videos of students modeling a good and bad example of a dialogue about this story, which were observed and discussed in the preparatory unit.

Table 4.3. Stories used in TDLT-2

Story	Author	Description	Unit
1. Excerpt from <i>Flight behavior</i>	Bertram Koeleman	See Table 4.2.	1
2. <i>Dood (Death)</i>	Martin Brill	A girl realizes her love interest has died because of senseless violence, and thinks about all that will never happen anymore.	2 & 5*
3. <i>Blood</i>	Gerard Reve	See Table 4.2.	3 & 5*
4. <i>Volgens de regels (Following the rules)</i>	Mirjam Bonting	A father has always set strict rules for his daughter; when he falls during mountaineering, she follows the rules and leaves him behind.	4 & 5*
5. <i>Van geluk spreken (Count oneself lucky)</i>	Marga Minco	A woman meets an acquaintance after World War II; they have a painful conversation about who survived the war and who didn't.	6
6. <i>The freshwater steak</i>	Hans Dorresteyn	See Table 4.2.	6**
7. <i>Een najaarsdag (An autumn day)</i>	Thomas Heerma van Voss	A guard feels compassion for a summer camp host on death row, who murdered children; after the execution, the guard tells his son a bedtime story about a friendly summer camp host, but his son falls asleep before the end.	6**
8. <i>The right</i>	Annelies Verbeke	See Table 4.2.	6**
9. <i>Hoela (Hula)</i>	Cees Nooteboom	From behind a window, a boy at a birthday party watches his little nephew drown in a garden pond, without doing anything to help.	6**

Note. \* In unit 5, students were asked to compare these three stories. \*\* In unit 6, students chose one of these four stories for the final task and read it at home.

### 3.3.2 Reoperationalization of internal dialogue principle

Overall, teachers evaluated internal dialogue tasks in TDLT-1 positively: they were clear, well-structured, and practical to work with. A relatively high on task percentage endorsed this. However, students were not too appreciative of the tasks and hardly commented on them in the qualitative data. As suggested by

various teachers, it may be important to explicate for students why and how they learn to engage in internal dialogues, as they were not used to this approach.

For TDLT-2, we therefore designed a rubric with learning objectives, among which a) learning to notice responses during reading, b) learning to deal with incomprehension while reading, and c) gaining insights into reading experiences, for which three levels were described (see Appendix D, p. 280). Students indicated their starting level in unit 1 and evaluated their progress at the end of the intervention.

In addition, we postponed explicit attention for transformative reading experiences until unit 3. We included a unit focused on evaluative responses (e.g., "fun", "unexpected", "boring", "challenging"; unit 2), as students were more familiar with such responses than with responses related to transformative reading: evaluating fictional texts is usually attended to in lower grades of secondary education.

Whereas students read stories in unit 1 and 2, attention for internal dialogues remained implicit. Unit 3, then, was dedicated to noticing responses during reading. After discussing explicitly why this could be beneficial for students, teachers explicated personal responses by thinking aloud during reading, while being observed by students. Only after these explicit preparations, students were asked to focus on their own responses and to reflect on transformative reading experiences. Moreover, dealing-with-difficulties strategies (i.e., stop reading for a moment, write a question mark, think about possible meanings, ask the teacher or classmates for help) were introduced and summarized on a 'first aid card', which was available to students at all times.

### *3.3.3 Reoperationalization of external dialogue principle*

Whereas students considered it a strength of TDLT-1 that they could talk in groups during literature class, they also indicated it was not always clear why this was relevant and how they could do it well. Teachers indicated that students asked too few follow-up questions and were quickly done talking. The low on task percentage substantiated their impression. Teachers felt their students needed more time to practice and get used to external dialogues, as well as more feedback on and scaffolding of dialogues. They also questioned the need for whole-class dialogues at the end of each unit in TDLT-1.

First, as indicated, we extended TDLT-2 from four to six units. This enabled a buildup in how challenging and unfamiliar external dialogues were. In unit 1, dialogues concerned famous quotes about literature and reading (e.g., "We

read to know we're not alone", William Nicholson). Students identified what a quote meant, evaluated if they agreed with it, and explained their opinion. In unit 2 they talked about their opinions about a story, explained them by referring to literary devices in the text (e.g., flashbacks, gaps), and presented conclusions on a poster. These units prepared students for subsequent ones, in which they were asked to identify, evaluate and explain, for instance, connections between story themes, themselves and others. For example, in unit 4 they read part of a story, were instructed to imagine to be in the protagonist's position, and talked from that point of view about possible story ends.

In TDLT-2, more time was available for small-group dialogues because whole-class dialogues were no longer implemented in all lessons: because units crossed the boundaries of lessons, these dialogues were sometimes shifted to a next lesson. In other occasions, whole-class dialogues were replaced by individual reflections, for example, when students indicated how well they applied dialogue guidelines and wrote down how they might improve in a next unit. Finally, more 'incubation time' was available because TDLT-2 taught in four to eight weeks, rather than two to four weeks.

Furthermore, teachers explicated why and how students should engage in external dialogues, in line with reoperationalizing the internal dialogue principle. First, the rubric contained the objective "contributing actively to dialogues about stories and reading experiences". Second, teachers discussed in unit 3 the purpose and relevance of sharing reading experiences. Moreover, dialogue guidelines were summarized on the first aid card and were thus available to students at all times. The card also included examples of follow-up questions and topics to bring up in dialogues. Finally, we used peer modeling videos in unit 1, to show examples of what dialogues about reading experiences should (not) be like.

Finally, the teacher guideline was adapted. We implemented more feedback moments (e.g., when students presented opinions and support on a poster in unit 2, teachers wrote comments and questions on post-its attached to it). We also added specific instructions for providing students with process-oriented feedback, including a list of example responses and questions teachers could use to prompt students to deepen a dialogue. During the workshop for new teachers, such feedback situations were imagined, discussed and rehearsed.

### 3.4 *Second Assessment Phase*

We examine the validity and practicality of TDLT-2. If applicable, we indicate whether significant differences compared to TDLT-1 were found.

### 3.4.1 *Intervention as a whole*

*Content and construct validity.* Teachers indicated in their logs that students were generally interested and engaged in the intervention ( $M = 3.7$ ,  $SD = .8$ ) and that it was clear for them ( $M = 4.0$ ,  $SD = .7$ ). Scores did not differ significantly from TDLT-1.

*Practicality.* Teachers indicated that of all phases, 77.6% was fully completed as intended, 14.4% was partly completed, and 8.0% was not completed. These percentages differed significantly from TDLT-1 ( $\chi^2(2) = 12.43$ ,  $p = .002$ ), in which more phases were fully completed (88%) and fewer were partly and not completed (9.4% and 2.7%). As in TDLT-1, teachers had spent somewhat less time than planned on partly completed phases. Non-completed phases were due to time constraints or deliberate decisions (e.g., one teacher felt her students were too agitated and distracted at the end of the day to discuss life lessons). In general, however, teaching was deemed attainable ( $M = 4.1$ ,  $SD = .7$ ) and there was order and discipline in the classrooms ( $M = 3.9$ ,  $SD = .8$ ). Scores did not differ significantly from TDLT-1.

Students were on task in 85.2% of the observed time, and off task in 14.7% of the time. The on task percentage was well above the standard of 80% (Kauchak & Eggen, 1993; Mujs & Reynolds, 2010) and was significantly higher than in TDLT-1 (85.2% versus 72.5%,  $\chi^2(1) = 39.36$ ,  $p < .001$ ). On task percentages varied significantly across activities ( $\chi^2(8) = 17.44$ ,  $p = .026$ ), as illustrated in subsequent sections. In contrast with TDLT-1, percentages did not vary significantly across units, nor across teachers. All teachers seemed to be equally able to engage their students in efficiently spent learning time.

### 3.4.2 *Selected stories*

*Content and construct validity.* For the four stories they could choose from for the final intervention task, students' mean appreciation score was 3.5 ( $SD = .8$ ), ranging from 2.9 for *Hula* to 3.8 for *An autumn day*. Students thus evaluated the stories neutrally to positively. Even though we used a different scale than in assessing TDLT-1, story appreciation was rather alike. In terms of relevance, students indicated repeatedly that simply reading these stories helped them to achieve their goals (15.2% of all indications; see Table 4.4). We collected no data that indicated the construct validity of selected stories, as we only adjusted the central theme of the intervention.

Table 4.4. Distribution of TDLT-2 elements perceived as helpful ( $N = 1,784$ )

Intervention element	%	% per objective*			
		1	2	3	4
1. Observing peer modeling videos: good and less good dialogue about reading experiences	5.8	2.0	0.9	0.8	2.1
2. Dialogue guidelines on first aid card, e.g., listening carefully, asking follow-up questions	9.5	2.5	1.6	1.8	<u>3.6</u>
3. Other suggestions on card, e.g., strategies to deal with incomprehension, follow-up questions	9.6	2.4	2.1	2.3	2.9
4. Reading the selected stories	15.2	3.6	<u>4.1</u>	<u>4.1</u>	3.4
5. Observing the teacher modeling to notice responses during reading (teacher think-aloud)	14.5	<u>4.9</u>	3.1	<u>3.4</u>	3.1
6. Focus on noticing own responses through reading instructions (e.g., annotating)	13.7	<u>4.5</u>	3.1	<u>3.4</u>	2.7
7. External dialogue tasks in pairs, groups or class	14.7	3.1	<u>3.4</u>	3.1	<u>5.1</u>
8. Teacher's explanations about noticing responses, sharing responses, and other topics	16.9	<u>4.0</u>	<u>4.3</u>	<u>4.1</u>	<u>4.4</u>
Totals	100	27.0	22.6	23.0	27.4

*Note.* \* Elements helpful for objectives: 1) noticing responses while reading; 2) dealing with incomprehension while reading; 3) gaining insights in reading experiences; 4) contributing actively to dialogues about stories and reading experiences (see Appendix D). For each objective, the three elements most frequently perceived as helpful are underlined.

*Practicality.* Students indicated that the story they read for their final intervention task was not too difficult ( $M = 2.5$ ;  $SD = 1.1$ ). The mean value for three of the four stories was below the scale mean of 3; only *Hula* was considered more difficult ( $M = 3.5$ ,  $SD = 1.1$ ). As an additional practicality indicator, the on task percentage for story reading was 84.6%, which was only slightly lower than in TDLT-1 (91%).

### 3.4.3 Internal dialogues

*Content and construct validity.* Students indicated that learning to notice responses relatively often was helpful for them (13.7% of all indications; see Table 4.4). Thus, at least for part of the students, it was relevant to engage in internal dialogues with stories. However, suggestions on the first aid card, which included those for dealing with difficulties, were somewhat less often considered helpful (9.6%). In contrast, teachers' explanations about, amongst other topics, literary reading and noticing responses were most often deemed helpful (16.9%). The same was true for the teacher modeling how to notice responses

(14.5%). As Table 4.4 indicates, the latter was most often considered helpful for learning to notice responses, which suggested a coherent connection between the reoperationalization of the design principle and the objective.

*Practicality.* During individual internal dialogue activities, students were on task in 86.4% of the time, which suggested the activities were completed as intended.

### 3.4.4 External dialogues

*Content and construct validity.* Students indicated that external dialogue tasks relatively often helped them to achieve progress on intervention objectives (14.7% of all indications; see Table 4.4). In particular – and not surprisingly – they felt these tasks helped them to learn to actively contribute to dialogues about stories and reading experiences. This indicated a coherent connection between the reoperationalization of the design principle and the objective. Support for external dialogues was less often seen as helpful (peer modeling videos: 5.8%; dialogue guidelines: 9.5%). Teachers' explanations, which also concerned how to engage in small-group dialogues, were most helpful to students, as indicated above (16.9%). Clearly, the role of the teacher in the intervention cannot be underestimated.

*Practicality.* The on task percentage for whole-class activities (76%), that included classroom dialogues, was lower than for the other activities. However, for dialogues in pairs or small groups, it was clearly higher (90.1%). This was in stark contrast with TDLT-1, where the on task percentage for such activities was only 65.5%.

### 3.5 From Initial Design Principles to Subprinciples for Operationalization

The design process described in this paper started from three rather broad design principles. Here, we complement them with subprinciples that represent how the initial principles were eventually operationalized. By specifying the initial principles, we offer practical suggestions for designing literary instruction that focuses on transformative reading and gaining insight into human nature:

1. Fictional texts should be selected that are thematically relevant for the intervention aim.
  - a. These texts should coherently center around a single relevant theme. The purpose of thematic coherence (here: justice and injustice) was to help students to identify relations between stories, themselves and

their perceptions of others. Thereby, rather than introducing a new theme in each unit, students could build on outcomes of previous dialogues.

- b. For 10<sup>th</sup> grade students, texts can be selected that teachers consider to be *literary texts*. We used stories that teachers considered to be appropriate in terms of literary quality, for which on task percentages during reading were high. Whereas previous studies were indecisive as to the role of literary reading in fostering students' insight into human nature (Schrijvers et al., 2018), the current study indicated at least that 10<sup>th</sup> grade students and their teachers considered reading literary texts to be appropriate and helpful for the objective "gaining insight into reading experiences", even though its effects on students' insight into human nature requires further analysis.
2. Writing tasks should be designed that stimulate students to activate previous personal experiences before reading, to notice and annotate responses during reading, and/or to write down or reflect on responses directly after reading, to stimulate an internal dialogue between reader and text as preparation for external dialogues.
    - a. Teachers should explicitly explain the purpose and importance of internal dialogue tasks. Students appeared to need guidance in determining the relevance of engaging in internal dialogues with a text. In TDLT-2, we found that they appreciated their teacher's explanations about this topic as well as other topics.
    - b. Students should observe the teacher modeling how to notice responses while reading. Students in 10<sup>th</sup> grade are not necessarily used to paying attention to the responses evoked by a literary text. The redesigned intervention showed that students found it helpful to observe their teachers, who modeled noticing responses during reading.
    - c. Students should be taught strategies for dealing with difficulties during reading a story, which they may consult at any time via specific support tools. Incomprehension is a rather common response that students may notice while reading. Attending to strategies for dealing with difficulties stimulates students to not simply ignore them, but to work toward finding solutions for them. A support card or other tool that summarizes such strategies may remind students of applying them.
    - d. Internal dialogue tasks should first focus on more familiar responses (i.e., opinions about a story) before turning to transformative reading experiences. We implemented a buildup in the intervention: students

- first engaged in tasks they recognized from lower grades of secondary education (i.e., thinking about their overall opinion about a story) and then gradually moved to less familiar internal dialogues in which they were asked to attend to responses related to transformative reading.
3. Students should engage in exploratory dialogic activities that stimulate them to verbally share their personal experiences related to texts and text themes.
    - a. Teachers should explicitly explain the purpose and importance of external dialogue tasks. Similar to internal dialogues, students were given guidance in establishing the relevance of external dialogues. Teachers thus attended to the importance of talking about multiple reading experiences and to the characteristics of a good dialogue.
    - b. Students should be taught guidelines and other suggestions for external dialogues. The first aid card in the intervention included dialogue guidelines and other suggestions for deepening the talk about a story (e.g., examples of open-ended questions to be asked, suggestions for what to tell about the own reading experience). Students may be familiarized with such guidelines in a first, preparatory unit.
    - c. Teachers should be given guidelines for providing students with feedback and guiding their dialogic processes. For teachers who are new to dialogic approaches in the literature classroom, it may be challenging to assist students without being too involved in the content of their dialogue. Therefore, teachers were given prompts and questions to stimulate small-group dialogues (e.g., "What (else) does this story make you think about?", "Why might anyone have different opinions?", "I hear you struggle with [x]. Could we think about how to solve that issue?")
    - d. Students should be given enough time to learn to engage in external dialogues about stories. Four units, designed for four 50-minute lessons, appeared to be too few for students to get used to talking about stories and reading experiences; in TDLT-2, we opted for six units. Ideally, however, attention for external dialogues and preparation for them by internal dialogues should not be confined to several units or lessons, but may be interwoven in the regular literature curriculum in secondary schools.

## 4 DISCUSSION

This study has resulted in a literature classroom intervention that aims to foster 10<sup>th</sup> grade students' insight into human nature. The intervention (TDLT-2) is characterized by an emphasis on learning to notice responses in internal dialogues that are related to transformative reading, and share these in external dialogues. For example, students are asked to reflect on and talk about whether they picture the setting and characters in a story in their minds, or feel sympathy and compassion for characters.

In adult readers, such experiences have been shown to result in insights into themselves and others (Fialho, 2018). The present study has paved the way for investigating whether these effects also occur in adolescents in the context of the literature classroom, as it resulted in a valid and practical intervention that can function as a dependent variable in future effect studies. In addition, the study has offered concrete suggestions for how the three initial design principles are to be operationalized in teaching practice. In what follows, we reflect on the validity and practicality of the eventual intervention (TDLT-2) and evaluate the design process, before turning to the conclusions.

### 4.1 *Validity and Practicality of the Intervention*

The study confirmed, first, the validity of the selected stories. The main outcome of the iterative design process was to use stories that centered around a single relevant theme in TDLT-2, to increase coherence and offer students multiple opportunities to reflect on the theme and consider how it connects to themselves and the lives of others. Teachers felt the stories were appropriate in terms of difficulty and literary quality. Students moderately appreciated them and indicated they were relevant for achieving progress on the intervention objectives. In terms of practicality, teachers preferred short stories over excerpts from novels and felt that organizing various learning activities before and after the story was manageable. For students, the stories were not too difficult to read.

The main adjustment in redesigning internal dialogues was to discuss their purpose and the strategies to establish them more explicitly with students. In TDLT-2, students gained insights into different kinds of responses and reading experiences via, for example, the rubric used in unit 1, teachers' explanations, and observing their teacher who modeled reading with purposeful attention for the responses a story evoked. Overall, students felt that the attention for noticing responses evoked by a text and the teacher modeling what this looked like were helpful, relevant intervention elements.

The main challenge in the design process was to establish external dialogues in which students would share responses and reading experiences in such a way that it would stimulate their insight into human nature. Even though students felt that the opportunity to talk in groups was a strength of TDLT-1, they questioned the purpose of external dialogues, showed remarkably low on task percentages during small-group dialogues, and demonstrated artificial and superficial talk. Teachers ascribed this to students not being used to talking about stories, let alone about transformative reading experiences. They also felt there was not enough time in TDLT-1 for students to learn how to engage in external dialogues and considered it a challenge to guide student groups toward more profound dialogues.

To remedy these validity and practicality concerns, we extended the intervention from four to six units and let go of the strict structure of pre-reading task, noticing responses while reading, individual reflection, small-group dialogue, and classroom dialogue in a single 50-minute lesson; in particular, whole-class dialogues were shifted to next lessons or substituted by individual reflections on small-group dialogues. More attention was paid to the potential benefits of sharing reading experiences, as teachers explicitly discussed this with students, and to how students could achieve progress in contributing actively to external dialogues (rubric, support tools, teacher feedback). In addition, teachers were provided with more extensive guidelines for providing students with feedback.

After these adjustments, the on task percentage during small-group dialogues in TDLT-2 was considerably higher than in TDLT-1, and students indicated that teachers' explanations and dialogic tasks were helpful for learning to contribute actively to dialogues. As a whole, TDLT-2 was generally used as it was intended, as teacher logs and students' overall on task percentage indicated. All in all, the iterative design process resulted in a valid and practical intervention that may be implemented in literature classrooms to potentially foster students' insight into human nature.

#### *4.2 Evaluation of the Design Process*

The design process has several strengths, but is also subject to limitations. First of all, only a few teachers were able to commit to the project for a longer period of time. This led us to recruiting new teachers for the various steps and phases of the design process, which is both a strength and a limitation. On the one hand, the number of teachers involved yielded a large variety of experiences and perspectives on TDLT, for example, during expert consultation, trial

studies and interviews. On the other hand, some teachers mentioned that it was challenging to teach the intervention precisely because they had not been involved in developing it.

The instruments used in the assessment phases yielded valuable information regarding the validity and practicality of TDLT. We recommend to consider both teachers and students as stakeholders of an intervention, and to collect evaluation and implementation data of both groups. In the present study, teacher interviews and student evaluation forms and tasks were complementary and thereby offered a nuanced picture of validity and practicality. Moreover, digital teacher logs and time on task observations proved to be practical and time-efficient instruments for collecting implementation data. The high response rates for teacher logs indicated that it was feasible for teachers to complete them after each unit. Moreover, in the first assessment phase, they provided useful input for teacher interviews.

A potential criticism on the design process and our conclusions about validity and practicality is that they strongly depend on comparisons of data collected in the first and the second assessment phases. We purposefully compared teacher log data and time on task results of both TDLT versions, but these findings must be considered cautiously as the groups of teachers and students may not have been fully comparable.

Moreover, the instruments used in the assessment phases of the design process tell us little about how students talked about stories and reading experiences. Therefore, a process-oriented perspective on student dialogues and analysis of how students would express considerations of themselves and others in literary dialogues may be important to further develop the dialogic approach presented in this paper. While qualitative studies into dialogic reading discussions are available (Janssen & Pieper, 2009), to date they have not focused on students' insights into themselves and others that potentially emerge during such dialogues.

Finally, we assessed not only the validity and practicality of TDLT-1 and 2, but also their effectiveness, on which we will report in future publications. To this end, the units were embedded in quasi-experimental studies with strict research designs and data collection procedures. This may have influenced teachers, who were sometimes concerned about the time frame, and students, who may have felt to be involved in "just" a research project which was not part of their regular curriculum. Even though TDLT was taught by the students' own teachers and was embedded as optimally as possible in the curriculum, its ecological validity may have been affected. In future design projects, researchers may choose to focus in a first assessment phase solely on implementation

and evaluation. Once a valid and practical intervention has been established, a subsequent assessment phase can take the form of an effect study with a more rigorous research design.

### 4.3 *Conclusions*

All in all, this study has shown that iterative cycles of development and assessment may result in the design of a valid and practical domain-specific intervention. Triangulation of data sources explicated a wide range of experiences of teachers and students, from which valuable information for redesigning the intervention was derived. This, in turn, resulted in a set of design principles in which the initial ones were complemented by subprinciples for operationalization in the classroom. Our next step is to assess the effects of both versions of TDLT, to shed further light on their quality and to contribute to ongoing research on transformative reading by expanding its scope to adolescents in the literature classroom. In the meantime, we hope that this paper offers an example of how a comprehensive description of intervention development may be conceptualized, in order to open the metaphorical “black box” and to enhance the validity and replicability of domain-specific intervention research.



## CHAPTER 5

### EFFECTS OF DIALOGIC LITERARY INSTRUCTION ON 10<sup>TH</sup> GRADE STUDENTS' INSIGHT INTO HUMAN NATURE

In this study, we aimed to investigate the effects of dialogic literary instruction on adolescent students' insights into human nature and other transformative reading experiences such as identification, imagery and sympathy for characters. A four-unit intervention centered around short literary stories with various social-moral themes. Students engaged in five types of tasks: 1) tasks to activate prior knowledge about and personal experiences with the theme of a story to be read, 2) internal dialogues with the story, focused on transformative reading experiences, 3) external dialogues in small groups or pairs about the story theme and their reading experiences, 4) external dialogues at the classroom level, and 5) reflection on their learning. Participants were 603 students in grade 10 (13 teachers; 22 classes) of six secondary schools in the Netherlands. In a quasi-experimental pretest-posttest design with switching replications, we assessed students' perceptions of learning via a written learner report task, and their learning outcomes in terms of insight into human nature via a questionnaire and a written story response task. In their learner reports, students repeatedly indicated that the intervention affected their skills to talk about stories, transformative reading experiences, and thinking and opinionating skills. However, questionnaire and story task responses showed no consistent intervention effects on students' insight into human nature. We discuss possible explanations, for example, that the intervention has been too skills-oriented or that students were not fully engaged in what they were reading, as well as implications for practice and future research.

#### 1 INTRODUCTION

I hope that my fear will overwhelm my senses. That I won't be conscious when the airplane crashes. [...] I knew this was going to happen. I knew it. I try to get my wife's attention, but mortal fear has fossilized her. In her eyes, I see that I won't be able to reach her anymore. I cry and I scream that I love her. She doesn't respond. Her gaze is far away. A crucifix peaks from her folded hands. (Koeleman, 2016; translation MS).

The short story *Flight behavior* immediately draws the reader into a frightening situation: facing an inevitable death in a plane crash. We may relive the fear we felt during bad turbulence, consider how we would react in such a situation, or think about why religion offers comfort to some of us when we face our worst fears, as it does for the protagonist's wife, while it does not for others. Like

many short stories, *Flight behavior* offers starting points for considering ideas about human nature: how would we react to a story situation that might happen to us in real life? What kind of human being do our reactions make us? Would others, different from ourselves, react differently? How do we feel about how they think and behave?

Empirical research suggests that one of the merits of reading fictional and literary texts – novels, stories, poetry – is that it offers readers insight into themselves and others (Hakemulder, Fialho, & Bal, 2016; Koopman & Hakemulder, 2015). Fictional worlds have been theorized to be simulations and abstractions of social experience (Mar & Oatley, 2008). Reading fictional and literary texts may thus be a *transformative* experience (Fialho, 2012; 2018), because, in simulating social experiences, it may alter readers' perceptions of themselves and themselves in relation to others. In other words, it may develop their insight into what it means to be human, into human nature.

In the present study, we apply these insights from empirical literary studies to the context of the literature classroom. We investigated whether an intervention in secondary school literature classrooms in the Netherlands could foster students' insight into human nature. We focused on 15-year-old students in 10<sup>th</sup> grade of the higher general secondary education track, which is the second highest track in the Dutch educational system. It prepares for higher vocational education, but not for university studies. In this track, formal literature teaching starts in 10<sup>th</sup> grade. Literary instruction is a subdomain of Dutch language class, like writing and grammar instruction. This particular educational context offers opportunities to develop students' insight into human nature, but it also poses a number of constraints. These will be discussed after we have described how we conceptualize "insight into human nature" and "transformative reading" in this study.

### 1.1 *Conceptualizing Insight into Human Nature*

Philosophers and literary scholars alike have argued that reading fictional and literary texts may affect how we perceive both ourselves and others, or in our terms, impact readers' insight into human nature (e.g., Keen, 2007; Nussbaum, 1995; Zunshine, 2006). It may include – but is not limited to – insight into previously unrecognized personal qualities or shortcomings, insight into self-other relations, understandings of and altered attitudes toward individual others and groups of people, and considerations of difficulties or moral dilemmas that people may face. Empirical studies support the potential of reading fictional and literary texts as a catalyst for developing insight into human nature, a pro-

cess also described as transformative reading (Fialho, 2012; Hakemulder et al., 2016; Koopman & Hakemulder, 2015).

In search of components that underlie transformative reading, Fialho (2018) found that adult readers' responses reflected: vivid imagery of the setting and characters in a story (*imagery*); recognition of aspects of self or others in characters (*identification*); enactment and embodiment of characters' experiences (*experience-taking*); evaluations of characters, positively or negatively (*character evaluation*); feelings of sympathy and compassion for characters (*sympathy*); awareness of particularly striking words, phrases or sentences (*aesthetic awareness*); and new or deeper insights into themselves and others (*self-other insights*). For adult readers, the first six components appeared to precede insight into oneself and others, or what we call in this paper, insight into human nature. Together, the seven components form a preliminary model of transformative reading (Fialho, 2018). Transformative reading, thus, is considered a mode of reading in which insight into human nature comes about.

This model of transformative reading has not yet been validated for adolescent readers. Yet, studies with adolescent participants suggest that they may have similar experiences. Adolescents expressed in interview studies that they, as a result of fiction reading, reflected on who they would or would not like to become (i.e., possible future selves, Richardson & Eccles, 2007), that they better understood other people's experiences and therefore felt connected to others or saw new possibilities for their own lives (Rothbauer, 2011), and that they compared their own lives to story situations and experienced empathetic engagements with characters' feelings (Charlton, Pette, & Burbaum, 2004; Ross, McKechnie, & Rothbauer, 2006). All in all, it appears as if adolescents may also gain insight into human nature as a result of reading.

### 1.2 *Developing Insight into Human Nature in the Literature Classroom*

Research suggests that literary instruction must meet certain preconditions to enable students' development of insight into human nature. First, *engaged reading* may positively affect students' personal, social, and moral development (Ivey & Johnston, 2015; Lysaker, Tonge, Gauson, & Miller, 2011). Because engaged reading is considered to be relational and dialogic, it is assumed to offer opportunities for 'self- and other-construction' (Ivey & Johnston, 2015, p. 301; Rosenblatt, 1938/1983). Indeed, evidence suggests that engaged reading influences readers' social imagination and social behavior (Kaufman & Libby, 2012; Lysaker et al., 2011).

However, whereas Ivey and Johnston (2015) indicated that freedom in choosing reading materials is important for engaged reading, and by extension for personal, social, and moral development, a recent review of empirical intervention studies (Schrijvers, Janssen, Fialho, & Rijlaarsdam, 2018) suggested otherwise. In only one literature classroom intervention that fostered students' insight into human nature, students selected reading materials from a preselected list (Vezzali, Stathi, & Giovannini, 2012); in other interventions, students did not choose their texts themselves, but nonetheless developed insight into human nature. The review suggested that, rather than freedom of choice, *text themes* are important for fostering insight into human nature: fictional texts were used that were thematically relevant for intervention aims, such as texts addressing social relations or moral dilemmas (e.g., Adler & Foster, 1997; Malo-Juvera, 2014; 2016; White, 1995).

The importance of dialogue in the literature classroom for fostering insight into human nature, as suggested by Ivey and Johnston (2015), was supported by the recent review study (Schrijvers et al., 2018). In reviewed intervention studies, reading was of a dialogic nature in two ways. First, writing tasks were implemented that prompted students to activate previous personal experiences before reading, to notice their experiences during reading, and/or to write (reflective) responses directly after reading (Eva-Wood, 2004; Halász, 1991; Malo-Juvera, 2014; 2016; White, 1995). As such, they were encouraged to engage in an *internal dialogue* with the text that generated an initial process of meaning-making or, in Rosenblatt's (1938/1983) terms, a transaction between reader and text. Because secondary school students have been found to often read fictional and literary texts superficially, in a uni-dimensional and rather closed way (Andringa, 1995; Earthman, 1992; Janssen, Braaksma, Rijlaarsdam, & Van den Bergh, 2012), learning to actively interact with texts in internal dialogues may help them to construct interpretations, and subsequently, to develop insight into human nature beyond the process of reading.

Second, initial processes of meaning-making were combined with or followed by verbally sharing personal responses related to texts and text themes with peers, in *external dialogues* (e.g., Adler & Foster, 1997; Eva-Wood, 2004; Malo-Juvera, 2014; 2016; White, 1995). Such dialogues were organized in pairs, small groups and/or at the classroom level. Students could explore multiple perspectives on a text and the themes or issues it addresses. Via such explorations, they may come to understand that a variety of interpretations can be valid, as fictional and literary texts are assumed to be, in essence, multi-interpretable. This process may, by extension, elicit alternative or elaborated views, ideas and perceptions of self and others in relation to texts and themes.

Likewise, Schrijvers, Janssen, Fialho, and Rijlaarsdam (2016) showed that if literature teachers reported to allow for more interaction and dialogue in the classroom, their students more often reported to have learned about themselves and others.

### 1.3 *Constraints in Dutch Literature Classrooms*

Although a dialogic approach to literature teaching seems promising for developing students' insight into human nature, implementing such an approach in Dutch literature classrooms is subject to several constraints. These partly determined the design of the intervention in this study.

First, even though teachers and curriculum organizations may strive for students' personal and social development (Dutch Institute for Curriculum Development, 2015; Janssen, 1998; Oberon, 2016), it is not explicitly mentioned as one of the three global objectives that are formulated in the Dutch literature curriculum: in examinations, teachers assess students' literary-historical knowledge, their structural-analytical skills, and their ability to reflect on their literary reading experiences and development (Dutch Institute for Curriculum Development, 2012). Moreover, attending to objectives additional to those mentioned above may be complicated by time constraints. In upper general secondary education (grade 10 and 11), on average only 30 to 40 minutes per week are devoted to literature teaching (Oberon, 2016).

A second constraint pertains to limited choice in text materials for students (Ivey & Johnston, 2015; Lenters, 2006). Up to 9<sup>th</sup> grade, Dutch students mostly read children's and young adult literature, but from 10<sup>th</sup> grade onwards most teachers require students to read increasingly complex literary texts intended for adult readers. Students in 10<sup>th</sup> grade have little experience with such texts, but their freedom to choose otherwise is limited. In adhering to teachers' objective to familiarize students with literary texts of a certain complexity, students' freedom in choosing text materials was also limited in the intervention designed for the present study. Text selection was mainly informed by thematic considerations, as the review study by Schrijvers et al. (2018) suggested that text themes are important in fostering students' insight into human nature. In most Dutch schools, there are no constraints for text choice on the thematic level: it is possible to select literary texts intended for adult readers that address social-moral themes.

Third, a dialogic approach in the literature classroom may be new and unfamiliar for students. This applies to actively attending to their own responses and meaning-making processes in internal dialogues during reading, as well as

to sharing responses with others. In Dutch literature classrooms, responding to literary texts mostly happens *after* reading, individually, and in writing, for instance by writing a book report or review, collecting these reports and other writing tasks in a portfolio, completing a written test on a book (Dirksen, 2007; Oberon, 2016). Responding verbally mostly happens individually as well: students give book presentations in class, or take oral examinations which are often shaped as analytical, reflective conversations that mostly take place between a single student and the teacher, who hardly can be considered an equal conversation partner. All in all, it may be challenging for students to monitor their reading experiences and responses during reading in internal dialogues, as well as to exploratively share these responses with peers in external dialogues.

#### 1.4 *Research Questions*

The aim of the present study was to test the effects of a literature classroom intervention, characterized by attention for transformative reading experiences and by dialogic learning tasks, on students' transformative reading experiences which include their insight into human nature. To yield a first indication of whether the intervention aims were achieved, our first research question concerned students' perceptions of learning; second, we focused on the effects of the intervention in comparison to a control condition:

1. Which learning experiences do 10<sup>th</sup> grade students report at the end of an intervention that is based on a model of transformative reading and that stimulates them to engage in internal and external dialogues with and about short stories with social-moral themes?
2. Does the intervention have a positive effect on students' transformative reading experiences, such as insight into human nature, imagery, identification, and sympathy, in comparison to students who participate in an untreated control condition?

## 2 METHOD

### 2.1 *Research Design*

In a quasi-experimental pretest-posttest design with switching replications, group A (11 classes,  $n = 311$ ) first participated in the intervention and then in an untreated control condition, and group B (11 classes,  $n = 292$ ) vice versa (see Figure 5.1). The strengths of this design lie in the fact that it allows for assessing

the sustainability of potential intervention effects (measurement 3 for group A), as well as their replicability (measurement 3 for group B; see Shadish, Cook & Campbell, 2002).

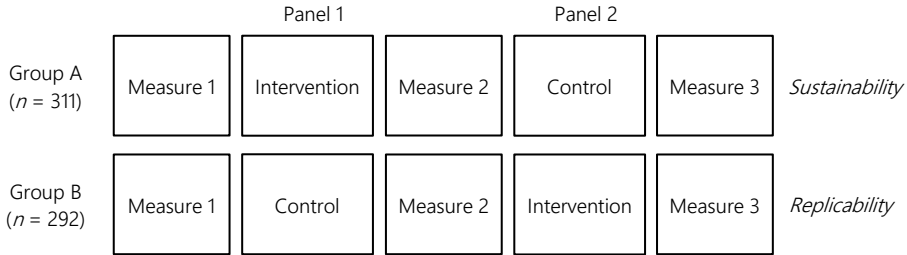


Figure 5.1. Research design.

## 2.2 Participants

We recruited teachers via social media groups for teachers of Dutch and via our network. We found 13 teachers willing to participate, who worked in six different schools across the Netherlands. They participated with either one or two classes. On average, they were 46.2 years old and they had 13.3 years of teaching experience. Two teachers were male. Teachers and their school leaders gave active consent for participation. We randomly assigned teachers to group A or B.

Initially, 23 classes with 634 students participated in the study. Students' parents received an informed consent letter and could object to their child's participation. None of them withheld their consent. Class size ranged from 13 to 33 students ( $M = 27.4$ ,  $SD = 4.2$ ). In one class – of a teacher who initially participated with three classes – data collection procedures were not executed according to plan, which led to a large amount of missing data at the second measurement. Therefore, we removed this class from the data. Thus, 22 classes remained, with 603 students. Students in group A and B did not differ in terms of gender (50.2% female in group A; 56% in group B;  $\chi^2(1) = 2.06$ ,  $p = .151$ ), age ( $M = 15.9$  years in group A;  $M = 15.8$  in group B;  $t(575) = 1.36$ ,  $p = .176$ ), grade for literature ( $M = 6.2$  for group A on a grading scale of 1-10;  $M = 6.3$  for group B;  $t(531) = 1.19$ ,  $p = .235$ ), and familiarity with fiction as measured by an Author Recognition Test (see p. 135;  $M = 4.9$  for group A;  $M = 5.1$  for group B;  $t(537) = .82$ ,  $p = .411$ ).

### 2.3 *Intervention*

Taking into account theory and previous empirical research on developing insight into human nature in the literature classroom, transformative reading, and dialogic literary instruction, as well as constraints that Dutch literature classrooms are subject to, we developed a four-unit intervention, titled Transformative Dialogic Literature Teaching, or TDLT for short. Teachers of Dutch language and literature were consulted in the design process. This process, as well as the validity and practicality of the intervention, are described in another paper (see Chapter 4).

During the intervention, students received four units of 50 minutes of instruction, over a period of two to four weeks in April and May 2017, depending on the schedules of the participating teachers. TDLT was restricted to four units due to the time constraints as described in Section 1.3 of this chapter; teachers who were consulted in the design process also expressed their concerns about implementing an intervention of a longer duration.

The overall objective of TDLT was to help students to identify connections between short literary stories and themselves (i.e., their personalities, the way they are, the way they think) and their views on the social world (i.e., how other people are, behave and think). TDLT consisted of one preparatory unit and three reading-and-dialogue units. Appendix C provides a more detailed description of the four units and the stories that were read during these units.

#### 2.3.1 *Preparatory unit*

The purpose of this unit was to familiarize students with guidelines for small-group dialogues. These external dialogues were a major part of the subsequent units, but we assumed students were not by definition familiar with engaging in them. Students were therefore introduced to a strategy for conducting small-group dialogues about stories and reading experiences, via explicit instruction, modeling, and practice (e.g., McKeown, Beck, & Blake, 2009; Merrill, 2002; Guthrie & Wigfield, 2000). Students first were asked to consider what characterizes good and less good dialogues and were then given five guidelines that were used throughout the intervention: Listen carefully, Ask follow-up questions, Postpone a first judgment, Equally distribute speaking time, and Deepen the content of the talk (in Dutch, these formed an acronym that translates as “fluent”; students were encouraged to engage in a “fluent dialogue”). Next, students observed their teacher who modeled how to ask follow-up questions. Subsequently, students discussed a reading-related topic in small groups, that is, which considerations they took into account when choosing a book to read.

They used cards with follow-up questions to practice using such questions. Finally, experiences that were discussed in each group were shared in class.

### 2.3.2 *Reading-and-dialogue units*

In each of the three subsequent units, students read a short literary story (see Table 4.2, p. 103). The stories centered around various social-moral themes like family and parent-child relationships, mistreatment of children, religion, and prejudice and racism. All stories could be read in 7-15 minutes. The dialogic approach in these units was operationalized by including five types of learning activities, in a fixed order, that were implemented after a short teacher-led introduction. Thus, in each unit, the same basic structure was applied:

1. Activating prior knowledge about and personal experiences with a story theme, prior to reading, to prepare for internal dialogues. For example, before reading the story *She was everywhere* in unit 2, students were asked to write down their thoughts about how someone might react when a relationship ends (time: 2-3 minutes).
2. Internal dialogue with a story: noticing responses during reading and individual reflection on the reading experience directly after reading; in line with the theoretical-empirical model of transformative reading, attention focused in particular on experiences like imagery, identification and sympathy. For example, in unit 2, students were asked to focus on the responses a story evoked in them during reading. Directly after reading, they were asked to indicate to which extent they had noticed experiences such as imagery, identification and sympathy. Thereby, they determined what kind of reading experience was prominent to them, to prepare for external dialogues (time: 10-15 minutes).
3. External dialogue in small groups or pairs: completing tasks that stimulated sharing personal responses, for example, by comparing and contrasting. For instance, in unit 2, students who had indicated that they felt sympathy for a character were asked to compare the moments in the story when each of them experienced this, and to share in their group what they thought and felt at those moments. Next, they were asked to brainstorm about what kind of help would be of avail to the protagonist. As a third step, they reached a conclusion about what kind of help they would offer the protagonist, by talking about issues like: how feasible would the ideas be? What would be best for the protagonist? How would you take action? How would the protagonist respond? Students were asked to take notes of dialogic tasks; in this case, for instance, one student in the group would

write down the ideas that emerged during the brainstorm (time: 15-20 minutes).

4. External dialogue at the classroom level, to explore an additional layer of perspectives. In small-group dialogues, students had shared and compared their responses to a story. In the final learning activity, they were asked to share their group's outcomes with the rest of the class. For example, in unit 2, this enabled students to realize that they were entitled to deepen a particular reading experience (e.g., imagery), but that other groups of students had explored other kinds of experiences (e.g., identification, sympathy). We assumed that this second layer of exploration would help students to consider alternative or elaborated self-other perceptions in relation to texts and themes (time: 5-10 minutes).
5. Short written reflection: students reflected individually on their learning in the unit. For example, in unit 1, students were asked to indicate what went well during the small-group dialogue and what could be improved; in unit 3, they reflected on whether their views on a particular issue (as activated before reading) had changed after reading and talking about a story.

### *2.3.3 Role of the teacher*

Teachers were asked, in preparatory face-to-face meetings with the first author as well as in a written teacher manual, to minimize their interference during small-group dialogues, to allow for a safe space in which students would be free to express and share their ideas, thoughts, experiences, and emotions. When observing a need for guidance or feedback, teachers were asked to provide this in a non-directive way, for instance, by prompting questions instead of explaining what a story is about.

## *2.4 Instruments*

We collected data on four types of variables: students' backgrounds, their perceptions of learning, their learning outcomes, and implementation fidelity of TDLT.

### *2.4.1 Background variables*

*Demographics.* Students indicated their gender, age, exam subjects (educational profile), and most recent grade for Dutch and literature. If no separate grade had been received for literature, we asked students to estimate their grade.

*Author Recognition Test.* We assessed students' familiarity with fiction prior to the intervention with an adapted version of the Author Recognition Test (originally developed by Stanovich & West, 1989). An adapted version of the original test was used, which has been validated for Dutch-speaking teenagers (Koek, Janssen, Hakemulder, & Rijlaarsdam, 2016; Schrijvers et al., 2016). The ART has predictive validity for real-world reading, while avoiding socially desirable answering to questions about reading frequency and motivation (Mar, Oatley, Hirsh, dela Paz & Peterson, 2006; Rain & Mar, 2014). The test presents a list of contemporary and canonical authors' names as well as foils. Participants encircle what they believe to be real authors' names. Its instruction read: "Encircle those names which you know for sure are authors' names. Some of these people are not authors, so do not guess." The number of correctly recognized names minus the wrongly encircled names indicates one's familiarity with fiction. We created three versions, varying the order of the listed names.

The ART was administered as pretest only. It was completed by 539 students (response rate 89.4%). Scores could range from -40 (all foils but no real names encircled) to 40 (all real names and no foils encircled); students scored on average 5.0 ( $SD = 3.3$ ), with a range from -7 to 21. Their familiarity with fiction, thus, seemed to be rather low.

#### 2.4.2 *Perceptions of learning: learner report task*

To measure students' perceptions of learning, we asked them, at the end of the last TDLT unit, to write in their workbooks what they had learned from the project. This task was inspired by the open-ended learner report used by Schrijvers et al. (2016), which was originally developed by De Groot (1980a; 1980b) and has been validated in studies in literature education (Janssen, 1998; Janssen & Rijlaarsdam, 1996) and arts education (Van der Kamp, 1980). Students were given writing prompts such as "I learned that..." and "I now know that I (do not)..." and were asked to write down two responses.

#### 2.4.3 *Learning outcomes*

*Transformative Reading Experiences Questionnaire (TREQ).* First, to assess students' learning outcomes, including insight into human nature, we administered the newly developed TREQ. Items were not story-specific but referred to reading experiences in general (i.e., "When I read stories, I tend to experience X"). A similar approach has been taken, for example, in studies in which partici-

pants completed Miall and Kuiken's (1995) Literary Response Questionnaire (Eva-Wood, 2004; Fialho, Zyngier, & Miall, 2011).

The TREQ included eight scales, each indicating a transformative reading component (Fialho, 2018; see Table 5.1): 1) imagery, 2) identification, 3) experience-taking, 4) character evaluation, 5) sympathy, 6) aesthetic awareness, and self-other insights which was split into: 7) self-insight, and 8) insight into others. The instruction read that the questionnaire was about fictional stories (i.e., short stories and books or novels; not expository books or articles) and explained how items should be scored, using a 5-point scale (*completely disagree – completely agree*).

An initial version of the questionnaire was tested in a pilot study (57 items,  $N = 198$ ). After analyzing internal consistency, we shortened the questionnaire by removing items from scales if a smaller item set maintained sufficient internal consistency. To ensure that students would not complete three identical measures, we created three versions by varying the items in the scales. For example, if an internally consistent scale contained items A-B-C-D, three-item combinations were selected (A-B-C, B-C-D, and A-C-D) if these were internally consistent as well. We were able to apply this procedure to four scales; the remaining four scales were identical in the three versions (see Table 5.1). We ordered items differently in the versions to reduce remembrance effects as a threat to internal validity. Depending on the version, three or four items were formulated negatively.

Table 5.1 includes an example item of each scale and shows that internal consistency for all scales was sufficient across versions and measurements. Therefore, we calculated scale scores. For each measurement point, correlations between imagery on the one hand and aesthetic awareness, self-insights and insights into others on the other hand were rather low (around .20), whereas correlations between self-insights and insight into others were high (around .75; see Appendix E, p. 281).

*Table 5.1. Transformative Reading Experiences Questionnaire: subscales, number of items, example items, and internal consistency per version (A, B, C) and measurement occasion (1, 2, 3)*

Subscale	Items	Example item	Measure	Internal consistency in Cronbach's $\alpha$ ( <i>n</i> students)		
				A	B	C
Imagery	3 <sup>a</sup>	When reading, I clearly picture in my mind the setting where the story takes place.	1	.75 (188)	.69 (183)	.83 (168)
			2	.77 (166)	.73 (175)	.81 (173)
			3	.78 (174)	.75 (161)	.76 (186)
Identification	4 <sup>a</sup>	When I read stories, I recognize something of myself in the protagonist or other characters.	1	.80 (188)	.83 (182)	.69 (168)
			2	.84 (166)	.87 (179)	.88 (174)
			3	.84 (172)	.86 (161)	.82 (186)
Experience-taking	4 <sup>b</sup>	While reading stories, I see, think and feel what a character sees, thinks and feels.	1	.79 (186)	.77 (180)	.75 (166)
			2	.76 (166)	.77 (179)	.77 (173)
			3	.73 (172)	.81 (161)	.75 (183)
Character evaluation	3 <sup>a</sup>	When I read stories, I notice that I evaluate characters positively or negatively.	1	.74 (188)	.74 (183)	.76 (168)
			2	.81 (166)	.75 (179)	.85 (172)
			3	.80 (172)	.83 (161)	.82 (185)
Sympathy	3 <sup>b</sup>	If something bad happens to the protagonist in stories, that doesn't do much to me.	1	.86 (187)	.86 (181)	.76 (169)
			2	.80 (167)	.84 (178)	.80 (173)
			3	.78 (173)	.82 (159)	.76 (186)
Aesthetic awareness	3 <sup>a</sup>	When I read stories, particular words, sentences or passages really stand out for me.	1	.80 (188)	.74 (183)	.73 (167)
			2	.79 (166)	.76 (177)	.83 (173)
			3	.82 (174)	.84 (161)	.80 (185)
Self-insight	5 <sup>b</sup>	I have the idea that I understand myself better because of story reading.	1	.82 (186)	.78 (183)	.77 (167)
			2	.81 (164)	.83 (177)	.87 (172)
			3	.84 (172)	.81 (163)	.82 (186)
Insight into others	4 <sup>b</sup>	Story reading offers me more insight into how other people are.	1	.83 (187)	.84 (182)	.78 (168)
			2	.83 (165)	.83 (177)	.87 (172)
			3	.85 (172)	.85 (161)	.83 (185)

*Note.* For missing values at one measurement point, data was excluded listwise. <sup>a</sup> In all three versions of the questionnaire, items were identical. <sup>b</sup> Items varied per version.

*Story response task.* Second, we collected responses evoked during reading a particular story. Students were asked to read a story or excerpt from a novel and to write their first responses next to it during reading, but they were also allowed to add responses after finishing the story. We provided examples of responses: “something the story reminds you of”, “you find strange or unclear”, “that evokes a feeling in you”, and so forth. Next, students wrote in response to three open-ended questions about 1) what “message” the story conveyed to them, 2) what they expected would happen to the characters and how they relate to each other, and 3) what was, according to them, the value or significance of story reading, for which they could list as little or as many values as came to mind.

To warrant generalizability and avoid remembrance effects, we used three texts at each measurement, all featuring characters who struggled to communicate in a traumatic situation: 1) the opening excerpt of the young adult novel *Birk* by Jaap Robben (2014), in which a boy seems to struggle to tell his mother that he saw his father drown; 2) a short story titled *Something needs to happen* by Maartje Wortel (2015), in which a couple attempts to communicate after the – suggested – suicide of their son; 3) the opening excerpt of the novel *The asylum seeker* by Arnon Grunberg (2003), in which a man’s girlfriend is incurably ill, announces she wants to marry an asylum seeker, and asks her boyfriend to be her witness at the wedding.

#### 2.4.4 Implementation fidelity

To monitor the implementation fidelity of TDLT (O’Donnell, 2008), we used three data sources: teacher logs, time on task observations and students’ workbooks.

*Teacher logs.* To assess the extent to which units were completed as intended, teachers completed logs online after each TDLT unit. For each phase as described in the teaching plans, they indicated whether it was fully, partly or not completed. If their answer was “fully” or “partly”, they evaluated on 5-point scales how feasible it was to teach the phase, how clear the phase was for students, how much order there was in the classroom, and how interested and engaged students were. If they did not complete a phase, they indicated the reason for this by selecting options such as “I did not have enough time” or “I forgot about it”. The response rate was high (96.6% of administered logs). In total, we collected 446 indications of the completeness of a phase (16 missing).

*Time on task observations.* We used time on task observations as an indicator of the proportion of available learning time that students were engaged in the tasks assigned to them (e.g., Karweit, 1984). In each class, either the first author or one of six trained research assistants observed one of the four TDLT units. Observations were scheduled in consultation with teachers. In seven classes, we observed unit 1; in five classes, unit 2; in four classes, unit 3; and in six classes, unit 4.

Students' behavior was coded "on task" if they worked on the given task and did what was asked of them, for example, listening to the teacher or a peer, talking about a task, reading a story, talking to the teacher, or asking questions. It was coded "off task" when students were not working on given tasks or subject matter, but were, for example, looking at cell phones, talking about something else than their task, or being disruptive. If a student's task behavior could not be determined, for example, when another student got into the line of sight, we coded it as "unclear".

We selected six students in the classroom, evenly distributed over the seating arrangement (one toward the front on the right, one toward the back on the right, etcetera). If possible, we selected three boys and three girls, except for classes where gender was unevenly distributed. These six students were observed in multiple rounds of six minutes. The focus was on a single student for one minute, during which this student was observed twice for twenty seconds. This allowed for sufficient time for coding and taking notes. After six minutes (one minute per student), there was a one-minute break, after which the next round started. This process continued until the end of the lesson. In total, we collected 1690 units of observations.

*Students' workbooks.* We randomly selected workbooks of one third of the students ( $n = 198$ ; nine from each class) and checked to which extent tasks were completed. We selected 15 out of 19 tasks to be eligible for coding: four in unit 1, 2, and 3, and three in unit 4. Not eligible were dialogic tasks that did not involve any writing, nor tasks in which one group member took notes, because, if left blank, we could not determine if a student had not completed the task or was not the one appointed to take notes.

We coded the tasks as completed, partly completed, or not completed. If a student had been absent during a unit and thus left blank all tasks in that unit, we treated these cases as missing data, because students' absence told us little about the implementation of the units. In contrast, assessing single tasks in comparison to other tasks in a particular unit or the intervention as a whole might point to implementation issues such as the difficulty of the task or the

time frame in which it was to be completed. Therefore, we deemed it necessary to make a distinction between truly noncompleted tasks and tasks left blank because a student did not attend the unit at all. Of 2,970 observations (15 tasks for 198 students), 138 were missing due to students' absence, which left 2,832 tasks to assess.

### *2.5 Procedures*

The first author met with each teacher before the study started, to familiarize teachers with the materials, structure of the units, and theoretical-empirical background of TDLT. Each teacher then scheduled the units within a maximum of four weeks: some teachers distributed the units over four weeks, whereas others needed to teach them in two weeks, depending on school schedules (excursions, test weeks). TDLT was preceded or followed by an untreated control period.

Trained research assistants or the first author administered the TREQ and story response task three times during regular hours of Dutch class: first, shortly before the start of panel 1 in the research design; second, in between panel 1 and 2; and third, directly after panel 2. Students completed the measures individually on paper. When students completed the learner report task, which was in their workbook, no representative of the research team was present.

### *2.6 Control Condition*

We opted for an untreated control condition in which no literary instruction took place, in view of the time constraints that teachers faced. Consequently, teachers followed their regular program for Dutch language class. They could teach any language domain (e.g., writing skills, grammar), except for literature. To monitor the control condition, teachers completed two digital logs during this period. In total, 40 logs were collected (response rate 90.1%). On average, 3.5 lessons of 50 minutes were taught in the control period. Mostly, reading skills of nonfictional, expository texts (32% of all indications) and rhetoric and argumentation (e.g., formulating opinions, substantiating them with arguments, recognizing sophisms; 22%) were taught. The majority of teachers indicated that students read literary novels at home and completed a task subsequently, which is conventional in Dutch upper secondary schools. In the preparatory meetings, teachers indicated that these tasks were administered several months before the study started (usually in September or October 2016). The tasks, thus, were not adapted to TDLT aims and continued as usual.

## 2.7 Data Analysis

### 2.7.1 Perceptions of learning

We randomly selected learner report tasks of 198 students (nine from each class). As the task did not apply to the control condition, these data are of a descriptive nature. In total, 420 responses were included; occasionally, students wrote down more responses than the two they were asked for. We coded responses inductively, looking for those that occurred more than once. We distinguished ten categories, as shown in Table 5.2. The first author coded all learning experiences. A trained research assistant coded 10% of the data (42 learning experiences). Agreement was 83.4%,  $\kappa = .78$ .

*Table 5.2. Coding scheme for students' perceptions of learning*

Code	Description	Example
Human nature	Learning about one's own or others' personality, life, ways of thinking, behavior, and views on the world	"I learned that people's reactions in certain situations are very different from mine."
Transformative reading experiences (other)	Learning that stories may evoke imagery, identification, experience-taking, evaluations of characters, sympathy, and aesthetic awareness	"I learned to put myself in a character's experience, that became more important to me during reading."
Talking about stories	Learning to talk to peers about stories and reading experiences, apply dialogue guidelines, listen carefully to others, ask follow-up questions, keep a dialogue going	"I learned how you can have a good dialogue because of the dialogue guidelines."
Noticing responses during reading	Learning to focus on emerging thoughts and feelings while reading, to write down responses while reading	"I discovered that I feel quite a lot of responses during reading a story."
Multiple interpretations and responses	Learning that peers express different responses evoked by the same story, that others interpret stories differently, that multiple opinions about stories exist	"You can understand stories in various ways, as became clear from the different opinions in class."
In-depth processing of stories	Learning to consider a topic profoundly, to discuss "deeper layers" in stories, to think "deeply" about a story theme	"I learned to think more in-depth about stories and to consider the deeper layers."
Thinking skills and forming opinions	Learning to formulate an opinion about stories or topics, support ideas by referring to story characteristics, postpone a judgment, reconsider a first impression	"I learned how to formulate my responses and my opinion about a story."

Code	Description	Example
Literary reading and interpretation skills	Learning how to interpret stories, finding out which stories are (not) appreciated, evaluating reading in general	"I learned how to interpret stories."
No learning, negative comments	When students wrote to have learned nothing from TDLT or wrote down a negative remark	"I didn't learn anything. It wasn't interesting."
Unclear	Incomprehensible responses	"Supah."

### 2.7.2 Learning outcomes

*Questionnaire data.* Descriptive statistics of TREQ scale scores for both groups across measurements are included in Appendix E (see p. 284). Analysis of these data consisted of three steps. First, we investigated whether TREQ scales might represent underlying factors. Via principal components analysis with Varimax rotation for all measurements, we extracted two factors (see Table 5.3). Factor 1, *Insights*, contained self-insights, insights into others, and aesthetic awareness as exclusive factor loadings. In addition, identification loaded high on this component. Factor 2, *Imagery*, contained imagery as an exclusive factor loading, but experience-taking also yielded high loadings on this factor. Factors scores were used for subsequent analyses.

Table 5.3. Factor loadings of PCA for TREQ scales, per measurement

Scale	Measurement 1		Measurement 2		Measurement 3	
	Factor 1 <i>Insights</i>	Factor 2 <i>Imagery</i>	Factor 1 <i>Insights</i>	Factor 2 <i>Imagery</i>	Factor 1 <i>Insights</i>	Factor 2 <i>Imagery</i>
Imagery		.85		.85		.87
Identification	.68	.56	.68	.58	.77	.44
Experience-taking	.34	.75	.31	.79	.30	.81
Character evaluation	.39	.58	.37	.63	.41	.58
Sympathy	.56	.52	.51	.64	.52	.64
Aesthetic awareness	.69		.70		.72	
Self-insight	.88		.91		.90	
Insight into others	.87		.88		.86	
Eigenvalue	4.30	1.08	4.63	1.06	4.42	1.21
Variance explained (%)	53.8	13.4	57.9	13.2	55.3	15.2

*Note.* Factor loadings < .30 not displayed. Sampling adequacy and sphericity assumptions were met: KMO values .87, .87 and .84; Bartlett's test of sphericity all  $p$ 's < .001.

Second, we identified potential outliers. We first inspected boxplots for both factors at the three measurement points and identified 17 outliers. As a next step, we applied case wise diagnostics, using regression analyses and graphs to plot students' scores of measurement 1 against 2, and measurement 2 against 3, on both components, using group (A or B) as a factor in the analyses. We looked for cases deviating more than 2.5 standard deviations from the regression line and found 32 outliers, including the 17 cases we already detected using boxplots: 20 in group A and 12 in group B (5.3% of total  $N$ ). We created a filter variable to exclude the outliers from the data.

Third, we treated the remaining data ( $n = 571$ ) as those of two repeated quasi-experiments: we first performed analyses for panel 1, in which group A ( $n = 291$ ) represented the experimental condition and group B ( $n = 280$ ) the control condition, and then repeated the analyses for the second panel in which the order was switched. We took the hierarchical data structure into account: as measurements were nested within students, we applied mixed models analyses with Student as random factor. For both panels and factors (*Insights* and *Imager*), we tested four models. We started with a basic null model in which only the intercept was included (Model 1). In Model 2, we added the effect of pretest as a fixed factor to test to which extent posttest scores can be explained by pretest scores (measurement 1 as pretest in the first panel, and measurement 2 as pretest in the second panel). In Model 3, we added condition as a fixed factor to test whether the average posttest scores differed between the experimental and control condition. Finally, in Model 4, we added the interaction effect of condition and pretest, to test if effects of condition on posttest scores were dependent of pretest scores.

*Story task data.* At each measurement, 198 story tasks were selected (nine from each class, three of each version). When exploring the data inductively, we noticed differences in the extensiveness of responses written next to the stories, as well as in what kind of responses they were, for which we developed two coding schemes. Similarly, we developed coding schemes for the questions about Message, Prediction, and Values of stories. Incomprehensible responses were coded as "unclear" and were not included in further analyses. A trained assistant coded 10% of the selected tasks to assess inter-rater reliability. Appendix E (see p. 282) presents all coding schemes and reliability values.

Next, we treated story task data as those of two repeated quasi-experiments. At the posttest of panel 1, group A ( $n = 99$ ) represented the experimental condition and group B ( $n = 99$ ) the control condition. At the posttest of panel 2, conditions had been switched. Descriptive statistics for story task data

are included in Appendix E (see p. 284). For scale-coded variables (Extensiveness, Message, and Prediction), we used mixed-models analyses with individual students as random factor, testing four models: 1) a basic null model; 2) a model in which we added the effect of pretest as a fixed factor; 3) a model in which we added condition as a fixed factor; and 4) a model in which we added a pretest\*condition interaction effect.

For Story responses, we first analyzed *self-related* and *socially-related* responses (see p. 282). As students could have written down multiple of these responses, we initially focused on whether they had done so at all (1 for "yes, at least once", 0 for "no", for both response types). We used loglinear analyses to test for significant three-way interaction effects of pretest, posttest and condition, which would indicate that relations between pretest and posttest are different for both conditions. If so, we would perform additional analyses on the number of responses written down by the students. If an intervention effect would occur, we would further specify the analysis for additional codes assigned to these responses. For Values of stories, we applied loglinear analysis to test for significant three-way interaction effects on the six Insights categories (see p. 283).

### 3 RESULTS

#### 3.1 Implementation Fidelity

Teacher logs indicated that no units were skipped in any of the classes. Overall, units were completed as intended: 88% of the phases was fully completed, 9.4% partly, and only 2.7% was not completed. Teachers evaluated feasibility ( $M = 4.2$ ,  $SD = .5$ ), clarity for students ( $M = 4.2$ ,  $SD = .5$ ), order in the classroom ( $M = 4.0$ ,  $SD = .6$ ), and students' interest and engagement ( $M = 3.8$ ,  $SD = .8$ ) overall positively. Phases that were not completed mostly occurred toward the end of a unit. Teachers indicated ten times there was too little time left, once that there had been a misunderstanding of the teacher guidelines, and once that a phase simply had been forgotten.

Students were on task in 72.5% of the observed time, and off task in 24.9% of the time (2.5% coded "unclear"). There was a large variability between classes ( $\chi^2(42) = 155.02$ ,  $p < .001$ ), with on task percentages ranging from 54.1 to 88.6. The threshold for effective teaching lies around 80% time on task (e.g., Kauchak & Eggen, 1993; Muijs & Reynolds, 2010). In only five classes, the percentage was above 80%, which was cause for concern.

Of 2,832 workbook tasks, 76.7% was fully completed, 10.1% was partly completed, and 13.2% was not completed. Two individual reflection tasks, at the

end of units 2 and 3, were remarkably often not completed: the task in unit 2 was left blank by 90% of the students, and the task in unit 3 by 54%. In light of teachers' logs, we suspect there was too little time left at the end of these units for completing the tasks.

All in all, implementation fidelity seemed to be satisfactory in terms of the number of units and unit phases that were completed, as well as the tasks in students' workbooks. However, task-oriented behavior among students was not up to standard. The difference among classes suggests that TDLT has been implemented effectively in some, but not in all classes.

### 3.2 *Students' Perceptions of Learning*

Students' responses about what they felt they learned from TDLT offered a first indication of whether the intervention aims were achieved. Figure 5.2 shows the percentages of responses in each category.

The largest share of learning experiences (28.1%) concerned talking about stories. Examples included: "I learned how you can have a good dialogue because of the dialogue guidelines", "I learned to talk about stories, you never really talk to anyone about a book or story, but now we did and I think that's good", and "I always thought I was not good at talking about literature. But after the lessons, I know I can do it."



Figure 5.2. *Distribution of learning experiences, in percentages of N = 420 responses.*

Next, 15% of the learning experiences concerned transformative reading experiences other than insight into human nature, such as: "I learned how to put

myself in the experience of a character and that has become more important to me during reading", "I found out that I can recognize things in stories, and then I compare it to real life, so I can better put myself in the story situation", and "I learned that I can really identify with most characters." Third, 12.6% concerned thinking skills and forming opinions, for instance: "I discovered how a good dialogue can make you think differently and reconsider a story", "I learned to postpone my first judgments", and "I learned how to formulate my responses and my opinion about a story." Responses in other categories occurred less frequently.

### 3.3 Learning Outcomes: TREQ Data

Results of the fit and comparison for the four models, for Insight and Imagery in panel 1 and 2, are shown in Table 5.4. All Models 2 fit the data significantly better than Models 1, which indicated that students' average posttest scores were partly explained by their pretest scores. Effects of condition and interaction effects were not consistent (see Table 5.5 for estimates).

*Panel 1.* A main effect of condition on Insights in panel 1 occurred: mean posttest scores of TDLT students differed significantly from mean posttest scores of control students. Contrary to our expectations, the regression coefficient for the control condition was larger than for the TDLT condition (pretest  $\beta = .744$ ; control condition added  $\beta = .148$ ). Thus, at the posttest, control students scored significantly higher than TDLT students. The pretest\*condition interaction for Insights was not significant. For Imagery, only Model 2 had a better fit than Model 1: there was no effect of condition, nor a pretest\*condition interaction effect.

*Panel 2.* Table 5.4 shows there was no main effect of condition on Insights in panel 2. However, adding a pretest\*condition interaction effect significantly improved the model. Parameter estimates (see Table 5.5) indicated that for TDLT students, posttest scores were more strongly dependent on the pretest than for control students: in the TDLT condition, the effect of condition was larger for those students who had relatively high pretest scores ( $\beta = .117$ ; Cohen's  $d = .10$ ). For Imagery in panel 2, there was a main effect of condition (see Table 5.4), in favor of TDLT (pretest  $\beta = .756$ , experimental condition added  $\beta = .124$ ; see Table 5.11). Thus, at the posttest in panel 2, TDLT students scored significantly higher than control students ( $d = .10$ ). The interaction effect of pretest\*condition was not significant.

Table 5.4. Fit and comparison of models 1-4 for Insights and Imagery in panel 1 and 2

Panel and variable	Model	-2LL	$N_{\text{pars}}$	Models	Comparison		$p$
					$\chi^2$	$df$	
Panel 1 <i>Insights</i>	1. null	1252.30	3				
	2. + pretest	854.02	4	2 vs 1	398.28	1	< .001
	3. + condition	847.77	5	3 vs 2	6.25	1	.012 <sup>a</sup>
	4. + pretest*condition	844.42	6	4 vs 3	3.45	1	.067
Panel 1 <i>Imagery</i>	1. null	1239.13	3				
	2. + pretest	868.79	4	2 vs 1	370.34	1	< .001
	3. + condition	866.27	5	3 vs 2	2.52	1	.112
	4. + pretest*condition	864.37	6	4 vs 3	1.90	1	.168
Panel 2 <i>Insights</i>	1. null	1216.67	3				
	2. + pretest	794.87	4	2 vs 1	421.80	1	< .001
	3. + condition	792.12	5	3 vs 2	2.75	1	.097
	4. + pretest*condition	787.99	6	4 vs 2	6.88	1	.001 <sup>b</sup>
				4 vs 3	4.13	1	.042 <sup>b</sup>
Panel 2 <i>Imagery</i>	1. null	1201.84	3				
	2. + pretest	810.09	4	2 vs 1	391.75	1	< .001
	3. + condition	805.50	5	3 vs 2	4.59	1	.032 <sup>a</sup>
	4. + pretest*condition	804.10	6	4 vs 3	1.40	1	.237

Note. <sup>a</sup> Main effect of condition. <sup>b</sup> Pretest\*condition interaction effect.

Table 5.5. Parameter estimates for main and interaction effects on Insights and Imagery

Panel, variable, model	Parameter	B	SE	$df$	$t$	$p$
Panel 1 <i>Insights</i>	Intercept	-.070	.041	448	-1.71	.088
	Pretest	.744	.029	448	25.25	< .000
Model 3	Control condition	.148	.059	448	2.51	.012
Panel 2 <i>Insights</i>	Intercept	-.015	.040	439	-.39	.698
	Pretest	.709	.040	439	17.54	< .001
Model 4	Intervention*pretest	.117	.057	439	2.04	.042
Panel 2 <i>Imagery</i>	Intercept	-.044	.040	439	-1.08	.280
	Pretest	.756	.030	439	25.15	< .000
Model 3	Intervention condition	.124	.058	439	2.15	.032

*Conclusions.* In short, the questionnaire data yielded incoherent results. We found a significant interaction effect for Insights and a main effect of condition for Imagery in panel 2, both in favor of the experimental condition, but these outcomes were not consistent with panel 1. Moreover, effect sizes were small. To look for additional explanations, we systematically added background variables to the models for both panels: gender, educational track, grade for Dutch, estimated grade for literature, and familiarity with fiction. None of the variables significantly improved the models in either panel; therefore, they could not further explain the results.

### 3.4 Learning Outcomes: Story Task Data

*Panel 1.* A main effect of condition on Extensiveness occurred in panel 1 (see Table 5.6), but in favor of the control condition (pretest  $\beta = .256$ ; control condition added  $\beta = .349$ ; see Table 5.7). Thus, taking pretest into account, control students responded more extensively to stories at the posttest than TDLT students. For Self-related and Socially-related story responses, loglinear analyses revealed no significant three-way interaction effects, indicating that relations between pretest and posttest had not been affected by TDLT.

Table 5.6. Fit and comparison of models 1-4 for Extensiveness in panel 1

Model	-2LL	$N_{\text{pars}}$	Models	Comparison		$p$
				$\chi^2$	$df$	
1. null	546.25	3				
2. + pretest	527.94	4	2 vs 1	18.31	1	< .001
3. + condition	520.86	5	3 vs 2	7.08	1	.008*
4. + pretest*condition	520.56	6	4 vs 3	.30	1	.584

Note. \* Main effect of condition.

Table 5.7. Parameter estimates under model 3 on Extensiveness in panel 1

Parameter	$\beta$	$SE$	$df$	$t$	$p$
Intercept	-.070	.041	448	-1.71	.088
Pretest	.744	.029	448	25.25	< .000
Control condition	.148	.059	448	2.51	.012

For Message and Prediction, mixed-models analyses indicated no main effects of condition, nor significant pretest\*condition interactions. For Values of stories, loglinear analysis revealed one significant three-way interaction effect, for the number of students who reported 'gaining insights in life and the world' as a value ( $\chi^2(1) = 5.81, p = .016$ ). For this variable, thus, the relation of pretest and posttest differed significantly between conditions. Yet, two separate chi-square tests showed that this difference occurred for control students ( $\chi^2(1) = 9.40, p = .002$ ), but not for TDLT students. The odds ratio indicated that control students were 2.2 times more likely to mention this value than TDLT students.

*Panel 2.* For all scale variables, mixed-models analyses indicated no main effects of condition, nor significant pretest\*condition interaction effects. Likewise, for categorically coded variables, loglinear analyses revealed no significant three-way interactions, indicating that the relations between pretest and posttest had not been affected by the intervention.

## 4 DISCUSSION

We set out to investigate whether a dialogically oriented literature classroom intervention that was based upon a model of transformative reading would foster students' transformative reading experiences – most importantly, their insight into human nature. Our main conclusion is that, whereas students reported relevant perceptions of learning, we were not able to demonstrate any effects of the intervention on their insight into human nature, nor on other transformative reading components. Students seemed to perceive the intervention as a project that taught them how to engage in dialogues about stories, whereas this was only partly what the intervention aimed for. In this section, we contextualize our findings in an attempt to explain them. We address the limitations of this study and consider next steps for future intervention studies in the literature classroom.

### 4.1 *Explaining the Absence of Effects*

In this study, we set up a quasi-experimental pretest-posttest design with switching replications. Ideally, intervention effects would be demonstrated at the posttest of panel 1, which would sustain for group A and be replicated for group B at the posttest of panel 2. However, the main effects or interaction effects we found were either in favor of the control condition or, when in favor of TDLT, negligibly small.

Effects may be absent if an intervention is not implemented as intended in its original design (O'Donnell, 2008). Yet, implementation data – teacher logs, time on task observations, and completed workbook tasks – indicated that TDLT was, overall, implemented as intended, even though the overall percentage of observed time that students were on task (72.5%) was below the 80% threshold for effective teaching and learning. Furthermore, our findings cannot be ascribed to insufficient sample sizes, nor to insufficient reliability of the questionnaire or the coding systems used to analyze the story task data. Finally, as we took pretest scores and the hierarchical data structure into account, these aspects cannot explain the absence of effects. However, other limitations, in terms of intervention design and methodology, may function as explanations for our findings.

#### *4.1.1 Explanations based on intervention design*

First, TDLT may have been too skills-oriented. Students felt that they learned what they were most explicitly trained to do: engaging in external dialogues about short stories. Theoretically, we expected that sharing responses and considering multiple perspectives on relevant story themes would result in insight into human nature, but attention for gaining such insights may have remained too implicit for this group of students. Indeed, when assessing the practicality of TDLT, students as well as teachers indicated that its ultimate purpose was not clear enough and might have been made more explicit (see Chapter 4; TDLT-1).

Second, we are unsure whether transformative reading (Fialho, 2018) may at all occur in 10<sup>th</sup> grade students, who can be considered novice readers of literary texts. Students reported some transformative reading experiences in the learner report task, but they did so rather infrequently and seldomly reported about insight into human nature. Perhaps, TDLT would have been more effective if it were implemented in grade 11, the final year of the higher general secondary education track. Moreover, Dutch secondary school students do not often read for enjoyment, have a neutral to negative attitude toward fiction, and their motivation for literary reading decreases in higher grades of secondary school (OECD, 2010; Van Schooten, 2005). Their relatively low ART scores suggest that the participants in this study were no exception. As a consequence, students may not have had sufficient experience with literary story reading to reflect profoundly on their own reading experiences and their insight into human nature. This was confirmed by several of their teachers, who indicated that the focus on noticing transformative reading experiences was

completely new to their students and that students sometimes felt confused or even frustrated when being asked to reflect upon what a literary text evoked in them.

Furthermore, as a potential consequence of their limited freedom of text choice, students may not have been fully engaged in reading. As outlined in the Introduction of this chapter, it was challenging to bridge the gap between teachers' and curricular demands on the one hand, and students' autonomy that may foster engaged reading on the other hand. Moreover, the research literature is ambiguous: studies have suggested that self-selected reading may foster students' insight into human nature (e.g., Ivey & Johnston, 2015; Richardson & Eccles, 2007; Rothbauer, 2011), but similar results have been found in a review of intervention studies in which students were hardly given any autonomy in selecting reading material (Schrijvers et al., 2018). The role of freedom of choice in reading materials for developing insight into human nature, thus, is yet to be further illuminated.

In addition, the number of stories or units may have been too few. Time constraints allowed for only four intervention units of 50 minutes; in three of these, a story was read. Reading more stories or interacting with a story more in-depth may result in different learning outcomes. Moreover, as we opted for variety and differentiation (i.e., using stories with various themes), there was no accumulated reflection on one particular social-moral theme. Selecting stories around a single social-moral theme might offer students the opportunity to consider more extensively, in multiple units, how such a theme affects and relates to themselves and other human beings.

Finally, the outcomes of our study raise the intriguing question why students scored higher on various posttests when they had just participated in the untreated control condition, compared to those who had received the intervention. The control condition was unrelated to the intervention: no instructional attention was paid to literary texts and literary responses. Why attention for reading skills of expository texts and rhetoric and argumentation have resulted in higher posttest scores remains unclear. As effects in favor of the control condition occurred not only in panel 2 but also in panel 1, they cannot be explained as 'delayed' intervention effects: on the contrary, raw TREQ scores (see Appendix E, p. 284) indicated that intervention students' mean scores mostly decreased from pretest to posttest. We can only speculate why this happened: perhaps students experienced reluctance, boredom or frustration, even though their perceptions of learning did not support this assumption.

### 4.1.2 *Methodological explanations*

One might argue that potential reluctance, boredom or frustration may also have applied to repeated administration of measurements in a short period of time. Even though a research design with switching replications is, theoretically, a strong design to assess intervention effects (Shadish, Cook, & Campbell, 2002), in this study it involved three measurement points in at most eight weeks. However, if repetition would have negatively affected students' posttest scores, we would have expected to see a similar effect in the control conditions of both panels.

Specifically for story task data, an explanation for the absence of effects might be found in the role of students' writing abilities. We collected no information about their writing skills, for example, in the form of their most recent results on writing tests. If writing is difficult for students, think-aloud tasks (e.g., as applied by Janssen, Braaksma, & Couzijn, 2009) may be a more valid way of gaining insights into students' responses to stories.

### 4.2 *Implications for Practice and Research*

This study was the first to empirically investigate, in the specific context of 10<sup>th</sup> grade of higher general secondary education in the Netherlands, whether dialogically oriented literature lessons would foster students' insight into human nature. Because we did not find considerable effects of condition, implications for practice can only be based on descriptive learner report data. This study suggests that strategy instruction helps students to engage in external dialogues about stories and reading experiences. In particular, students found practical guidelines for dialogues helpful, as illustrated by a large share of learning experiences such as: "I learned how to engage in a good dialogue, due to the dialogue guidelines (FLUENT)" and "I learned 'FLUENT', so how to follow a good conversation. And how to go more into depth". Research into the particular needs of students in the higher general education track in the Netherlands confirms that they prefer clear structures and guidelines in their learning materials: they feel the need to know what to do and how to do it (Vermaas & Van der Linden, 2007). This may also imply that the ultimate purpose of the intervention – developing insight into human nature – remained too implicit for this population of students.

Therefore, our first step in future research is to assess whether redesigning the intervention would result in the outcomes that could not be demonstrated in this paper. The redesign should operationalize several points for improvement as discussed here, among which putting more emphasis on the purpose

of eliciting transformative reading experiences and developing insight into human nature, a longer duration that allows for this group of relatively novice readers to read and interact with a larger selection of stories, and centering those stories and interactions around a single social-moral theme.

In addition, as our measures relied on individual self-reports and could therefore not capture external dialogues among students, in future studies researchers might choose to administer dialogue tasks as pretest and posttest. Analysis of observations might demonstrate changes in how students refer in their group talk to transformative reading experiences, including insight into human nature. Finally, such studies may also add to existing knowledge about transformative reading by conducting interview and think-aloud studies with adolescent participants. This would offer valuable insights into their reading and reflection processes, which may offer relevant input for the design of interventions like the one described here.

#### 4.3 *Conclusion*

The current literature classroom intervention study demonstrated that students, as a result of only four 50-minute units, reported that they learned to talk about their responses to literary stories, and, to a lesser extent, to notice transformative reading experiences and to develop their thinking and opinionating skills. However, the study could not demonstrate any consistent increase of students' transformative reading experiences, including their insight into human nature. This was true for their reflections on their reading experiences in general, as well as for story-specific responses as expressed in writing during and directly after reading.

Nonetheless, our study provides a number of suggestions for how future interventions might successfully foster students' insight into human nature. After all, the absence of statistical effects does not necessarily mean that such outcomes cannot be achieved in the literature classroom, as is suggested by a response written by a student in the workbook: "During the lessons, I learned that the meaning of life for other people is much more complex than I thought it was." Responses such as this one suggest that it is worthwhile, for teachers and researchers alike, to pursue the systematic design and empirical testing of literature classroom interventions that aim to foster students' insights into who they are, how they relate to others, and how they position themselves in life and in the world around them.



## CHAPTER 6

### TRANSFORMATIVE DIALOGIC LITERATURE TEACHING FOSTERS ADOLESCENTS' INSIGHT INTO HUMAN NATURE

This quasi-experimental study assessed the effects of the newly developed Transformative Dialogic Literature Teaching (or TDLT) intervention on 15-year-old students' insight into human nature, reasons for reading, use of literary reading strategies, and motivation for literature education. Six TDLT lessons centered around short stories about "justice and injustice". Students were stimulated to engage in internal dialogues with stories and in external dialogues with peers about stories and reading experiences. TDLT students ( $n = 166$ ) were compared to students who received lessons focused on analysis of literary texts ( $n = 166$ ). Results showed that TDLT fostered students' insight into human nature, their support for eudaimonic reasons for reading, their reported use of strategies to deal with difficulties in literary texts, and their motivation for literature education. Strategy use and two basic needs for motivation, competence and relatedness, mediated effects of TDLT. Limitations and implications for future work are discussed.

#### 1 INTRODUCTION

"What part of the story really stood out to you?

The part where the man was executed.

How come?

Well, he got the death penalty and not even begged for sentence reduction.

Maybe it was simply because he acknowledges that he did wrong.

But still, wouldn't you beg for sentence reduction? Nobody wants to die.

I think he had so much regret that he thought the death penalty did him justice.

Yeah, that could be.

After all, he was a kind man, according to his friends and colleagues."

This response, of a 15-year-old student, is part of a written dialogue about a short story with an imaginary classmate. The student reasons about behavior and motives: why does a character act in a certain way? She considers possible scenarios ("Maybe", "But still", "I think") and compares the character to herself ("Wouldn't you beg for reduction of the sentence?"). Thereby, she explicates her ideas about how people may respond to complex situations that evoke

social-moral questions. This example, from data collected in 10<sup>th</sup> grade literature classrooms in the Netherlands, illustrates that reading and responding to literary texts may be a catalyst for developing insight into what it means to be human, or into human nature.

Developing and deepening insight into human nature is not a formal objective of the Dutch literature curriculum. However, it appears to be more important than ever for students to learn to reflect on who they are and how they relate to other human beings, considering the globalizing society in which they grow up. Indeed, Dutch teachers regard fostering personal development as an important objective of literature teaching (Janssen, 1998; Oberon, 2016), and suggested that literary reading may familiarize students with other worlds, contributes to moral development, and helps them to think about people's choices, about themselves, others and the world (Curriculum.nu, 2018a). A synthesis of intervention research suggests that such claims may be valid (Schrijvers, Janssen, Fialho, & Rijlaarsdam, 2018), but few studies address the Dutch curriculum.

Therefore, we designed an intervention, titled Transformative Dialogic Literature Teaching (TDLT), and investigated its effects on Dutch students' insight into human nature, including insight into themselves and others. In addition, we investigated whether this intervention would alleviate two particular challenges in Dutch literature classrooms: students' struggle to deal with difficulties during literary reading and their low motivation for literature education.

### 1.1 *Insight into Human Nature*

Reading fictional and literary texts is assumed to affect readers' perceptions of self and others (Keen, 2007; Nussbaum, 1995), or in our terms, their insight into human nature. Results of empirical studies support these claims (Hakemulder, Fialho, & Bal, 2016; Koopman & Hakemulder, 2015). Insight into human nature is theorized to arise from simulated social experiences that readers live through when they read fictional texts, for example, novels, stories, or poems (Mar & Oatley, 2008). This process has also been conceptualized as "transformative reading" (Fialho, 2012). In phenomenological studies, Fialho (2018) found that adult readers may experience self-other perceptual depth, which resembles what we call "insight into human nature", as it entails both self and others. Fialho showed that self-other perceptual depth was predicted by six other experiences: vividly imagining story setting and characters (*imagery*); recognizing aspects of self or others in characters (*identification*); enacting and embodying the experiences of a character (*experience-taking*); evaluating characters, posi-

tively or negatively (*character evaluations*); feeling sympathy and compassion for characters (*sympathy*); and being aware of striking words, phrases or sentences (*aesthetic awareness*).

For adolescent readers, it has yet to be determined whether these experiences are precedents of self-other insights. Empirical studies suggest that processes and outcomes similar to transformative reading may occur amongst adolescents. For instance, they were found to construct their possible future selves when reading fiction (Richardson & Eccles, 2007), compared their own lives to stories and engaged empathetically with characters (Charlton, Pette, & Burbaum, 2004), and better understood experiences of others, which offered them new options for their own lives (Rothbauer, 2011). In addition, adolescents developed insight into human nature when reading was school-based rather than a leisure activity (e.g., Malo-Juvera, 2014; Schrijvers, Janssen, Fialho, & Rijlaarsdam, 2016; White, 1995).

The literature classroom, thereby, seems a promising domain for adolescents to develop insight into themselves, fictional others, and real-world others. Examples include insight into previously unrecognized personal qualities or shortcomings, insight into self-other relations, understandings of why characters think, feel and behave in a certain way, understandings of individual others or groups of people, and insight into moral dilemmas that people may face. Moreover, readers may consider gaining insight into human life to be a meaningful reason for reading (eudaimonic reasons for reading), in addition to reading for pleasure and enjoyment (hedonic reasons), or for plot (Koopman, 2016; Miall & Kuiken, 1995). But how might we best design literary instruction when the aim is to foster students' insight into human nature?

## 1.2 *Design Principles*

In a review, we identified a set of instructional design principles based on intervention studies for which empirical support was found (Schrijvers, Janssen, Fialho, & Rijlaarsdam, 2018). The first principle suggests that insight into human nature may be fostered when fictional texts that are read are thematically relevant for an intervention aim, such as texts addressing social relations or moral dilemmas. Malo-Juvera (2014), for instance, used a young adult novel about sexual harassment to affect students' attitudes toward such behavior.

Second, exploratory dialogic activities appeared to be important to foster insight into human nature (e.g., Adler & Foster, 1997; Eva-Wood, 2004; Malo-Juvera, 2014; White, 1995). If reading would remain an individual activity, readers only explore their own reading experiences. Talking to peers offers the op-

portunity to consider a broader range of thoughts, questions, feelings, ideas, and perspectives in response to texts and the social-moral themes they address. This may happen in small-group or whole-class dialogues, or in a combination where the latter follows the former. Such a buildup creates multiple layers of sharing responses and interpretations.

The third design principle indicates that students, to prepare for external dialogues, should engage in internal dialogues with texts. In such dialogues, students establish awareness of the responses that texts evoke in them as well as of how these responses are related to the outer-textual world, which may stimulate them to engage in transactional processes of meaning-making (see Rosenblatt, 1938/1983). Internal dialogues may be stimulated by writing tasks that prompt students to activate previous personal experiences before reading (White, 1995), to notice and annotate responses during reading (Eva-Wood, 2004), and/or to write (reflective) responses directly after reading (Malo-Juvera, 2014). However, implementing these three instructional design principles may not be sufficient, as students face challenges that potentially interfere with developing insight into human nature.

### 1.3 *Challenges for Students*

First, secondary school students may struggle to deal with difficulties in fictional and literary texts. As relatively novice readers of literature, they may doubt their own abilities as readers (Levine & Horton, 2013). Their responses tend to be confined to literal reiterations, character descriptions, or simple evaluations (see the introduction by McCarthy & Goldman, 2017, for an overview of the literature). To facilitate comprehension, attending solely to personal responses, relying on background knowledge, and making real-world inferences may not suffice (see McMaster et al., 2012, for research with younger children). Rather, there is a need to guide students toward making connections between the initial responses they notice during internal dialogues with texts, and textual elements (Eva-Wood, 2004). In 10<sup>th</sup> grade, this may concern literary devices (e.g., flashbacks, focalization, psychological suspense, motifs); students may consider how these devices evoke personal experiences and responses.

Moreover, students may yet have to learn literary reading strategies, such as monitoring their reading process, thinking about what they read, and actively considering questions that arise (Peskin, 1998). This was illustrated by what teachers reported in a recent study in 10<sup>th</sup> grade Dutch literature classrooms: students who struggle with comprehension oftentimes simply ignore aspects of texts that are difficult to understand (see Chapter 4). Thus, students' metacog-

nitive awareness of applying reading strategies seems limited (Mokhtari & Reichard, 2002). All in all, an instructional approach that explicitly attends to dealing with difficulties in literary reading may increase the likelihood that students engage in meaningful dialogues with texts.

A second challenge is Dutch students' low motivation for literature education (Stokmans, 2009; Van Schooten, 2005). As a remedy, a student-centered literature curriculum may contribute to students' motivation for literary reading, compared to a traditional teacher-centered curriculum (Verboord, 2005). Moreover, student-centered and affect-oriented interventions implemented in Dutch and other literature curricula were found to positively influence outcomes related to motivation, such as task interest (Henschel, Meier, & Roick, 2016), appreciation of literary texts (Janssen, Braaksma, & Couzijn, 2009), involvement in texts (Fialho, Zyngier, & Miall, 2012), and contributions to classroom talk (Eva-Wood, 2004). An instructional approach oriented toward student-readers' authentic responses to texts may therefore increase motivation for literature education, which may be determined by the extent to which students' needs for competence, autonomy and relatedness are satisfied (Deci & Ryan, 2000).

#### 1.4 *Hypotheses*

We designed a reader- and affect-oriented intervention for 10<sup>th</sup> grade literature classrooms (Transformative Dialogic Literature Teaching, or TDLT), based on a theoretical-empirical model of transformative reading and a set of design principles concerning thematically relevant texts, internal dialogues, and external dialogues. First, as Figure 6.1 shows, we expect that TDLT fosters students' insight into human nature and their eudaimonic reasons for reading. Second, we expect that TDLT positively affects students' self-reported use of strategies to deal with difficulties during literary reading (for short: "strategy use") and their motivation for literature education (operationalized as feelings of autonomy, competence, and relatedness). Finally, we hypothesize that strategy use and motivation function as mediators of TDLT effects on insight into human nature and eudaimonic reasons for reading. Strategy use may positively affect students' internal dialogues with texts, via which they may develop insight into human nature. Motivation for literature class may enhance engagement in response tasks, which may be a precondition to develop insight into human nature and to consider this insight a reason for reading.

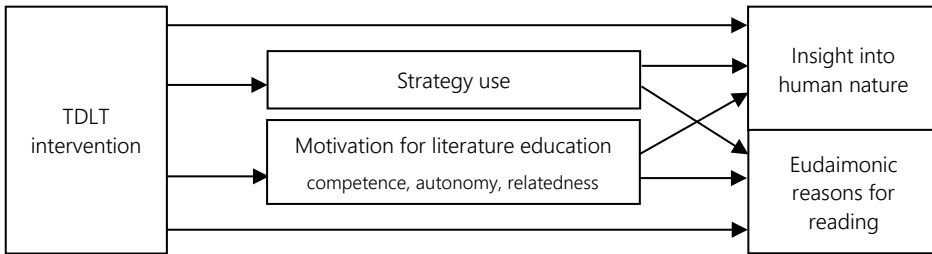


Figure 6.1. Hypothesized direct and mediated effects of TDLT.

## 2 METHOD

### 2.1 Research Design

We implemented a quasi-experimental design with a pretest, posttest and delayed posttest. Six classes were assigned to TDLT and six to a control condition. TDLT teachers scheduled the units in six weeks, in the Fall semester of 2017. The control condition also lasted for six weeks. It involved regular literary instruction ("business as usual"), focused on identifying literary devices by analyzing literary texts (see section 2.8 in this chapter).

### 2.2 Participants

Ten Dutch teachers from five schools volunteered to participate in the study (see Table 6.1). Three teachers had already been involved in designing TDLT (see Chapter 4), and were therefore assigned to the TDLT condition. The remaining teachers signed up for either TDLT or the control condition. To avoid contamination of conditions, teachers were involved in one of both conditions. Six teachers taught a TDLT class; four taught either one or two control classes. All TDLT teachers were female, who had on average 18.7 years of teaching experience ( $SD = 12.3$ ). Five control classes were taught by females; one by a male teacher. Their experience ( $M = 13.3$  years;  $SD = 5.6$ ) did not differ significantly from TDLT teachers.

Teachers' classes were 10<sup>th</sup> grade classes in the higher general secondary education track, which prepares for higher vocational education but not for university. In both conditions, 166 students participated ( $N = 332$ ). Conditions did not differ significantly in gender (53.6% females in TDLT, 45.2% in control condition), average age (15.5 years old in both), and average grade for the subject Dutch language and literature (6.6 out of 10 for both). In addition, we assessed students' Familiarity with fiction and Trait empathy (see Section 2.4 in

this chapter); no significant differences between conditions were found. Students' parents received an informed consent letter and could object to their child's participation. None of them withheld their consent.

*Table 6.1. Participating teachers*

Condition	Teacher ID and status	School	Location	Classes	Students
TDLT	101 Involved in TDLT design	A	Provincial town	1	30
TDLT	102 Involved in TDLT design	A		1	29
TDLT	103 Involved in TDLT design	B	Major city	1	26
TDLT	201 Signed up for TDLT	B		1	28
TDLT	202 Signed up for TDLT	C	Provincial city	1	25
TDLT	203 Signed up for TDLT	D	Provincial city	1	28
Control	301 Signed up for control	D		1	26
Control	302 Signed up for control	D		1	26
Control	303 Signed up for control	E	Major city	2	27, 28
Control	304 Signed up for control	A	(see above)	2	32, 27

### 2.3 Intervention

TDLT consisted of one preparatory and five reading-and-dialogue units. In total, TDLT included about 300 minutes of classroom work; teachers scheduled the units in 50- or 60-minute lessons. In addition, students completed homework assignments. Short stories were read that centered around "justice and injustice" (see Table 4.3, p. 112). Teachers involved in designing TDLT classified the stories as literary texts in terms of language use that was unconventional for students (Van Peer, Zyngier, & Hakemulder, 2007) and gaps readers needed to fill in (Iser, 1980). We included canonical and more recently published stories.

The primary aim of TDLT was for students to learn to express, orally and in writing, a) the responses that stories evoked in them, b) which new insights into themselves, others and social life these stories offered them, and c) which literary devices evoked these responses and insights. To achieve these aims, students were taught strategies for both internal and external dialogues. Table 6.2 presents an overview of the teaching and learning activities in each unit (see Appendix D for full TDLT overview). In the preparatory unit strategies for external dialogues were introduced. Students observed and evaluated videos of peers talking about a story, received explicit instruction about the strategy –

which was summarized on a “first aid card” they used throughout TDLT – and applied the strategy in a small-group dialogue about famous quotes about literature and reading that were printed on small cards (e.g., “A good book has no ending”, R.D. Cumming).

In all subsequent reading-and-dialogue units, external dialogues were applied in combination with internal dialogues; together, they formed the two-step basic structure that was central in TDLT (see Table 6.2).

Internal dialogues remained implicit in unit 1 and 2, were explicitly introduced in unit 3, and were applied during reading in units 3 to 6. The first aid card also included strategies for dealing with difficulties during reading that the teacher introduced in unit 3, for example, writing down question marks, pausing to think, and asking for help. Moreover, from unit 3 onwards, internal and external dialogues focused on transformative reading experiences: students considered, for example, experiences of imagery and sympathy. The activities in the units were miscellaneous, short, and high-paced to keep students engaged and motivated. Students were stimulated to monitor their progress by working with a rubric (see Appendix D, p. 280).

Teachers were given guidelines for providing students with feedback and guiding their dialogic processes, such as prompts and questions for stimulating dialogues in a student group (e.g., “What else does this story make you think about?”; “Could someone have another opinion?”; “I hear you struggling with this. Let’s talk about how to solve that issue”).

#### *2.4 Instruments*

Table 6.3 provides an overview of all questionnaires and subscales used to measure various indicators of students’ insight into human nature (e.g., transformative reading experiences, empathy for characters, moral competence), their reasons for reading, strategy use, and motivation for literature education, including example items, internal consistency, scoring scales, and references. We randomized the order of items included at two or more measurement occasions. We used existing, validated questionnaires, except for the Transformative Reading Experiences Questionnaire (TREQ), which was developed based on Fialho’s (2018) transformative reading model. Descriptives for quantitative data are included in Appendix F (see p. 288).

*Table 6.2. Operationalization of internal and external dialogues, in sequence (1, 2, ...)*

Unit	Internal dialogue	External dialogues
1	1. Implicit	2. Learning-by-observation: video of peers talking about story 1 3. Explicit instruction 4. Practice
2	2. Implicit	1. Apply to theme: talk about "justice and injustice" 3. Apply to theme and story 2: talk about injustice in story, opinion about story, and support with literary devices
3	1. Preparation: write response to moral statement relevant to story 3 2. Explicit instruction 3. Learning-by-observation: teacher thinks aloud, annotates part of story 3 4. Apply to rest of story 3: annotate responses, reflect on prominent responses	5. Apply to story 3: small-group talk, deepen prominent responses 6. Apply in class: share experiences, teacher-led
4	1. Apply to part of story 4: annotate responses, reflect on prominent responses 3. Apply to part of story 4: write story ending as response 5. Apply to end of story 4: annotate responses	2. Apply to part of story 4: imagine characters' position, small-group talk about just and unjust story ends 4. Apply to written response: share feedback on story endings in pairs 6. Apply in class: teacher-led talk about justice and injustice in original and written story endings
5	1. Recall of previous internal dialogues	2. Apply to stories 2-4: compare reading experiences, justice and injustice, formulate life lesson in small groups 3. Apply in classroom: share life lessons, teacher-led talk
6	1. Apply to story 5: annotate responses 3. Apply to story 6 (homework): annotate responses, write dialogue with imaginary peer	2. Apply to story 5: speed dates in pairs about character, about responses and literary devices; about injustice

*Note.* For an overview of all phases in each unit, see Appendix D.

Table 6.3. Variables and quantitative effect measures: Subscales, example items, number of items, internal consistency, per measurement occasion

Variable	Instrument	Subscale: Example item	Items	Cronbach's $\alpha$			
				T1	T2	T3	
Human nature: self, fictional others, real-world others	Transformative Reading Experiences Questionnaire (developed for this project; in Dutch; see Chapter 5) <sup>a</sup>	<i>Imagery</i> When I'm reading, I clearly picture in my mind the setting where the story takes place.	3	.82	.77	.75	
		<i>Identification</i> When I'm reading, I recognize something of myself in the protagonist or other characters.	4	.79	.80	.77	
		<i>Experience-taking</i> When I read stories, I see, think, and feel what a character sees, thinks, and feels.	4	.78	.79	.88	
		<i>Character evaluation</i> When I read stories, I notice that I evaluate characters positively or negatively.	3	.80	.78	.79	
		<i>Sympathy</i> If something bad happens to a story character, that doesn't do much to me.	3	.87	.81	.78	
		<i>Aesthetic awareness</i> During reading, particular words or sentences really stand out for me.	3	.79	.79	.77	
		<i>Self-insights</i> I have the idea that I understand myself better because of story reading.	5	.86	.85	.81	
		<i>Insights into others</i> Story reading offers me more insight into how other people are.	4	.85	.82	.80	
Human nature: fictional others	Literary Response Questionnaire (Miall & Kuiken, 1995; translation by Van Schooten, 2005) <sup>a</sup>	<i>Empathy for characters</i> Sometimes it is as if story characters almost become real people that I know.	7	.77	.83	.84	

Variable	Instrument	Subscale: Example item	Items	Cronbach's $\alpha$		
				T1	T2	T3
Human nature: real-world others	Civic Attitudes and Skills Questionnaire (Moely et al., 2002; our translation) <sup>a</sup>	<i>Diversity attitudes</i> <sup>†</sup> I like meeting people who have a totally different background than I have myself.	5	.63	.66	-
Human nature: self-real-world others	Moral Competence Test (Lind, 2016; translation by Duriez & Demarez, 2000) <sup>b</sup>	Two moral dilemma scenarios; indicate agreement with decision made in scenarios; evaluate acceptability of six pro and six contra arguments for the decisions.	26	Pretest-posttest only; $\alpha$ not applicable		
Reasons for reading	Literary Response Questionnaire (see above) <sup>b</sup>	<i>Story-driven reading</i> <sup>†</sup> I like it when events build tension in the story.	8	.72	.67	.56
	Motivations for Reading Scale (Oliver & Raney, 2011; translation and adaptation for literary reading by Koopman, 2016) <sup>c</sup>	<i>Eudaimonic reasons</i> I like books that focus on meaningful human conditions.	6	.89	.88	.86
		<i>Hedonic reasons</i> It is important for me to have fun when I read a story.	6	.75	.68	.73
Strategy use	Metacognitive Awareness of Reading Strategies Inventory (Mokhtari & Reichard, 2002; our translation and adaptation for story reading) <sup>a</sup>	No subscales I stop from time to time to think about what I'm reading.	10	.76	.77	-
	Basic Need Satisfaction Scale (Van den Broeck et al., 2010; our translation and adaptation for literature education) <sup>a</sup>	<i>Autonomy</i> In literature class, I feel free to express my ideas and opinions. <i>Competence</i> I master the skills I need in literature class, like reading and discussing stories and doing response tasks. <i>Relatedness</i> In literature class, I feel part of the group.	6 5 8	.87 .76 .72	.71 .83 .78	-

Note. <sup>a</sup> 5-point scales; <sup>b</sup> 7- and 9-point scales; <sup>c</sup> 7-point scales. <sup>†</sup> Excluded due to insufficient Cronbach's alpha.

Correlations (see Appendix F, p. 287) suggested that TREQ subscales might represent underlying components. Via principal components analysis with Varimax rotation for the three measurement occasions, two components were extracted (see Table 6.4). Factor 1, Insight beyond story worlds, distinguished itself by high factor loadings of self-insights, insights into others, and aesthetic awareness. Generally speaking, it seemed to represent reflections that go “beyond” story worlds. Identification loaded on both factors, but clearly higher on Factor 1. Factor 2, Experiences within story worlds, represented reflections on experiences “within” story worlds, that is, imagery, experience-taking and evaluations of how characters think, feel and behave. Sympathy loaded on both factors, without clearly higher factor loadings for either of both factors. Factor scores were used for subsequent analyses.

*Table 6.4. Factor loadings of PCA for TREQ scales, per measurement*

	T1		T2		T3	
	Factor 1	Factor 2	Factor 1	Factor 2	Factor 1	Factor 2
Imagery		.87		.87		.83
Identification	.72	.51	.75	.41	.76	.39
Experience-taking		.70		.76	.63	
Character evaluation		.64		.63		.69
Sympathy	.54	.57	.34	.64	.62	.42
Aesthetic awareness	.73		.73		.65	
Self-insights	.91		.92		.90	
Insights into others	.85		.79	.34	.83	
Eigenvalue	4.18	1.17	4.11	1.19	3.90	1.01
Variance explained (%)	52.2	14.7	51.4	14.9	48.7	12.7

*Note.* Factor 1 = Insight beyond story worlds; Factor 2 = Experiences within story worlds. Factor loadings < .30 not displayed. Sampling adequacy and sphericity assumptions were met: KMO values .84, .85 and .85; Bartlett’s test of sphericity all  $p$ ’s < .001.

To look for additional indicators of insight into human nature, we analyzed students’ final TDLT task (see Appendix D, p. 279). Control teachers administered this writing task at the end of the control period. Students wrote a dialogue with an imaginary peer in response to a story they read, which they selected from four options. The instruction read: “Imagine you are having a dialogue about the story with a classmate. You talk, for example, about how you experienced the story, about its theme, the characters, things you found unclear... Write this dialogue on the next pages, as a comic. You start with the

sentence that is already given. Try to make it a real dialogue, not a question-and-answer interview. Use at least two pages." The dialogue started with: "What part of the story really stood out to you?" Students completed this task individually; they did not actually talk to a peer.

### 2.5 *Background variables*

As background variables, we assessed familiarity with fiction and trait empathy. Familiarity with fiction was assessed by administering an Author Recognition Test (Stanovich & West, 1989), as adapted by Schrijvers et al. (2016). Students scored relatively low, although there were considerable differences among individual students ( $M = 4.3$ ,  $SD = 3.0$ ; range -2 to 13). We measured trait empathy with two scales of the Interpersonal Reactivity Index (Davis, 1983): *Empathic concern*, which assessed students' feelings of sympathy and concern for unfortunate others (e.g., "I am often worried about people who are less well off than I am", Cronbach's  $\alpha = .79$ ), and *Perspective-taking* (e.g., "When someone upsets me, I try to put myself in his or her position for a moment", Cronbach's  $\alpha = .74$ ) which assessed their tendency to adopt the point of view of others. Items were scored on a 7-point agreement scale.

### 2.6 *Procedures*

Pretests were administered one week prior to the start of the lessons, posttests maximum one week after intervention or control lessons had finished, and delayed posttests approximately four months after the posttest. Instruments were administered on paper during regular hours of Dutch class, by the first author or a trained research assistant.

### 2.7 *Implementation Fidelity*

Implementation fidelity is important to examine the extent to which teachers implemented an intervention as intended, and to assess whether implementation differences might influence the study's outcomes (O'Donnell, 2008). We measured implementation fidelity via teacher logs and time on task observations.

Teachers were asked to complete an online log after each TDLT unit. In total, we administered 36 logs (6 classes \* 6 units). The response rate was 94%. The logs were organized according to the phases of each unit as described in the teacher guidelines (in total 31 phases in 6 units). For each phase, teachers

indicated whether it was completed, partly completed or not completed. Teachers also evaluated the partly and fully completed phases on 5-point scales, from *completely disagree* to *completely agree*.

Furthermore, we conducted time on task observations twice in each TDLT class. These functioned as an indicator of the proportion of available learning time that students are engaged in the tasks assigned to them (e.g., Karweit, 1984). Students' behavior was coded "on task" if they worked on the given task or subject matter and did what was asked of them, for example, listening to the teacher or a peer, talking about a task, reading a story, talking to the teacher, or asking questions. It was coded "off task" when students were not working on given tasks or subject matter, but were, for example, looking at cell phones, talking about something else than their task, being disruptive, or waiting for a next task. If a students' task behavior could not be determined, for example, when another student got into the line of sight, we coded it as "unclear". We randomly selected six students in the classroom and observed them in multiple rounds. In one round, each student was observed twice for twenty seconds. After a one-minute break, the next observation round started. This process continued until the end of the lesson. In total, we collected 877 observations in 12 lessons.

## 2.8 Control Condition

The control condition involved regular literary instruction, focused on identifying literary devices by analyzing literary texts. Control teachers completed an online log for each class at the end of each week of the study. In total, 36 logs (6 classes \* 6 weeks) were administered; the response rate was 94%. On average, each control class was taught 5.7 lessons during the study, compared to 6 lessons in the TDLT condition.

In the logs, control teachers reported what they taught. Their descriptions indicated that they focused, as they announced before the study started, on learning to identify and apply literary devices by analyzing stories, for example, perspective, characters, chronological structure, motives, story lines, and gaps. In four classes, a textbook was used, combined with online resources and self-developed teaching materials; in two classes, online resources and self-developed materials but no textbook were used.

Teachers mainly worked with literary short stories and, occasionally, with excerpts from literary novels. In two classes, a story was used that was also read in the TDLT condition (*Blood*, by Gerard Reve). Mostly, students answered analytically oriented questions after reading a text, for instance, about perspec-

tive, setting, and focalization. Some argumentative and oral skills were practiced, such as substantiating an evaluation of a story, analyzing a literary review, and exchanging answers to assignments in small groups or in teacher-led classroom conversations. Teachers also indicated that students read a self-selected novel at home.

## 2.9 *Data Analysis*

### 2.9.1 *Quantitative data*

To detect outliers, we first inspected boxplots of dependent variables across measurement occasions. We then applied casewise diagnostics, using regression analyses and graphs to plot students' scores (T1 on T2; T2 on T3), with Condition as a factor. Cases deviating more than 3 SD from the regression line were considered outliers. Across variables and measurement occasions, we identified 18 outliers (5.4% of total  $N$ , varying from 0,3 to 1% per variable). For each variable, we created a filter variable for outlier exclusion.

Indicators of insight into human nature (Insight beyond story worlds, Experiences within story worlds, and Empathy for characters) and reasons for reading (Eudaimonic and Hedonic reasons) were analyzed using mixed models growth curve analysis, with Student as subject variable for correlated random effects and Time as repeated variable for correlated residuals within random effects. We also tested models with Class as random factor. As its inclusion did not significantly improve the model fit, we excluded Class from the models to optimize statistical power. For each variable, we tested three linear models: (1) a model with Time as fixed effect, to test whether change over time occurred regardless of condition; (2) a model in which we added Condition as a fixed effect, to check whether intercepts of the two conditions differed significantly; and (3) a model in which we added a Time\*Condition interaction effect, to examine whether change over time differed between conditions.

For variables measured at T1 and T2 only (Moral competence, Strategy use, Autonomy, Competence, and Relatedness), we applied linear mixed models analysis. Again, Student was the subject variable for correlated random effects. We tested four models: (1) a basic null model, intercept only; (2) a model in which we added Pretest as a fixed effect, to test to which extent pretest scores explained posttest scores; (3) a model in which we added Condition as a fixed effect, to test whether average posttest scores differed between conditions; and (4) a model in which we added a Pretest\*Condition interaction effect, to test if effects of condition on posttest scores depended on pretest scores.

Next, we tested whether strategy use and motivation mediated TDLT effects on dependent variables at T2. We specified dependent variables in separate mediation analyses, with Condition as independent variable and Autonomy, Competence, Relatedness, and Strategy use as parallel mediators. Pretest scores of the dependent and mediating variables involved in the analysis were added as statistical controls.

### 2.9.2 *Qualitative data*

Students' written dialogues were assumed to contain indicators of insight into human nature and/or other transformative reading experiences. However, the total response rate was rather low (69%): not all TDLT students handed in their task and only four out of six control classes completed the task due to scheduling issues. Group sample size was therefore unequal. We checked for significant differences on relevant variables between TDLT ( $n = 134$ ) and control ( $n = 94$ ) subgroups. The groups did not differ in gender ( $\chi^2 = 2.74, p = .10$ ), Familiarity with fiction and Trait empathy (Wilk's  $\Lambda = .997, p = .91$ ), pretest scores on dependent variables (Wilk's  $\Lambda = .973, p = .48$ ), and pretest scores on mediating variables (Wilk's  $\Lambda = .990, p = .75$ ). We inferred that the data were admissible for analysis.

Students' written dialogues consisted on average of 267 words ( $SD = 121$ ; for two examples of dialogues, see Appendix F, p. 289). Dialogues were split into segments: whenever a new topic was addressed, we distinguished a new segment. In total, we distinguished 1,686 segments. Segments were coded for relevant response types, such as indicators of insight into human nature and other transformative reading experiences (see Appendix F, p. 291, for the coding scheme). Each segment could contain multiple response types and could thus be assigned multiple codes. An independent researcher coded 100 segments (6% of the data). Agreement was acceptable:  $\kappa = .72$  (for calculation procedure, see Appendix F, p. 295). We compared for both conditions the number of student dialogues in which a response type occurred at least once, using chi-square analyses, and how often each response type occurred overall, using  $t$ -tests. Further, we analyzed numbers of words and segments, as indicators of extensiveness of students' responses.

### 3 RESULTS

#### 3.1 *Implementation Fidelity*

TDLT was, overall, well-implemented. According to teachers' logs, 77.6% of the TDLT phases was completed as planned; 14.4% was partly completed, and 8% was not completed. They indicated that phases were feasible to teach ( $M = 4.1$ ,  $SD = .7$ ), proceeded orderly ( $M = 3.9$ ,  $SD = .8$ ), were clear for students ( $M = 4.0$ ,  $SD = .7$ ) and were interesting and engaging for students ( $M = 3.7$ ,  $SD = .8$ ).

Of the 877 time on task observations collected in 12 lessons, students were off task in 14.7% and on task in 85.2% of the cases, which was above the standard of 80% (e.g., Muijs & Reynolds, 2010; one observation was coded as "unclear", 0.1%). More specifically, for teacher-led activities (83.3%), individual student activities (86.4%) and story reading (84.6%), there was little variation in the on task percentage. For whole-class activities, it was slightly lower (76.0%), whereas it was clearly higher for dialogic activities in pairs or small groups (90.1%). Control teachers' logs showed that some oral and argumentative skills were practiced (e.g., supporting opinions about stories, analyzing literary reviews, sharing answers in small groups or the classroom), but that this mostly focused on analysis of literary devices.

#### 3.2 *Effects of TDLT*

Table 6.5 presents model comparisons for all dependent variables, with parameter estimates for significant best-fit models. Looking at indicators of insight into human nature, Model 3 fitted the data best for Insights beyond story worlds, Experiences within story worlds, and Eudaimonic reasons for reading. Significant interaction effects indicated that the intervention affected change over time on these indicators of insight into human nature (see Figures 6.2, 6.3, and 6.4). At T2, effects were medium for Insights beyond story worlds ( $d = .59$ ) and Eudaimonic reasons ( $d = .54$ ), and small for Experiences within story world (.19). At T3, all effects were small ( $d = .23$  for Insights beyond story world,  $d = .24$  for Experiences within story worlds, and  $d = .34$  for Eudaimonic reasons). We found no effects on Empathy for characters, Hedonic reasons, and Moral competence.

Table 6.5. Mixed growth curve (MGC) and linear mixed (LM) model comparisons, with parameter estimates and effect sizes for significant best-fit models

Variable	Analysis	Model	-2LL	$N_{\text{pars}}$	Models	Comparison		Parameter estimates		Cohen's $d$		
						$\chi^2$	$df$	$\beta$	$SE$	$p$	T2	T3
Insight beyond story worlds	MGC	1 Time	2111.45	5								
		2 Conditional	2105.78	6	2 vs 1	5.67	1	.017				
		3 Interaction	2097.97	7	3 vs 2	7.81	1	.005	.139	.05	.005	.59
Experiences within story worlds	MGC	1 Time	2120.34	5								
		2 Conditional	2117.43	6	2 vs 1	2.91	1	.088				
		3 Interaction	2113.44	7	3 vs 2	3.99	1	.046	.106	.05	.046	.19
Empathy for characters	MGC	1 Time	1704.43	5								
		2 Conditional	1704.37	6	2 vs 1	.06	1	.810				
		3 Interaction	1704.34	7	3 vs 2	.03	1	.856				
Eudaimonic reasons	MGC	1 Unconditional	2475.82	5								
		2 Conditional	2469.07	6	2 vs 1	6.78	1	.009				
		3 Interaction	2457.00	7	3 vs 2	12.04	1	.001	.205	.06	.001	.54
Hedonic reasons	MGC	1 Unconditional	1948.50	5								
		2 Conditional	1948.11	6	2 vs 1	.39	1	.533				
		3 Interaction	1945.89	7	3 vs 2	2.23	1	.136				

Variable	Analysis	Model	-2LL	$N_{\text{pairs}}$	Models	Comparison		Parameter estimates		Cohen's <i>d</i>		
						$\chi^2$	<i>df</i>	<i>p</i>	$\beta$	<i>SE</i>	T2	T3
Moral competence	LM	1 Null	2070.83	3								
		2 + pretest	2007.31	4	2 vs 1	63.52	1	.000				
		3 + condition	2006.43	5	3 vs 2	.35	1	.347				
		4 + interaction	2002.54	6	4 vs 3	3.88	1	.500				
Strategy use	LM	1 Null	447.11	3								
		2 + pretest	313.41	4	2 vs 1	133.71	1	< .001				
		3 + condition	300.49	5	3 vs 2	12.91	1	< .001	.206	.06	< .001	.37
		4 + interaction	299.88	6	4 vs 3	.62	1	.430				
Autonomy	LM	1 Null	454.28	3								
		2 + pretest	422.48	4	2 vs 1	31.80	1	< .001				
		3 + condition	360.41	5	3 vs 2	62.07	1	< .001	.541	.06	< .001	.98
		4 + interaction	360.03	6	4 vs 3	.38	1	.537				
Competence	LM	1 Null	541.35	3								
		2 + pretest	443.68	4	2 vs 1	97.70	1	< .001				
		3 + condition	428.45	5	3 vs 2	15.23	1	< .001	.290	.07	< .001	.57
		4 + interaction	428.45	6	4 vs 3	.001	1	.975				
Relatedness	LM	1 Null	400.52	3								
		2 + pretest	307.53	4	2 vs 1	93.00	1	< .001				
		3 + condition	273.34	5	3 vs 2	34.19	1	< .001	.330	.05	< .001	.64
		4 + interaction	272.76	6	4 vs 3	.58	1	.445				

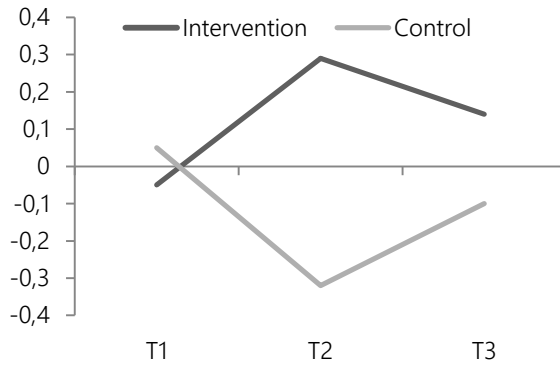


Figure 6.2. Factor scores for Insights beyond story worlds.

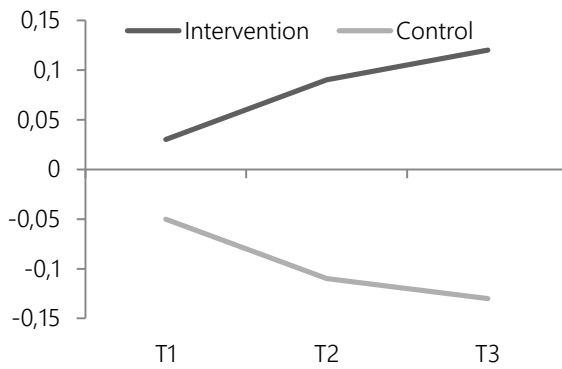


Figure 6.3. Factor scores for Experiences within story worlds.

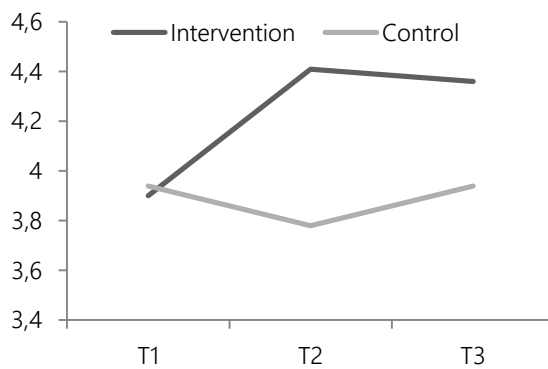


Figure 6.4. Adjusted mean scores for Eudaimonic reasons for reading.

Furthermore, TDLT students wrote more extensive dialogues,  $M = 295$  words,  $SD = 134$ , than control students,  $M = 227$ ,  $SD = 87$  ( $t(225) = 4.58$ ,  $p < .001$ ). No difference was found for the number of segments (TDLT  $M = 7.1$ ,  $SD = 3.3$ ; control  $M = 7.8$ ,  $SD = 3.4$ ). Students in both conditions thus addressed equal numbers of topics, but TDLT students appeared to do so more elaborately. Descriptive-evaluative statements, such as reiterations of story events, simple evaluations, or expressions of incomprehension without attempts to solve it, occurred most frequently (see Table 6.6), but significantly less often in TDLT than in the control group. Two reasoning response types closely related to insight into human nature occurred more often in TDLT: Reasoning to understand and interpret characters' acts, thoughts, feelings and motives, and Moral reasoning confined to the story world (see Table 6.6). In addition, Reasoning to substantiate evaluations, and Reasoning to understand and interpret story events occurred significantly more often in TDLT than in the control group. Finally, four transformative reading experiences – Character evaluations, Aesthetic awareness, Imagery, and Identification – occurred more often in TDLT students' dialogues than in those of control students, although the latter two were, overall, mentioned rather infrequently.

For Strategy use, Autonomy, Competence, and Relatedness, main effects of condition were found (see Table 6.6). Non-significant interaction models indicated that these effects did not depend on pretest scores. The intervention had a small effect on students' self-reported Strategy use ( $d = .37$ ), a large effect on satisfying their need for Autonomy ( $d = .98$ ), and medium effects on Competence ( $d = .57$ ) and Relatedness ( $d = .64$ ).

### 3.3 *Strategy Use and Motivation as Mediators*

We performed mediation analyses on dependent variables for which we found an effect of condition. The mediation model for Insight beyond story worlds explained 58% of the variance in students' scores ( $R^2 = .58$ ,  $F = 31.72$ ,  $p < .001$ ). Results indicated partial mediation. There was a direct effect of Condition ( $\beta = .539$ ,  $p < .001$ ) as well as indirect effects of three mediating variables: Competence ( $\beta = .127$ , 95% CI [.06, .22]), Relatedness ( $\beta = -.083$ , 95% CI [-.19, -.01]), and Strategy use ( $\beta = .052$ , 95% CI [.01, .13]).

Table 6.6. *Response types in students' written dialogues*

Response type	% of dialogues where response occurs at least once		$\chi^2$ -value ( <i>p</i> -value)	M occurrences per dialogue		<i>t</i> -value ( <i>p</i> -value)
	TDLT	Control		TDLT	Control	
Descriptive evaluative statements	79.1	92.6	7.69 (.006)	2.60	4.61	1.70 (.000)
Reasoning						
Evaluative	65.7	34.0	22.17 (.000)	1.14	0.38	6.46 (.000)
To understand and interpret characters	58.2	40.4	6.99 (.008)	1.04	0.68	2.45 (.015)
To understand and interpret story events	30.6	8.5	15.97 (.000)	0.38	0.12	3.76 (.000)
Moral: story world	31.3	10.6	13.45 (.000)	0.43	0.13	4.15 (.000)
Moral: beyond story world	12.7	11.7	n.s.	0.16	0.12	n.s.
Hypothetical	11.9	14.9	n.s.	0.13	0.19	n.s.
Transformative reading components						
Character evaluation	51.5	37.2	4.53 (.033)	0.80	0.59	n.s.
Aesthetic awareness	17.9	6.4	6.43 (.011)	0.19	0.06	2.88 (.009)
Imagery	9.0	2.1	4.47 (.035)	0.10	0.02	2.45 (.015)
Identification	7.5	0.0	7.34 (.007)	0.07	0.00	3.28 (.001)
Sympathy	31.3	21.3	n.s.	0.37	0.28	n.s.
Experience-taking	16.4	8.5	n.s.	0.20	0.10	n.s.
Referentiality						
Self-references	25.4	26.6	n.s.	0.40	0.41	n.s.
Real-world	13.4	8.5	n.s.	0.15	0.09	n.s.
Content-irrelevant						
Other	35.8	45.7	n.s.	0.40	0.70	2.72 (.007)
Unclear	1.5	1.1	n.s.	0.01	0.01	n.s.

Note. For examples of response types, see Appendix F (p. 291).

The second mediation model explained 55% of variance in Experiences within story worlds ( $R^2 = .59$ ,  $F = 28.29$ ,  $p < .001$ ). Whereas in mixed growth curve analysis a direct effect of condition on Experiences was found, it disappeared in the mediation model ( $\beta = .006$ ,  $p = .97$ ), which indicated complete mediation via other variables. Strategy use was the only significant mediator ( $\beta = .095$ , 95% CI [.04, .18]).

Finally, the mediation model for Eudaimonic reasons explained 59% of variance ( $R^2 = .59$ ,  $F = 32.36$ ,  $p < .001$ ). The relationship between Condition and Eudaimonic reasons was partly mediated: in addition to a direct effect ( $\beta = .609$ ,  $p < .001$ ), Competence functioned as a mediator ( $\beta = .072$ , 95% CI [.01, .17]).

### 3.4 *Summary of Results*

Our first hypothesis was largely confirmed: TDLT positively affected students' insight into human nature and eudaimonic reasons for reading, although effects were not fully consistent across all indicators insight into human nature that we assessed. The second hypothesis was confirmed: TDLT enhanced students' self-reported strategy use and motivation for literature education. Our third hypothesis was partly confirmed: Competence, Relatedness and/or Strategy use mediated effects of TDLT on Eudaimonic reasons for reading and indicators of insight into human nature.

## 4 DISCUSSION

### 4.1 *Comparing TDLT to "Business as Usual"*

In this quasi-experimental intervention study, we compared the effects of Transformative Dialogic Literature Teaching on 10<sup>th</sup> grade students' insight into human nature to a "business as usual" approach that mainly focused on learning to identify literary devices in analyzing short stories. Analysis of students' questionnaire data and writing task data indicated that TDLT fostered insight into human nature, which included insight into themselves, fictional others, and real-world others. For example, TDLT had a medium effect on Insights beyond story worlds, the first TREQ component that included the scales Self-insights and Insights into others.

However, results for insight into moral dilemmas that people may face were ambiguous. No conditional differences were found on the Moral Competence Test, but written imaginary dialogues suggested that moral reasoning confined

to the story world occurred more frequently in TDLT than in the control condition. For example, a TDLT student wrote:

“The end really struck me. What about you?  
 Yes, me too. It’s not just. [*referring to a father who slaps his son*]  
 Oh? I thought it was quite fair.  
 Ooh, why then?  
 Well, the boy should just know that such things aren’t allowed, so he deserved it.  
 He only made a joke.  
 Yes, I think that’s disrespectful.  
 So, therefore it’s okay to slap your child?  
 Well, okay, it could have gone differently.  
 Yes, the father shouldn’t slap his child because of this.  
 Why do you think that?  
 You wouldn’t slap your child for a joke he made days ago, right?  
 Yes, but the father wanted to make clear he doesn’t allow that behavior.  
 Come off it, that’s the wrong way to raise his child.”

The effects on Insight beyond story worlds, Experiences within story worlds, and Eudaimonic reasons sustained up until four months after TDLT, although T3 effects for Insight and Eudaimonic reasons became smaller than at T2. As Figures 6.2 and 6.4 illustrate, this may partly be explained by the control condition, where scores dropped at T2 and – nearly – returned to pretest level at T3. Thus, focusing on identifying literary devices by analyzing short stories may have negatively affected students’ insight into human nature, whereas TDLT seemed to foster it.

This finding does not imply that identifying literary devices had no value. On the contrary: TDLT students also identified literary devices, as a means to reflect on transformative reading experiences and insight into themselves, others, and moral considerations. This is illustrated by an excerpt of a dialogue written by a TDLT student, who refers to “psychological suspense” to reason about a character’s thoughts and feelings, resulting in insight into “why a character does something”:

“Did you also think this story had much suspense? Because I noticed some sort of suspense, but I’m not sure how.  
 I think you mean psychological suspense, because you really feel along with the thoughts and emotions of the protagonist. You notice he gets different feelings about the man who is executed, and that it confuses him.  
 Yeah, that’s what I meant. Because of that, I really get a sense of knowing why a character does something. If I wouldn’t have known the protagonist’s thoughts, I wouldn’t have understood why he is so kind to the convicted man.”

This excerpt also exemplifies the extensiveness of TDLT students’ dialogues. We found that those more elaborate dialogues contained significantly less descrip-

tive-evaluative statements than the dialogues of control students. Such rather trivial statements, thus, did not seem to function as "fillers" in TDLT students' more extensive dialogues, which might indicate that their explorations of social-moral themes were more in-depth than those of control students. Moreover, in the excerpt above, the student mentions to "feel along" with the protagonist, which indicates experience-taking. Although no differences were found for Experience-taking and Sympathy, other transformative reading experiences were mentioned more often by TDLT students than by control students. As such experiences have been shown to precede self-other insights in adult readers (Fialho, 2018), these findings are promising in terms of fostering adolescents' insight into human nature.

Furthermore, we considered Eudaimonic reasons for reading to be a relevant indicator of students' willingness to develop insight into human nature in the literature classroom. TDLT students more strongly agreed to read for meaningful insights into human conditions than control students. This did not mean that reading for pleasure and enjoyment decreased: for Hedonic reasons, no difference was found. Thus, once TDLT students became more aware of the potential of eudaimonic reasons, both types of reasons appeared to co-exist. This finding adds to Oliver and Raney's (2011) work on reasons for watching movies, by expanding their conclusion to another type of media and research population. In addition, we have shown that adolescents seem to benefit from instructional guidance to develop eudaimonic reasons for reading. TDLT appeared to guide students toward recognizing and endorsing that they may read literary texts "to search for and ponder life's meaning, truths, and purposes" (Oliver & Raney, 2011, p. 985).

#### *4.2 The Role of Strategy Use and Motivational Factors*

We further investigated students' strategy use for dealing with difficulties in literary reading and their motivation for literature education in relation to TDLT. Whereas the effect on strategy use was small, effects on motivation were medium to large. TDLT thus appeared to alleviate prominent challenges in literature teaching. These findings are in line with other studies that found positive effects of reader- and affect-oriented approaches on student engagement in literature classrooms (e.g., Eva-Wood, 2004; Fialho et al., 2012; Janssen et al., 2009; Levine & Horton, 2013).

As an indicator of motivation for literature education, Autonomy did not function as a mediator. Strategy use, Competence, and Relatedness explained the effect of TDLT on insight into human nature and eudaimonic reasons for read-

ing to a small extent;  $\beta$ 's indicated that mediating effects were small compared to direct effects. For Insights beyond story worlds, the role of a mediator was most obvious: in addition to a prominent direct effect ( $\beta = .539$ ), students' feelings of Competence played a mediating role ( $\beta = .127$ ). Thus, the more competent students felt after TDLT, the higher their Insight posttest scores. This suggests we should be responsive to students' abilities in the literature classroom. In TDLT, this was operationalized by explicit strategy instruction about internal and external dialogues, by use of support tools such as a first aid card, and as illustrated by the rubric in Appendix D (see p. 280), by monitoring progress when moving toward new and challenging ways of interacting with and about literary texts.

#### 4.3 *Limitations and Directions for Future Studies*

As in most intervention studies, conclusions concern TDLT as a whole. Future experimental studies may investigate the effects of text selection, internal dialogues, and external dialogues in separate conditions. Likewise, the effect of attending to transformative reading experiences could not be assessed separately from other intervention elements, but this study is a first step in validating the transformative reading model (Fialho, 2018) for adolescents.

Furthermore, outcomes of this study can neither be generalized to other educational tracks, grade levels, or foreign literature curricula, nor to other genres than short stories, such as novels or poetry. Future work may address how TDLT can be adapted to other educational contexts and other text genres. For example, as it may be beneficial to develop insight into human nature prior to adolescence, research may focus on how to accommodate TDLT to reading activities in primary school.

In addition, conditions were taught by different groups of teachers. Therefore, although students in both conditions were well-comparable, teacher differences may have affected the outcomes of this study. When we explored potential differences among classes on six relevant background variables, using analyses of variance, differences were found for students' average for grade for Dutch class ( $p = .032$ ) and their familiarity with fiction ( $p = .048$ ). However, for both, post-hoc Bonferroni analyses revealed no significant differences between particular classes. Therefore, differences among classes appeared to be negligible, which somewhat counterbalances the lack of random assignment to conditions. To avoid undesirable teacher effects altogether, future studies may, for example, apply a switching replications design.

Finally, written dialogues were collected at T2 only. Therefore, the question remains whether effects would have sustained at T3. Nonetheless, written dialogues added to the quantitative data we collected, as students explicated their thoughts in response to a story. In addition, transformative reading experiences and other response types were well-distinguishable in the dialogues. The instrument may be of added value for future studies that aim to explicate students' thinking in response to fictional and literary texts.

#### 4.4 *Conclusion and Implications*

In societies facing globalization, migration, and polarization, it is perhaps more important than ever for people to reflect on what it means to be 'human'. In this paper, we have shown that literature education may be a promising domain for affecting adolescents' insight into human nature. This study suggests that TDLT may guide students toward developing such insight. The instructional differences between the experimental and control condition imply that we may want to move away from formalist, knowledge-oriented instruction that may still exist in literature classrooms, as aptly described by Wilhelm (2016):

"Teachers [...] may emphasize knowing and recognizing literary devices, getting at the "internal logic" of a text's construction [...], and relating a work's central "organic" meaning to how this meaning was expressed. There may be an emphasis on "rightness" of literary interpretation. Interpretative questions about the text will be answered after reading [...], and discussions mediated by the teacher, who acts as the authority on the text" (p. 25).

If developing insight into human nature is acknowledged as one of the objectives of literature teaching, such a formalist, knowledge-oriented approach appears not to be helpful. In contrast, as in TDLT, instruction should encourage students to explore their personal responses in dialogic interactions with and about literary texts, by completing purposefully designed combinations of pre-, during- and post-reading tasks in which analysis of literary devices is a means to reason about reading experiences, themes, characters, and moral implications.

As a reader- and affect-oriented approach, TDLT further appears to alleviate recurring challenges in the literature classroom, in terms of students' motivation and the ways they handle difficulties in literary texts. All in all, if the demonstrated effects are supported or expanded in future studies, TDLT may be a promising approach for fostering students' insight into human nature, their awareness and use of reading strategies, and their motivation for literature education.



## CHAPTER 7

### GENERAL DISCUSSION

The purpose of this dissertation was to investigate which approach to literature teaching in Dutch upper secondary education would be appropriate for fostering students' insight into human nature. In this final chapter, we first summarize our main findings, by providing answers to the research questions addressed in Studies 1 to 5. Subsequently, we discuss four key concepts that guided these studies: insight into human nature, transformative reading, literariness, and dialogic learning in literature classrooms. Next, we address potential validity issues in our studies regarding the intervention-as-designed and the intervention-as-implemented, the instruments and the research designs we applied, followed by discussing the external validity of our studies and the risk of a potential researcher bias. Finally, we discuss directions for future studies as well as the implications of our research for educational practice.

#### 1 MAIN FINDINGS

##### 1.1 *Learning Experiences about Self and Others, Related to Teacher Approach*

We started this research project by exploring students' learning experiences about themselves and others and their teachers' practices in upper secondary school literature classrooms. In Study 1, we attempted to answer two research questions:

1. What kind of learning experiences about themselves and others do students in upper secondary literature classrooms in the Netherlands report?
2. Do different teachers' approaches to literary instruction generate different kinds of learning experiences?

Exploring students' learning experiences provided initial directions for the project: if relevant learning experiences were found, this would indicate that gaining insight into self and others would be a feasible outcome of literary instruction in the Netherlands. We concluded that this was the case for a sample of 297 students in various grade levels of the higher general and pre-university education track. In learner reports, students reported a variety of learning experiences, in which "learning about oneself" and "learning about others" were

closely interwoven. Students in some classes reported more of such learning experiences than students in other classes. We were particularly interested in verifying whether the variety across classrooms was related to the instructional approach taken by the students' teachers.

The approaches of the thirteen teachers in the study were rather eclectic. Some of them reported to emphasize an analytical-interpretative, more teacher-led perspective, whereas others emphasized a personal-experiential, student-oriented approach, or positioned themselves somewhere in the middle of the continuum. Focusing on various aspects of teaching (i.e., attitude toward literary reading, students' roles in the classroom, and intended teaching content) enabled us to establish relations between teaching characteristics and students' learning experiences. We found that students of teachers who emphasized students' personal experiences with literary texts – in terms of student autonomy and interaction in the classroom – were found to report learning experiences about self and others more frequently than students of teachers who reported more teacher-led practices in their literature classrooms.

Despite several limitations – the sample of teachers was small, differences were not found for all categories of learning about self and others, and conclusions were based on multiple univariate tests – we considered these findings to be initial parameters for the design of an instructional approach. Offering students opportunities for autonomy and personal engagement in expressing their reading experiences with literary texts and to interact about texts and reading experiences appeared to stimulate learning about self and others.

### 1.2 *Instructional Approaches in Previous Intervention Studies*

Study 1 provided first indications of elements of an instructional approach. To create a solid design framework, we reviewed previous empirical studies in the field, aiming to identify instructional design principles of effective interventions. We examined instructional approaches for which – via experimental and quasi-experimental testing – empirical support was found. In Study 2, we addressed the question:

In previous intervention studies aimed at fostering students' insight into human nature (including self and others), what effects did researchers expect to achieve, to what extent were these expected effects empirically supported, and which instructional approaches were implemented in interventions for which empirical support was found?

Via a systematic literature search we identified thirteen intervention studies that met our inclusion criteria. Researchers of these studies expected that their instructional approaches would foster students' insight into themselves, their understanding of fictional others, and/or their understanding of, views on, or intended behavior toward real-world others. At this point, we concluded that these categories were not always clearly distinguishable, which led us to reformulate "insight into self and others" into the more broadly formulated concept of "insight into human nature".

After quality assessment, which functioned as a gatekeeper for the validity of researchers' conclusions, we found empirical support for nine out of the thirteen intervention studies that were included in the review. Analyses of the instructional approaches as they were reported in these studies yielded three instructional design principles. First, studies revealed that insight into human nature may be fostered when fictional texts are selected that are thematically relevant for an intervention aim, such as texts addressing social relations or moral dilemmas. For example, White (1995) asked students to read short stories which centered around parent-child and dating relations to affect students' understanding of fictional characters, and Malo-Juvera (2014) used a young adult novel about sexual harassment to affect students' attitudes toward such behavior.

Second, in the majority of the intervention studies, students were asked to share their reading experiences in exploratory dialogic activities (e.g., Adler & Foster, 1997; Eva-Wood, 2004; Malo-Juvera, 2014; White, 1995). Thus, dialogues appeared to be relevant to foster insight into human nature. Talking to peers may open up a wider array of thoughts, questions, feelings, ideas, and perspectives in response to texts and the social-moral themes they address. In the analyzed intervention studies, this happened in small-group or whole-class dialogues, or in combinations where the latter followed the former. Such build-ups, we argue, create multiple layers of sharing responses that allow for a variety of interpretations, nuances, and generalizations to be brought into the open.

A precondition for sharing reading experiences in exploratory dialogic activities is that students are aware of those experiences. To prompt this awareness and to prepare for external dialogues with peers, in most intervention studies students were stimulated to engage in internal dialogues with texts. To this end, writing tasks were designed that prompted students to activate previous personal experiences before reading (e.g., White, 1995), to notice and annotate

responses during reading (e.g., Eva-Wood, 2004), and/or to write down (reflective) responses directly after reading (e.g., Malo-Juvera, 2014; 2016).

In short, to increase the likelihood that adolescent students gain insight into human nature in the literature classroom, instructional designers are advised to:

1. Select fictional texts such as novels, short stories, passages, or poems, that are thematically relevant for the intended outcomes of the intervention;
2. Design writing tasks related to fictional texts and text themes that prompt students to activate previous personal experiences before reading, notice and annotate their experiences during the reading process, and/or write down (reflective) responses directly after reading;
3. Design exploratory dialogic activities that stimulate students to verbally share their personal experiences related to fictional texts and text themes.

In addition, the review study yielded insights for future intervention studies. Based on methodological quality assessment of the studies, we recommend in particular that researchers report on implementation fidelity and elaborate on the role of the teacher in classroom practices.

### 1.3 *Designing a Valid and Practical Intervention*

In addition to the insights gained in Study 1 and the design principles concerning text selection, internal dialogues, and external dialogues identified in Study 2, a model of transformative reading informed the design of an intervention. Fialho (2012; 2018) conceptualized “transformative reading” as a mode of reading in which readers experience self-other perceptual depth. As it entails both self and others, self-other perceptual depth complies with what we call in this dissertation insight into human nature.

In a phenomenological study with adult readers, Fialho (2018) showed that transformative reading included not only self-other perceptual depth but also six other experiences: vividly imagining story setting and characters (*imagery*); recognizing aspects of self or others in characters (*identification*); enacting and embodying the experiences of a character (*experience-taking*); evaluating characters positively or negatively (*character evaluations*); feeling sympathy and compassion for characters (*sympathy*); and being aware of striking words, phrases or sentences (*aesthetic awareness*). This model of transformative reading thus informed the instructional design. More specifically, in the design process we considered how students could be guided toward reflecting upon these particular kinds of reading experiences.

In Study 3, we aimed to design a literature classroom intervention for 10<sup>th</sup> grade students, in which the model of transformative reading, the text selection principle, the internal dialogue principle, and the external dialogue principle were integrated. Because validity – both at the content and construct level – and practicality are considered to be amongst the indicators of the quality of an intervention (Nieveen, 1999; O'Donnell, 2008), the study was guided by the following question:

To what extent is the intervention that is designed a valid and practical instructional approach for upper secondary literature classrooms, according to both students and teachers?

We named the intervention Transformative Dialogic Literature Teaching, or TDLT. It resulted from a design process in which two iterations were carried out, both consisting of a development phase and an assessment phase. Consequently, the design process yielded two versions of TDLT. In line with Study 3, in the remainder of this chapter we will refer to the first version as "TDLT-1" and to the second version as "TDLT-2".

We designed TDLT-1 in collaboration with various teachers and ran two trial studies to optimize the design. TDLT-1 was then taught by 13 teachers in 22 classes. From these teachers and their students, we collected implementation and evaluation data, via teacher logs, time on task observations, teacher interviews, and evaluation forms and tasks for students. This enabled us to draw informed conclusions about the validity and practicality of the TDLT-1 intervention, which appeared to be suboptimal. For example, students struggled to see why internal and external dialogues with and about stories were relevant (content validity), found it unclear how they could engage in these dialogues (construct validity), were too often off task, and needed, according to their teachers, more time to get used to dialogic response practices in the literature classroom (practicality). We aimed to remedy the validity and practicality issues of TDLT-1 by setting up a second design iteration, in which three teachers who taught TDLT-1 cooperated with us to redesign the intervention. This second development phase resulted in TDLT-2. In Table 7.1, we present an overview of the main adjustments from TDLT-1 to TDLT-2. Next, we describe TDLT-2 in closer detail.

*Table 7.1. Adjustments from TDLT-1 to TDLT-2*

	TDLT-1	TDLT-2
Overall	<ul style="list-style-type: none"> <li>- Four units</li> <li>- 200 minutes of classroom work</li> <li>- Exact same structure in all units</li> </ul>	<ul style="list-style-type: none"> <li>- Six units</li> <li>- 300 minutes of classroom work</li> <li>- About 45 minutes of homework</li> <li>- More variety in structure of units</li> </ul>
Text selection	<ul style="list-style-type: none"> <li>- Reading three stories (literary fiction)</li> <li>- Each story addresses a different social-moral theme</li> </ul>	<ul style="list-style-type: none"> <li>- Reading six stories (literary fiction)</li> <li>- All stories address the same social-moral theme ("justice and injustice")</li> </ul>
Internal dialogues	<ul style="list-style-type: none"> <li>- Applied in three units (2-4)</li> <li>- No explicit attention for importance</li> <li>- No explicit strategy instruction</li> </ul>	<ul style="list-style-type: none"> <li>- Applied in six units (1-6)</li> <li>- Explicit attention for importance</li> <li>- Explicit strategy instruction in unit 3</li> <li>- Monitoring progress with rubric</li> </ul>
External dialogues	<ul style="list-style-type: none"> <li>- Applied in four units (1-4)</li> <li>- Some attention for importance in unit 1</li> <li>- Strategy instruction with dialogue guidelines in unit 1</li> </ul>	<ul style="list-style-type: none"> <li>- Applied in six units (1-6)</li> <li>- More explicit attention for importance in units 1 and 3</li> <li>- Strategy instruction with dialogue guidelines and additional prompts in unit 1, available on "first aid card"</li> <li>- Monitoring progress with rubric</li> </ul>
Support for teachers	<ul style="list-style-type: none"> <li>- Teacher guidelines and lesson plans</li> <li>- Face-to-face preparatory meeting with each teacher individually</li> </ul>	<ul style="list-style-type: none"> <li>- Teacher guidelines and lesson plans, including prompts for guiding talk and giving feedback</li> <li>- Workshop for new teachers, including walk-through of materials and feedback practice</li> </ul>

TDLT-2 consisted of one preparatory and five reading-and-dialogue units. It included about 300 minutes of classroom work, complemented by about 45 minutes of homework assignments. In line with the text selection principle, in TDLT-2 short stories with a social-moral theme were read. This principle was further operationalized by a single-theme approach: all stories centered around "justice and injustice", for two reasons. First, we expected that this particular theme would trigger students to consider their personal responses to complex social situations and moral dilemmas, which may offer them insight into how they themselves as well as others would be affected by and act upon such situations. Second, we assumed that a single theme in multiple stories would help students to identify relations between the story situations, as well as to reconsider and/or deepen their responses each time the theme was addressed in a new story. The latter is in line with findings from previous studies,

which have shown that deepening perceptions of self and others unfolds over time (Fialho, 2012; Kuiken, Miall, & Sikora, 2004; Kuiken, Phillips, Gregus, Miall, Verbitsky, & Tonkonogy, 2004). The text selection principle was further operationalized by using fictional stories that teachers regarded as being "literary texts" for their students, as TDLT was meant to serve as a start of the formal literature curriculum in grade 10. The teachers in our studies substantiated this classification by referring to the complexity and "depth" of the stories, for example, when language use was unconventional for students or gaps needed to be filled in to interpret a story.

The primary aim of TDLT-2 was for students to learn to express, orally and in writing, a) the responses that stories evoked in them, b) which new insights into themselves, others and social life these stories offered them, and c) which literary devices evoked these responses and insights. To achieve these aims, students were taught strategies for both external and internal dialogues. In the preparatory unit (unit 1) strategies for external dialogues were introduced. Students observed and evaluated videos of peers talking about a story, received explicit instruction about the strategy – which was summarized on a "first aid card" they used throughout TDLT – and applied the strategy in a small-group dialogue about famous quotes about literature and reading that were printed on small cards. The external dialogue principle was further operationalized from the perspective of the teacher. Teachers were asked to take on a guiding, non-authoritative role when students engaged in external dialogues. TDLT-2 offered guidelines for teachers that addressed how they could provide guidance and feedback on students' dialogic processes. These guidelines included, for example, prompts and questions that stimulated students to continue and deepen their talk.

In all subsequent reading-and-dialogue units (2 to 5), external dialogues were applied in combination with internal dialogues; together, the internal and external dialogue formed the two-step basic structure that was central in TDLT. Internal dialogues remained implicit in units 1 and 2, when students read stories but were not given a particular reading instruction. The purpose of and strategies for internal dialogues were explicitly introduced in unit 3: the teacher explained that students could attend to "the voice in their mind" during reading, and modeled "noticing and annotating responses" by thinking aloud while reading. As incomprehension was considered to be a legitimate response that may come up during an internal dialogue, the teacher also introduced strategies for dealing with difficulties during reading, for example, writing down question marks, pausing to think, and asking for help. These were summarized

on the 'first aid card' as well. Next, students for the first time applied the internal dialogue strategies to a story they read, as preparation for their external dialogues. This two-step structure was repeated in all following units. Moreover, from unit 3 onwards, internal and external dialogues focused on transformative reading experiences: students considered, for example, experiences of imagery and sympathy – in unit 1 and 2, they had focused on responses that were likely to be more familiar to them, such as their initial opinion about a story. The activities in all units were miscellaneous, short, and high-paced to keep students engaged and motivated. Students were stimulated to monitor their progress in engaging in internal and external dialogues by working with a rubric.

The three teachers who were involved in the redesign process taught TDLT-2 to one of their classes, as did three new teachers (six classes in total). Here, we also collected implementation and evaluation data. From these data, we inferred, for example, that teachers felt that TDLT-2 was generally practical to work with, and that students were on task for a larger proportion of time in TDLT-2 than in TDLT-1, in particular during external dialogues. Moreover, students considered their teachers' explanations about internal and external dialogues particularly helpful, as well as the teacher modeling internal dialogue strategies. Overall, we concluded that TDLT-2 was a valid and practical operationalization of the transformative reading model and the set of design principles.

#### 1.4 *Testing the Effects of TDLT-1*

In Study 4 we tested the effects of TDLT-1 on students' transformative reading experiences, including insight into human nature. We addressed the following questions:

1. Which learning experiences do 10<sup>th</sup> grade students report after participating in TDLT-1?
2. Does TDLT-1 have a positive effect on students' transformative reading experiences, such as insight into human nature, imagery, identification, and sympathy, in comparison to students who participate in an untreated control condition?

We assessed the effects of TDLT-1 in a quasi-experimental research design with pretest, posttest, and delayed posttest, with switching replications. The three tests were administered within a time frame of maximum eight weeks. Of 22

classes, 11 first participated in TDLT-1 and subsequently in an untreated control condition in which regular Dutch lessons were taught (but no literature); in the other 11 classes, the order of conditions was switched.

Learning experiences were measured by a written learner report at the posttest, directly after the intervention. Transformative reading experiences, including insight into human nature, were measured via a questionnaire and a story response task administered at pretest, posttest and delayed posttest. Results indicated that students most often reported to have learned how to talk about stories (28% of all reported learning experiences). In addition, they reported some transformative reading experiences (15%), and reported that they had developed their thinking and opinionating skills (13%). However, from the data collected via the questionnaire and story task we detected no consistent effects of TDLT-1 on students' transformative reading experiences and insight into human nature. On the contrary: students who had been involved in the untreated control condition scored higher on several indicators of transformative reading, including their insight into human nature, than students who had been involved in the intervention.

There may be several explanations for these findings. First, although students felt that they learned to participate in dialogues about short stories – “talking about stories” was reported in almost one third of students' learning experiences – these dialogues remained rather short and superficial, as observed by various teachers and endorsed by the rather low on task percentage during small-group dialogues. Furthermore, students may have had too little experience with literary reading to be able to notice their responses during reading, as suggested by the smaller share of learning experiences that concerned, for instance, transformative reading (15%), in-depth processing of stories (9%), and noticing responses during reading (7%). In addition, students and teachers indicated that the ultimate purpose of TDLT-1 – gaining insight into human nature – remained too implicit, which may have caused students to not have been sufficiently motivated for and engaged in the lessons and stories. Finally, the instruction time may have been too short. Based on these findings, adjustments were made that resulted in TDLT-2.

### 1.5 *Testing the Effects of TDLT-2*

We expected that adjustments made in the second design iteration – such as explicit attention for relevance and importance of internal and external dialogues, and strategy instruction – would contribute to alleviating two prominent challenges in Dutch literature classrooms, as repeatedly expressed by

teachers involved in teaching TDLT-1: students' limited ability to deal with difficulties during literary reading, and their rather low motivation for literature education. In Study 5, we therefore aimed to shed light not only on the effect of TDLT-2 on transformative reading, but also on the role of strategy use and motivation. To this end, we measured students' reported use of strategies and, as indicators of motivation, the extent to which they felt that their needs for autonomy, competence, and relatedness were satisfied (Ryan & Deci, 2000).

In addition, we operationalized "developing insight into human nature", the transformative reading component that was central in our studies, more broadly than in Study 4: we defined additional indicators for it (i.e., students' empathy for fictional characters and their moral competence) and included it as a potential reason for reading (eudaimonic reasons) that students might become more aware of as a result of TDLT-2. We addressed the following research questions:

1. Does TDLT-2 have a positive effect on a) students' transformative reading experiences and other indicators of insight into human nature (empathy, moral competence), b) their eudaimonic reasons for reading, c) their use of strategies to deal with difficulties during literary reading, and d) their motivation for literature education, in comparison to students in a control condition focused on identifying literary devices and analysis of short stories?
2. To which extent do strategy use and motivation function as mediators for the effect of TDLT-2 on students' insight into human nature and eudaimonic reasons for reading?

We assessed the effects of TDLT-2 by applying a quasi-experimental design with pretest, posttest and delayed posttest (four months after the intervention), in which six classes participated in TDLT-2 and six classes in the control condition in which students followed their teachers' regular literature curriculum, focused on literary devices and analysis.

As Table 7.2 shows, both in Study 4 and 5 we complemented quantitative data with data from a writing task in response to a story: a story response task as pretest, posttest, and delayed posttest to assess the effects of TDLT-1, and a task in which students wrote a dialogue with an imaginary peer in response to a story, as posttest-only to assess the effects of TDLT-2.

*Table 7.2. Assessed effects of TDLT-1 and/or TDLT-2*

Variable	Instrument	TDLT-1			TDLT-2		
		T1	T2	T3	T1	T2	T3
Indicators of insight into human nature	Transformative Reading Experiences Questionnaire (TREQ)	x	x	x	x	x	x
	Literary Response Questionnaire: empathy for characters				x	x	x
	Moral Competence Test				x	x	
	Written story response task	x	x	x			
	Written imaginary dialogue task					x	
Reasons for reading	Motivations for Reading Scale: eudaimonic, hedonic reasons				x	x	x
Reported learning experiences	Learner report task		x				
Strategy use	Metacognitive Awareness of Reading Strategies Inventory				x	x	
Motivation for literature education	Basic Need Satisfaction Scale: autonomy, competence, relatedness				x	x	

Findings indicated that TDLT-2 had positive effects on several indicators of students’ insight into themselves, fictional others, and real-world others. For example, TDLT-2 positively affected students’ Insight beyond story worlds, a factor score derived from the TREQ that included the transformative reading components self-insights, insights into real-world others, and aesthetic awareness, as well as their Experiences within story worlds, a factor that included imagery, experience-taking, and evaluations of how characters think, feel and behave. Likewise, TDLT-2 had a positive effect on students’ eudaimonic reasons for reading. Four months after the intervention, these effects were still statistically significant, although they were smaller than directly after the intervention. Additionally, TDLT-2 positively affected students’ strategy use, as well as their feelings of autonomy, competence, and relatedness, as indicators of their motivation for literature education. Finally, we found that students’ strategy use and feelings of competence and relatedness functioned as mediators for the effects of TDLT-2 on several indicators of insight into human nature and on eudaimonic reasons for reading. These mediating effects were statistically significant but relatively small. Students’ feelings of competence played the larg-

est mediating role on Insight beyond story worlds, the factor score that included insight into themselves and real-world others.

All in all, the findings indicated that dialogic literary instruction that centers around a single social-moral theme and focuses on transformative reading experiences enhanced 15 year-old students' insight into themselves, fictional others, and real-world others, as well as their endorsement of gaining such insights as a reason for reading. These results can be achieved in as little as five hours of classroom work. Furthermore, the study demonstrated that TDLT-2 had a small effect on students' use of reading strategies and medium to large effects on their motivation for literature education, which indicated that TDLT-2 may alleviate prominent challenges in the literature classroom that students and teachers face. Finally, mediation analysis suggested that teachers and educational designers should be responsive to students' feelings of competence in the literature classroom: the more competent students felt, the more they indicated to have gained insight into themselves, fictional others and real-world others.

## 2 CONCEPTUAL CONSIDERATIONS

This PhD research was part of the project *Uses of Literary Narrative Fiction in Social Contexts*, funded by the Dutch Organization for Scientific Research, which focused on developing and validating a theoretical-empirical model of transformative reading (Fialho, 2018) and investigating how this model could inform learning about self and others in the workplace and in the literature classroom. As a consequence, this dissertation is of an interdisciplinary nature: it is predominantly positioned as domain-specific educational research, with empirical literary studies, psychology, and sociology as adjacent fields. As an interdisciplinary work, this dissertation aimed to integrate knowledge and methods from various disciplines, utilizing a synthesis of approaches (see Repko, 2008). Therefore, it necessarily addressed conceptual and terminological multiplicities; even within a particular discipline, such as educational or literary studies, definitions of central concepts may be ill-defined, or understood differently by various researchers and other stakeholders. In this section, we reflect on four key concepts that are underlying our studies: insight into human nature, transformative reading, literariness, and dialogic learning in literature classrooms.

## 2.1 *Insight into Human Nature*

Even within the rather narrow field of Dutch literature education, teacher organizations and policy makers use a variety of terms that are all relevant to this dissertation, such as “putting reading experiences in a societal context”, “developing citizenship” and “empathic capabilities”, and “exploring one’s own perspectives, values, and assumptions” (see Chapter 3, p. 42). One of the main endeavours in this dissertation was to capture these and other relevant terms in a single concept. The term *human nature*, albeit part of the title of this dissertation, had not been established from the start of the project. Rather, the outcomes of Study 1 and 2 made us reconsider the terminology.

In Study 1, we borrowed the term *perceptions of self and others* from the grant application for the overarching project *Uses of Literary Narrative Fiction in Social Contexts*. However, the findings of Study 1 suggested that “self” and “other” are not entirely separate concepts, an insight already presented in Fialho’s (2012) work on transformative reading. She noted that, rather than by “clear differentiation between self and other”, this mode of reading is characterized by “total blurring of boundaries between self (reader) and other” (p. 273). This is illustrated by our analysis of students’ learning experiences: we found it to be difficult to discriminate between learning experiences about oneself on the one hand, and learning experiences about others on the other hand. For example, one student wrote: “I learned that I don’t like it when people are being selfish” (see Table 2.5, p. 26). This student not only realized that other people sometimes act selfishly, but also discovered something about herself: that she disliked that kind of behavior.

Similarly, some intervention studies we analyzed in Study 2 made us reconsider not only the distinction between self and other, but also between fictional and real-world others (Mar & Oatley, 2008). For instance, in an intervention study in which students were stimulated to write and talk about the theme of sexual orientation (Malo-Juvera, 2016), students were asked to consider how a character might think and feel about his sexual orientation, and were simultaneously prompted to rely on their own ideas about it, for which they were likely to rely on their own or other people’s real-world experiences. In Study 2, literature classroom interventions were found to focus on fostering students’ insight into themselves, their understanding of fictional characters (including empathy for characters, as measured in Study 5), and their understanding of, views on, or intended behavior toward real-world others (including moral competence, as measured in Study 5). In our attempt to capture these different manifestations in a single concept, we arrived at “insight into human nature”. Additional

terminological issues occurred during the design of TDLT and the development of the teacher guidelines, as we needed to work with Dutch terms to capture the central purpose of TDLT. In Dutch, one could refer to terms such as “de mens” (literally: “the human”, as an entity), “mensbeeld” (literally: “human image”), or “de aard van de mens” (literally: “the nature of mankind”). We argue that in English “insight into human nature” approached these notions as closely as possible. Moreover, as this term includes both self and others, it complies with Fialho’s most recent work, in which she distinguished “self-other perceptual depth” as an outcome of transformative reading (Fialho, 2018; Fialho, Hakemulder, & Hoeken, 2018).

## 2.2 *Transformative Reading*

In the studies included in this dissertation, we relied on Fialho’s (2012; 2018) descriptive model of transformative reading, as described for the first time in Chapter 4 (p. 90). A central question in this dissertation was whether the model would also apply to adolescent readers in the literature classroom. To investigate its applicability, the model informed both the design of TDLT-1 and TDLT-2 in Study 3 and the development of the quantitative research instrument and coding schemes for qualitative data used in Study 4 and 5.

From the findings of our studies, we concluded that transformative reading could be measured both quantitatively and qualitatively, and that the TDLT approach guided students toward noticing, reflecting on, and talking about transformative reading experiences that include, for instance, imagery, sympathy, and insight into self and others. For example, students were found to report transformative reading experiences in writing as a result of TDLT. Students’ perceptions of learning in Study 4 included transformative reading experiences, such as: “I found out that I can recognize things in stories [...] so I can better put myself in the story situation” (indicators of identification and experience-taking). Similar outcomes were found in students’ written dialogues with an imaginary peer in Study 5. For example, they referred occasionally to experiencing imagery (“Can you also picture the story in your mind?” “Yes, I already pictured it when he caught that fish.”) and, more frequently, they evaluated characters (“I think Arthur’s nephew is a rather insensitive person.”). In addition, validity indices for the Transformative Reading Experiences Questionnaire were satisfactory: both in Study 4 and 5, internal consistency of questionnaire scales as well as results of principle components analysis were consistent across measurement occasions. Its application in the quasi-experimental re-

search design of Study 5 led us to conclude that students' transformative reading experiences were fostered as a result of participating in TDLT-2.

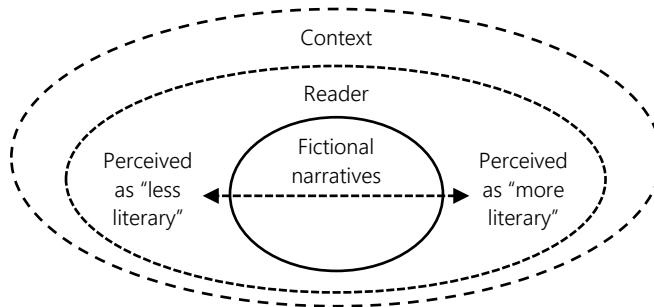
In our view, thus, this dissertation is innovative in two ways. First, the findings provide initial indications that the concept of transformative reading may apply to adolescent readers, by which it has expanded previous research into transformative reading that mainly focused on adult readers. Second, it has demonstrated that transformative reading, as originated in the field of empirical literary studies, can be meaningful in an educational context, as we were able to specify an instructional approach that fosters this mode of reading in the literature classroom. Yet, the setup of the intervention studies did not enable us to actually validate the model of transformative reading for this population of young readers. To which extent experiences like imagery, identification, and sympathy may predict self-other insights, as has been shown for adult readers, needs to be investigated further. As a next step, phenomenological and experimental studies with adolescent readers may be conducted, both in the literature classroom and in a leisure reading context, to validate the transformative reading model for these readers. The instruments developed in this project, such as the TREQ and the dialogue writing task including its coding scheme, may potentially come to aid for measuring adolescents' transformative reading experiences in experimental studies or for designing of interview schedules in future studies with adolescent readers.

### 2.3 *Literariness*

TDLT was situated in the context of *literature* education in Dutch 10<sup>th</sup> grade classrooms. Although one of the initial design principles identified in Study 2 referred only to fictional texts – based on the reviewed intervention studies, we could not draw conclusions about literariness – we used fictional narratives (i.e., short stories) that students' teachers considered to be literary texts. Here, we present and discuss our perspective and choices regarding the literariness of the texts used in TDLT.

The literariness of fictional narratives, we argue, runs gradually from "less" to "more" literary, and is influenced by three elements. First, the text itself may have characteristics that contribute to its literariness, such as the use of perspective and focalization, theme and motifs, gaps that must be filled in by the reader (Iser, 1980), and stylistic features that deviate from conventional language use (e.g., Van Peer, Hakemulder, & Zyngier, 2007). In the Dutch literature curriculum, such text characteristics are referred to as "literary devices", a term that stems from the field of linguistics. Second, from a sociological point

of view (e.g., Schmidt, 1982; Steen & Schram, 2001), it may be argued that literariness is influenced by conventions in a social context, largely established by literary institutions such as publishers, critics, academia, and schools. Third, literariness is affected by how readers in a social context perceive the text (e.g., Bourdieu, 1996; Ellis, 1974; Holland, 1975; Jauss, 1982). For these reasons, we assumed that 15 year-old students have different notions of literariness than teachers and researchers. For example, if teachers and researchers perceive a text as “more literary” due to its textual characteristics and the way it is positioned by literary critics and academics, a student reader may still perceive the text merely as difficult (see Chapter 6). In Figure 7.1, we visualize our conceptualization of literariness.



*Figure 7.1. Conceptualization of “literariness” in the domain of literature education.*

The initial design principle about fictional texts was operationalized by selecting texts that met Dutch teachers’ requirements for literariness for 10<sup>th</sup> grade students. Usually, teachers expect their 10<sup>th</sup> grade students to read texts of a certain complexity in terms of literary devices and with a certain status as established by literary institutions; more often than not, the implied reader (Booth, 1983; Iser, 1974) of these texts is an adult reader. The design of TDLT adhered thereto, as our aim was to design an instructional approach that could be implemented in the regular curriculum. We thus relied on teachers’ expertise in selecting texts that they considered to be appropriate for their students in terms of literariness, as these teachers in their daily practice constantly negotiate between their “professional” perceptions of literariness on the one hand, and students’ perceptions of it on the other hand.

In TDLT-2, particular attention was paid to the relation between literary devices and the way they affected the responses of individual readers. Students were asked to refer to literary devices to clarify and substantiate the responses

that a text evoked in them. The following excerpt, from a written dialogue by a TDLT student in Study 5 in response to the story *Hula* by Cees Nooteboom, illustrates that students were capable of such clarifications:

"I thought it was rather vague that [the protagonist] didn't want to go outside and that he just stood by the window, as if he was expecting something.

I also didn't really get an impression of him.

Yeah, I had the same thing, that's because his emotions and what he thinks are not really clearly expressed.

I really got the feeling that he was empty inside.

Exactly. But none of the characters has been described. The author mainly used scent and sound to describe the events."

This student referred to a gap caused by stylistic features ("his emotions and what he thinks are not really clearly expressed", "The author mainly used scent and sound") in substantiating the impression that the character was "empty inside". In Dutch literature classrooms, the focus on literary devices often serves an analytical purpose: students are asked to "analyze" a text, in search of such devices. Sometimes, tasks and (test) questions are confined to locating literary devices, when students are asked, for example, "Does the story contain a flashback? Where?" or "Who is telling the story?" We argue that focusing on literary devices may serve purposes that go beyond locating them, as the dialogue excerpt above suggests: students may consider how literary devices affect their reading experiences and interpretations, for instance, in developing and substantiating their understanding of a character in the text.

#### *2.4 Dialogic Learning in Literature Classrooms*

As the name TDLT – Transformative Dialogic Literature Teaching – constitutes, dialogic teaching and learning was pivotal in this instructional approach. As Higham, Brindley, and Van der Pol (2014) note, researchers refer in various ways to dialogic practices in classrooms, for example, using the terms "dialogic teaching" (Alexander, 2008), "exploratory talk" (Wegerif, Mercer, & Dawes, 1999), and "dialogic inquiry" (Wells, 2000). Many of these conceptualizations can be traced back to Bakhtin's (1981) notion of dialogue in which thinking and language are rooted. Although it is equally relevant in secondary schools, most research on dialogic education has been carried out in primary schools (Higham et al., 2014; also see Murphy, Wilkinson, Soter, Hennessey, & Alexander, 2009, for research into discussion-based reading programs). This dissertation contributes to the existing body of research on dialogic teaching by applying it in the secondary school setting. Furthermore, in our specific focus on

literature classrooms, we have argued that dialogues may take place on two levels: between the reader and the text (which we termed the internal dialogue), and among readers in response to the text (the external dialogue).

First, in terms of the internal dialogue, Bakhtin (1963/1984) suggested that fictional and literary texts offer opportunities for dialogue: among characters, between the reader and the characters, or between the reader and the author (Oatley, 1999). This point of view resonates with Rosenblatt's (1938/1983) notion of transactional reading: meaning is not just "in the text" or "in the mind of the reader", but emerges in the transaction between the reader and the text. Rosenblatt formulated several principles that may enable this transactional process of meaning-making, which included giving students freedom to deal with their own reactions and offering them opportunities for "an initial crystallization of a personal sense of the work" (1938/1983, p. 69). In line with Rosenblatt's work, Probst (1988) suggested that "instruction in literature should enable readers to find the connections between their experience and the literary work" (p. 34). Tasks that attended to identifying these connections were frequently implemented in the intervention studies that we reviewed in Study 2. Consequently, internal dialogue tasks in TDLT-2 prompted students to focus on their initial, highly personal reactions and responses to literary texts. Studies 4 and 5, however, demonstrated that internal dialogues did not come naturally in our sample of 10<sup>th</sup> grade students: for example, they needed "assistance in identifying the elements in the text that have contributed most powerfully to shaping their responses" (Probst, 1988, p. 35). In TDLT-2, this assistance took the form of strategy instruction, via which students learned how to engage in internal dialogues with the text.

The relevance of external dialogues in the classroom can be explained from both a cognitive and a social constructivist perspective (Frijters, Ten Dam, & Rijlaarsdam, 2008). First, social interaction has been found to affect cognitive elaboration processes, because language functions such as explaining, reasoning, and asking questions stimulate thinking and the development of knowledge. Second, from a social constructivist perspective, learning is considered to be a "dialogue", or a way of collaborative meaning-making. Thus, learners who engage in dialogues may be stimulated to take the perspective of others into account, engage in active learning, and develop their higher-order thinking skills (Renshaw, 2004; Salomon & Perkins, 1998). However, as Probst (1988) noted specifically for dialogues about literary texts, students "[...] are likely to need a great deal of assistance in learning the difficult process of talking with others". Similar to learning to engage in internal dialogues, we applied strategy

instruction in TDLT-2 to guide students toward engaging in external dialogues: they observed example dialogues, received explicit instruction about dialogues, and practiced dialogues of increasing complexity.

Most dialogue tasks in TDLT-2 were peer-led. However, researchers have debated the pros and cons of peer-led as well as teacher-led dialogues. As Lewis (1997) points out in her research with primary school students, peer-led talk about literature may bring students to challenging and negotiating positions of power, as “the nature of these peer groups [brings] to the surface the competing identities students must address within themselves and others” (p. 198). In this sense, peer-led dialogues seemed fit for TDLT-2, in which the purpose was to gain insight into self and others. However, Lewis ultimately concluded that the absence of the teacher in such dialogues may lead dominant students toward taking up the position of power. Even though the strategy instruction in TDLT-2 attempted to alleviate this potential risk – for example, by implementing dialogue guidelines such as “Listen carefully to others”, “Postpone your first judgment”, and “Equally distribute speaking time” – some students may have made their presence more felt, talked more, and/or directed the dialogue more than others. On the other hand, this is not to say that students who linger in the background of a group are not learning; they may well benefit from listening to and thinking about various perspectives their peers bring into the dialogue.

Furthermore, Chinn, Anderson, and Waggoner (2001) showed that moving from teacher-led toward student-led talk about literature resulted in changes in patterns of discourse in the classroom: students “gained greater control over when to speak, how long to speak, and what to speak about” (p. 403), the stance in the classroom shifted from efferent to critical-analytical and aesthetic (also see Murphy et al., 2009), and patterns of discourse suggested greater engagement and intellectual productivity. Likewise, Janssen, Braaksma, and Couzijn (2009) found that students who received a self-questioning instruction when reading and responding to short stories afterwards appreciated short stories more than students who received instructor-prepared questions. With TDLT-2, we attempted to adhere to these insights by asking teachers to move away from a monologic initiation – response – evaluation pattern of discourse when guiding whole-class discussions and students’ talk in small groups. Instead, they were asked to offer students prompts and open-ended questions to enable dialogic discourse (Applebee, Langer, Nystrand, & Gamoran, 2003; Eeds & Wells, 1989; Nystrand, 1997). By striving for such a discourse in the

classroom, teachers could avoid holding the authority of “the single correct answer” (Chinn et al., 2001, p. 403).

### 3 VALIDITY ISSUES

In this section, we first address potential threats to the validity the intervention design and its implementation. Next, we focus on the validity of the instruments and research designs that we applied in our studies. Finally, we address the external validity of the research project as a whole, and the potential influence of researcher bias.

#### 3.1 *Validity of the Intervention Design*

The design principles that informed the design of TDLT were identified via a review of previous intervention studies. The question was to which extent the review procedure was threatened by potential validity issues. Validity of systematic reviews is increased when not only research databases but also other sources are consulted during the search procedure (Moher, Liberati, Tetzlaff, Altman, & The PRISMA Group, 2009). Therefore, we applied citation tracking, conducted hand searches, and consulted experts.

We aimed to strengthen the validity of the design framework by formulating design principles that were based on two criteria: 1) valid empirical support for intervention effects that researchers expected; and 2) sufficient intervention descriptions in multiple studies. As for the first criterion, to ensure the validity of provided empirical support, we systematically assessed strengths and shortcomings of methodological characteristics of the included studies. All included studies were coded by two raters, for whom high inter-rater reliability scores were found. In some cases, presented empirical support was not deemed admissible, for example, when researchers overestimated their conclusions in view of the data and their analyses. As an example of the second criterion, themes of selected texts were often clearly reported in researchers’ intervention descriptions. Therefore, a design principle focusing on text theme could be identified. In contrast, insufficient information was reported about the role of teachers in the included interventions; consequently, we decided not to formulate a design principle about teachers’ roles and their interaction with students in their classrooms. Critical quality appraisal of methodologies and intervention descriptions in the included studies, thus, functioned as a gatekeeper for the overall validity of the review study and the design principles identified in it.

Grounding the intervention design in a theoretical-empirical framework that was based upon multiple sources – an empirical exploration of the field in Study 1, a systematic review of relevant intervention studies in Study 2, and a preliminary version of the model of transformative reading (Fialho, 2018) – contributed to the validity of TDLT (see Nieveen, 1999). Furthermore, the iterative design process contributed to the validity of the intervention design. The evaluation of TDLT-1 brought to light multiple important suggestions for improvement, upon which the redesign into TDLT-2 was based. In this sense, TDLT-2 can be considered a more valid operationalization of the design principles than TDLT-1, as TDLT-2 better complied with students' and teachers' needs and suggestions.

### *3.2 Validity of the Intervention-as-Implemented*

The implementation in practice of both TDLT-1 and TDLT-2 was closely monitored. In this process, as recommended by McKenney, Nieveen, and Van den Akker (2006), we ensured triangulation of data collection methods by applying multiple instruments, including logs, observations, interviews, and evaluation forms. Moreover, our conclusions about the validity of the interventions-as-implemented did not concern a single group of stakeholders, but were based on experiences of both students and teachers. Therefore, we appear to have gained a valid impression of the extent to which TDLT-1 and TDLT-2 were implemented in accordance with their original design (implementation fidelity; O'Donnell, 2008).

The data collected via implementation and evaluation measures suggested that implementation fidelity of TDLT-1 had been at stake. Despite the fact that teachers reported in their logs that both TDLT-1 and TDLT-2 were largely implemented as intended – in terms of the number of TDLT phases that were completed – time on task observations indicated that the threshold of 80% time on task (e.g., Muijs & Reynolds, 2010) was not met in all classes. In particular, students too often showed off task behavior during small-group dialogues, a finding that was supported by the impressions of various teachers, who mentioned in interviews that students were “quickly done talking” or “completed tasks superficially”. This validity issue appeared to be alleviated in TDLT-2, as time on task percentages were well above the 80% threshold.

When evaluating the validity of the interventions-as-implemented, ecological validity must be considered as well. If ecological validity is threatened, an intervention does not sufficiently resemble the real-world setting. We attempted to optimize ecological validity of TDLT-1 and TDLT-2 in various ways. Both

were taught by the students' own teachers, during regular hours of Dutch class. Moreover, by involving teachers in designing and redesigning TDLT, we guarded against too drastic deviations from usual practices in literary instruction. The problems and needs expressed by students and teachers during the evaluation of TDLT-1 – for example, concerning its relevance and its fit in the regular literature curriculum – formed the starting point for attuning TDLT-2 even more to “regular practice”. For instance, more emphasis was put on learning to refer to literary devices in writing and talking about reading experiences, and we implemented a final intervention task that could be included in students' literature portfolios that are often kept in the upper grades of literature education. As a consequence, TDLT-2 may have been more ecologically valid than TDLT-1.

### 3.3 *Validity of the Instruments*

In the intervention studies, we focused on measuring reading experiences. To collect data from the entire sample in both studies, we developed the Transformative Reading Experiences Questionnaire (TREQ), based upon Fialho's (2018) phenomenological work on transformative reading. The internal validity of the TREQ was supported by results of principal components analysis and internal consistency of subscales across measurements in Studies 4 and 5. These findings indicated that different samples of students generally responded to the TREQ items in a similar way.

We further aimed to validly bring subjective reading experiences into the open via data triangulation. In addition to the TREQ, we developed writing tasks to collect students' responses to stories. The transformative reading model (Fialho, 2012; 2018) informed the analysis of students' learning experiences and story responses in Study 4 and their written dialogues with an imaginary peer in Study 5. In these written responses, we found indicators of imagery, identification, sympathy, insight into oneself, and so forth. Students, thus, were not only capable of reflecting on their transformative reading experiences by indicating their agreement with questionnaire items, but also expressed such experiences in more open writing tasks. Moreover, transformative reading experiences were measured on two different levels: as general experiences, via items in the TREQ that were formulated as “When I read stories, I...”, and in response to a specific literary text. In short, even though reading experiences are highly subjective, our studies showed concurrent validity in the sense that these experiences were explicated in various instruments. In this way, we also avoided what Shadish et al. (2002) refer to as “monomethod bias” as a threat to construct validity.

In Study 4, students were first asked to write their initial responses to a story in the margin. In both the experimental and the control conditions, students' responses were non-extensive to medium extensive (see Appendix E, p. 284). Self- and socially-related responses did not occur frequently. The validity of this part of the instrument may have been questionable, for it may be in the nature of reading responses that more complex responses, such as life lessons, come up after reading a story, when readers reflect on a text as a whole (e.g., Fialho, 2012; 2018; Fialho, Miall, & Zyngier, 2012; Fialho, Zyngier & Miall, 2011; Koopman & Hakemulder, 2015). However, in the short writing tasks that students were asked to complete after reading, we found similar results: literal inferences were more frequent than life lessons, and predictions for characters remained rather naive and conformist. Perhaps, the stories were too difficult for students or the task was not interesting and engaging enough for students; in particular the short writing tasks after reading may have evoked the impression of "test questions about a text". In this sense, the instrument did not fit the TDLT principles and may therefore have failed to capture any effects. However, no data was collected that may support these assumptions: we did not ask students to indicate the difficulty of the stories, nor their appreciation of the task.

The written dialogue with an imaginary peer in Study 5 may have given a more valid impression of the intervention effects, since this task was presented – both in the experimental and the control condition – as part of literary instruction; as such, it was a more creative task than the one in Study 4, that merely functioned as a research instrument. Moreover, the task in Study 5 was an after-reading task, in contrast with the first part of the task in Study 4. Finally, writing a dialogue may accurately reflect what students learned from TDLT-2. Not only did the task prompt students to engage in an internal dialogue with a story as well as rely on their knowledge of external dialogues about stories, but writing a dialogue is also deemed a powerful way to generate new ideas (Scardamalia & Bereiter, 1987). All in all, whereas one of the instruments in Study 4 may have been subject to validity issues, we are rather confident that the effects of TDLT-2 that were found in Study 5 were valid results of the intervention.

Two instruments that were developed for Study 1 were used or adapted for later studies: the learner report and the Author Recognition Test. A potential threat to the validity of the learner report was the extent to which it actually reflected students' learning experiences. We may have gained access to only the metaphorical top of the iceberg of students' learning experiences: it is conceivable that part of their learning experiences remained implicit, and were not

reported in writing. On the other hand, students indicated that they did not find the learner report too difficult to complete. They reported on average 10 learning experiences, and only 2.2 percent of all learning experiences was irrelevant or incomprehensible (see Table 2.9, p. 30). In line with previous research (e.g., Janssen, 1998), we concluded that the learner report was a valid way of tapping into students' learning experiences. Consequently, we used a shorter version of the learner report in Study 4: it was presented as a reflection task in the final unit of TDLT-1. It was more open – asking students what they learned from TDLT-1 – than in Study 1, in which we specifically asked what students learned about themselves and others via book reading and attending literature class. Moreover, students were asked to write two learning experiences, in contrast with “as many learning experiences as they could think of” in Study 1. Still, the shorter version of the learner report appeared to be a valid instrument: learning experiences were well-distinguishable, as demonstrated by the inductively developed coding system for which inter-rater agreement was substantial, and were in line with various important aims of the intervention (e.g., learning to talk about stories). Moreover, similar to the more extensive version of the learner report, only few responses (3.3%) were unclear.

Second, we administered the Author Recognition Test (ART) in Studies 1, 4, and 5, to measure familiarity with fiction as a background variable. The test has been validated in previous studies, as it was found to predict real-world reading (e.g., Mar, Oatley, Hirsh, de la Paz, & Peterson, 2006; Rain & Mar, 2014). In Study 1, its validity was further supported for a sample of Dutch adolescents ( $N = 297$ ), as those in the pre-university track scored significantly higher than those in the higher general secondary education track, which is in accordance with the fact that pre-university students are required to read more literary works than students in the higher general track (Van Grootheest & Van Grinsven, 2016; SLO, 2012). Moreover, average ART scores on a scale of -40 to +40 were rather similar across studies for 10<sup>th</sup> grade students in the higher general education track, varying between 4.3 and 5.4. The ART, thus, appeared to be a valid instrument for measuring students' familiarity with fiction.

### 3.4 *Validity of the Research Designs*

In quasi-experimental studies, internal validity refers to “the validity of inferences about whether observed covariation between A (the presumed treatment) and B (the presumed outcome) reflects a causal relationship from A to B as those variables were manipulated or measured” (Shadish, Cook, & Campbell, 2002, p. 38). An essential question, then, is whether we can ascribe the effects

of TDLT to the treatment, or whether alternative explanations for these effects are more plausible.

As a first potential validity issue, in particular in small-scale intervention studies, undesired teacher effects may occur. For example, if one teacher is involved in the intervention condition and another teacher in the control condition, effects may actually be teacher effects rather than intervention effects. Therefore, in Studies 4 and 5, multiple teachers were involved in both the experimental and control condition. In Study 4, we applied a switching replications design, which is a strong design for detecting intervention effects because it allows for conclusions about sustainability and generalizability (Shadish et al., 2002). Teachers were randomly assigned to either of two orders of conditions (experimental-control or control-experimental). In Study 5, as teachers indicated they could not devote six lessons to the control condition in addition to the six TDLT-2 units, we applied a quasi-experimental design without switching replications.

Second, in Study 5, teachers and their classes ideally should have been randomly assigned to conditions. However, three teachers were already involved in redesigning TDLT; three others signed up as “new” TDLT teachers, and four other teachers volunteered to participate in the control condition. Thereby, we avoided contamination of conditions, but the absence of random assignment to conditions posed a threat to validity. However, upon comparing the students in these teachers’ classes on relevant background variables, we found no significant differences between particular classes. This finding counterbalances the lack of random assignment to conditions to some extent.

Selection bias may have been a third potential threat to validity: the teachers who were already familiar with TDLT may have given the experimental group an advantage. However, when we performed additional *t*-tests on dependent variables at T2 and T3, such an effect was not observed: students of teachers who were already familiar with the approach did not score significantly higher than students of the teachers who were new to TDLT-2.

Fourth, the control conditions in the intervention studies remained somewhat obscure. This may have posed a threat to the validity of the effects found in Study 5. Control teachers indicated, prior to the study, that they planned to focus on literary devices and analysis. They were asked to continue these literature lessons “as usual”. The logs they completed about their lessons generally complied with what they planned to teach as announced prior to the study. Occasionally, they reported that their students talked in groups about literary stories, as did the students in the experimental condition. Other than that,

there appeared to be no contamination of conditions: none of the control teachers reported to have paid attention to transformative reading experiences or explicit strategies for internal and external dialogues. Due to practical constraints, we did not perform time on task observations in their classrooms as additional fidelity checks. Thus, less information about the implementation of the control condition was gathered than about implementation of TDLT-2. On the other hand, there were no reasons to assume that control teachers would provide inaccurate information about their literature lessons.

Finally, a potential criticism might be that the outcomes of Study 5 were due to a so-called Hawthorne effect: a positive effect that occurs simply because participants engage in something new (Izawa, French, & Hedge, 2011). Indeed, the approach taken in TDLT-2 was new to students, in particular in comparison to the control condition: in lower grades of secondary education, there usually is already some attention for literary devices in texts. From a methodological point of view, therefore, we acknowledge that a Hawthorne effect cannot be ruled out as an alternative explanation for the positive effects of TDLT-2 on students' insight into human nature and eudaimonic reasons for reading. However, from an educational perspective the question rises to which extent such an effect is problematic. We would argue that if teachers implement a new and unfamiliar instructional approach that evokes students' curiosity, challenges them, and successfully engages them in learning, hardly any objections can be brought against it.

### 3.5 *External Validity*

External validity refers to the generalizability of research findings across treatment variations, populations, and settings. A first question is whether the effects of TDLT-2 would hold if changes were to be made in its operationalization (see Shadish et al., 2002, p. 87). The design principles that informed the design of TDLT were validly drawn from previous interventions, implemented in a variety of settings, for which empirical support was found. Therefore, the three initial design principles may be considered what Merrill (2002) calls "first principles": they may be operationalized in various ways, but the mere fact that the principles are operationalized increases the probability that the desired learning outcome will be achieved.

This dissertation in itself exemplifies that different operationalizations of first principles may result in different outcomes, as shown by Studies 4 and 5. Therefore, we cannot be sure that similar effects would occur with even the smallest variation of what happens in class. On the other hand, we are not

dealing with lab experiments that focus on effects of highly controlled treatments; rather, in our intervention studies we assessed the sum of a number of operationalizations that were implemented in an ecologically valid setting. Small variations in implementation are inevitable, as shown by teachers' logs: for example, a teacher might have skipped a particular phase in a unit, whereas other teachers taught it according to plan. We argue, therefore, that it is unlikely that small adjustments in TDLT units would drastically change its outcomes. For instance, as long as fictional stories center around a single social-moral theme, selecting different stories than those read in TDLT-2 is unlikely to result in different outcomes.

The findings of the research project as a whole cannot be generalized beyond students in upper secondary education. Within this setting, a strength of the studies that relied on student data (Studies 1, 3, 4, and 5) was that participants were students from a variety of schools and classes across the Netherlands. We have no reasons to assume that these students would be different from their peers in other upper secondary school classrooms. However, we cannot generalize the findings of the studies to other grade levels and educational tracks than the ones addressed in the studies, nor to literature curricula in other countries than the Netherlands. Yet, as suggested above, the design principles may be operationalized differently to tailor TDLT to grade levels and literature curricula other than 10<sup>th</sup> grade higher general education classrooms in the Netherlands.

The intervention studies relied not only on samples of students, but also on samples of teachers. Would the effects of TDLT-2 hold if different samples of teachers were involved? Although the variety in years of teaching experience was rather large among the six teachers who taught TDLT-2 (ranging from 5 to 32 years), they all seemed equally able to implement TDLT as it was intended. This was indicated by results of time on task observations that did not differ significantly across teachers (see Chapter 4, p. 115), as well as by teacher log data, for which an additional analysis of variance showed a significant difference for one of the implementation fidelity indicators (feasibility), between only two out of six teachers: one had five years of teaching experience, the other ten years. Years of teaching experience, thus, does not appear to impede the generalizability of TDLT-2. Similarly, as indicated on p. 115, effects of TDLT-2 were consistent across teachers who were already familiar with TDLT and those who were not.

A final aspect of generalizability that we discuss here, is "narrowing down generalizability" (Shadish et al., 2002, p. 83). Interventions that show positive

average outcomes are not necessarily equally effective for individual students. In both our intervention studies, although we specified repeated measures of individual students as the random factor in the statistical analyses, the outcomes still concern the experimental condition versus the comparison condition as a whole. Thus, TDLT-2 was found to be effective for the average student, but we cannot simply assume that the intervention effects are generalizable to individual students. However, when we performed additional explorative moderator analyses, effects of condition on insight into human nature and eudaimonic reasons for reading were not moderated by students' familiarity with fiction, average grade for Dutch class, and gender. This indicates that TDLT-2 was equally effective for subgroups within the sample, such as boys and girls, as well as more and less well-read students.

### 3.6 *Researcher Bias*

A potential limitation that has not yet been addressed but that may undermine the validity of our conclusions, is researcher bias. For example, as researchers, we were well aware of the conditions and pretests and posttests, which may have influenced our coding and analysis of students' written responses. We have taken several measures to counterbalance the risk of researcher bias. We designed TDLT in collaboration with teachers and relied on multiple data sources in the design process, such as teacher interviews, student evaluation forms, and observations. Although our analysis of written responses may have been somewhat biased due to our knowledge of conditions and measurement occasions, neither of the intervention studies merely relied on those data, but also included questionnaire data that was not subject to coding. In addition, data were not collected by a single researcher, but in collaboration with a team of research assistants, which reduced the risk of bias in, for example, time on task observations. Finally, when analyzing qualitative data, independent researchers were asked to code subsets of the data to assess inter-rater reliability, which was sufficient in all cases. All in all, it seems unlikely that the outcomes of this research project have mainly been caused by researcher bias.

## 4 FUTURE STUDIES

This dissertation provides several starting points for future studies. First, future studies may focus on assessing the respective roles of the core elements of TDLT-2: text selection (in particular in terms of the themes addressed in texts), internal dialogues, external dialogues, and attention to transformative reading

experiences. This would require experimental or quasi-experimental studies in which these elements are implemented in separate conditions. Such studies may illuminate, for example, whether external dialogues are essential for developing insight into human nature, or whether students may already gain such insights from internal dialogues with texts that they explicate in writing tasks; or vice versa: whether students may gain insights from external dialogues even when in the intervention no attention is paid to internal dialogues. Potentially, differences amongst students come to light: some students may learn best by writing, whereas others may benefit more from talking to their peers.

Second, future studies may provide further insight into cognitive and affective processes that are underlying students' development of insight into human nature. Although the written dialogue task indicated what students thought and felt during internal dialogues and to which extent they would express and share their responses in external dialogues, alternative methods may further elucidate these processes. For example, reflective response interviews about a story that is read may demonstrate to which extent and how students individually gain insight into human nature in relation to the text. Content analysis of students' talk in small groups may illuminate how issues related to human nature are addressed in such external dialogues. Learning more about these response processes would be particularly helpful for the design of additional guidelines or training for literature teachers, focused on providing students with process-oriented rather than outcome-oriented feedback or feedforward.

Finally, there is a need for further development and validation of instruments for measuring insight into human nature. The TREQ, for instance, may be administered to a larger and more varied sample of adolescents. The exploratory factor analyses performed in Studies 4 and 5 may be complemented with confirmatory factor analysis. In addition, the written dialogue with an imaginary peer may be developed further as a research instrument. For example, studies may focus on developing a scale for the extent to which (indicators of) insight into human nature can be detected in students' written dialogues, via comparative analysis procedures. Development of such a scale would, in contrast with the coding system that was inductively developed in Study 5, enable a more time-efficient analysis of the written dialogues. Triangulating the outcomes of such an analysis with data from other sources, such as the TREQ or data collected in interview studies with student readers, may further validate the written dialogue task as a research instrument.

## 5 IMPLICATIONS FOR EDUCATIONAL PRACTICE

Since the start of this project, in 2014, attention for the potential of literature education to foster students' insight in what we call in this dissertation "human nature" has become more prominent. In the context of an imminent nationwide curricular reform in the Netherlands that is to be implemented in 2021 (titled Curriculum.nu), teachers, policy makers, educational designers, and researchers debate what the curriculum for students of future generations should entail. In one of their interim conclusions, the National Curriculum Design Team for Dutch language and literature expressed the importance of literature education as follows:

'By talking about [books] with others [...] and by reading by themselves, students gain insight into ethical issues and political, socio-cultural, and societal discussions. They learn to explore and take on multiple perspectives, and learn to ask questions to texts, themselves and others. Furthermore, they learn to call into question their own stances and opinions, to postpone their judgments, and to weigh arguments based on inquiry and interaction' (Curriculum.nu, 2018b, p. 9, *translation MS*).

We believe that the main contributions of this dissertation to Dutch teaching practices are threefold. First, it offers research-based instructional design principles that meet the kind of approach to literature teaching as envisioned by Curriculum.nu. Second, it has shown how these design principles can effectively be operationalized in class, in as little as six units of about fifty minutes. Finally, it goes beyond claims at the rhetoric level: it has empirically demonstrated that this operationalization fosters 10<sup>th</sup> grade students' insight into human nature, their support for eudaimonic reasons for reading, their reported use of strategies to deal with difficulties during literary reading, and their motivation for literature education.

In this project, "developing insight into human nature" was the main objective of literary instruction. As we argued in the Introduction, developing this insight appears to be highly relevant in contemporary society. Simultaneously, it may potentially have positive implications for students' reading comprehension. Boerma, Mol, and Jolles (2017) showed for younger readers of 8 to 11 years old that their "ability to make inferences about characters' mental states and to infer and understand the complex social relations that are often present in narratives" (p. 181) was positively related to their reading comprehension, because they create a more complete mental model of a story. As students who participated in TDLT were found to develop understandings of fictional others, we hypothesize that this may also positively impact their comprehension of literary texts. In addition, Mol and Bus (2011), although focusing on

reading as a leisure activity, pointed to the potential of reciprocal causation in reading (cf. Stanovich, 1986): if children and adolescents enjoy reading, they read more frequently; this improves their reading comprehension, which in turn stimulates them to continue reading. In other words, an upward spiral in reading frequency may be created. As we have shown that TDLT-2 enhanced students' motivation for literature education, it may perhaps stimulate students to enter such a spiral.

TDLT offers an instructional specification for reader-oriented approaches to literature education, which oftentimes are distinguished from approaches that are more oriented toward texts-as-objects, their analysis, or their cultural-historical functions (e.g., Janssen, 1998; Verboord, 2005; Fialho et al., 2011; 2012; Henschel et al., 2016). Although previous studies as well as the current dissertation endorse the positive effects of reader- and affect-oriented approaches on students' insight into human nature and their motivation for literature education, we emphasize that teaching practices in reality tend to be more eclectic (e.g., Applebee, 1994; Janssen, 1998; Van de Ven & Doecke, 2011). Yet, we argue that, within this variety of teacher practices, the starting point of any instructional approach should be that students learn to become aware of their initial responses to literary texts and to put these into words. If students get a grip on their initial responses, this subsequently may create opportunities for achieving particular learning objectives, be it gaining insight into human nature or learning to substantiate interpretations of literary texts (e.g., Janssen, Braaksma, & Couzijn, 2009; Koek, Janssen, Hakemulder, & Rijlaarsdam, 2017; Levine, 2014).

In our view, the notion that students' initial responses to literary texts form the starting point for TDLT has two important implications. First, it answers to the potential criticism that TDLT, in its focus on developing insight into human nature, cannot truly be considered "literary instruction", but rather is a form of citizenship education or moral education in which stories or other texts are used instrumentally. We argue that precisely the attention for students' initial responses to texts, as well as the focus on substantiating those responses with references to literary devices identified in texts, differentiates TDLT as an approach to literary instruction from other disciplines. The second implication of taking students' initial responses to texts as a starting point is that monologic, teacher-centered literary instruction does not appear to be the way forward for future literature curricula. Rather than taking the lead and functioning as authorities on "the" meaning of a literary text, literature teachers may function as models who actively engage in interaction with and about literary texts, and

are the appointed experts-by-experience for guiding their students toward such interactions.

These implications raise at least three questions. First, one might ask how the implementation of TDLT and its principles may find its way into literature classrooms. Although we have demonstrated that teachers were generally able to implement TDLT-2 as it was intended, we have to bear in mind that they were either involved in its design or received a workshop prior to the study. Teachers, thus, may be in need of at least some training to implement an approach that emphasizes interaction with and about literary texts. In the current so-called knowledge bases for teacher training programs of Dutch language and literature, knowledge and skills that concern dialogic interaction with and about literary texts are not included as objectives (De Blauw, Bloemhoff, Nuijten, Severijnen, & Wegman, 2011/2012; Van der Borden, Van Dam-Helmig, Kniep, De Puit, & Stienen, 2017). Therefore, there may be a need to develop a professional development course for pre-service and in-service teachers. Moreover, as our studies have shown that teachers' input is indispensable in developing literature classroom interventions, it may be worth considering to set up teacher design teams for teachers' professional development (Binkhorst, Handzalts, Poortman, & Van Joolingen, 2015).

Second, when developing an instructional approach, questions about assessment and testing inevitably emerge. Although such questions were beyond the scope of this dissertation, they are highly relevant to teachers. Importantly, we should ask ourselves whether insight into human nature is a learning objective that can and should be tested and graded. The notion that reading experiences are personal, multifaceted, and not "right" or "wrong" lies at the heart of the approach developed in this dissertation. We would therefore advise against summative testing and grading in the context of TDLT and similar approaches. Yet, formative assessment (e.g., Sadler, 1989; Bennett, 2011) may help students to reflect on their own learning gains. The self-evaluation rubric that was used in TDLT (see Appendix F, p. 280) may be a preliminary example of a formative assessment tool, which would need further development and validation. In addition, teachers may evaluate and provide feedback on the quality of students' interactions with and about literary texts. For this, the task to write a dialogue with an imaginary peer as we implemented in Study 5 may be useful. Evaluation and feedback may focus, for example, on the range of topics students address, the extent to which they support their statements by referring to literary devices found in the text, and the form of the imagined conversation (e.g., the use of open-ended and follow-up questions). Similarly, teachers and peers may

provide feedback on such aspects when a small group of students engages in an actual external dialogue in response to a literary text. Lastly, as we have shown in Study 1, a learner report task is a valid way to map students' self-perceived learning outcomes. Therefore, such a task may function as another useful tool in formatively assessing the effects of the TDLT approach. As emphasized by De Groot (1980b), it is important to assess fundamental personal learning experiences, in addition to assessing those outcomes that can be tested objectively and reliably.

Finally, we may ask ourselves to which texts and in which other educational disciplines TDLT principles may be applied. In 10<sup>th</sup> grade as well as lower grades, adolescent or young adult literature may be selected for TDLT. Such texts often center around adolescent protagonists' identity development. Usually, characters' reflections on their adaptation to the world around them and their integration into society are addressed; sometimes, there is also a focus on the shift from childhood into adolescence and/or from adolescence into adulthood (Joosen & Vloeberghs, 2008; Van den Hoven & Van Lierop-Debrauwer, 2014). Conceivably, these texts may evoke transformative reading experiences such as identification and experience-taking in adolescent readers. On the other hand, Study 1 indicated that students may also gain insight into themselves and others from texts in which "the other" plays a prominent role, because "in books, suddenly you come very close to different people with other philosophies of life", as one of the students reported (see p. 26). Whether TDLT principles may also be applied to historical literary texts or texts in foreign language curricula remains to be seen: on the one hand, such texts may be thematically relevant, but on the other hand, a text in a historical or foreign language may hinder students' transformative reading experiences. Lastly, TDLT may offer opportunities for interdisciplinarity in schools, by connecting literary instruction to other disciplines in which the concept of human nature plays a role, such as history, citizenship education, social studies, and arts education.

## 6 CONCLUDING REMARKS

This dissertation opened with John Green's suggestion that books may offer us windows into the lives of others and mirrors so that we can better see ourselves. In our studies, we have developed an instructional approach to bring these potentials of reading fictional and literary texts into the limelight. In doing so, for the research community focusing on language and literature education we hope to have provided useful examples of how first principles for literary instruction may be identified, what an iterative design process in collabora-

tion with teachers may look like, and how the implementation and effects of interventions may be assessed.

One of our key findings is that, in addition to the recurring issue of what students in upper secondary school literature classrooms read, it is equally important to address how they read. If we aim for fostering their insight into human nature, students appear to benefit from dialogic interaction with and about texts to develop awareness of their responses to what they read, and to deepen their responses by sharing them with others and comparing their responses. In the Netherlands, the most recent public debate about literature education – in national newspapers, manifests, and on blogs and social media – mainly focused on “the required reading list” for students (e.g., Boogers, 2015; Pruis & De Vries, 2016; Weijts, 2016). The relevance of this debate is beyond dispute and had already been given empirical depth by Witte (2008), who developed a model for fostering students’ literary competence by focusing on text selection (also see Witte, Rijlaarsdam, & Schram, 2012). The model has been operationalized into a practice-oriented tool – widely embraced by Dutch literature teachers – which helps students to select “the right book at the right moment”. As an addition to Witte’s work, the present dissertation suggests that considering the instructional approach that is implemented in the literature classroom may be equally important: as researchers, teachers, and curriculum designers, we should discuss what it is that we aim for when we teach literature to young people, and, from an instructional point of view, what it takes for students and teachers to achieve those aims. Our research suggests that students’ insight into human nature may be a learning outcome that is worth striving for in the literature classroom. Therefore, it is our hope that this dissertation may contribute to the ongoing dialogues on what literature education is, may, or should be about.

## REFERENCES

- Adler, E. S., & Foster, P. (1997). A literature-based approach to teaching values to adolescents: Does it work? *Adolescence*, *32*(126), 275-286.
- Alexander, R. (2008). *Towards dialogic teaching: Rethinking classroom talk*. York, UK: Dialogos.
- Andringa, E. (1990). Leeservaring: wat maakt het uit? Leesprocessen bij het verhaal 'Eb en Vloed' van F. B. Hotz. [Reading experience: what does it matter? Reading processes by the story 'Eb en Vloed' by F. B. Hotz.] In E. Andringa & D. Schram (Eds.), *Literatuur in functie. Empirische literatuurwetenschap in didactisch perspectief* (pp. 27-49). Houten: Bohn Stafleu Van Loghum.
- Andringa, E. (1995). Strategieën bij het lezen van literatuur. [Strategies in reading literature.] *Spiegel*, *13*(3), 7-33.
- Andringa, E. (1996). Effects of "narrative distance" on readers' emotional involvement and response. *Poetics*, *23*, 431-452. doi:10.1016/0304-422x(95)00009-9
- Applebee, A. N. (1978). *The child's concept of story: Ages two to seventeen*. Chicago, IL: Chicago University Press.
- Applebee, A. N. (1994). *Literature in the secondary school. Studies of curriculum and instruction in the United States. NCTE Research Report 25*. Urbana, IL: NCTE.
- Applebee, A. N., Langer, J. A., Nystrand, M., & Gamoran, A. (2003). Discussion-based approaches to developing understanding: Classroom instruction and student performance in middle and high school English. *American Educational Research Journal*, *40*, 685-730. doi:10.3102/00028312040003685
- Appleyard, J. A. (1991). *Becoming a reader. The experience of fiction from childhood to adulthood*. Cambridge, MA: Cambridge University Press. doi:10.1017/CBO9780511527609
- Bakhtin, M. (1981). *The dialogic imagination*. Austin, TX: University of Texas Press.
- Bakhtin, M. (1984). *Problems of Dostoevsky's poetics*. London, MN: University of Minneapolis Press. (Original work published 1963)
- Bal, P. M., Butterman, O. S., & Bakker, A. B. (2011). The influence of fictional narrative experience on work outcomes: A conceptual analysis and research model. *Review of General Psychology*, *15*, 361-370. doi:10.1037/a0025416
- Bal, M., & Veltkamp, M. (2013). How does fiction reading influence empathy? An experimental investigation on the role of emotional transportation. *Plos One*, *8*, 1-12. doi:10.1371/journal.pone.0055341
- Banks, W. P. (2009). Literacy, sexuality, and the value(s) of queer young adult literatures. *English Journal*, *98*(4), 33-36.
- Barnes, D. (1976). *From communication to curriculum*. Harmondsworth, UK: Penguin.
- Beach, R. (1993). *Reader response theories*. Urbana, IL: National Council for Teacher Education.
- Beach, R., Appleman, D., Hynds, S., & Wilhelm, J. D. (2011). *Teaching literature to adolescents* (2<sup>nd</sup> ed.). New York, NY: Routledge.
- Bean, T. W. & Moni, K. (2003). Developing students' critical literacy: exploring identify construction in young adult fiction. *Journal of Adolescent & Adult Literacy*, *46*(8), 638-648.
- Benavides Buitrago, C. (2017). EFL students' social identities construction through genderbased short stories. *Colombian Applied Linguistics Journal*, *19*, 11-21. doi:10.14483/calj.v19n1.10641
- Bender-Slack, D. (2002). Using literature to teach global education: A humanist approach. *English Journal*, *91*, 70-75. doi:10.2307/821401
- Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice*, *18*, 5-25. doi:10.1080/0969594x.2010.513678
- Biesta, G. (2007). The education-socialisation conundrum, or 'Who is afraid of education?' *Utbildning & Demokrati*, *16*(3), 25-36.

- Binkhorst, F., Handelzalts, A., Poortman, C. L., & Van Joolingen, W. R. (2015). Understanding teacher design teams. A mixed methods approach to developing a descriptive framework. *Teaching and Teacher Education, 51*, 213-224. doi:10.1016/j.tate.2015.07.006
- Bintz, W. P. (1993). Resistant readers in secondary education: Some insights and implications. *Journal of Reading, 36*(8), 604-615.
- Blackie, M., & Wear, D. (2015). Three things to do with stories: Using literature in medical, health professions, and interprofessional education. *Academic Medicine, 90*, 1309-1313. doi:10.1097/acm.0000000000000761
- Bloome, D. & Green, J. L. (2002). Directions in the sociolinguistic study of reading. In P.D. Pearson, R. Barr, M.L. Kamil, & P. Mosenthal (Eds.), *Handbook of Reading Research* (pp. 395-422). London, UK / Mahwah, NJ: Lawrence Erlbaum. (Original work published 1984)
- Boerma, I. E., Mol, S. E., & Jolles, J. (2017). The role of home literacy environment, mentalizing, expressive verbal ability, and print exposure in third and fourth graders' reading comprehension. *Scientific Studies of Reading, 21*, 179-193. doi:10.1080/10888438.2016.1277727
- Boogers, A. (2015). *De lezer is niet dood. [The reader is not dead.]* Amsterdam, Netherlands: Podium.
- Boostrom, R. (1998). "Safe Spaces": Reflections on an educational metaphor. *Journal of Curriculum Studies, 30*, 397-408. doi:10.1080/002202798183549
- Booth, W. C. (1983). *The rhetoric of fiction* (2<sup>nd</sup> Ed.). Chicago, IL: University of Chicago Press.
- Bourdieu, P. (1996). *The rules of art: Genesis and structure of the literary field*. Stanford, CA: Stanford University Press.
- Brokerhof, I. (2018). Uses of narrative fiction in the workplace: Changes in self and social perceptions (doctoral dissertation). Manuscript in preparation.
- Charlton, M., Pette, C., & Burbaum, C. (2004). Reading strategies in everyday life: Different ways of reading a novel which make a distinction. *Poetics Today, 25*, 241-263. doi:10.1215/03335372-25-2-241
- Chinn, C. A., Anderson, R. C., & Wagonner, M. A. (2001). Patterns of discourse in two kinds of literature discussion. *Reading Research Quarterly, 36*, 378-411. doi:10.1598/rrq.36.4.3
- Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment, 6*, 284-290. doi:10.1037//1040-3590.6.4.284
- Cummins, A. (2013). Border crossings: Undocumented migration between Mexico and the United States in contemporary young adult literature. *Children's Literature in Education, 44*, 57-73. doi:10.1007/s10583-012-9176-1
- Curriculum. (2017). *Nederlands – Vakgebonden eindtermen* [Dutch – Domain specific objectives]. Retrieved from <http://eindtermen.vlaanderen.be/secundair-onderwijs/derdegraad/aso/vakgebonden/nederlands/eindtermen.htm>
- Curriculum.nu. (2018a). *Tweede tussenproduct Nederlands [Second intermediate report Dutch]*. Retrieved from <https://curriculum.nu/ontwikkelteam/nederlands/>
- Curriculum.nu (2018b). *Derde tussenproduct Nederlands [Third intermediate report Dutch]*. Retrieved from <https://curriculum.nu/ontwikkelteam/nederlands/>
- Curwood, J. S. (2013). Redefining normal: A critical analysis of (dis)ability in young adult literature. *Children's Literature in Education, 44*, 15-28. doi:10.1007/s10583-012-9177-0
- Darragh, J. J. (2015). Exploring the effects of reading young adult literature that portrays people with disabilities in the inclusion classroom. *Electronic Journal for Inclusive Education, 3*(4), 1-48.
- Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology, 44*, 113-126. doi:10.1037/0022-3514.44.1.113

- Davis, O. L., Yeager, E. A., & Foster, S. J. (2001). *Historical empathy and perspective taking in the social studies*. London, UK / Lanham, MD: Rowman & Littlefield.
- De Blauw, E., Bloemhoff, H., Nuijten, K., Severijnen, O., & Wegman, H. (2011/2012). *Kennisbasis docent Nederlands: master. [Knowledge base teacher Dutch: master.]* Den Haag, Netherlands: HBO-raad / Vereniging van hogescholen.
- De Groot, A. D. (1980a). Learner reports as a tool in the evaluation of psychotherapy. In W. De Moor & H. R. Wijngaarden (Eds.), *Psychotherapy, Research and Training* (pp. 177-182). Amsterdam: Elsevier/North-Holland Biomedical Press.
- De Groot, A. D. (1980b). Over Leerervaringen en Leerdoelen. [On Learning Experiences and Learning Goals.] *Handboek voor de Onderwijspraktijk [Handbook for Educational Practice]* (B1-B.18). Deventer, The Netherlands: Van Loghum Slaterus.
- De Leon, C. (2017). Engaging adolescents in mental health discussions through young adult literature: A literature review. *Queen's University Graduate Student Symposium*. Retrieved from <http://qspace.library.queensu.ca/handle/1974/15468>
- Deci, E. L., & Ryan, R. M. (2000). The 'what' and 'why' of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, *11*, 227-268. doi:10.1207/s15327965pli1104\_01
- Derricott, R. (2014). *Citizenship for the 21st century: An international perspective on education*. London, UK: Routledge.
- Dijkstra, K., Verkoeijen, P., Van Kwijk, I., Yee Chow, S., Bakker, A., & Zwaan, R. (2015). Leidt het lezen van literaire fictie tot meer empathie? Een replicatiestudie. [Does reading literary fiction lead to more empathy? A replication study.] *De Psycholoog*, October, 10-21.
- Dirksen, J. (2007). Leerlingen, literatuur en literatuuronderwijs. [Students, literature and literature education.] In H. Goosen (Ed.), *Forum of Arena: Opvattingen over Literatuuronderwijs* (pp. 12-24). Amsterdam, Netherlands: VON.
- Dochy, F., Segers, M., & Buehl, M. M. (1999). The relation between assessment practices and outcomes of studies: The case of research on prior knowledge. *Review of Educational Research*, *69*, 145-186. doi:10.3102/00346543069002145
- Duriez, B., & Demarez, P. (2000). *Voorstelling en validering van de Morele Oordeel Test (MOT), de Nederlandstalige versie van de Moral Judgment Test (MJT) [Proposal and validation of the Moral Competence Test (MCT) in Dutch]*. Leuven, Belgium: KU Leuven Center for Psychology and Religion.
- Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development*, *82*, 405-432. doi:10.1111/j.1467-8624.2010.01564.x
- Dutch Institute for Curriculum Development. (2012). *Handreiking schoolexamen Nederlands havo/vwo. [Guide school exam Dutch higher general/pre-university education.]* Enschede, Netherlands: SLO. Retrieved from <http://downloads.slo.nl/>
- Dutch Institute for Curriculum Development. (2015). *Curriculumspiegel Deel B: Vakspecifieke Trendanalyse [Curricular Description Part B: Domain-specific trend analysis]*. Enschede, Netherlands: SLO. Retrieved from <http://downloads.slo.nl/>
- Earthman, E.A. (1992). Creating the virtual work: Readers' processes in understanding literary texts. *Research in the Teaching of English*, *26*(4), 351-384.
- Eccleston, P., Werneke, U., Armon, K., Stephenson, T., & MacFaul, R. (2001). Accounting for overlap? An application of mezzich's  $\kappa$  statistic to test interrater reliability of interview data on parental accident and emergency attendance. *Journal of Advanced Nursing*, *33*, 784-790. doi:10.1046/j.1365-2648.2001.01718.x
- Eeds, M., & Wells, D. (1989). Grand conversations: an exploration of meaning construction in literature study groups. *Research in the Teaching of English*, *23*(1), 4-29.

- Elias, M. J., Zins, J. E., Weissberg, R. P., Frey, K. S., Greenberg, M. T., Haynes, N. M., . . . Shriver, T. P. (1997). *Promoting social and emotional learning: Guidelines for educators*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Ellis, J. M. (1974). *The theory of literary criticism: A logical analysis*. Berkeley, CA: University of California Press.
- Engberg, M. E. (2004). Improving intergroup relations in higher education: A critical examination of the influence of educational interventions on racial bias. *Review of Educational Research, 74*, 473-524. doi:10.3102/00346543074004473
- Eva-Wood, A. L. (2004). How think-and-feel-aloud instruction influences poetry readers. *Discourse Processes, 38*, 173-192. doi:10.1207/s15326950dp3802\_2
- Faust, M. (2000). Reconstructing familiar metaphors: John Dewey and Louise Rosenblatt on literary art as experience. *Research in the Teaching of English, 35*(1), 9-34.
- Fialho, O. (2012). *Self-modifying experiences in literary reading: A model for reader response* (Doctoral dissertation). Retrieved from <http://era.library.ualberta.ca>
- Fialho, O. (2018). *Deepening readers' perceptions of self and others: The role of enactment-imagery, resonance and sympathy*. Paper presented at the IGEL Conference, Stavanger, Norway, July 27, 2018.
- Fialho, O., Hakemulder, F., & Hoeken, H. (2018). *Transformative reading: An empirical model*. Manuscript in preparation.
- Fialho, O., Miall, D. S., & Zyngier, S. (2012). Experiencing or interpreting literature: wording instructions. In M. Burke, S. Csabi, L. Week, & J. Zerkowitz (Eds.), *Pedagogical Stylistics: Current Trends in Language, Literature and ELT* (pp. 58-74). London, UK / New York, NY: Continuum.
- Fialho, O., Zyngier, S., & Burke, M. (2016). Empirical approaches to the study of literature in learning environments. In M. Burke, O. Fialho, & S. Zyngier (Eds.), *Scientific approaches to literature in learning environments* (pp. 1-16). Amsterdam, Netherlands: John Benjamins.
- Fialho, O., Zyngier, S., & Miall, D. S. (2011). Interpretation and experience: Two pedagogical interventions observed. *English in Education, 45*, 236-253. doi:10.1111/j.1754-8845.2011.01103.x
- Fredricks, L. (2012). The benefits and challenges of culturally responsive EFL critical literature circles. *Journal of Adolescent & Adult Literacy, 55*, 494-504. doi:10.1002/jaal.00059
- Frijters, S., Ten Dam, G., & Rijlaarsdam, G. (2008). Effects of dialogic learning on value-loaded critical thinking. *Learning and Instruction, 18*, 66-82. doi:10.1016/j.learninstruc.2006.11.001
- Galda, L., & Beach, R. (2001). Response to literature as a cultural activity. *Reading Research Quarterly, 36*, 64-73. doi:10.1598/rrq.36.1.4
- Garrod, A. (1989). Promoting moral development through a high school English curriculum. *Alberta Journal of Educational Research, 35*(1), 61-79.
- Gersten, R., Fuchs, L. S., Compton, D., Coyne, M., Greenwood, C., & Innocenti, M. S. (2005). Quality indicators for group experimental and quasi-experimental research in special education. *Exceptional Children, 71*, 149-164. doi:10.1177/001440290507100202
- Goldman, S. R., McCarthy, K. S., & Burkett, C. (2015). Interpretive inferences in literature. In E. J. O'Brien, A. E. Cook, & R. F. Loch, Jr. (Eds.), *Inferences during reading* (pp. 386-415). Cambridge, MA: Cambridge University Press. doi:10.1017/cbo9781107279186.018
- Graff, G. (2007). *Professing literature: An institutional history*. (20th anniversary ed.). Chicago, IL: University of Chicago Press. doi:10.7208/chicago/9780226305257.001.0001
- Green, J. (2016, April 12). *On the banning of Looking for Alaska* [online video]. Retrieved from <https://www.youtube.com/watch?v=69rd-7vEF3s>
- Green, M. C., & Brock, T. C. (2000). The role of transportation in the persuasiveness of public narratives. *Journal of Personality and Social Psychology, 79*, 701-721. doi:10.1037//0022-3514.79.5.701

- Guthrie, J., & Wigfield, A. (2000). Engagement and Motivation in Reading. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research, volume III*. Mahwah, NJ: Lawrence Erlbaum.
- Hakemulder, F. (2000). *The moral laboratory: Experiments examining the effects of reading literature on social perception and moral self-knowledge*. Amsterdam: John Benjamins. doi:10.1075/upal.34
- Hakemulder, F. (2008). Imagining what could happen: Effects of taking the role of a character on social cognition. In S. Zyngier, M. Bortolussi, A. Chesnokova, & J. Auracher (Eds.), *Directions in empirical literary studies: In honor of Willie van Peer* (pp. 139-160). Amsterdam, Netherlands: John Benjamins.
- Hakemulder, F., Fialho, O., & Bal, P. M. (2016). Learning from literature: Empirical research on readers in schools and at the workplace. In M. Burke, O. Fialho, & S. Zyngier (Eds.), *Scientific approaches to literature in learning environments* (pp. 19-38). Amsterdam, Netherlands: John Benjamins.
- Halász, L. (1991). Emotional effect and reminding in literary processing. *Poetics*, 20, 247-272. doi:10.1016/0304-422x(91)90009-e
- Halstead, J. M., & Taylor, M. J. (Eds.). (1996). *Values in education and education in values*. London, UK: Falmer.
- Hanauer, D. (1999). Attention and literary education: A model of literary knowledge development. *Language Awareness*, 8, 15-29. doi:10.1080/09658419908667114
- Hart, D., Atkins, R., & Ford, D. (1998). Urban America as a context for the development of moral identity in adolescence. *Journal of Social Issues*, 54, 513-530. doi:10.1111/0022-4537.801998080
- Hebert, M., Simpson, A., & Graham, S. (2013). Comparing effects of different writing activities on reading comprehension: A meta-analysis. *Reading and Writing*, 26, 111-138. doi:10.1007/s11145-012-9386-3
- Henschel, S., Meier, C., & Roick, T. (2016). Effects of two types of task instructions on literary text comprehension and motivational and affective factors. *Learning and Instruction*, 44, 11-21. doi:10.1016/j.learninstruc.2016.02.005
- Higham, R. J. E., Brindley, S., & Van de Pol, J. (2014). Shifting the primary focus: Assessing the case for dialogic education in secondary classrooms. *Language and Education*, 28, 86-99. doi:10.1080/09500782.2013.771655
- Hillocks, G. (2011). Commentary on "Research in secondary English, 1912-2011: Historical continuities and discontinuities in the NCTE imprint." *Research in the Teaching of English*, 46(2), 187-192.
- Holland, N. N. (1975) *Five readers reading*. New Haven, CT: Yale University Press.
- Holland, N. N., & Schwartz, M. (1975). The Delphi seminar. *College English*, 36, 789-800. doi:10.2307/375176
- Iser, W. (1974). *The implied reader. Patterns of communication in prose fiction from Bunyan to Beckett*. Baltimore, MD: Johns Hopkins University Press.
- Iser, W. (1980). Interaction between text and reader. In S. R. Suleiman & I. Crosman (Eds.), *The reader in the text* (pp. 106-120). Princeton, NJ: Princeton University Press.
- Ives, D. (2012). Kristina's "ghetto family": Tensions and possibilities at the intersection of teacher and student literacy agendas. *Research in the Teaching of English*, 47(1), 39-63.
- Ivey, G., & Johnston, P. H. (2015). Engaged reading as a collaborative transformative practice. *Journal of Literacy Research*, 47, 297-327. doi:10.1177/1086296x15619731
- Izawa, M. R., French, M. D., & Hedge, A. (2011). Shining new light on the Hawthorne illumination experiments. *Human Factors*, 53, 528-547. doi:10.1177/0018720811417968
- Jauss, H. R. (1982). *Toward an aesthetic of reception* (T. Bahti, Trans.). Minneapolis, MN: University of Minnesota Press.

- Janssen, T. (1998). *Literatuuronderwijs bij benadering [Approaches to literature education]* (Doctoral dissertation). Retrieved from <http://dare.uva.nl/home>
- Janssen, T., & Braaksma, M. (2016). Effects of creative writing on adolescent students' literary response. In M. Burke, O. Fialho, & S. Zyngier (Eds.), *Scientific approaches to literature in learning environments* (pp. 193-211). Amsterdam, Netherlands: John Benjamins
- Janssen, T., Braaksma, M., & Couzijn, M. (2009). Self-questioning in the literature classroom: Effects on students' interpretation and appreciation of short stories. *L1 – Educational Studies in Language and Literature, 9*, 91-116. doi:10.17239/L1esll-2009.09.01.05
- Janssen, T., Braaksma, M., Rijlaarsdam, G., & Van den Bergh, H. (2012). Flexibility in reading literature: Differences between good and poor adolescent readers. *Scientific Study of Literature, 2*, 83-107. doi:10.1075/ssol.2.1.05jan
- Janssen, T., & Pieper, I. (2009). Empirical studies on verbal interaction and literary understanding: An annotated list of references. *L1 – Educational Studies in Language and Literature, 9*, 117-137. doi:10.17239/L1ESLL-2009.09.01.06
- Janssen, T., & Rijlaarsdam, G. (1996). Students as self-assessors: learning experiences of literature teaching in secondary schools. In E. Marum (Ed.), *Children and Books in the Modern World: Contemporary Perspectives on Literacy* (pp. 98-115). London, UK / Washington D.C.: Falmer. doi:10.4324/9780203046487
- Janssen, T., & Rijlaarsdam, G. (2007). Describing the Dutch literature curriculum: A theoretical and empirical approach. In W. Martyniuk (Ed.), *Towards a Common Framework of Reference for Languages of School Education? Proceedings of a Conference* (pp. 205-226). Kraków, Poland: Universitas.
- Johnson, D. R. (2013). Transportation into literary fiction reduces prejudice against and increases empathy for Arab-Muslims. *Scientific Study of Literature, 3*, 77-92. doi:10.1075/ssol.3.1.08joh
- Joosen, V., & Vloeberghs, K. (2008). *Uitgelezen jeugdliteratuur. Ontmoetingen tussen traditie en vernieuwing. [Exquisite children's literature. Encounters of tradition and renewal.]* Zoetermeer, Netherlands / Leuven, Belgium: NDB Biblion / LannooCampus.
- Karweit, N. (1984). Time-on-task reconsidered. Synthesis of research on time and learning. *Educational Leadership, 41*(8), 32-35.
- Kauchak, D., & Eggen P. (1993). *Learning and Teaching*. Boston, MA: Allyn and Bacon.
- Kaufman, G. F., & Libby, L. K. (2012). Changing beliefs and behavior through experience-taking. *Journal of Personality and Social Psychology, 103*, 1-19. doi:10.1037/a0027525
- Kaywell, J. F. (1993). *Adolescents at risk: A guide to fiction and nonfiction for young adults, parents, and professionals*. Westport, CT: Greenwood.
- Keen, S. (2006). A theory of narrative empathy. *Narrative, 14*, 207-236. doi:10.1353/nar.2006.0015
- Keen, S. (2007). *Empathy and the novel*. Oxford, UK: Oxford University Press. doi:10.1093/acprof:oso/9780195175769.001.0001
- Keen, S. (2013): Narrative empathy. In P. Hühn, et al. (Eds.): *The Living Handbook of Narratology*. Hamburg, Germany: Hamburg University Press. Retrieved from <http://hup.sub.uni-hamburg.de/lhn/index.php?title=Narrative+Empathy&oldid=2044>
- Kidd, D. C., & Castano, E. (2013). Reading literary fiction improves Theory of Mind. *Science, 342*, 377-380. doi:10.1126/science.1239918
- Klein, P. D., Boscolo, P., Kirkpatrick, L. C., & Gelati, C. (Eds.) (2014). *Writing as a learning activity*. Leiden, Netherlands: Brill.
- Koek, M., Janssen, T., Hakemulder, F., & Rijlaarsdam, G. (2017). *Critical thinking in the literature classroom*. Paper presented at the ARLE Conference, Tallinn, Estonia, June 15, 2017.
- Koek, M., Janssen, T., Hakemulder, F., & Rijlaarsdam, G. (2016). Literary reading and critical thinking. *Scientific Study of Literature, 6*, 243-277. doi:10.1075/ssol.6.2.04koe

- Kohlberg, L., & Gilligan, C. (1971). The adolescent as a philosopher: The discovery of the self in a postconventional world. *Daedalus*, *100*, 1051-1086.
- Koopman, E. M. (2016). *Reading suffering: An empirical inquiry into empathic and reflective responses to literary narratives* (Doctoral dissertation). Rotterdam, Netherlands: Erasmus Research Center for Media, Communication and Culture.
- Koopman, E. M., & Hakemulder, F. (2015). Effects of literature on empathy and selfreflection: A theoretical-empirical framework. *Journal of Literary Theory*, *9*, 79-111. doi:10.1515/jlt-2015-0005
- Kooy, M. (2006). The telling stories of novice teachers: Constructing teacher knowledge in book clubs. *Teaching and Teacher Education*, *22*, 661-674. doi:10.1016/j.tate.2006.03.010
- Kuiken, D., Miall, D. S., & Sikora, S. (2004). Forms of self-implication in literary reading. *Poetics Today*, *25*, 171-203. doi:10.1215/03335372-25-2-171
- Kuiken, D., Phillips, L., Gregus, M., Miall, D. S., Verbitsky, M., & Tonkonogy, A. (2004). Locating self-modifying feelings within literary reading. *Discourse Processes*, *38*, 267-286. doi:10.1207/s15326950dp3802\_6
- Lenters, K. (2006). Resistance, struggle, and the adolescent reader. *Journal of Adolescent & Adult Literacy*, *50*, 136-146. doi:10.1598/jaal.50.2.6
- Levine, S. (2014). Making interpretation visible with an affect-based strategy. *Reading Research Quarterly*, *49*, 283-303. doi:10.1002/rrq.71
- Levine, S., & Horton, W.S. (2013). Using affective appraisal to help readers construct literary interpretations. *Scientific Study of Literature*, *3*, 105-136. doi:10.1075/ssol.3.1.10lev
- Lewis, C. (1997). The social drama of literature discussions in a fifth/sixth-grade classroom. *Research in the Teaching of English*, *31*(2), 163-204.
- Lind, G. (2016). *How to teach morality. Promoting deliberation and discussion, reducing violence and deceit*. Berlin, Germany: Logos.
- Liu, A., & Want, S. (2015). Literary fiction did not improve affective ToM. Retrieved from <http://PsychFileDrawer.org/replication.php?attempt=Mj11>
- Lysaker, J. T., Tonge, C., Gauson, D., & Miller, A. (2011). Reading and social imagination: What relationally oriented reading instruction can do for children. *Reading Psychology*, *32*(6), 520-566.
- Malo-Juvera, V. (2014). *Speak*: The effect of literary instruction on adolescents' rape myth acceptance. *Research in the Teaching of English*, *48*(4), 407-427.
- Malo-Juvera, V. (2016). The effect of an LGBTQ themed literary instructional unit on adolescents' homophobia. *Study and Scrutiny: Research on Young Adult Literature*, *2*, 1-34. doi:10.15763/jissn.2376-5275.2016.2.1.1-34
- Mar, R. A., & Oatley, K. (2008). The function of fiction is the abstraction and simulation of social experience. *Perspectives on Psychological Science*, *3*, 173-192. doi:10.1111/j.1745-6924.2008.00073.x
- Mar, R. A., Oatley, K., Hirsh, J., dela Paz, J., & Peterson, J. B. (2006). Bookworms versus nerds: Exposure to fiction versus non-fiction, divergent associations with social ability, and the simulation of fictional social worlds. *Journal of Research in Personality*, *40*, 694-712. doi:10.1016/j.jrp.2005.08.002
- Markus, H., & Nurius, P. (1986). Possible selves. *American Psychologist*, *41*, 954-969. doi:10.1037//0003-066x.41.9.954
- McCarthy, K. S. & Goldman, S. R. (2017). Constructing interpretive inferences about literary text: The role of domain-specific knowledge. *Learning and Instruction*. Advance online publication. doi:10.1016/j.learninstruc.2017.12.004
- McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods*, *1*, 30-46. doi:10.1037/1082-989X.1.1.30

- McKenney, S., Nieveen, N., & van den Akker, J. (2006). Design research from a curriculum perspective. In: J. van den Akker, K. Gravemeijer, S. McKenney, & N. Nieveen (Eds.), *Educational design research* (pp. 62-90). London, UK: Routledge.
- McKeown, M. G., Beck, I. L., & Blake, R. G. (2009). Rethinking reading comprehension instruction: A comparison of instruction for strategies and content approaches. *Reading Research Quarterly*, 44, 218-253. doi:10.1598/rrq.44.3.1
- McMaster, K. L., Van den Broek, P., Espin, C. A., White, M. J., Rapp, D. N., Kendeou, P., ..., Carlson, S. (2012). Making the right connections: Differential effects of reading intervention for subgroups of comprehenders. *Learning and Individual Differences*, 22, 100-111. doi:10.1016/j.lindif.2011.11.017
- Mercer, N., & Dawes, L. (2008). The value of exploratory talk. In N. Mercer & S. Hodgkinson (Eds.), *Exploring talk in school* (pp. 55-71). London, UK: Sage.
- Merrill, M. D. (2002). First principles of instruction. *Educational Technology Research and Development*, 50, 43-59. doi:10.1007/bf02505024
- Miall, D. S., & Kuiken, D. (1994). Foregrounding, defamiliarization, and affect: Response to literary stories. *Poetics*, 22, 389-407. doi:10.1016/0304-422x(94)00011-5
- Miall, D. S., & Kuiken, D. (1995). Aspects of literary response: A new questionnaire. *Research in the Teaching of English*, 29(1), 37-58.
- Miall, D. S., & Kuiken, D. (1998). The form of reading: Empirical studies of literariness. *Poetics*, 25, 327-341. doi:10.1016/s0304-422x(98)90003-1
- Miall, D. S., & Kuiken, D. (1999). What is literariness? Three components of literary reading. *Discourse Processes*, 28, 121-138. doi:10.1080/01638539909545076
- Miall, D. S., & Kuiken, D. (2002). A feeling for fiction: Becoming what we behold. *Poetics*, 30, 221-241. doi:10.1016/s0304-422x(02)00011-6
- Moely, B. E., Mercer, S. H., Ilustre, V., Miron, D., & McFarland, M. (2002). Psychometric properties and correlates of the Civic Attitudes and Skills Questionnaire (CASQ). *Michigan Journal of Community Service Learning*, 8(2), 15-22.
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & The PRISMA Group (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Med* 8(7): e1000097. doi:10.1371/journal.pmed1000097
- Mokhtari, K., & Reichard, C. A. (2002). Assessing students' metacognitive awareness of reading strategies. *Journal of Educational Psychology*, 94, 249-259. doi:10.1037//0022-0663.94.2.249
- Mol, S. E., & Bus, A. G. (2011). To read or not to read: A meta-analysis of print exposure from infancy to early adulthood. *Psychological Bulletin*, 137(2), 267-296. doi:10.1037/a0021890
- Muijs, D., & Reynolds, D. (2010). *Effective teaching: Research and practice*. London, UK: Paul Chapman.
- Mukařovský, J. (1976). *On poetic language* (J. Burbank & P. Steiner, Trans.). Lisse, Netherlands: Peter de Ridder.
- Murphy, P. K., Wilkinson, I., Soter, A., Hennessey, M. N., & Alexander, J. F. (2009). Examining the effects of classroom discussion on students' comprehension of text: A meta-analysis. *Journal of Educational Psychology*, 101, 740-764. doi:10.1037/a0015576
- National Council of Teachers of English. (2003-2014). *Annotated bibliography of research in the teaching of English*. Retrieved from <http://www.ncte.org/journals/rte/biblios>
- National Council of Teachers of English. (2012). *NCTE/IRA standards for the English language arts*. Retrieved from <http://www.ncte.org/standards/ncte-ira>
- Newell, G. E. (1996). Reader-based and teacher-centered instructional asks: Writing and learning about a short story in middle-track classrooms. *Journal of Literacy Research*, 28, 147-172. doi:10.1080/10862969609547914

- Nieveen, N. (1999). Prototyping to reach product quality. In: J. van den Akker, R. Branch, K. Gustafson, N. Nieveen, & T. Plomp (Eds.), *Design approaches and tools in education and training* (pp. 125-136). Dordrecht, Netherlands: Kluwer.
- Nucci, L., Krettenauer, T., & Narváez, D. (Eds.). (2014). *Handbook of moral and character education*. New York, NY: Routledge.
- Nussbaum, M. (1995). *Poetic justice: The literary imagination and public life*. Boston, MA: Beacon.
- Nussbaum, M. (2010). *Not for profit: Why democracy needs the humanities*. New Jersey, NJ: Princeton University Press.
- Nystrand, M. (1997). *Opening dialogue: Understanding the dynamics of language and learning in the English classroom*. New York, NY: Teachers College.
- Nystrand, M. (2006). Research on the role of classroom discourse as it affects reading comprehension. *Research in the Teaching of English, 40*(4), 392-412.
- Oatley, K. (1999). Meetings of minds: Dialogue, sympathy, and identification, in reading fiction. *Poetics, 26*, 439-454. doi:10.1016/s0304-422x(99)00011-x
- Oberon (2016). *Lees- en literatuuronderwijs in havo/vwo [Reading and literature teaching in higher general and pre-university secondary education]*. Amsterdam, Netherlands: Stichting Lezen.
- O'Donnell, C. L. (2008). Defining, conceptualizing, and measuring fidelity of implementation and its relationship to outcomes in K-12 curriculum intervention research. *Review of Educational Research, 78*, 33-84. doi:10.3102/0034654307313793
- OECD (2010). *PISA 2009 results: Learning to learn – Student engagement, strategies and practices* (Volume III). Retrieved from <http://www.oecd.org/pisa/pisaproducts/48852630.pdf>
- Oliver, M. B. & Raney, A. A. (2011). Entertainment as pleasurable and meaningful: Identifying hedonic and eudaimonic motivations for entertainment consumption. *Journal of Communication, 61*, 984-1004. doi:10.1111/j.1460-2466.2011.01585.x
- Palmer, A. (2004). *Fictional Minds*. London, UK / Lincoln, NE: University of Nebraska Press.
- Peskin, J. (1998). Constructing meaning when reading poetry: An expert-novice study. *Cognition and Instruction, 16*, 235-263. doi:10.1207/s1532690xci1603\_1
- Peskin, J. (2007). The genre of poetry: Secondary school students' conventional expectations and interpretive operations. *English in Education, 41*, 20-36. doi:10.1111/j.1754-8845.2007.tb01162.x
- Pieper, I., & Wieser, D. (2012). Understanding metaphors in poetic texts: Towards a determination of interpretative operations in secondary school students' engagement with imagery. *L1 – Educational Studies in Language and Literature, 12*, 59-84. doi:10.17239/L1ESLL-2012.01.03
- Plomp, T. (2013). Educational design research: An introduction. In T. Plomp & N. Nieveen (Eds.), *Educational design research, part A: An introduction* (pp. 10-51). Enschede, Netherlands: SLO.
- Pressley, M., Wood, E., Woloshyn, V. E., Martin, V., King, A., & Menke, D. (1992). Encouraging mindful use of prior knowledge: Attempting to construct explanatory answers facilitates learning. *Educational Psychologist, 27*, 91-109. doi:10.1207/s15326985ep2701\_7
- Probst, R. E. (1988). Dialogue with a text. *The English Journal, 77*, 32-38. doi:10.2307/818021
- Pruis, M., & De Vries, J. (2016). Wat moeten scholieren lezen en waarom? [What do students have to read and why?]. *De Groene Amsterdammer, 140*(16). Retrieved from <https://www.groene.nl/artikel/wat-moeten-scholieren-lezen-en-waarom>
- Pyle, D., Pyle, N., Lignugaris/Kraft, B., Duran, L., & Akers, J. (2017). Academic effects of peer-mediated interventions with English language learners: A research synthesis. *Review of Educational Research, 87*, 103-133. doi:10.3102/0034654316653663
- Rain, M., & Mar, R. A. (2014). Measuring reading behavior: Examining the predictive validity of print-exposure checklists. *Empirical Studies of the Arts, 32*, 93-108. doi:10.2190/em.32.1f

- Reigeluth, C. M. (1999). What is instructional design theory and how is it changing? In C. M. Reigeluth (Ed.), *Instructional design theories and models: A new paradigm of instructional theory* (Vol. 2, pp. 5-29). Mahwah, NJ: Lawrence Erlbaum.
- Renshaw P. D. (2004) Dialogic learning, teaching and instruction: theoretical roots and analytical frameworks. In J. van der Linden, & P. Renshaw (Eds.), *Dialogic learning* (pp. 1-15). Dordrecht, Netherlands: Springer. doi:10.1007/1-4020-1931-9\_1
- Repko, A. F. (2008). *Interdisciplinary research: Process and theory*. Thousand Oaks, CA: Sage.
- Richardson, P. W., & Eccles, J. S. (2007). Rewards of reading: Toward the development of possible selves and identities. *International Journal of Educational Research*, 46, 341-356. doi:10.1016/j.ijer.2007.06.002
- Rietdijk, S., Janssen, T., Van Weijen, D., Van den Bergh, H., & Rijlaarsdam, G. (2017). Improving writing in primary schools through a comprehensive writing program. *The Journal of Writing Research*, 9, 173-225. doi:10.17239/jowr-2017.09.02.04
- Rijlaarsdam, G., Janssen, T., Rietdijk, S., & Van Weijen, D. (2017). Reporting design principles for effective instruction of writing: Interventions as constructs. In R. Fidalgo & T. Olive (Series Eds.), & R. Fidalgo, K. Harris, & M. Braaksma (Vol. Eds.), *Studies in writing series: Vol. 34. Design principles for teaching effective writing* (pp. 280–313). Leiden, Netherlands: Brill.
- Rosenblatt, L. M. (1968). *Literature as exploration*. Champaign, IL: NCTE. (Original work published 1938)
- Rosenblatt, L. M. (1983). *Literature as exploration*. New York, NY: The Modern Language Association of America. (Original work published 1938)
- Rosenblatt, L. M. (1994). *The reader, the text, the poem*. Carbondale, IL: Southern Illinois University Press. (Original work published 1978)
- Ross, C. S., McKechnie, L., & Rothbauer, P. M. (2006). *Reading matters: What the research reveals about reading, libraries, and community*. Westport: Libraries Unlimited.
- Rothbauer, P. M. (2011). Rural teens on the role of reading in their lives. *Journal of Research on Libraries and Young Adults*, 1(2). Retrieved from [www.yalsa.ala.org/jrlya/2011/02/rural-teens-on-the-role-of-reading-in-their-lives](http://www.yalsa.ala.org/jrlya/2011/02/rural-teens-on-the-role-of-reading-in-their-lives)
- Runge, L. L. (2012). *Annotated bibliography on literature pedagogy*. Retrieved from [http://chuma.cas.usf.edu/~runge/Class\\_Bib.htm](http://chuma.cas.usf.edu/~runge/Class_Bib.htm)
- Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional science*, 18, 119-144. doi:10.1007/bf00117714
- Salomon, G., & Perkins, D. N. (1998). Individual and social aspects of learning. In P. D. Pearson, & A. Iran-Nejad (Eds.), *Review of research in education* (pp. 1-24). Washington, DC: American Educational Research Association.
- Scardamalia, M., & Bereiter, C. (1987). Knowledge telling and knowledge transforming in written composition. In S. Rosenberg (Red.), *Advances in applied psycholinguistics*, Vol. 2 (pp. 142-175). Cambridge, NY: Cambridge University Press.
- Schmidt, S. J. (1982). Die empirische Literaturwissenschaft: Ein neues Paradigma [The empirical study of literature: A new paradigm]. *SPIEL*, 1, 5-25.
- Schrijvers, M., Janssen, T., Fialho, O., & Rijlaarsdam, G. (2016). The impact of literature education on students' perceptions of self and others: Exploring personal and social learning experiences in relation to teacher approach. *L1 – Educational Studies in Language and Literature*, 17, 1-37. doi:10.17239/L1ESLL-2016.16.04.01
- Schrijvers, M., Janssen, T., Fialho, O., & Rijlaarsdam, G. (2018). Gaining insight into human nature: A review of literature classroom intervention studies. *Review of Educational Research*. Advance online publication. doi:10.3102/0034654318812914

- Selman, R. L. (1975). Level of social perspective taking and the development of empathy in children: Speculations from a social-cognitive viewpoint. *Journal of Moral Education, 5*, 35-43. doi:10.1080/0305724750050105
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. (2<sup>nd</sup> Ed.) Boston, MA: Houghton, Mifflin and Company.
- Shen, Y. (2008). Two levels of foregrounding in literary narratives. In S. Zyngier, M. Bortolussi, A. Chesnokova, & J. Auracher (Eds.), *Directions in empirical literary studies: In honor of Willie van Peer* (pp. 103-111). Amsterdam, Netherlands: John Benjamins.
- Sigvardsson, A. (2016). Teaching poetry reading in secondary education: Findings from a systematic literature review. *Scandinavian Journal of Educational Research, 61*, 1-16. doi:10.1080/00313831.2016.1172503
- Sikora, S., Kuiken, D., & Miall, D.S. (2010). An uncommon resonance: The influence of loss on expressive reading. *Empirical Studies of the Arts, 28*, 135-153. doi:10.2190/em.28.2.b
- Sikora, S., Kuiken, D., & Miall, D. S. (2011). Expressive reading: A phenomenological study of readers' experience of Coleridge's *The rime of the ancient mariner*. *Psychology of Aesthetics, Creativity, and the Arts, 5*, 258-268. doi:10.1037/a0021999
- Slone, M., Tarrasch, R., & Hallis, D. (2000). Ethnic stereotypic attitudes among Israeli children: Two intervention programs. *Merrill-Palmer Quarterly, 46*(2), 370-389.
- Smith, M.W., & Wilhelm, J.D. (2002). *"Reading don't fix no chevys": Literacy in the lives of young men*. Portsmouth, NH: Heinemann.
- Soter, A., Wilkinson, I., Connors, S., Murphy, P., & Shen, V. (2010). Deconstructing "aesthetic response" in small-group discussions about literature: A possible solution to the "aesthetic response" dilemma. *English Education, 42*(2), 204-225.
- Soter, A., Wilkinson, I., Murphy, P. K., Rudge, L., Reninger, K., & Edwards, M. (2008). What the discourse tells us: Talk and indicators of high-level comprehension. *International Journal of Educational Research, 47*, 372-391. doi:10.1016/j.ijer.2009.01.001
- Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21*, 360-407. doi:10.1598/RRQ.21.4.1
- Stanovich, K. E., & West, R. F. (1989). Exposure to print and orthographic processing. *Reading Research Quarterly, 24*(4), 402-433.
- Steen, G., & Schram, D. (2001). The empirical study of literature: Psychology, sociology and other disciplines. In D. Schram & G. Steen (Eds.), *The psychology and sociology of literature* (pp. 1-16). Amsterdam, Netherlands: John Benjamins.
- Stevahn, L., Johnson, D. W., Johnson, R. T., Green, K., & Laginski, A. M. (1997). Effects on high school students of conflict resolution training integrated into English literature. *Journal of Social Psychology, 137*, 302-315. doi:10.1080/00224549709595442
- Stevahn, L., Johnson, D. W., Johnson, R. T., & Real, D. (1996). The impact of a cooperative or individualistic context on the effectiveness of conflict resolution training. *American Educational Research Journal, 33*, 801-823. doi:10.3102/00028312033004801
- Stokmans, M. (2009). De invloed van literatuuronderwijs op leesattitude [The influence of literature education on reading attitude]. *Levende Talen Tijdschrift, 10*(2), 35-43.
- Tengberg, M., Olin-Scheller, C., & Lindholm, A. (2015). Improving students' narrative comprehension through a multiple strategy approach. Effects of dialogic strategy instruction in secondary school. *L1 – Educational Studies in Language and Literature, 15*, 1-25. doi:10.17239/L1ESLL-2015.15.01.01
- Turner, R. N., Hewstone, M., Voci, A., Paolini, S., & Christ, O. (2007). Reducing prejudice via direct and extended crossgroup friendship. *European Review of Social Psychology, 18*, 212-255. doi:10.1080/10463280701680297

- Turner, R. N., Hewstone, M., Voci, A., & Vonofakou, C. (2008). A test of the extended contact hypothesis: The mediating role of intergroup anxiety, perceived ingroup and outgroup norms, and inclusion of the outgroup in the self. *Journal of Personality and Social Psychology, 95*, 843-860. doi:10.1037/a0011434
- UK Literary Association. (2016). *Curriculum and assessment in English 3 to 19: A better plan. The essentials of English*. Retrieved from <https://ukla.org/resources/details/curriculum-and-assessment-in-english-3-to-19-a-better-plan>
- Van de Ven, P. H., & Doecke, B. (Eds.). (2011). *Literary praxis: A conversational inquiry into the teaching of literature*. Rotterdam, Netherlands: Sense.
- Van den Akker, J. (1999). Principles and methods of development research. In J. van den Akker, R. M. Branch, K. Gustafson, N. Nieveen, & T. Plomp (Eds.), *Design approaches and tools in education and training* (pp. 1-14). Dordrecht, Netherlands: Springer.
- Van den Akker, J. (2013). Curricular development research as a specimen of educational design research. In T. Plomp & N. Nieveen (Eds.), *Educational design research, part A: An introduction* (pp. 52-71). Enschede, Netherlands: SLO.
- Van den Akker, J., Gravemeijer, K., McKenney, S., & Nieveen, N. (Eds.). (2006). *Educational design research*. New York, NY: Routledge.
- Van den Broeck, A., Vansteenkiste, M., De Witte, H., Soenens, B., & Lens, W. (2010). Capturing autonomy, competence, and relatedness at work: Construction and initial validation of the work-related basic need satisfaction scale. *Journal of Occupational and Organizational Psychology, 83*, 981-1002. doi:10.1348/096317909X481382
- Van den Hoven, P., & Van Lierop-Debrauwer, H. (2014). Zoektochten zonder wegwijzers. Adolescentenliteratuur [Quests without signposts. Adolescent literature]. In: R. Ghesquière, V. Joosen, & H. Van Lierop-Debrauwer (Eds.), *Een land van waan en wijs. Geschiedenis van de Nederlandse jeugdliteratuur [A land of illusion and wisdom. History of Dutch children's literature]* (pp. 374-404). Amsterdam, Netherlands: Atlas Contact.
- Van der Bolt, L. (2000). *Ontroerend goed: Een onderzoek naar affectieve leeservaringen van leerlingen in het basis- en voortgezet onderwijs [Touching quality: A study into affective reading experiences in elementary and secondary school students]* (Doctoral dissertation). Retrieved from <http://dare.uva.nl/home>
- Van der Borden, J., Van Dam-Helmig, M., Kniep, J., De Puit, N., & Stienen, M. (2017). *Kennisbasis tweedegraadsopleiding Nederlands [Knowledge base second degree teacher training program Dutch]*. Den Haag, Netherlands: 10voordeleraar, Vereniging van Hogescholen.
- Van der Kamp, M. (1980). *Wat neemt de leerling mee van kunstzinnige vorming? [What does the student take away from arts education?]* (Doctoral dissertation). Retrieved from <http://www.lkca.nl/publicaties/informatiebank>
- Van Grootheest, A., & Van Grinsven, V. (2016). *Rapportage: Lezen op havo/vwo [Report: Reading in higher general/pre-university secondary education]*. Amsterdam, Netherlands: Stichting Lezen.
- Van Kesteren, B. J. (1993). Applications of De Groot's 'Learner report': A tool to identify educational objectives and learning experiences. *Studies in Educational Evaluation, 19*, 65-86. doi:10.1016/s0191-491x(05)80057-4
- Van Peer, W., Hakemulder, J., & Zyngier, S. (2007). Lines on feeling: Foregrounding, aesthetics and meaning. *Language and Literature, 16*, 197-213. doi:10.1177/0963947007075985
- Van Schooten, E. J. (2005). *Literary response and attitude toward reading fiction* (Doctoral dissertation). Retrieved from <http://www.rug.nl/research/portal/publications/search.html>
- Verboord, M. (2005). Long-term effects of literary education on book-reading frequency: An analysis of dutch student cohorts 1975-1998. *Poetics, 33*, 320-342. doi:10.1016/j.poetic.2005.09.007

- Verboord, M. & Van Rees, K. (2003). Do changes in socialization lead to decline in reading level? How parents, literary education, and popular culture affect the level of books read. *Poetics*, *31*, 283-300. doi:10.1016/s0304-422x(03)00036-6
- Vermaas, J. & Van der Linden, R. (2007). *Beter inspelen op havo-leerlingen [Facilitate students' learning in the higher general secondary education track]*. Tilburg, Netherlands: IVA.
- Vezzali, L., Stathi, S., & Giovannini, D. (2012). Indirect contact through book reading: Improving adolescents' attitudes and behavioral intentions toward immigrants. *Psychology in the Schools*, *49*, 148-162. doi:10.1002/pits.20621
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Weber, C. D. (2010). Literary fiction as a tool for teaching social theory and critical consciousness. *Teaching Sociology*, *38*, 350-361. doi:10.1177/0092055x10378816
- Wegerif, R., Mercer, N., & Dawes, L. (1999). From social interaction to individual reasoning: An empirical investigation of a possible sociocultural model of cognitive development. *Learning and Instruction*, *9*, 493-516. doi:10.1016/s0959-4752(99)00013-4
- Wei, L., Murphy, P. K., & Firetto, C. M. (2018). How can teachers facilitate productive small-group talk? An integrated taxonomy of teacher discourse moves. *Elementary School Journal*, *118*, 578-609. doi:10.1086/697531
- Weijts, C. (2016, January 18). Die boekenlijst is misdadig, fuck de canon [That reading list is criminal, fuck the canon]. *NRC*. Retrieved from <https://www.nrc.nl>
- Wells, G. (2000). Dialogic inquiry in education. Building on the legacy of Vygotsky. In C. D. Lee, & P. Smagorinsky (Eds.), *Vygotskian perspectives on literacy research. Constructing meaning through collaborative inquiry* (pp 51-85). Cambridge, UK: Cambridge University Press.
- White, B. F. (1995). Effects of autobiographical writing before reading on students' responses to short stories. *Journal of Educational Research*, *88*, 173-184. doi:10.1080/00220671.1995.9941296
- Wilhelm, J. D. (2016). *You gotta BE the book. Teaching engaged and reflective reading with adolescents* (3<sup>rd</sup> Ed.). New York, NY: Teachers College.
- Witte, T. (2008). *Het oog van de meester. Een onderzoek naar de literaire ontwikkeling van havo-en vwo-leerlingen in de tweede fase van het voortgezet onderwijs. [The eye of the master. A study into literary development of upper higher general education and pre-university students.]* Delft, Netherlands: Eburon.
- Witte, T., Rijlaarsdam, G., & Schram, D. (2012). An empirically grounded theory of literary development: Teachers' pedagogical content knowledge on literary development in upper secondary education. *L1 – Educational Studies in Language and Literature*, *12*, 1-33. doi:10.17239/l1esll-2012.01.01
- Yilmaz, K. (2007). Historical empathy and its implications for classroom practices in schools. *The History Teacher*, *40*(3), 331-337.
- Zahavi, D. (2014). *Self and other: Exploring subjectivity, empathy, and shame*. New York, NY: Oxford University Press.
- Zeitz, C. M. (1994). Expert-novice differences in memory, abstraction, and reasoning in the domain of literature. *Cognition and Instruction*, *12*, 277-312. doi:10.1207/s15322690xci1204\_1
- Zunshine, L. (2006). *Why we read fiction: Theory of Mind and the novel*. Columbus, OH: Ohio State University Press.
- Zunshine, L. (2015). The secret life of fiction. *PMLA*, *130*, 724-731. doi:10.1632/pmla.2015.130.3.724



## AUTHOR INDEX

- Adler, E. S., 55, 57, 58, 59, 60, 61,  
63, 64, 66, 67, 69, 70, 71, 73, 74,  
76, 77, 78, 79, 80, 81, 92, 128, 157,  
185, 267
- Akers, J., 56
- Alexander, J. F., 48, 199
- Alexander, R., 199
- Altman, D. G., 202
- Anderson, R. C., 69, 201, 271
- Andringa, E., 13, 51, 128
- Applebee, A. N., 14, 36, 38, 47, 201,  
213
- Appleman, D., 10, 46
- Appleyard, J. A., 13, 35, 36
- Armon, K., 284
- Atkins, R., 43
- Bakhtin, M., 11, 199, 200
- Bakker, A., 11
- Bakker, A. B., 12
- Bal, M., 2, 6, 10, 11, 12, 44
- Bal, P. M., 2, 6, 10, 12, 45, 80, 81, 90,  
126, 156
- Banks, W. P., 53
- Barnes, D., 82
- Beach, R., 8, 10, 14, 15, 37, 38, 46,  
55, 81, 82
- Bean, T. W., 36
- Beck, I. L., 132
- Benavides Buitrago, C., 53
- Bender-Slack, D., 53
- Bennett, R. E., 214
- Bereiter, C., 205
- Biesta, G., 1
- Binkhorst, F., 214, 308
- Bintz, W. P., 10
- Blackie, M., 53
- Blake, R. G., 132
- Bloemhoff, H., 214
- Bloome, D., 8, 14
- Boerma, I. E., 212
- Boogers, A., 216
- Boostrom, R., 80
- Booth, W. C., 198
- Boscolo, P., 81
- Bourdieu, P., 43, 198
- Braaksma, M., 13, 81, 87, 128, 152,  
159, 201, 213
- Brindley, S., 199
- Brock, T. C., 10, 11, 40
- Brokerhof, I., 6
- Buehl, M. M., 81
- Burbaum, C., 11, 44, 91, 127, 157
- Burke, M., 44
- Burkett, C., 13
- Bus, A. G., 212
- Butterman, O. S., 12
- Campbell, D. T., 52, 56, 131, 152,  
206
- Castano, E., 10, 11, 13, 35, 46
- Charlton, M., 11, 13, 35, 44, 91, 127,  
157
- Chinn, C. A., 201, 202
- Cicchetti, D. V., 57, 60

- Connors, S., 46  
Cook, T. D., 52, 56, 131, 152, 206  
Couzijn, M., 13, 87, 152, 159, 201, 213  
Cummins, A., 53  
Curriculum.nu., 212  
Curwood, J. S., 53
- Darragh, J. J., 55, 58, 59, 60, 64, 66, 68, 82, 267  
Davis, M. H., 66, 167, 270  
Davis, O. L., 36, 270  
Dawes, L., 47, 82, 199  
De Blauw, E., 214  
De Groot, A. D., 16, 135, 215  
De Leon, C., 55  
De Puit, N., 214  
De Vries, J., 216  
Deci, E. L., 159, 192  
dela Paz, J., 17, 44, 135  
Demarez, P., 165  
Derricott, R., 1, 42  
Dijkstra, K., 11, 306  
Dirksen, J., 9, 130  
Dochy, F., 81  
Doecke, B., 14, 15, 37, 213  
Duran, L., 56  
Duriez, B., 165  
Durlak, J. A., 55  
Dymnicki, A. B., 55
- Earthman, E.A., 13, 128  
Eccles, J. S., 2, 12, 13, 36, 44, 91, 127, 151, 157  
Eccleston, P., 284  
Eeds, M., 38, 201  
Elias, M. J., 1, 42  
Ellis, J. M., 43, 198  
Engberg, M. E., 56
- Eva-Wood, A. L., 13, 14, 57, 58, 60, 61, 62, 63, 64, 65, 66, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 80, 81, 83, 128, 136, 157, 158, 159, 179, 185, 186, 268
- Faust, M., 36  
Fialho, O., 4, 2, 5, 6, 7, 8, 10, 11, 12, 15, 19, 23, 41, 44, 45, 46, 48, 53, 80, 81, 88, 90, 91, 120, 126, 127, 128, 129, 136, 150, 156, 157, 159, 162, 179, 180, 186, 189, 194, 195, 196, 203, 204, 205, 213, 297, 298, 299, 301  
Firetto, C. M., 48  
Ford, D., 43  
Foster, P., 55, 57, 58, 59, 60, 61, 63, 64, 66, 67, 69, 70, 71, 73, 74, 76, 77, 78, 79, 80, 81, 92, 128, 157, 185, 267  
Foster, S. J., 36, 267  
Fredricks, L., 53  
French, M. D., 208, 308  
Frijters, S., 200
- Galda, L., 8, 38, 46, 55  
Gamoran, A., 14, 47, 201  
Garrod, A., 57, 58, 59, 60, 61, 62, 64, 66, 68, 82, 268  
Gauson, D., 2, 127  
Gelati, C., 81  
Gersten, R., 56  
Gilligan, C., 43  
Giovannini, D., 57, 92, 128  
Goldman, S. R., 13, 158  
Graff, G., 8  
Graham, S., 56  
Gravemeijer, K., 56  
Green, J., 1, 8, 14, 215

- Green, K., 55  
 Green, M. C., 10, 11, 40  
 Gregus, M., 189  
 Guthrie, J., 132
- Hakemulder, F., 2, 6, 10, 12, 35, 36, 44, 46, 55, 56, 57, 58, 60, 61, 64, 66, 68, 80, 81, 82, 90, 126, 127, 135, 156, 161, 196, 197, 205, 213, 269
- Halász, L., 55, 56, 58, 60, 63, 64, 69, 70, 71, 72, 75, 76, 77, 78, 80, 128, 269
- Hallis, D., 55  
 Halstead, J. M., 42  
 Hanauer, D., 13  
 Handelzalts, A., 214  
 Hart, D., 43  
 Hebert, M., 56  
 Hedge, A., 208  
 Hennessey, M. N., 48, 199  
 Henschel, S., 57, 58, 59, 60, 61, 62, 64, 65, 66, 82, 159, 213, 270  
 Hewstone, M., 79  
 Higham, R. J. E., 199  
 Hillocks, G., 37  
 Hirsh, J., 17, 44, 135  
 Hoeken, H., 196  
 Holland, N. N., 9, 14, 198  
 Horton, W.S., 87, 158, 179  
 Hynds, S., 10, 46
- Iser, W., 9, 80, 161, 197, 198  
 Ives, D., 37  
 Ivey, G., 127, 128, 129, 151  
 Izawa, M. R., 208
- Janssen, T., 4, 3, 6, 7, 8, 9, 13, 14, 15, 16, 19, 36, 37, 41, 45, 46, 81, 87, 88, 90, 92, 122, 128, 129, 135, 152, 156, 157, 159, 179, 201, 206, 213, 297, 298, 299, 301, 308
- Jauss, H. R., 9, 198  
 Johnson, D. R., 12, 13, 36, 44, 272  
 Johnson, D. W., 55, 272  
 Johnson, R. T., 55, 272  
 Johnston, P. H., 127, 128, 129, 151  
 Jolles, J., 212  
 Joosen, V., 215
- Karweit, N., 99, 139, 168  
 Kauchak, D., 107, 115, 144  
 Kaufman, G. F., 44, 127  
 Kaywell, J. F., 70  
 Keen, S., 8, 10, 11, 35, 126, 156  
 Kidd, D. C., 10, 11, 13, 35, 46  
 Kirkpatrick, L. C., 81  
 Klein, P. D., 81  
 Kniep, J., 214  
 Koek, M., 36, 135, 213  
 Kohlberg, L., 43, 68, 269  
 Koopman, E. M., 10, 12, 35, 43, 44, 46, 55, 90, 126, 127, 156, 157, 165, 205  
 Kooy, M., 12  
 Krettenauer, T., 1, 42  
 Kuiken, D., 8, 10, 12, 43, 45, 46, 62, 65, 70, 84, 136, 157, 164, 189, 268
- Laginski, A. M., 55  
 Langer, J. A., 14, 47, 201  
 Lenters, K., 10, 38, 40, 94, 129  
 Levine, S., 87, 158, 179, 213, 301  
 Lewis, C., 201  
 Libby, L. K., 44, 127  
 Liberati, A., 202  
 Lignugaris/Kraft, B., 56  
 Lind, G., 165

- Liu, A. Want, S., 11  
Lysaker, J. T., 2, 127
- MacFaul, R., 284
- Malo-Juvera, V., 37, 55, 57, 58, 59,  
60, 61, 62, 64, 66, 69, 70, 71, 72,  
73, 74, 75, 76, 77, 78, 79, 80, 81,  
83, 87, 92, 128, 157, 158, 185, 186,  
195, 271
- Mar, R. A., 1, 10, 17, 43, 44, 45, 48,  
70, 78, 79, 92, 126, 135, 156, 195,  
206
- Markus, H., 12, 13, 36
- Martin, V., 29, 32, 33, 112, 277
- McCarthy, K. S., 13, 158
- McGraw, K. O., 57
- McKechnie, L., 127
- McKenney, S., 56, 97, 203
- McKeown, M. G., 132
- McMaster, K. L., 158
- Meier, C., 57, 270
- Mercer, N., 47, 82, 199
- Merrill, M. D., 49, 69, 81, 132, 208
- Miall, D. S., 8, 10, 12, 43, 45, 46, 53,  
62, 65, 70, 84, 136, 157, 159, 164,  
189, 205, 268
- Miller, A., 2, 127
- Moely, B. E., 165
- Moher, D., 202
- Mokhtari, K., 159, 165
- Mol, S. E., 212
- Moni, K., 36
- Muijs, D., 107, 115, 144, 171, 203
- Mukařovský, J., 43, 70
- Murphy, P. K., 46, 48, 55, 76, 199,  
201, 301
- Narváez, D., 1, 42
- Newell, G. E., 81
- Nieveen, N., 56, 88, 93, 97, 187, 203
- Nucci, L., 1, 42
- Nuijten, K., 214
- Nurius, P., 12, 13, 36
- Nussbaum, M., 1, 44, 126, 156
- Nystrand, M., 14, 38, 47, 55, 82, 201
- Oatley, K., 1, 10, 17, 43, 44, 45, 48,  
70, 78, 79, 92, 126, 135, 156, 195,  
200
- Oberon, 2, 3, 92, 129, 130, 156
- OECD, 150
- Olin-Scheller, C. L., 14
- Oliver, M. B., 165, 179
- Palmer, A., 8
- Perkins, D. N., 200
- Peskin, J., 13, 158
- Peterson, J. B., 17, 44, 135
- Pette, C., 11, 44, 91, 127, 157
- Phillips, L., 189
- Pieper, I., 4, 13, 122, 299
- Plomp, T., 87, 88, 96
- Poortman, C. L., 214
- Pressley, M., 81
- Probst, R. E., 200
- Pruis, M., 216
- Pyle, D., 56
- Pyle, N., 56
- Rain, M., 17, 135, 206
- Raney, A. A., 165, 179
- Real, D., 55, 176
- Reichard, C. A., 159, 165
- Reigeluth, C. M., 49
- Repko, A. F., 194
- Reynolds, D., 107, 115, 144, 171, 203
- Richardson, P. W., 2, 12, 13, 36, 44,  
91, 127, 151, 157

- Rietdijk, S., 46, 88  
Rijlaarsdam, G., 4, 6, 7, 8, 9, 14, 15, 16, 36, 37, 41, 45, 46, 56, 60, 88, 90, 128, 129, 135, 156, 157, 200, 213, 216, 297, 298, 299, 301  
Roick, T., 57, 270  
Rosenblatt, L. M., 9, 45, 48, 81, 82, 92, 127, 128, 158, 200, 243  
Ross, C. S., 127  
Rothbauer, P. M., 11, 13, 35, 44, 91, 127, 151, 157  
Runge, L. L., 55  
Ryan, R. M., 159, 192  
  
Sadler, D. R., 214  
Salomon, G., 200  
Scardamalia, M., 205  
Schellinger, K. B., 55  
Schmidt, S. J., 198  
Schram, D., 4, 8, 82, 198, 216  
Schrijvers, M., 3, 7, 41, 45, 88, 90, 91, 92, 118, 128, 129, 135, 151, 156, 157, 167, 297, 298, 299, 301  
Schwartz, M., 14  
Segers, M., 81, 299  
Selman, R. L., 44  
Severijnen, O., 214  
Shadish, W. R., 52, 56, 131, 152, 204, 206, 207, 208, 209  
Shen, V., 46, 84  
Shen, Y., 46, 84  
Sigvardsson, A., 55  
Sikora, S., 10, 12, 13, 46, 189  
Simpson, A., 56  
Slone, M., 55  
Smith, M.W., 10, 38, 40  
Soter, A., 46, 48, 199  
Stanovich, K. E., 17, 135, 167, 213  
Stathi, S., 57, 92, 128  
Steen, G., 82, 198  
Stephenson, T., 284  
Stevahn, L., 55, 56, 58, 60, 61, 64, 66, 67, 69, 70, 73, 74, 76, 77, 78, 81, 272, 274  
Stienen, M., 214  
Stokmans, M., 159  
  
Tarrasch, R., 55  
Taylor, M. J., 42, 55  
Taylor, R. D., 42, 55  
Ten Dam, G., 200  
Tengberg, M., 13  
Tetzlaff, J., 202  
The PRISMA Group, 202  
Tonge, C., 2, 127  
Tonkonogy, A., 189  
Turner, R. N., 79  
  
Van Dam-Helmig, M., 214  
Van de Ven, P. H., 14, 15, 37, 213  
Van den Akker, J., 49, 56, 69, 88, 97, 203  
Van den Bergh, H., 88, 128  
Van den Broeck, A., 165  
Van den Hoven, P., 215  
Van der Bolt, L., 11, 35  
Van der Borden, J., 214  
Van der Kamp, M., 16, 135  
Van der Linden, R., 152  
Van Grinsven, V., 206  
Van Grootheest, A., 206  
Van Joolingen, W. R., 214  
Van Kesteren, B. J., 16  
Van Kuijk, I., 11  
Van Lierop-Debrauwer, H., 215  
Van Peer, W., 80, 161, 197  
Van Rees, K., 15, 19

- Van Schooten, E. J., 14, 150, 159, 164  
Van Weijen, D., 46, 88  
Veltkamp, M., 10, 11, 12, 44  
Verbitsky, M., 189  
Verboord, M., 15, 19, 159, 213  
Verkoeijen, P., 11  
Vermaas, J., 152  
Vezzali, L., 57, 58, 59, 60, 64, 66, 67, 69, 70, 71, 72, 73, 74, 75, 76, 78, 79, 80, 92, 94, 128, 273  
Vloeberghs, K., 215  
Voci, A., 79  
Vonofakou, C., 79  
Vygotsky, L. S., 82
- Wear, D., 53  
Weber, C. D., 53  
Wegerif, R., 199  
Wegman, H., 214  
Wei, L., 48, 76, 84  
Weijts, C., 216  
Weissberg, R. P., 55  
Wells, D., 38, 201  
Wells, G., 199
- Werneke, U., 284  
West, R. F., 17, 135, 167  
White, B. F., 57, 58, 59, 60, 64, 65, 69, 70, 71, 73, 75, 76, 77, 78, 79, 80, 81, 92, 128, 157, 158, 185, 273  
White, M. J., 57, 58, 59, 60, 64, 65, 69, 70, 71, 73, 75, 76, 77, 78, 79, 80, 81, 92, 128, 157, 158, 185, 273  
Wieser, D., 13  
Wigfield, A., 132  
Wilhelm, J. D., 10, 15, 37, 38, 40, 46, 47, 181  
Wilkinson, I., 46, 48, 199  
Witte, T., 8, 9, 10, 14, 39, 216  
Wong, S. P., 57
- Yeager, E. A., 36  
Yilmaz, K., 36
- Zahavi, D., 45, 78  
Zeitz, C. M., 13  
Zunshine, L., 8, 10, 126  
Zwaan, R., 11  
Zyngier, S., 12, 44, 53, 80, 136, 159, 161, 197, 205

## SUMMARY

### THE STORY, THE SELF, THE OTHER

#### DEVELOPING INSIGHT INTO HUMAN NATURE IN THE LITERATURE CLASSROOM

The aim of this dissertation was to investigate which instructional approach to literature teaching in Dutch upper secondary education would be appropriate for fostering students' insight into human nature, including their insight into themselves, fictional others, and real-world others. In the Introduction, we have argued that education has often been considered a place where people may learn to reflect on their own nature as well as the nature of others. Furthermore, for adults, adolescents and younger children alike, reading fictional and literary texts has been shown to result in such insights. Therefore, the literature classroom appears to offer a space of opportunity for adolescents to develop insight into human nature, which is valued as a potential outcome of literary instruction by teachers as well as policy makers in the Netherlands. However, in the Dutch context, there is no empirical support for the assumption that literary instruction may foster students' insight into human nature. With this dissertation, we aimed to fill this void.

#### STUDY 1 (CHAPTER 2)

In Study 1, we aimed to explore whether upper secondary school students gained any insight into themselves and others in the context of the literature classroom, as well as the relations between students' learning experiences and their teachers' classroom practices. Dutch students ( $N = 297$ , grades 10-12) wrote a learner report on what they learned about themselves and other people through literature education, and completed a measure on familiarity with fiction. Their teachers ( $N = 13$ ) completed a questionnaire, which indicated whether they used more analytical-interpretative or more personal-experiential

approaches to three aspects of teaching. Students of teachers with distinct approaches to the aspects were grouped to compare their learning experiences.

Findings showed that nearly all students reported to have learned something about themselves and others, mainly in terms of personal characterizations of oneself and others, oneself and others as literary readers, descriptions and evaluations of people's behaviors, and lessons for life. In addition, students of teachers who reported to allow for more classroom interaction and student autonomy – characteristics of a personal-experiential approach – were found to have reported learning about themselves and others more frequently. Although these findings may partly be explained by these students being more familiar with fiction and having a more positive attitude toward literary reading, we considered these findings as initial parameters for the design of an instructional approach.

### STUDY 2 (CHAPTER 3)

In Study 2, we aimed to identify design principles for an instructional approach. We conducted a systematic review of previous experimental and quasi-experimental literature classroom intervention studies that aimed to foster students' insights into themselves and others. Results from five research databases were screened, complemented with citation tracking, hand searches, and expert consultation. We included thirteen experimental and quasi-experimental intervention studies, for which we assessed methodological quality and the quality of the intervention descriptions. This process of quality appraisal functioned as a gatekeeper for the validity of researchers' conclusions.

Researchers of these thirteen studies expected that their instructional approaches would foster students' insight into themselves, their understanding of fictional others, and/ or their understanding of, views on, or intended behavior toward real-world others. Analysis of empirical support for expected intervention effects indicated that one intervention affected students' insight into themselves, two affected their understanding of fictional others, and six affected their understanding of, views on, or intended behavior toward real-world others. At this point, we concluded that these categories were not always clearly distinguishable, which led us to reformulate "insight into self and others" into the more broadly formulated concept of "insight into human nature". From the reviewed studies we inferred that, under certain conditions, literature education may foster students' insight into human nature.

Subsequent analysis of interventions with full or partial empirical support yielded three instructional design principles: 1) select fictional texts such as

novels, short stories, passages, or poems, that are thematically relevant for the intended outcomes of the intervention, such as texts with social-moral themes; 2) design writing tasks related to fictional texts and text themes that prompt students to activate previous personal experiences before reading, notice and annotate their experiences during the reading process, and/or write down (reflective) responses directly after reading; and 3) design exploratory dialogic tasks that stimulate students to verbally share and deepen their personal experiences related to fictional texts and text themes.

### STUDY 3 (CHAPTER 4)

In Study 3, we described the iterative design process via which we developed an instructional approach for 10<sup>th</sup> grade literature classrooms in the Dutch higher general secondary education track. The design was informed by a theoretical-empirical model of transformative reading. Transformative reading includes "self-other perceptual depth", which resembles "insight into human nature" as it includes both self and others. In addition, transformative reading includes six other experiences: vividly imagining story setting and characters (*imagery*); recognizing aspects of self or others in characters (*identification*); enacting and embodying the experiences of a character (*experience-taking*); evaluating characters, positively or negatively (*character evaluations*); feeling sympathy and compassion for characters (*sympathy*); and being aware of striking words, phrases or sentences (*aesthetic awareness*). In the design process, we considered how students could be guided toward reflecting upon these reading experiences. The design was further informed by the design principles identified in Study 2: students read short stories that addressed social-moral themes (principle 1), and were stimulated to engage in internal dialogues with texts via writing tasks (principle 2) to prepare for external dialogues with their peers about their responses to the text, its themes and the connections with their own lives and the lives of others (principle 3). In two iterations, we designed two subsequent versions of the intervention, which we named Transformative Dialogic Literature Teaching: TDLT-1 and TDLT-2. We assessed validity – both at the content and construct level – and practicality of both versions, as these are amongst the indicators of the quality of an intervention.

TDLT-1 was a four-unit intervention, designed in collaboration with teachers. It was taught by 13 teachers in 22 classes. From these teachers and students, we collected implementation and evaluation data. This enabled us to draw informed conclusions about the validity and practicality of the TDLT-1 intervention, which appeared to be suboptimal. For example, students strug-

gled to see why internal and external dialogues with and about stories were relevant (content validity), found it unclear how they could engage in these dialogues (construct validity), were too often off task, and needed, according to their teachers, more time to get used to dialogic response practices in the literature classroom (practicality). We aimed to remedy these issues by setting up a second design iteration, in which three teachers who taught TDLT-1 cooperated with us to redesign the intervention.

The second iteration resulted in TDLT-2, which consisted of six units: one preparatory and five reading-and-dialogue units. It included 300 minutes of classroom work, complemented by about 45 minutes of homework assignments. Short stories with a social-moral theme were read, but in contrast with TDLT-1, we used a single-theme approach: all stories centered around "justice and injustice". In TDLT-2, students were taught strategies for external and internal dialogues. Thus, TDLT-2 differed from TDLT-1 in its strategy instruction, which included, in unit 1, peer observations, explicit instruction about external dialogue strategies, and applying these strategies in a try-out dialogue in a small group. Teachers were asked to take on a guiding, non-authoritative role when students engaged in external dialogues. They followed guidelines for providing guidance and feedback on students' dialogic processes, which included, for example, prompts and questions that stimulated students to continue and deepen their talk.

In units 2 to 6, the internal and external dialogue together formed the two-step basic TDLT structure. Internal dialogues remained implicit in units 1 and 2, when students were not yet given a particular reading instruction. In unit 3, students received strategy instruction for internal dialogues, which included observation of their teacher thinking aloud during reading. As incomprehension was considered to be a legitimate response in internal dialogues, the teacher also introduced strategies for dealing with difficulties during reading. Moreover, from unit 3 onwards, internal and external dialogues focused on transformative reading experiences, whereas in unit 1 and 2 they had focused on more familiar responses such as their initial opinion about a story. Activities were miscellaneous, short, and high-paced to keep students engaged and motivated. Students were stimulated to monitor their progress by working with a rubric.

Six teachers, including the three teachers involved in the redesign, taught TDLT-2 to one of their classes. From implementation and evaluation data we inferred that teachers felt that TDLT-2 was generally practical to work with, and that students were on task for a larger proportion of time in TDLT-2 than in

TDLT-1, in particular during external dialogues. Moreover, students considered explanations about internal and external dialogues particularly helpful, as well as the teacher modeling internal dialogue strategies. Overall, we concluded that TDLT-2 was a valid and practical operationalization of the transformative reading model and the design principles.

#### STUDY 4 (CHAPTER 5)

In Study 4, we assessed the effects of TDLT-1 on students' transformative reading experiences, including – and most importantly – their insight into human nature, in a quasi-experimental design with pretest, posttest, and delayed posttest with switching replications ( $N = 603$ ). Students from 11 classes ( $n = 311$ ) first participated in TDLT-1 and subsequently in an untreated control condition in which regular Dutch lessons were taught (but no literature); in the other eleven classes ( $n = 292$ ), the order of conditions was switched. To gain first indications of what students learned from TDLT-1, we administered a written learner report directly after the intervention, as posttest-only. We developed the Transformative Reading Experiences Questionnaire (TREQ) and a story response task and administered both at the three measurement occasions. The TREQ included eight subscales: imagery, identification, experience-taking, character evaluations, sympathy, aesthetic awareness, self-insight, and insight into others. These scales were found to represent two underlying factors. For the story response task, three different texts were used. Students were asked to annotate their initial, spontaneous responses in the margin and were asked to complete three short writing tasks directly after reading.

Results from the learner reports indicated that students most often reported to have learned how to talk about stories. They further reported some transformative reading experiences, and indicated they had developed their thinking and opinionating skills. However, from the TREQ and story response task we detected no consistent effects of TDLT-1 on students' transformative reading experiences, including their insight into human nature. Several explanations are possible. First, although students relatively often indicated in their learner reports that they learned to talk about stories, dialogues remained rather short and superficial, as observed by various teachers and endorsed by a low on task percentage during small-group dialogues (see Study 3). Furthermore, students may have had too little experience with literary reading to be able to notice their responses during reading, as suggested by the smaller share of learning experiences that concerned, for instance, transformative reading, in-depth processing of stories, and noticing responses during reading. In addition, students

and teachers indicated that the purpose of TDLT-1 – gaining insight into human nature – remained too implicit. Therefore, students may not have been sufficiently motivated for and engaged in the units and stories. Finally, the instruction time may have been too short. Based on these findings, adjustments were made that resulted in TDLT-2.

#### STUDY 5 (CHAPTER 6)

In Study 5, we aimed to assess the effects of TDLT-2 on students' transformative reading experiences and other indicators of insight into human nature, their reasons for reading, their use of strategies to deal with difficulties during literary reading, and their motivation for literature education. In addition, we explored the extent to which strategy use and motivation functioned as mediators for effects of TDLT-2 on students' insight into human nature and eudaimonic reasons for reading.

We applied a quasi-experimental design with pretest, posttest and delayed posttest (four months after the intervention), in which six classes participated in TDLT-2 ( $n = 166$ ) and six classes in the control condition ( $n = 166$ ). In the control condition, students followed their teachers' regular literature curriculum, mainly focused on literary devices and analysis. As instruments, we used the TREQ as well as scales from other validated questionnaires. In addition, we analyzed a task in which students wrote a dialogue with an imaginary peer in response to a short story for transformative reading experiences and other indicators of insight into human nature.

Findings indicated that TDLT-2 had positive effects on students' insight into human nature. For example, TDLT-2 had a medium effect on their "insight beyond story worlds", a factor score derived from the TREQ that included self-insights and insights into real-world others, and a small effect on their "experiences within story worlds", a factor that included imagery, experience-taking, and evaluations of how characters think, feel and behave. Likewise, TDLT-2 had a medium effect on students' eudaimonic reasons for reading: these became more important to TDLT-2 students, but not to students in the control condition. Analysis of students' written dialogues supported these findings. For example, more TDLT-2 students reasoned to understand and interpret story characters significantly often than students in the control condition, and fewer of them wrote superficial descriptive-evaluative statements. The positive effects of TDLT-2 on students' insight into human nature and eudaimonic reasons for reading were still statistically significant four months after the intervention, although they were then smaller than directly afterwards.

In addition, TDLT-2 positively affected students' strategy use, as well as their feelings of autonomy, competence, and relatedness, as indicators of their motivation for literature education. Students' strategy use and feelings of competence and relatedness functioned as mediators for the effects of TDLT-2 on several indicators of insight into human nature and on eudaimonic reasons for reading. These mediating effects were statistically significant but relatively small. Students' feelings of competence played the largest mediating role for "insight beyond story worlds". This finding suggested that teachers and educational designers should be responsive to students' feelings of competence in the literature classroom: the more they indicated to have gained insight into themselves, fictional others and real-world others.

#### GENERAL DISCUSSION (CHAPTER 7)

In the general discussion we considered four concepts that are central in this dissertation. First, we reflected upon capturing multiple relevant terms concerning self and others in a single concept, which we ultimately termed "insight into human nature". We addressed the entanglement of the terms "self" and "other" as well as of fictional and real-world others, and we indicated how terminological issues were partly based upon the language setting in which this dissertation came about: TDLT was developed for Dutch language classrooms, but reported on in English.

Second, we discussed the application of the transformative reading model in the context of Dutch literature classrooms. We concluded that transformative reading can be measured both quantitatively and qualitatively, and that TDLT guided students toward noticing, reflecting on, and talking about transformative reading experiences. This dissertation has thus expanded previous research into transformative reading from adult readers to adolescents, and has demonstrated that transformative reading can be a meaningful concept in an educational context. Third, we discussed the possibility of using, in 10<sup>th</sup> grade classrooms, fictional narratives that are likely to be perceived by students as "literary texts", as students were capable to refer to literary devices in texts, to clarify and substantiate their responses and the insights they gain from reading and talking about stories.

Lastly, in terms of dialogic teaching and learning, we concluded that dialogues in literature classrooms may take place on two levels: between the reader and the text (internal dialogue), and among readers in response to the text (external dialogue). Building upon the work of Bakhtin and Rosenblatt, we explained the relevance of internal dialogues; likewise, we attempted to illumi-

nate the relevance of external dialogues from a cognitive and a social constructivist perspective. We have also discussed the pros and cons of peer-led and teacher-led dialogues in the literature classroom, including the consequences for students' power positions and teachers' roles as non-authoritative figures who avoid to provide "the right answer".

Subsequently, we discussed potential validity issues of our studies. We concluded that the systematic review (Study 2) contributed to the validity of the intervention design, because critical quality appraisal of methodologies and intervention descriptions functioned as a gatekeeper for the design principles that we identified. Furthermore, the iterative design process contributed to the validity of the intervention-as-designed. Simultaneously, the data collected in this process contributed to the validity of the intervention-as-implemented: by using multiple instruments to collect implementation and evaluation data from both students and teachers, we gained a valid impression of how TDLT-1 and TDLT-2 were implemented. Whereas implementation fidelity of TDLT-1 had been at stake, this validity issue appeared to be alleviated in TDLT-2. Moreover, we attempted to optimize ecological validity of TDLT-1 and TDLT-2 by asking students' own teachers to participate in the studies and by involving teachers in designing and redesigning TDLT.

The validity of the TREQ was supported because different samples of students generally responded to the items in a similar way. Because students also expressed transformative reading experiences in response to specific texts, via writing tasks, our studies showed concurrent validity and avoided a "monomethod bias" as a threat to construct validity. In addition, we concluded that the learner report was a valid instrument to tap into students' learning experiences, as students found the task feasible and reported few irrelevant or incomprehensible learning experiences; moreover, learning experiences were well-distinguishable as demonstrated by coding systems with substantial interrater agreement. Finally, this dissertation supports the validity of the Author Recognition Test for measuring Dutch adolescents' familiarity with fiction: in Study 1, students in pre-university scored significantly higher than those in higher general secondary education, and average scores were similar across our studies.

We further concluded that we could rather confidently ascribe the effects of TDLT-2 to the treatment. As a strength, multiple teachers were involved in both the experimental and control conditions. However, as a validity threat, teachers and classes were not randomly assigned to conditions in Study 5; yet, as we found no significant differences between particular classes on relevant back-

ground variables, the lack of random assignment to conditions was counter-balanced to some extent. Furthermore, we ruled out selection bias as a threat to validity in Study 5: students of teachers who were already familiar with TDLT did not score significantly higher than students of the teachers who were new to TDLT. Lastly, we addressed the possibility of a Hawthorne effect in Study 5, which could not be ruled out, but may not be all too problematic from an instructional point of view.

In terms of external validity, we indicated that the three initial design principles may be considered "first principles": the mere fact that they are operationalized increases the probability that the desired learning outcome will be achieved. It is unlikely that small adjustments in TDLT units would drastically change its outcomes. The findings of our intervention studies cannot be generalized beyond students in Dutch upper secondary education. Considering teaching experience, effects of TDLT-2 may hold if different samples of teachers are involved, as teachers with a wide range of teaching experience seemed equally able to implement TDLT-2 as intended, and effects were consistent across teachers who were and were not previously familiar with it. Furthermore, we should be careful to generalize the effects of TDLT-2 to individual students, although additional explorative analyses indicated that the intervention was equally effective for subgroups within the sample. Finally, we concluded that the outcomes of this research project do not seem to be caused by researcher bias, as we collaborated with teachers, relied on multiple data sources in the design process, triangulated data in the intervention studies, and avoided that data were collected by a single researcher.

Next, we addressed starting points for future studies, which may focus on the respective roles of text selection, internal dialogues, external dialogues, and attention to transformative reading experiences; on the cognitive and affective processes that underlie students' development of insight into human nature; and on further developing and validating instruments for measuring insight into human nature.

Finally, we addressed implications for educational practice. In view of an imminent nation-wide curricular reform in the Netherlands, we believe that the main contributions of this dissertation to Dutch teaching practices are threefold. It offers research-based instructional design principles that fit the kind of approach to literature teaching as envisioned for the curricular reform; it has shown how these design principles can effectively be operationalized in class; and it has empirically demonstrated that this operationalization fosters stu-

dents' insight into human nature, their support for eudaimonic reasons for reading, their reported use of strategies to deal with difficulties during literary reading, and their motivation for literature education. Further, we suggested that the approach may potentially have positive implications for students' reading comprehension, or may stimulate them to enter an upward spiral of reading frequency.

We further argued that the starting point of any instructional approach in literature classrooms should be that students learn to become aware of their initial responses to literary texts and to put these into words. With this notion, we have attempted to answer to the potential criticism that TDLT cannot truly be considered "literary instruction": precisely the attention for students' initial responses to texts, as well as the focus on substantiating those responses with references to literary devices identified in texts, differentiates TDLT as an approach to literary instruction from other educational disciplines. Taking students' initial responses to texts as a starting point further implies that monologic, teacher-centered literary instruction does not appear to be the way forward for future literature curricula.

These implications raised at least three questions. First, how may the implementation of TDLT and its principles find its way into literature classrooms? We indicated that there may be a need to develop a professional development course for pre-service and in-service teachers, and that it may be worth considering to set up teacher design teams for teachers' professional development. Second, we addressed questions about assessment and testing which are highly relevant to teachers. We argued in favor of formative assessment, for which the TDLT rubric, written dialogue, and learner report task may be helpful tools. Finally, we raised the question to which texts and educational domains TDLT principles may be applied. We discussed the potential of adolescent and young adult literature, and pointed out that, although they may be thematically relevant, the language in historical literary texts or texts in foreign language curricula may hinder their implementation in TDLT. Lastly, we have indicated that TDLT may offer opportunities for transdisciplinarity in schools, by connecting literary instruction to other disciplines in which human nature plays a role, such as history, citizenship education, social studies, and arts education.

# SAMENVATTING

## HET VERHAAL, IKZELF EN DE ANDER

### INZICHT IN “DE MENS” VERWERVEN IN LITERATUURONDERWIJS

Deze dissertatie had ten doel te onderzoeken welke didactische aanpak voor literatuuronderwijs in de bovenbouw van het voorgezet onderwijs geschikt zou zijn om het inzicht van leerlingen in zichzelf en anderen – zowel anderen in verhalen als in de echte wereld – te bevorderen. In de Inleiding hebben we betoogd dat het onderwijs vaak wordt gezien als een plek waar je kunt leren reflecteren op je eigen menselijke natuur en die van anderen. Dat het lezen van fictionele en literaire teksten tot zulke inzichten kan leiden is eerder aange- toond voor zowel volwassenen als adolescenten en jongere kinderen. Het lite- ratuuronderwijs lijkt jongeren daarom bij uitstek mogelijkheden te bieden om inzicht in “de mens” te ontwikkelen. Zowel docenten als beleidsontwikkelaars in Nederland zien dit als een waardevolle potentiële opbrengst van literatuuron- derwijs. In de Nederlandse onderwijscontext is er echter nog geen empirische ondersteuning voor de aanname dat literatuuronderwijs inzicht in de menselij- ke natuur kan bevorderen. Met dit proefschrift hebben we getracht deze leem- te op te vullen.

#### STUDIE 1 (HOOFDSTUK 2)

Studie 1 was exploratief van aard. Onze doelen waren verkennen of boven- bouwleerlingen überhaupt inzicht in zichzelf en anderen opdoen in de context van hun literatuuronderwijs, en of er relaties zijn tussen de leerervaringen van leerlingen en de didactische aanpak van hun docenten. Leerlingen ( $N = 297$ , klas 4-6) schreven een leerverslag waarin ze rapporteerden wat ze hadden ge- leerd over zichzelf en anderen via literatuuronderwijs, en maakten een test die een indicatie gaf van hun belezeneid. Hun docenten ( $N = 13$ ) vulden een vra-

genlijst in, die aangaf of zij een meer analytisch-interpretatieve of een meer persoonlijk-ervaringsgerichte benadering hanteerden ten aanzien van drie aspecten van hun lespraktijk. Leerlingen van docenten met uitgesproken benaderingen werden gegroepeerd, zodat we hun leerervaringen konden vergelijken.

Uit de resultaten bleek dat vrijwel alle leerlingen rapporteerden iets over zichzelf en anderen te hebben geleerd. Bij deze leerervaringen ging het voornamelijk om leren over jezelf en anderen als personen, over jezelf en anderen als literaire lezers, om leren over het gedrag van anderen, en het leren van levenslessen. Leerlingen van docenten die aangaven meer interactie in de klas en meer autonomie voor leerlingen toe te staan – kenmerken van een persoonlijk-ervaringsgerichte benadering – bleken meer leerervaringen over zichzelf en anderen te rapporteren. Hoewel deze resultaten deels verklaard kunnen worden door het feit dat deze leerlingen ook meer belezen waren en positiever tegenover literair lezen stonden, beschouwden we de bevindingen als eerste parameters voor het ontwerp van een didactische aanpak.

## STUDIE 2 (HOOFDSTUK 3)

Het doel van Studie 2 was om ontwerpregels voor een didactische aanpak te formuleren. Hiertoe voerden we een systematische reviewstudie van (quasi-) experimentele interventiestudies in het literatuuronderwijs die het bevorderen van zelfinzicht en sociaal inzicht als doel hadden. We screenen resultaten van vijf onderzoeksdatabases, aangevuld met de sneeuwbalprocedure (nagaan of citerende en geciteerde studies ook relevant zijn), het handmatig screenen van relevante onderzoeksoverzichten en tijdschriften, en het bevragen van experts. We namen dertien (quasi-)experimentele studies op in de review. We beoordeelden de kwaliteit van methodologische aspecten en interventiebeschrijvingen, waarmee we de validiteit van de getrokken conclusies bewaakten.

De onderzoekers van deze dertien studies verwachtten dat hun interventies bij leerlingen zelfinzicht, begrip voor verhaalpersonages, en/of begrip voor of positief voorgenomen gedrag ten aanzien van anderen in de echte wereld zouden bevorderen. Negen interventies toonden de verwachte effecten ook daadwerkelijk aan. Eén interventie bevorderde het zelfinzicht van leerlingen, twee interventies vergrootten hun begrip voor verhaalpersonages, en zes interventies hadden positieve effecten op hun begrip voor of voorgenomen gedrag ten aanzien van anderen in de echte wereld. Omdat deze categorieën niet altijd goed onderscheiden konden worden, herformuleerden we inzicht in "jezelf en anderen" tot inzicht in "de mens". De onderzochte studies lieten zien

dat literatuuronderwijs, tot op zekere hoogte, het inzicht van leerlingen in de mens kan bevorderen.

Vervolgens analyseerden we de didactische aanpak in de negen interventiestudies die positieve effecten hadden op inzicht in de mens. Deze analyse leverde drie ontwerpregels op: 1) selecteer fictionele teksten die thematisch relevant zijn voor het doel van de interventie, zoals romans, korte verhalen of fragmenten met sociaal-morele thema's; 2) ontwerp tekst- en thema-gerelateerde schrijftaken die leerlingen ertoe aanzetten om relevante persoonlijke ervaringen te activeren voorafgaand aan het lezen, hun reacties en ervaringen op te merken en te annoteren tijdens het lezen, en (reflectieve) responsen te schrijven direct na het lezen; en 3) ontwerp exploratieve dialogische taken die leerlingen stimuleren om hun persoonlijke reacties op fictionele teksten en de thema's daarin mondeling uit te wisselen en uit te diepen.

### STUDIE 3 (HOOFDSTUK 4)

In Studie 3 beschreven we het iteratieve ontwerpproces waarin we een didactische aanpak voor literatuuronderwijs in havo 4 ontwikkelden. Het ontwerp was gebaseerd op een theoretisch-empirisch model van *transformative reading* (hierna: TR). TR kan gekarakteriseerd worden als lezen waarbij de lezer diepere inzichten in zichzelf en anderen opdoet, wat overeenkomt met inzicht in "de mens". Naast het opdoen van zelfinzicht en inzicht in anderen bestaat TR uit zes andere leeservaringen: de setting en personages in een verhaal levendig voor je zien (*verbeelding*); aspecten van jezelf of anderen in personages herkennen (*herkenning*); je inleven in de situatie van een personage door "in diens huid te kruipen" (*inleving*); een mening vormen over personages (*oordelen over personages*); medeleven en compassie voelen voor personages (*sympathie*); en je bewust zijn van welke woorden, zinsdelen, zinnen of passages "eruit springen" of een bepaalde ervaring teweeg brengen (*bewustzijn van taal en stijl*). In het ontwerpproces stelden we ons de vraag hoe leerlingen ertoe bewogen kunnen worden om te reflecteren op deze specifieke leeservaringen. Daarnaast baseerden we de didactische aanpak op de ontwerpregels zoals geformuleerd in Studie 2: leerlingen lazen korte verhalen waarin sociaal-morele thema's aan bod kwamen (ontwerpregel 1), en werden gestimuleerd om met behulp van schrijftaken interne dialogen met teksten aan te gaan (ontwerpregel 2), waarmee ze zich voorbereidden op externe dialogen met medeleerlingen over hun reacties op de tekst, de thema's daarin, en de relaties daarvan met hun eigen en andermans leven (ontwerpregel 3). In twee iteraties ontwierpen we twee opeenvolgende versies van de interventie, die we Transformatief

Dialogisch Literatuuronderwijs noemden: TDL-1 en TDL-2. We onderzochten de inhouds- en constructvaliditeit en bruikbaarheid van beide versies, als indicatoren van de kwaliteit van een interventie.

TDL-1 bestond uit vier modules, ontworpen in samenwerking met docenten, waarin korte verhalen werden gelezen. 13 docenten implementeerden de interventie in 22 klassen. We verzamelden implementatie- en evaluatiedata van docenten en leerlingen. Op basis hiervan konden we empirisch onderbouwde conclusies trekken over de validiteit en bruikbaarheid van TDL-1, die niet optimaal bleken. Leerlingen hadden bijvoorbeeld moeite om in te zien waarom interne en externe dialogen met en over verhalen relevant kunnen zijn (inhoudsvaliditeit), vonden het onduidelijk hoe ze zulke dialogen konden aangaan (constructvaliditeit), vertoonden te weinig taakgericht gedrag, en hadden volgens hun docenten meer tijd nodig om te wennen aan dialogische taken bij literatuuronderwijs (bruikbaarheid). We stelden ons ten doel om deze problemen te verhelpen in een tweede ontwerpronde. Drie docenten die TDL-1 gegeven hadden, werkten met ons samen aan het herontwerp.

Deze tweede iteratie resulteerde in TDL-2, bestaande uit zes modules: één voorbereidende en vijf lees-en-dialoog-modules. Tijdens TDL-2 werkten leerlingen in totaal 300 minuten in de klas (bijvoorbeeld gedurende zes lessen van 50 minuten) en besteedden ze ongeveer 45 minuten aan huiswerk. We kozen, in tegenstelling tot bij TDL-1, voor een thematische aanpak: alle verhalen hadden iets te maken met (on)rechtvaardigheid. In TDL-2 leerden leerlingen strategieën voor het aangaan van externe en interne dialogen, door middel van observatie van andere leerlingen op video, expliciete instructie, en het oefenen van strategieën; TDL-2 onderscheidde zich daarin van TDL-1. Bij deze gesprekken namen docenten een begeleidende rol op zich, waarmee ze probeerden zich niet op te stellen als "de autoriteit" op het gebied van literatuur. Ze volgden richtlijnen voor het begeleiden van en geven van feedback op de dialogische processen in kleine groepjes, waarbij ze bijvoorbeeld prompts en vragen inzetten die leerlingen stimuleerden om hun gesprekken voort te zetten en verder uit te diepen.

In module 2 tot en met 6 vormden de interne en externe dialoog samen, in twee stappen, de basisstructuur van TDL-2. In module 1 en 2 bleven de interne dialogen nog impliciet: leerlingen kregen geen specifieke leesinstructie. In module 3 volgde strategie-instructie voor interne dialogen, inclusief het observeren van de docent die tijdens het lezen van een verhaal hardop nadacht. Aangezien we onbegrip als een authentieke reactie beschouwden, introduceerde de docent ook strategieën om daarmee om te gaan. Daarnaast richtten de in-

terne en externe dialogen zich vanaf module 3 op TR-ervaringen in plaats van, zoals in module 1 en 2, op meer bekende reacties zoals het vormen en verwoorden van je mening over een verhaal. Alle taken waren kort, gevarieerd, en wisselden elkaar in hoog tempo af om de betrokkenheid en motivatie van leerlingen te stimuleren. Tot slot hielden leerlingen hun ontwikkeling in het oog met behulp van een rubric.

Zes docenten, waaronder de drie docenten die meewerkten aan het herontwerp, implementeerde TDL-2 in een van hun klassen. Op basis van implementatie- en evaluatiedata concludeerden we onder meer dat deze docenten TDL-2 bruikbaar vonden in de praktijk en dat leerlingen meer taakgericht gedrag vertoonden, met name tijdens de dialogische taken in kleine groepjes. Bovendien gaven leerlingen aan instructie over interne en externe dialogen waardevol te vinden, evenals het observeren van hun docent. We concludeerden dat TDL-2 kon worden beschouwd als een valide en bruikbare operationalisering van het TR-model en de ontwerpregels.

#### STUDIE 4 (HOOFDSTUK 5)

In Studie 4 onderzochten we de effecten van TDL-1 op de TR-ervaringen van leerlingen, waaronder hun inzicht in de mens. TR-ervaringen werden gemeten in een quasi-experimenteel design met een voormeting, nameting en uitgestelde nameting en *switching replications*. In totaal namen 603 leerlingen deel aan het onderzoek. Leerlingen uit 11 van de 22 klassen ( $n = 311$ ) namen eerst deel aan TDL-1 en daarna aan een controleconditie waarin reguliere lessen Nederlands, maar geen literatuurlessen werden gegeven; in de andere 11 klassen ( $n = 292$ ) was de volgorde van de condities omgekeerd.

Om een eerste indruk te krijgen van wat leerlingen van TDL-1 hadden geleerd, vroegen we hen direct na de interventie een leerverslag te schrijven; dit instrument fungeerde alleen als nameting. We ontwikkelden een vragenlijst om TR-ervaringen te meten (afgekort TREQ) en een schrijftaak in reactie op een gelezen verhaal. Beide instrumenten werden afgenomen op de drie meetmomenten. De TREQ bevatte acht subschalen: verbeelding, herkenning, inleving, oordelen over personages, sympathie, bewustwording van taal en stijl, zelfinzicht, en inzicht in anderen. Deze schalen bleken twee onderliggende factoren te representeren. Voor de verhaaltaak werden drie verschillende verhalen gebruikt. We vroegen leerlingen om hun eerste, spontane reacties op het verhaal tijdens het lezen naast de tekst te noteren en om direct na het lezen drie korte schrijftaken te maken.

De leerverslagen lieten zien dat leerlingen het vaakst aangaven dat ze hadden geleerd hoe ze met elkaar over verhalen in gesprek konden gaan. Daarnaast rapporteerden ze TR-ervaringen en gaven ze aan dat ze hun gedachten en meningen over verhalen hadden leren verwoorden. Op basis van de TREQ en de verhaaltaak vonden we echter geen consistente effecten van TDL-1 op de TR-ervaringen van leerlingen, waaronder hun inzicht in de mens. Hiervoor zijn verschillende verklaringen denkbaar. Hoewel leerlingen in hun leerverslagen relatief vaak aangaven dat ze over verhalen hadden leren praten, bleven deze dialogen tamelijk kort en oppervlakkig. Dit bleek uit de reacties van docenten en uit het weinig taakgerichte gedrag van leerlingen tijdens dialoogtaken (zie Studie 3). Daarnaast hadden de leerlingen wellicht te weinig ervaring met literair lezen om hun ervaringen tijdens het lezen op te merken, getuige diverse leerervaringen die relatief weinig voorkwamen, waaronder TR-ervaringen, leerervaringen over diepere verwerking van verhalen, en over het opmerken van reacties tijdens het lezen. Leerlingen en docenten gaven tevens aan dat het doel van TDL-1 – inzicht in de mens verwerven – te impliciet bleef, waardoor de motivatie en betrokkenheid van de leerlingen wellicht te wensen over liet. Tot slot is het mogelijk dat leerlingen onvoldoende tijd hadden om dit inzicht te ontwikkelen. Op basis van deze bevindingen werden aanpassingen gedaan die resulteerden in TDL-2.

#### STUDIE 5 (HOOFDSTUK 6)

In Studie 5 onderzochten we of TDL-2 positieve effecten had op TR-ervaringen en andere indicatoren van inzicht in de mens, op redenen om te lezen, op strategiegebruik om met moeilijkheden in literaire teksten om te gaan, en op motivatie voor literatuuronderwijs. Tevens onderzochten we in hoeverre strategiegebruik en motivatie de effecten van TDL-2 op inzicht in de mens en eudaimonische redenen om te lezen medieerden.

In een quasi-experimenteel design met een voormeting, nameting en uitgestelde nameting (vier maanden na de interventie) namen zes klassen deel aan TDL-2 ( $n = 166$ ) en zes klassen aan een controleconditie ( $n = 166$ ). In de controleconditie gaven docenten regulier literatuuronderwijs dat met name gericht was op literaire begrippen en verhaalanalyse. Als instrumenten gebruikten we de TREQ en diverse andere schalen van gevalideerde vragenlijsten. Daarnaast analyseerden we of er TR-ervaringen en andere indicatoren van inzicht in de mens voorkwamen een schrijftaak: leerlingen lazen een kort verhaal en schreven in reactie daarop een dialoog met een denkbeeldige medeleerling.

De resultaten lieten zien dat TDL-2 positieve effecten had op inzicht in de mens. Zo had de interventie een middelgroot positief effect op Inzicht in de wereld buiten het verhaal, een eerste factorscore van de TREQ. Het effect op Inzicht in ervaringen van personages, een tweede factorscore, was eveneens significant, maar kleiner. TDL-2 had daarnaast een middelgroot effect op eudaimonische redenen om te lezen: deze redenen werden belangrijker voor leerlingen die aan TDL-2 hadden deelgenomen, maar niet voor leerlingen in de controleconditie. De geschreven dialogen ondersteunden deze bevindingen. In de groep die TDL-2 had gevolgd schreven bijvoorbeeld significant meer leerlingen een redenering waarin ze probeerden om verhaalpersonages beter te begrijpen dan in de controleconditie; minder leerlingen schreven oppervlakkigere, beschrijvend-evaluatieve responsen. De positieve effecten van TDL-2 op inzicht in de mens en eudaimonische redenen om te lezen bleken vier maanden na de interventie nog steeds statistisch significant, al waren ze wel kleiner dan direct na de interventie.

Daarnaast bleek TDL-2 een klein positief effect te hebben op strategiegebruik, en middelgrote tot grote effecten op gevoelens van autonomie, competentie en sociale verbondenheid, die we in navolging van de zelfdeterminatietheorie beschouwden als indicatoren van motivatie voor literatuuronderwijs. Bovendien bleken strategiegebruik, en gevoelens van competentie en sociale verbondenheid te fungeren als mediators van de effecten van TDL-2 op diverse indicatoren van inzicht in de menselijke natuur en op eudaimonische redenen om te lezen. Deze mediërende effecten waren klein, maar statistisch significant. Het gevoel van competentie had het sterkste mediërende effect, namelijk op Inzicht in de wereld buiten het verhaal. Deze bevinding suggereerde dat het belangrijk is dat docenten en andere didactici in acht nemen hoe competent leerlingen zich voelen tijdens hun literatuurlessen: hoe competentere leerlingen zich voelden in TDL-2, hoe vaker ze aangaven inzicht in zichzelf, verhaalpersonages en anderen in de echte wereld te hebben opgedaan.

## DISCUSSIE (HOOFDSTUK 7)

In Hoofdstuk 7 presenteerden we de voornaamste bevindingen, en bespraken we vier concepten die centraal staan in deze dissertatie. We hebben gereflecteerd op onze pogingen om diverse relevante termen met betrekking tot "jezelf" en "de ander" samen te nemen in één concept. Uiteindelijk kozen we voor de Engelse term "insight into human nature". We zijn ingegaan op hoezeer de concepten "zelf" en "ander" met elkaar verbonden zijn; datzelfde geldt voor "de ander" in een verhaal en in de echte wereld. Daarnaast hebben we aange-

geven voor welke terminologische uitdagingen we onszelf gesteld zagen als gevolg van de talige situatie waarin deze dissertatie geschreven is: TDL werd ontwikkeld voor het Nederlandstalige literatuuronderwijs, maar we rapporteerden erover in Engelstalige artikelen.

Ten tweede bediscussieerden we de toepassing van het TR-model in de context van het Nederlandse literatuuronderwijs. We concludeerden dat TR-ervaringen op zowel kwantitatieve als kwalitatieve wijze in kaart konden worden gebracht, en dat TDL leerlingen hielp om TR-ervaringen op te merken, erop te reflecteren, en erover te praten. Doordat we ons richtten op adolescenten heeft deze dissertatie het bestaande onderzoek naar TR uitgebreid, dat tot nog toe voornamelijk op volwassen lezers was gericht. Bovendien hebben we laten zien dat TR een betekenisvol concept kan zijn in een didactische context. Ten derde bespraken we de mogelijkheid om in havo-4 verhalen te gebruiken die leerlingen waarschijnlijk als "literair" beschouwen. Leerlingen bleken immers in staat om te verwijzen naar literaire kenmerken van deze teksten, waarmee zij hun reacties op de teksten en de inzichten die ze opdeden konden onderbouwen.

Ten slotte zijn we ingegaan op het concept "dialogisch leren". We concludeerden dat dialogen in literatuuronderwijs kunnen plaatsvinden op twee niveaus: tussen lezer en tekst (de interne dialoog) en tussen lezers onderling over hun ervaringen met de tekst (de externe dialoog). We hebben verwezen naar het werk van Bakhtin en Rosenblatt om de relevantie van interne dialogen te duiden, en hebben het belang van externe dialogen verklaard vanuit een cognitief en sociaal-constructivistisch perspectief. Tot slot hebben we de voor- en nadelen besproken van leerling- en docentgestuurde dialogen in het literatuuronderwijs, waaronder de gevolgen die deze keuze heeft voor de machtspositie van leerlingen en voor de rol van docenten die kunnen proberen begeleider te zijn, in plaats van de autoriteit die "het juiste antwoord" kent.

Vervolgens benoemden we mogelijke validiteitsproblemen van ons onderzoek. We concludeerden dat de systematische reviewstudie (Studie 2) bijdroeg aan de validiteit van het interventieontwerp: een kritische beoordeling van de methodologie en interventiebeschrijvingen fungeerde als "poortwachter" bij het formuleren van ontwerpregels op basis van deze studies. Daarnaast droeg het iteratieve ontwerpproces bij aan de validiteit van de interventie-als-ontworpen. Tegelijkertijd ondersteunde de data die we in het ontwerpproces verzamelden de validiteit van de interventie-als-geïmplementeerd: door diverse instrumenten te gebruiken om implementatie- en evaluatiedata te verzamelen van zowel leerlingen als docenten, kregen we een valide beeld van hoe TDL-1 en TDL-2 in

de praktijk werden gebracht. Waar TDL-1 niet optimaal werd geïmplementeerd in vergelijking met hoe de interventie was bedoeld, bleek dit validiteitsprobleem in TDL-2 niet langer aan de orde. Daarnaast trachtten we de ecologische validiteit van TDL-1 en TDL-2 te optimaliseren: TDL werd gegeven door de eigen docenten van leerlingen, en docenten waren betrokken bij het ontwerp en herontwerp ervan.

De validiteit van de TREQ werd ondersteund doordat verschillende groepen leerlingen over het algemeen op dezelfde manier op de items reageerden. Omdat leerlingen ook in staat waren te reflecteren op TR-ervaringen in schrijftaken, was er sprake van voortschrijdende validiteit van de meetinstrumenten. Een constructvaliditeitsprobleem, dat zou ontstaan wanneer we slechts één instrument zouden gebruiken om TR-ervaringen te meten, was daarmee niet aan de orde. We concludeerden tevens dat het leerverslag een valide instrument was om toegang te krijgen tot de leerervaringen van leerlingen. Leerlingen vonden deze taak niet te moeilijk en rapporteerden relevante en begrijpelijke leerervaringen; bovendien waren hun leerervaringen te onderscheiden met codeersystemen waarvoor de interbeoordelaarsbetrouwbaarheid voldoende was. Tot slot ondersteunt ons onderzoek de validiteit van de Auteursherkenningstest, die belezenheid meet, voor Nederlandse adolescenten: in Studie 1 scoorden vwo-leerlingen significant hoger dan havo-leerlingen, en de gemiddelde scores in de verschillende studies lagen dicht bij elkaar.

We concludeerden daarnaast dat we de effecten van TDL-2 met relatief grote zekerheid aan de interventie konden toeschrijven. Een sterk punt in de onderzoeksopzet was dat er meerdere docenten deelnamen aan zowel de experimentele als de controleconditie. Anderzijds was er sprake van een validiteitsprobleem omdat docenten en klassen niet willekeurig aan condities waren toegewezen. De gevolgen daarvan werden enigszins ondervangen doordat we geen significante verschillen tussen specifieke klassen vonden op relevante achtergrondvariabelen. De validiteit van de studie bleek niet ondermijnd te worden door *selection bias*, omdat leerlingen van docenten die al bekend waren met TDL-2 niet significant hoger scoorden dan leerlingen van docenten voor wie TDL-2 nieuw was. Tot slot hebben we besproken dat de effecten van TDL-2 mogelijk toegeschreven kunnen worden aan een Hawthorne-effect. Hoewel we dit niet konden uitsluiten, hebben we ook aangegeven dat een dergelijk effect vanuit een didactisch perspectief niet per se problematisch hoeft te zijn.

Met het oog op externe validiteit hebben we aangegeven dat de drie ontwerpregels als *first principles* gezien kunnen worden: het feit dat ze geoperati-

onaliseerd worden vergroot de kans dat de beoogde leeropbrengsten behaald worden. Het is onwaarschijnlijk dat kleine aanpassingen in de TDL-modules leiden tot drastische veranderingen in de uitkomsten. De bevindingen van de interventiestudies hebben alleen betrekking op leerlingen in havo-4 in Nederland. Wel kunnen de uitkomsten gegeneraliseerd worden naar meer en minder ervaren docenten: docenten met meer en minder jaren werkervaring implementeerden TDL-2 over het algemeen zoals bedoeld, en de effecten ervan waren consistent voor zowel docenten die al bekend waren met TDL als voor degenen die dat niet waren. Voorzichtigheid is geboden bij generalisatie van effecten naar individuele leerlingen, hoewel aanvullende exploratieve analyses hebben laten zien dat de interventie even effectief was voor subgroepen binnen de groep deelnemende leerlingen. Tot slot concludeerden we dat onze bevindingen niet lijken te zijn veroorzaakt door vooringenomenheid van de onderzoekers, aangezien we samenwerkten met docenten, het interventieontwerp baseerden op data uit verschillende bronnen, data-triangulatie toepasten in de interventiestudies, en vermeden dat data verzameld werden door slechts één onderzoeker.

We bespraken tevens aanknopingspunten voor vervolgonderzoek, dat zich zou kunnen richten op de afzonderlijke rol van respectievelijk tekstkeuze, interne dialogen, externe dialogen, en aandacht voor TR-ervaringen; op de cognitieve en affectieve processen die ten grondslag liggen aan de ontwikkeling van inzicht in de mens; en op het verder ontwikkelen en valideren van instrumenten om inzicht in de mens te meten.

Tot slot zijn we ingegaan op implicaties voor de onderwijspraktijk. Met het oog op de op handen zijnde curriculumherziening (Curriculum.nu), zijn we van mening dat deze dissertatie op drie manieren bijdraagt aan de praktijk van het literatuuronderwijs in Nederland: ze biedt ontwerpregels die gebaseerd zijn op eerder onderzoek; laat zien hoe deze ontwerpregels effectief kunnen worden geoperationaliseerd in de klas; en toont aan dat deze operationalisering het inzicht van leerlingen in de mens bevordert, evenals het belang dat zij hechten aan eudaimonische redenen om te lezen, hun gerapporteerde gebruik van strategieën om met moeilijkheden in literaire teksten om te gaan, en hun motivatie voor literatuuronderwijs. Daarnaast hebben we aangegeven dat deze aanpak wellicht positieve implicaties heeft voor de leesvaardigheid van leerlingen en ertoe zou kunnen leiden dat zij in een "positieve leesspiraal" terecht komen.

We hebben daarnaast gesteld dat het vertrekpunt van elke didactische aanpak in het literatuuronderwijs zou moeten zijn dat leerlingen zich bewust worden van hun eerste, authentieke reacties op literaire teksten, en dat zij leren deze onder woorden te brengen. Met dit idee hebben we geprobeerd een antwoord te formuleren op de potentiële kritiek dat TDL geen "echt literatuuronderwijs" zou zijn: juist door de aandacht voor initiële reacties op literaire teksten, evenals voor het onderbouwen van die reacties met verwijzingen naar literaire elementen in teksten, kan TDL als een *literatuurdidactiek* onderscheiden worden van andere onderwijsdomeinen. De initiële reacties van leerlingen als vertrekpunt nemen impliceert daarnaast dat er in het literatuuronderwijs van de toekomst wellicht minder ruimte is voor monologische, sterk docentgestuurde praktijken.

Deze implicaties roepen ten minste drie vragen op. Hoe kan de implementatie van TDL en de onderliggende ontwerpregels zijn weg vinden naar de onderwijspraktijk? We hebben aangegeven dat er wellicht professionaliseringstrajecten voor docenten-in-opleiding en zittende docenten ontwikkeld moeten worden, mogelijk in de vorm van docent-ontwikkelteams (DOTs). Ten tweede hebben we vragen met betrekking tot toetsing en beoordeling benoemd. Formatieve evaluatie past wellicht het best bij TDL, waarbij de ontwikkelde rubric en geschreven dialoogtaak behulpzaam kunnen zijn, evenals het leerverslag. Tot slot hebben we de vraag gesteld op welke teksten en in welke andere onderwijsdomeinen de principes van TDL zouden kunnen worden toegepast. We hebben het potentieel van adolescenten- en young adult-literatuur besproken en hebben opgemerkt dat, hoewel ze thematisch passend kunnen zijn, de taal in historische literaire teksten en literaire teksten die bij vreemde talen aan bod komen de implementatie van TDL in de weg zouden kunnen staan. Ten slotte hebben we besproken dat TDL mogelijkheden biedt voor domeinoverstijgende praktijken in het onderwijs, door literatuuronderwijs te verbinden met andere domeinen waarin "de aard van de mens" een rol speelt, zoals geschiedenis, burgerschapsonderwijs, maatschappijleer en kunst.



## APPENDICES

APPENDIX A	Search syntax per database	261
APPENDIX B	Overview of studies included in review	267
APPENDIX C	Description of TDLT-1	275
APPENDIX D	Description of TDLT-2	277
APPENDIX E	Descriptive data and coding schemes to Ch. 5	281
APPENDIX F	Descriptive data and coding schemes to Ch. 6	287



## APPENDIX A. SEARCH SYNTAX PER DATABASE

### ERIC

#### Cluster 1. Secondary education

(grade 9 OR grade 10 OR grade 11 OR grade 12 OR high schools OR secondary education).el. OR adolescents/ OR adolescent development/ OR grade 9/ OR grade 10/ OR grade 11/ OR grade 12/ OR secondary school students/ OR high school students/ OR (9th-grade\* OR ninth-grade\* OR grade 9 OR grade nine OR 10th-grade\* OR tenth-grade\* OR grade 10 OR grade ten OR 11th-grade\* OR eleventh-grade\* OR grade 11 OR grade eleven OR 12th-grade\* OR twelfth-grade\* OR grade 12 OR grade twelve OR adolescen\* OR highschool\* OR high school\* OR preuniversity OR pre-university OR secondary education OR secondary school OR student\* OR classroom\* OR young adult\*).ti,ab.

#### Cluster 2. Texts and literature teaching

adolescent literature/ OR "classics (literature)"/ OR english literature/ OR literary genres/ OR literature/ OR literature appreciation/ OR novels/ OR poetry/ OR prose/ OR twentieth century literature/ OR united states literature/ OR world literature/ OR (adolesc\* literature OR fiction OR literary OR literature class\* OR literature education OR literature lesson\* OR multicultural literature OR multi-ethnic literature OR novels OR poem\* OR poetry OR prose OR short stor\* OR teaching literature OR young adult literature).ti,ab.

#### Cluster 3. Learning about self

"adjustment (to environment)"/ OR aspiration/ OR coping/ OR emotional experience/ OR emotional response/ OR emotional development/ OR empowerment/ OR ethnicity/ OR persistence/ OR personal autonomy/ OR personality/ OR personality change/ OR personality development/ OR personality traits/ OR self actualization/ OR self concept/ OR self determination/ OR "self disclosure (individuals)"/ OR self esteem/ OR self expression/ OR sexual identity/ OR self motivation/ OR (((character\* OR evalua\* OR identit\* OR image\* OR individual\* OR personal\* OR self) ADJ2 (accept\* OR actuali#ation OR autonomy OR aware\* OR change\* OR concept\* OR confidence OR construct\* OR criticism OR determination OR develop\* OR disclosure OR discover\* OR esteem OR experienc\* OR express\* OR insight\* OR know\* OR learn\* OR perception\* OR reflect\* OR respect\* OR trait\* OR understand\*)) OR coping OR cultural capital OR emotional development\* OR emotional experience\* OR emotional response\* OR empower\* OR future time orient\* OR future consequence\* OR future orientation OR (future ADJ3 time perspective\*) OR (identit\* ADJ3 (academic OR ethnic\* OR formation\* OR gender OR sexual OR social)) OR human capital OR myself OR possible sel\* OR ((self OR selves) ADJ1 (desired OR feared OR future OR ideal OR imagined OR inner OR true OR wished for)) OR reflectiveness OR self motivat\* OR sense of coherence OR social capital OR volition OR yourself).ti,ab.

#### Cluster 4. Learning about others

altruism/ OR behavior standards/ OR cultural awareness/ OR cultural differences/ OR cultural influences/ OR citizenship/ OR citizenship education/ OR empathy/ OR ethical instruction/ OR ethnic stereotypes/ OR gender bias/ OR gender discrimination/ OR gender issues/ OR "identification (psychology)"/ OR interpersonal competence/ OR interpersonal relationship/ OR moral develop-

ment/ OR moral values/ OR perspective taking/ OR racial attitudes/ OR racial bias/ OR racial differences/ OR racial discrimination/ OR racial identification/ OR racial relations/ OR religious discrimination/ OR sex stereotypes/ OR social attitudes/ OR social bias/ OR social class/ OR social cognition/ OR social change/ OR social differences/ OR social discrimination/ OR social experience/ OR social influences/ OR social justice/ OR social problems/ OR social responsibility/ OR social status/ OR social values/ OR stereotypes/ OR "theory of mind"/ OR world views/ OR (altruism\* OR ((antisemitism\* OR antisemit\* OR arab\* OR bisexual\* OR black OR christian\* OR cultural\* OR ethnic\* OR gay OR gender OR homosexual\* OR homoerotic\* OR homophobic\* OR islam\* OR intercultural\* OR jew\* OR lesbi\* OR LGBT\* OR muslim\* OR person OR racial OR religion\* OR sex OR sociocultural\* OR transgender\* OR transsexual\* OR whiteness OR queer\*)) ADJ3 (awareness OR attitude\* OR bias\* OR competence\* OR difference\* OR discrimination OR equality OR issue\* OR norm\* OR perception\* OR perspective\* OR sensitive\* OR value\*)) OR citizenship\* OR cultural knowledge OR compassion OR empathy\* OR ethic\* education OR ethic\* instruction OR human right\* OR identification\* OR ingroup\* OR intercultural knowledge OR ((interpersonal OR social) ADJ1 (awareness OR approval OR bias\* OR competence\* OR class\* OR difference\* OR discrimination OR equality OR issue\* OR norm\* OR perception\* OR perspective OR value\* OR cognit\* OR experience\* OR knowledge OR status\* OR relationship\* OR responsibility\* OR justice OR acceptance OR skill\* OR stigma\* OR comparison\* OR influence\* OR insight\*)) OR moral\* OR outgroup\* OR ((perspective\* OR role\*) ADJ3 tak\*) OR prejudice\* OR racism\* OR sexism\* OR stereotypical\* OR sympathy\* OR taboo\* OR "theory of mind\*" OR tolerance OR (understand\* ADJ3 (character OR characters OR culture\* OR minority\* OR other\* OR people)) OR world view\*).ti,ab.

## PSYCINFO

### Cluster 1. Secondary education

(adolescence 13 17 yrs).ag. OR secondary education/ OR high schools/ OR high school students/ OR high school education/ OR students/ OR (9th-grade\* OR ninth-grade\* OR grade 9 OR grade nine OR 10th-grade\* OR tenth-grade\* OR grade 10 OR grade ten OR 11th-grade\* OR eleventh-grade\* OR grade 11 OR grade eleven OR 12th-grade\* OR twelfth-grade\* OR grade 12 OR grade twelve OR adolescen\* OR highschool\* OR high school\* OR preuniversity OR pre-university OR secondary education OR secondary school OR student\* OR classroom\* OR young adult\*).ti,ab,id.

### Cluster 2. Texts and literature teaching

literature/ or poetry/ or prose/ OR (adolesc\* literature OR fiction OR literary OR literature class\* OR literature education OR literature lesson\* OR multicultural literature OR multiethnic literature OR novels OR poem\* OR poetry OR prose OR short stor\* OR teaching literature OR young adult literature).ti,ab,id.

### Cluster 3. Learning about self

"adaptability (personality)"/ OR affective education/ OR agency/ OR aspirations/ OR assertiveness/ OR conscientiousness/ OR coping behavior/ OR emotional adjustment/ OR emotional development/ OR emotional intelligence/ OR emotional responses/ OR empowerment/ OR ethnic identity/ OR gender identity/ OR goal orientation/ OR identity formation/ OR "independence (personality)"/ OR openness to experience/ OR persistence/ OR personality/ OR personality traits/ OR personality development/ OR reflectiveness/ OR "self monitoring (personality)"/ OR "perceptiveness (personality)"/ OR "emotionality (personality)"/ OR personality change/ OR "sensitivity (personality)"/ OR role models/ OR self actualization/ OR self concept/ OR self confidence/ OR self determination/ OR "self disclosure (individuals)"/ OR self esteem/ OR self expansion/ OR self expression/ OR self per-

ception/ OR "sense of coherence"/ OR social identity/ OR volition/ OR (((character\* OR evalua\* OR identit\* OR image\* OR individual\* OR personal\* OR self) ADJ2 (accept\* OR actuali#ation OR autonomy OR aware\* OR change\* OR concept\* OR confidence OR construct\* OR criticism OR determination OR develop\* OR disclosure OR discover\* OR esteem OR experienc\* OR express\* OR insight\* OR know\* OR learn\* OR perception\* OR reflect\* OR respect\* OR trait\* OR understand\*)) OR coping OR cultural capital OR emotional development\* OR emotional experience\* OR emotional response\* OR empower\* OR future time orient\* OR future consequence\* OR future orientation OR (future ADJ3 time perspective\*) OR (identit\* ADJ3 (academic OR ethnic\* OR formation\* OR gender OR sexual OR social)) OR human capital OR myself OR possible sel\* OR ((self OR selves) ADJ1 (desired OR feared OR future OR ideal OR imagined OR inner OR true OR wished for)) OR reflectiveness OR self motivat\* OR sense of coherence OR social capital OR volition OR yourself).ti,ab,id.

#### Cluster 4. Learning about others

altruism/ OR antisemitism/ OR citizenship/ OR cross cultural differences/ OR cultural sensitivity/ OR ethnic values/ OR empathy/ OR "homosexuality (attitudes toward)"/ OR human rights/ OR ingroup outgroup/ OR interpersonal influences/ OR interpersonal relationships/ OR moral development/ OR morality/ OR prejudice/ OR "race and ethnic discrimination"/ OR "racial and ethnic attitudes"/ OR "racial and ethnic differences"/ OR "racial and ethnic relations"/ OR racism/ OR religious prejudices/ OR role taking/ OR sex discrimination/ OR sexism/ OR sex role attitudes/ OR social class/ OR social cognition/ OR social comparison/ OR social discrimination/ OR social equality/ OR social influences/ OR social issues/ OR social justice/ OR social skills/ OR social norms/ OR social perception/ OR social responsibility/ OR social values/ OR stereotyped attitudes/ OR taboos/ OR "theory of mind"/ OR world view/ OR (altruism\* OR ((anti-semit\* OR antisemit\* OR arab\* OR bisexual\* OR black OR christian\* OR cultural\* OR ethnic\* OR gay OR gender OR homosex\* OR homoerot\* OR homophobic\* OR islam\* OR intercultural\* OR jew\* OR lesbi\* OR LGBT\* OR muslim\* OR person OR racial OR religio\* OR sex OR sociocultur\* OR transgender\* OR transsex\* OR whiteness OR queer\*)) ADJ3 (awareness OR attitude\* OR bias\* OR competence\* OR difference\* OR discrimination OR equality OR issue\* OR norm\* OR perception\* OR perspective\* OR sensitiv\* OR value\*)) OR citizenship\* OR cultural knowledge OR compassion OR empath\* OR ethic\* education OR ethic\* instruction OR human right\* OR identification\* OR ingroup\* OR intercultural knowledge OR ((interpersonal OR social) ADJ1 (awareness OR approval OR bias\* OR competence\* OR class\* OR difference\* OR discrimination OR equality OR issue\* OR norm\* OR perception\* OR perspective OR value\* OR cognit\* OR experienc\* OR knowledge OR status\* OR relationship\* OR responsibilit\* OR justice OR acceptance OR skill\* OR stigma\* OR comparison\* OR influence\* OR insight\*)) OR moral\* OR outgroup\* OR ((perspective\* OR role\*) ADJ3 tak\*) OR prejudice\* OR racis\* OR sexism\* OR stereotyp\* OR sympath\* OR taboo\* OR "theory of mind\*" OR tolerance OR (understand\* ADJ3 (character OR characters OR cultur\* OR minorit\* OR other\* OR people)) OR world view\*).ti,ab,id.

## WEB OF SCIENCE

#### Cluster 1. Secondary education

TS=("9th-grade\*" OR "ninth-grade\*" OR "grade 9" OR "grade nine" OR "10th-grade\*" OR "tenth-grade\*" OR "grade 10" OR "grade ten" OR "11th-grade\*" OR "eleventh-grade\*" OR "grade 11" OR "grade eleven" OR "12th-grade\*" OR "twelfth-grade\*" OR "grade 12" OR "grade twelve" OR "adolescenc\*" OR "highschool\*" OR "high school\*" OR "preuniversity" OR "pre-university" OR "secondary education" OR "secondary school" OR "student\*" OR "classroom\*" OR "young adult\*")

## Cluster 2. Texts and literature teaching

TS=("adolesc\* literature" OR "fiction" OR "literary" OR "literature class\*" OR "literature education" OR "literature lesson\*" OR "multicultural literature" OR "multiethnic literature" OR "novels" OR "poem\*" OR "poetry" OR "prose" OR "short stor\*" OR "teaching literature" OR "young adult literature")

## Cluster 3. Learning about self

TS=((("character\*" OR "evalua\*" OR "identit\*" OR "image\*" OR "individual\*" OR "personal\*" OR "self") NEAR/1 ("accept\*" OR "actuali#ation" OR "autonomy" OR "aware\*" OR "change\*" OR "concept\*" OR "confidence" OR "construct\*" OR "criticism" OR "determination" OR "develop\*" OR "disclosure" OR "discover\*" OR "esteem" OR "experiec\*" OR "express\*" OR "insight\*" OR "know\*" OR "learn\*" OR "perception\*" OR "reflect\*" OR "respect\*" OR "trait\*" OR "understand\*")) OR "coping" OR "cultural capital" OR "emotional development\*" OR "emotional experience\*" OR "emotional response\*" OR "empower\*" OR "future time orient\*" OR "future consequence\*" OR "future orientation" OR ("future" NEAR/2 "time perspective\*") OR ("identit\*" NEAR/2 ("academic" OR "ethnic\*" OR "formation" OR "gender" OR "sexual" OR "social")) OR "human capital" OR "myself" OR "possible sel\*" OR (("self" OR "selves") NEAR/0 ("desired" OR "feared" OR "future" OR "ideal" OR "imagined" OR "inner" OR "true" OR "wished for")) OR "reflectiveness" OR "self motivat\*" OR "sense of coherence" OR "social capital" OR "volition" OR "yourself")

## Cluster 4. Learning about others

TS=("altruis\*" OR ("anti-semit\*" OR "antisemit\*" OR "arab\*" OR "bisex\*" OR "black" OR "christian\*" OR "cultural\*" OR "ethnic\*" OR "gay" OR "gender" OR "homosex\*" OR "homoerot\*" OR "homophob\*" OR "islam\*" OR "intercultural\*" OR "jew\*" OR "lesbi\*" OR "LGBT\*" OR "muslim\*" OR "person" OR "racial" OR "religio\*" OR "sex" OR "sociocultur\*" OR "transgender\*" OR "transsex\*" OR "whiteness" OR "queer\*") NEAR/2 ("awareness" OR "attitude\*" OR "bias\*" OR "competence\*" OR "difference\*" OR "discrimination" OR "equality" OR "issue\*" OR "norm\*" OR "perception\*" OR "perspective\*" OR "sensitiv\*" OR "value\*")) OR "citizenship\*" OR "cultural knowledge" OR "compassion" OR "empath\*" OR "ethic\* education" OR "ethic\* instruction" OR "human right\*" OR "identification\*" OR "ingroup\*" OR "intercultural knowledge" OR (("interpersonal" OR "social") NEAR/1 ("awareness" OR "approval" OR "bias\*" OR "competence\*" OR "class\*" OR "difference\*" OR "discrimination" OR "equality" OR "issue\*" OR "norm\*" OR "perception\*" OR "perspective" OR "value\*" OR "cognit\*" OR "experiec\*" OR "knowledge" OR "status\*" OR "relationship\*" OR "responsibilit\*" OR "justice" OR "acceptance" OR "skill\*" OR "stigma\*" OR "comparison\*" OR "influence\*" OR "insight\*")) OR "moral\*" OR "outgroup\*" OR (("perspective\*" OR "role\*") NEAR/2 "tak\*") OR "prejudice\*" OR "racis\*" OR "sexis\*" OR "stereotyp\*" OR "sympath\*" OR "taboo\*" OR "theory of mind\*" OR "tolerance" OR ("understand\*" NEAR/2 ("character" OR "characters" OR "cultur\*" OR "minorit\*" OR "other\*" OR "people")) OR "world view\*")

## LINGUISTICS AND LANGUAGE BEHAVIOR ABSTRACTS (LLBA)

## Cluster 1. Secondary education

ALL("9th-grade\*" OR "ninth-grade\*" OR "grade 9" OR "grade nine" OR "10th-grade\*" OR "tenth-grade\*" OR "grade 10" OR "grade ten" OR "11th-grade\*" OR "eleventh-grade\*" OR "grade 11" OR "grade eleven" OR "12th-grade\*" OR "twelfth-grade\*" OR "grade 12" OR "grade twelve" OR "adolescenc\*" OR "highschool\*" OR "high school\*" OR "preuniversity" OR "pre-university" OR "secondary education" OR "secondary school" OR "student\*" OR "classroom\*" OR "young adult\*")

## Cluster 2. Texts and literature teaching

ALL("adolesc\* literature" OR "fiction" OR "literary" OR "literature class\*" OR "literature education" OR "literature lesson\*" OR "multicultural literature" OR "multiethnic literature" OR "novels" OR "poem\*" OR "poetry" OR "prose" OR "short stor\*" OR "teaching literature" OR "young adult literature")

## Cluster 3. Learning about self

ALL(("character\*" OR "evalua\*" OR "identit\*" OR "image\*" OR "individual\*" OR "personal\*" OR "self") NEAR/1 ("accept\*" OR "actuali#ation" OR "autonomy" OR "aware\*" OR "change\*" OR "concept\*" OR "confidence" OR "construct\*" OR "criticism" OR "determination" OR "develop\*" OR "disclosure" OR "discover\*" OR "esteem" OR "experienc\*" OR "express\*" OR "insight\*" OR "know\*" OR "learn\*" OR "perception\*" OR "reflect\*" OR "respect\*" OR "trait\*" OR "understand\*") OR "coping" OR "cultural capital" OR "emotional development\*" OR "emotional experience\*" OR "emotional response\*" OR "empower\*" OR "future time orient\*" OR "future consequence\*" OR "future orientation" OR ("future" NEAR/2 "time perspective\*") OR ("identit\*" NEAR/2 ("academic" OR "ethnic\*" OR "formation" OR "gender" OR "sexual" OR "social")) OR "human capital" OR "myself" OR "possible sel\*" OR (("self" OR "selves") NEAR/0 ("desired" OR "feared" OR "future" OR "ideal" OR "imagined" OR "inner" OR "true" OR "wished for")) OR "reflectiveness" OR "self motivat\*" OR "sense of coherence" OR "social capital" OR "volition" OR "yourself")

## Cluster 4. Learning about others

ALL("altruis\*" OR ("anti-semit\*" OR "antisemit\*" OR "arab\*" OR "bisex\*" OR "black" OR "christian\*" OR "cultural\*" OR "ethnic\*" OR "gay" OR "gender" OR "homosex\*" OR "homoerot\*" OR "homophob\*" OR "islam\*" OR "intercultural\*" OR "jew\*" OR "lesbi\*" OR "LGBT\*" OR "muslim\*" OR "person" OR "racial" OR "religio\*" OR "sex" OR "sociocultur\*" OR "transgender\*" OR "transsex\*" OR "whiteness" OR "queer\*") NEAR/2 ("awareness" OR "attitude\*" OR "bias\*" OR "competence\*" OR "difference\*" OR "discrimination" OR "equality" OR "issue\*" OR "norm\*" OR "perception\*" OR "perspective\*" OR "sensitiv\*" OR "value\*")) OR "citizenship\*" OR "cultural knowledge" OR "compassion" OR "empath\*" OR "ethic\* education" OR "ethic\* instruction" OR "human right\*" OR "identification\*" OR "ingroup\*" OR "intercultural knowledge" OR (("interpersonal" OR "social") NEAR/1 ("awareness" OR "approval" OR "bias\*" OR "competence\*" OR "class\*" OR "difference\*" OR "discrimination" OR "equality" OR "issue\*" OR "norm\*" OR "perception\*" OR "perspective" OR "value\*" OR "cognit\*" OR "experienc\*" OR "knowledge" OR "status\*" OR "relationship\*" OR "responsibilit\*" OR "justice" OR "acceptance" OR "skill\*" OR "stigma\*" OR "comparison\*" OR "influence\*" OR "insight\*")) OR "moral\*" OR "outgroup\*" OR (("perspective\*" OR "role\*") NEAR/2 "tak\*") OR "prejudice\*" OR "racis\*" OR "sexis\*" OR "stereotyp\*" OR "sympath\*" OR "taboo\*" OR "theory of mind\*" OR "tolerance" OR ("understand\*" NEAR/2 ("character" OR "characters" OR "cultur\*" OR "minorit\*" OR "other\*" OR "people")) OR "world view\*")

## SCOPUS

## Cluster 1. Secondary education

TITLE-ABS-KEY("9th-grade\*" OR "ninth-grade\*" OR {grade 9} OR {grade nine} OR "10th-grade\*" OR "tenth-grade\*" OR {grade 10} OR {grade ten} OR "11th-grade\*" OR "eleventh-grade\*" OR {grade 11} OR {grade eleven} OR "12th-grade\*" OR "twelfth-grade\*" OR {grade 12} OR {grade twelve} OR adolescenc\* OR highschool\* OR "high school\*" OR preuniversity OR {pre-university} OR {secondary education} OR {secondary school} OR student\* OR classroom\* OR "young adult\*")

## Cluster 2. Texts and literature teaching

TITLE-ABS-KEY("adolesc\* literature" OR {fiction} OR {literary} OR {literature class\*} OR {literature education} OR "literature lesson\*" OR {multicultural literature} OR {multiethnic literature} OR {novels} OR poem\* OR {poetry} OR {prose} OR "short stor\*" OR {teaching literature} OR {young adult literature})

## Cluster 3. Learning about self

TITLE-ABS-KEY({coping} OR {cultural capital} OR "emotional development\*" OR "emotional experience\*" OR "emotional response\*" OR empower\* OR "future time orient\*" OR "future consequence\*" OR {future orientation} OR {human capital} OR {myself} OR "possible sel\*" OR {reflectiveness} OR "self motivat\*" OR {sense of coherence} OR {social capital} OR {volition} OR {yourself}) OR TITLE-ABS-KEY((character\* OR evalua\* OR identit\* OR image\* OR individual\* OR personal\* OR {self}) W/1 (accept\* OR actuali?ation OR {autonomy} OR aware\* OR change\* OR concept\* OR {confidence} OR construct\* OR {criticism} OR {determination} OR develop\* OR {disclosure} OR discover\* OR {esteem} OR experienc\* OR express\* OR insight\* OR know\* OR learn\* OR perception\* OR reflect\* OR respect\* OR trait\* OR understand\*)) OR TITLE-ABS-KEY({future} W/2 "time perspective\*") OR TITLE-ABS-KEY(identit\* W/2 ({academic} OR ethnic\* OR {formation} OR {gender} OR {sexual} OR {social})) OR TITLE-ABS-KEY(({self} OR {selves}) W/0 ({desired} OR {feared} OR {future} OR {ideal} OR {imagined} OR {inner} OR {true} OR {wished for}))

## Cluster 4. Learning about others

TITLE-ABS-KEY(altruis\* OR citizenship\* OR {cultural knowledge} OR {compassion} OR empath\* OR "ethic\* education" OR "ethic\* instruction" OR "human right\*" OR identification\* OR ingroup\* OR {intercultural knowledge} OR moral\* OR outgroup\* OR prejudice\* OR racis\* OR sexism\* OR stereotyp\* OR sympath\* OR taboo\* OR "theory of mind\*" OR {tolerance} OR "world view\*") OR TITLE-ABS-KEY(("anti-semit\*" OR antisemit\* OR arab\* OR bisex\* OR {black} OR christian\* OR cultural\* OR ethnic\* OR {gay} OR {gender} OR homosex\* OR homoerot\* OR homophob\* OR islam\* OR intercultural\* OR jew\* OR lesbi\* OR LGBT\* OR muslim\* OR {person} OR {racial} OR religio\* OR {sex} OR sociocultur\* OR transgender\* OR transsex\* OR {whiteness} OR queer\*) W/2 ({awareness} OR attitude\* OR bias\* OR competence\* OR difference\* OR {discrimination} OR {equality} OR issue\* OR norm\* OR perception\* OR perspective\* OR sensitiv\* OR value\*) OR TITLE-ABS-KEY(((interpersonal} OR {social}) W/1 ({awareness} OR {approval} OR bias\* OR competence\* OR class\* OR difference\* OR {discrimination} OR {equality} OR issue\* OR norm\* OR perception\* OR {perspective} OR value\* OR cognit\* OR experienc\* OR {knowledge} OR status\* OR relationship\* OR responsibilit\* OR {justice} OR {acceptance} OR skill\* OR stigma\* OR comparison\* OR influence\* OR insight\*)) OR TITLE-ABS-KEY((perspective\* OR role\*) W/2 tak\*) OR TITLE-ABS-KEY(understand\* W/2 ({character} OR {characters} OR cultur\* OR minorit\* OR other\* OR {people}))

## APPENDIX B. OVERVIEW OF STUDIES INCLUDED IN REVIEW

Study	Characteristics: 1. Participants; 2. Intervention; 3. Methodology; 4. Main findings.
Adler & Foster (1997)	<ol style="list-style-type: none"> <li>1.           <ul style="list-style-type: none"> <li>- <math>N = 57</math></li> <li>- Grade 7, <math>M_{\text{age}} = 13.1</math> years</li> <li>- United States</li> </ul> </li> <li>2.           <ul style="list-style-type: none"> <li>- Category: affect understanding of, views on, or intended behavior toward real-world others (support for the value 'caring for others').</li> <li>- Texts: three novels: <i>Friends are Like That</i> (Hermes, 1984), <i>Red Cap</i> (Wisler, 1984) and <i>The Clay Marble</i> (Ho, 1991). Thematically relevant texts; no reflections on literariness.</li> <li>- Tasks: classroom discussions and exercises to reinforce the theme presented in the books. Exercises included writing favorite quotes and feelings in a journal, making a group collage that expressed the theme and the feelings in each novel, and webbing how the characters in the books were connected to each other.</li> <li>- Duration: 10 weeks (no further information).</li> </ul> </li> <li>3.           <ul style="list-style-type: none"> <li>- Design: experimental with pretest and posttest.</li> <li>- Comparison condition: same activities, but reading three books from the regular 7<sup>th</sup> grade curriculum.</li> <li>- Group sample sizes: not given.</li> <li>- Instruments: completing three essays as pretest and the same three essays as posttest, about caring for a family member, caring for strangers who lost their home to a fire, and friends as stand-in for family. Essays were coded and pre- and posttest essays were compared for change in how caring for others was valued (coding schemes reported).</li> </ul> </li> <li>4.           <ul style="list-style-type: none"> <li>- Empirical support: partial.</li> <li>- Essay about friends: significantly more students in experimental group showed positive change in valuing 'caring for others' than in comparison group. Effect sizes could not be calculated.</li> <li>- No significant differences between groups for the other two essays.</li> </ul> </li> </ol>
Darragh (2015)	<ol style="list-style-type: none"> <li>1.           <ul style="list-style-type: none"> <li>- <math>N = 229</math></li> <li>- Grade 8, <math>M_{\text{age}} = 14</math> years</li> <li>- United States</li> </ul> </li> <li>2.           <ul style="list-style-type: none"> <li>- Category: affect understanding of, views on, or intended behavior toward real-world others (views on disabled people).</li> <li>- Texts: one young adult novel portraying a disabled character, student-selected from <i>Al Capone Does my Shirts</i> (Choldenko, 2004); <i>Things Not Seen</i> (Clements, 2002); <i>Rules</i> (Lord, 2006); <i>Hurt Go Happy</i> (Rorby, 2006); <i>Small Steps</i> (Sachar, 2006); <i>Stuck in Neutral</i> (Trueman, 2000). Two stories read aloud by teacher from <i>Owning It: Stories about Teens with Disabilities</i> (Gallo, 2008). Thematically relevant texts; no reflections on literariness.</li> </ul> </li> </ol>

- Tasks: individual reading of the novel; responding to three writing prompts: 1) during reading, to make text-to-self connections; 2) after reading, to express appreciation of the novel; 3) after listening to short stories, about challenges for characters and comparisons to own life.
  - Duration: 5 weeks (no further information).
- 3.
- Design: experimental with pretest, posttest and delayed posttest.
  - Comparison condition: same writing tasks and read-aloud short stories, but choosing and reading a novel not portraying disabled character.
  - Group sample sizes: not given.
  - Instruments: *Shared Activities Questionnaire* (SAQ; Morgan, Walker, Bieberich & Bell, 1996); *Adjective Checklist* (ACL; Siperstein, 1980; Siperstein & Bak, 1977); responses to writing prompts used for additional analysis (no coding scheme reported).
- 4.
- Empirical support: no.
  - No differences between experimental and control condition.
  - Only exemplary written responses of experimental group discussed.
- Eva-Wood (2004)
- 1.
- $N = 40$
  - Grade 11,  $M_{\text{age}} =$  not given
  - United States
- 2.
- Categories: affect insight into oneself (in personal qualities and own world), affect understanding of fictional others (empathy for characters, understanding poems' speakers).
  - Texts: poems by American authors; no titles mentioned. No particular thematic considerations; reflection on poetry as a literary genre.
  - Tasks: responding to poems by thinking-and-feeling-aloud; observing teacher modeling the strategy; writing a peer's responses; focus on identifying emotions, interpretative questions and comments after reading poetry.
  - Duration: 4 weeks (no further information).
- 3.
- Design: quasi-experimental with pretest and posttest.
  - Comparison condition: literary analysis, using the same poems.
  - Group sample sizes:  $n = 19$  vs.  $n = 21$ .
  - Instrument: *Literary Response Questionnaire*, including relevant scales Insight and Empathy (Miall & Kuiken, 1995).
- 4.
- Empirical support: partial.
  - No differences between experimental and control condition on relevant *LRQ* scales on self-insights and empathy for characters.
  - Analysis of contributions to classroom discussions: more expressions of identification with and understandings of poems' speakers in experimental condition than in control condition. Effect size not applicable.
- Garrod (1989)
- 1.
- $N = 44$
  - Grade 10,  $M_{\text{age}} = 15.9$  years
  - United States
- 2.
- Category: affect understanding of, views on, or intended behavior toward real-world others (understanding moral dilemmas people face).
  - Texts: examples of literary texts given (e.g., *Huckleberry Finn*, *To Kill a Mockingbird*); also other texts used (e.g., drama and expository readings). Texts considered 'literature of high caliber'. No further reflections on themes or literariness.

- Tasks: Socratic discussions about moral dilemmas in texts; small-group and whole-class discussions; writing activities; drama activities. No details reported.
  - Duration: 1 year (no further information).
- 3.
- Quasi-experimental design with pretest and posttest.
  - Comparison condition: Socratic pedagogy using the same material as in experimental condition, without the specified literary texts.
  - Group sample sizes:  $n = 21$  vs.  $n = 23$ .
  - Instrument: written version of *Kohlberg's Moral Judgment Interview*. Reference to scoring manual, no coding scheme reported.
- 4.
- Empirical support: no.
  - No overall difference between experimental and control condition.
  - Subgroup of students with lowest pretest scores: more growth in experimental condition ( $n = 5$ ) than in control condition ( $n = 12$ ).
- Hakemulder (2008)
- 1.
- $N = 99$
  - Grade not given,  $M_{\text{age}} = 16.7$  years
  - The Netherlands
- 2.
- Category: affect understanding of, views on, or intended behavior toward real-world others (views on immigrants).
  - Text: excerpt from multi-cultural novel *Turkish gold* about honour violence among Turkish immigrants (Mukhtar, 2007). Thematical considerations reported, no reflections on literariness.
  - Task: following a reading instruction focused on role-taking.
  - Duration: 1 session (no further information).
- 3.
- Design: experimental, posttest-only.
  - Comparison conditions: reading the same text (diversion instruction, focus on text structure); reading essay about same multi-cultural issue (no instruction); reading unrelated control text (no instruction).
  - Group sample sizes: not given.
  - Instrument: self-constructed task to imagine being a Turkish girl growing up in the Netherlands and scoring five items from that perspective: 'Would you 1) want to choose your own husband? 2) always obey your parents? 3) consider it normal that your parents would punish you physically? 4) accept the way women are treated? 5) resist unequal rights for men and women?' Analysis on item-level (differences of group means).
- 4.
- Empirical support: no.
  - Students in experimental condition and essay condition: lower mean score on item 1 compared to students in unrelated control condition; students in essay condition: lower mean score on item 5 compared to students in unrelated control condition. Effect sizes could not be calculated.
  - No other differences between experimental and other conditions.
- Halász (1991)
- 1.
- $N = 72$
  - Grade not given;  $M_{\text{age}} = 17$  years
  - Hungary
- 2.
- Category: affect insight into oneself (personal memories evoked by reading).
  - Text: short story *The Vulture* (Kafka, n.d.). No reflections on story theme; metaphorical complexity and literariness considered.
  - Tasks: while reading the story, documenting in writing personal memories that

- are evoked (experiment 1); after reading the story, writing down personal memories evoked by given salient words from the text (experiment 2).
- Duration: 1 session without time limit.
3.
    - Design: experimental, posttest-only.
    - Comparison conditions: same task while reading an essay, same task while reading an expository text, all texts featured a bird of prey (experiment 1); same task after reading essay and expository texts, or same task without reading the texts first (experiment 2).
    - Group sample sizes: three groups of  $n = 24$  (experiment 1); six groups of  $n = 24$  (experiment 2).
    - Instrument: written responses qualitatively analyzed. Responses coded as primary (actor, observer or hearsay) or secondary sources (fiction, non-fiction).
  4.
    - Empirical support: full.
    - Literary reading condition evoked more affective and personal responses than expository and essay conditions, in both experiments; in experiment 2, responding to salient words after reading evoked more affective and personal responses than responding without reading first. Effect sizes could not be calculated.
    - Responses to literary text mostly refer to secondary sources (fiction, music, art).
- Henschel,  
Meier &  
Roick (2016)
1.
    - $N = 226$
    - Grade 9,  $M_{\text{age}} = 15.8$  years
    - Germany
  2.
    - Category: affect understanding of fictional others (empathy for characters).
    - Texts: three short literary texts. Excerpt from the *Daughterlove* (Von Schach, 1988), the song text *The Huge Awakening* (Louisian, 2005) and the short story *They are Eating a Schnitzel* (Kästner, 1923). No reflections on text theme; texts considered 'literary' without explanation.
    - Tasks: reading each text was followed by completing creative, reader-oriented response tasks; focused on imagination, construction, transformation, e.g., writing how to transform story into movie; writing what protagonist would say; depicting metaphor in drawing; working both individually and in pairs, both verbally and in writing.
    - Duration: 3 sessions of each 45 minutes, on one day.
  3.
    - Design: experimental with pretest (after reading but prior to tasks) and posttest (after tasks).
    - Comparison conditions: reading the same texts, followed by completing text-based, analytical tasks; reading the same texts without completing any tasks (the last group only completed the posttest).
    - Group sample sizes:  $n = 81$  vs.  $n = 84$  vs.  $n = 61$ .
    - Instrument: three items measuring the affective dimension of fantasy empathy (based on Davis, 1983).
  4.
    - Empirical support: no.
    - No difference between reader-based and text-based comparison condition on posttest; control condition not included in analysis.
    - Mean score of experimental group increased from pretest to posttest; however, comparison group had higher pretest mean score, which decreased over time; no conditional difference at posttest.

- Malo-Juvera (2014)
1. -  $N = 139$   
- Grade 8,  $M_{\text{age}} = 14.3$  years  
- United States
  2. - Category: affect understanding of, views on, or intended behavior toward real-world others (views on sexual harassment behavior).  
- Text: young-adult novel *Speak* (Anderson, 1999) about particular form of sexual harassment (date rape). Thematically relevant; no reflection on literariness.  
- Tasks: four dialogic sequences consisting of: 1) individual response writing, 2) small-group dialogues with minimal teacher interference and 3) teacher-led whole-class conversations. Two sequences after finishing particular excerpts; two after finishing the novel.  
- Duration: 12 lessons of 1 hour and 45 minutes, during 5 weeks.
  3. - Design: quasi-experimental with pretest and posttest.  
- Comparison condition: same activities, using text by Shakespeare.  
- Group sample sizes:  $n = 82$  vs.  $n = 57$ .  
- Instrument: *Adolescent Rape Myth Scale* (self-developed, based on previously validated instruments); two factors of beliefs about rape: 'She wanted it' and 'She lied'.
  4. - Empirical support: full.  
- Main effect of condition (with pretest scores as covariate;  $d = .84^a$ ); lower acceptance scores on posttest in experimental condition.  
- Effect mainly due to items on factor 'She wanted it' (representing that a rape victim asked for it).
- Malo-Juvera (2016)
1. -  $N = 138$   
- Grade 8,  $M_{\text{age}} = 14.1$  years  
- United States
  2. - Category: affect understanding of, views on, or intended behavior toward real-world others (views on sexual orientation, reduce homophobia).  
- Text: young-adult novel *Geography Club* (Hartinger, 2003) about the coming-out of a homosexual teenager. Thematic considerations; no reflection on literariness.  
- Tasks: see Malo-Juvera (2014). Seven dialogic sequences; one before reading, three after finishing particular excerpts, three after finishing the novel.  
- Duration: 5 weeks (no further information)
  3. - Design: quasi-experimental with pretest and posttest.  
- Comparison condition: waitlist (no further information).  
- Group sample sizes: not given.  
- Instrument: *Adolescent Homophobia Index* (self-developed). Three factors of beliefs about homosexuality: 'Interpersonal homophobia', 'Attitudes toward homophobia' and 'Sexuality as a choice'.
  4. - Empirical support: full.  
- Main effect of condition (with pretest scores as covariate;  $d = .87^a$ ); lower homophobia scores on posttest in experimental condition.  
- Effect mainly due to items on 'Attitudes toward homophobia' and 'Sexuality as a choice'.

- Stevahn et al. (1997)
1. -  $N = 40$   
- Grade 9,  $M_{\text{age}}$  = not given  
- Canada
  2. - Category: affect understanding of, views on, or intended behavior toward real-world others (improve conflict resolution strategies).  
- Aim: improve conflict-resolution strategies.  
- Text: coming-of-age novel *Crabbe* (Bell, 1986). Conflicts in novel considered thematically relevant; no reflection on literariness.  
- Tasks: Peacemakers program with focus on conflicts in novel. Discussing which common conflict teenagers face, identifying conflicts in chapters, discussing and enacting how characters would use conflict resolution strategy, observing teacher modeling the resolution strategy, observing peers enacting a conflict resolution.  
- Duration: 8 sessions, total of 9 hours and 30 minutes.
  3. - Design: quasi-experimental with pretest, posttest, delayed posttest.  
- Comparison condition: studying the same novel traditionally, with written notebook entries and classroom discussions, without focus on conflict resolution.  
- Group sample sizes:  $n = 20$  vs.  $n = 20$ .  
- Instruments: achievement test about the novel, with answers coded as literal, interpretive or insightful (posttest only); *How I Manage Conflicts Measure*, with number of correctly recalled negotiation steps scored (0-6) (Johnson, Johnson & Dudley, 1992; posttest and delayed posttest); *Conflict Scenario Written Measure*, with application of strategies scored on two dimensions: 'most destructive' to 'most constructive' (0-12) and 'forcing' to 'negotiating' (1-5) (pretest and posttest).
  4. - Empirical support: full.  
- Higher scores on both conflict negotiation measures in experimental condition than control condition, both on posttests and delayed posttests (no difference between conditions on pretest;  $d's > 1.00^b$ ).
- Stevahn et al. (1996)
1. -  $N = 111$   
- Grade 7 and 8,  $M_{\text{age}}$  = not given  
- Canada
  2. - Category: affect understanding of, views on, or intended behavior toward real-world others (improve conflict resolution strategies).  
- Text: historical fiction novel *Days of Terror* (Smucker, 1979). Conflicts in novel considered thematically relevant; no reflection on literariness.  
- Tasks: Peacemakers program with focus on conflicts in novel, implemented in two versions. Cooperative training: role-playing negotiation strategies using conflicts from the novel; negotiation lesson related to event in the novel; conceptual analysis and reflection on the novel; explaining analyses to peers (other activities mentioned but no further information given). Individualistic training: similar focus, but with writing out conflict and strategies and explaining to teacher instead of peers.  
- Duration: 4 weeks, total of 22 hours and 50 minutes.
  3. - Design: experimental with pretest, posttest, delayed posttest.  
- Comparison conditions: studying the novel cooperatively without a focus on conflicts (see Stevahn et al., 1997); studying the novel individualistically without a focus on conflicts.

- Group sample sizes:  $n = 29$  vs.  $n = 27$  vs.  $n = 27$  vs.  $n = 27$ .
  - Instruments: see Stevahn et al. (1997); plus *Conflict Word Association Measure* (Dudley, Johnson, & Johnson, 1995).
4. - Empirical support: full.
- All measures<sup>c</sup>: main effects of training and cooperative learning, as well as interaction effects; e.g., better achievement on novel test after training than non-training ( $d = .99^b$ ) and after cooperative training than individual training ( $d = .77^b$ ); interaction effect ( $d = .45^b$ ); more negotiation steps after training than non-training ( $d = 1.89^b$ ) and after cooperative training than individual training ( $d = 1.59^b$ ); interaction effect ( $d = 1.58^b$ ).
  - Delayed posttests for all measures: effects sustained, smaller effect sizes.
- Vezzali et al. (2012)
1. -  $N = 96$
  - Grade not given;  $M_{\text{age}} = 12.8$  years
  - Italy
  2. - Category: affect understanding of, views on, or intended behavior toward real-world others (views on immigrants).
  - Texts: reading a book with an intercultural theme, featuring immigrant characters, chosen from a school-assigned reading list. One example given: *Le Nuvole da Latte* (Frescura, 2002). Thematic considerations; no reflections on literariness.
  - Tasks: read the book as homework during summer holidays; writing a summary; evaluating how much you liked the book, to which extent it was interesting, pleasant, and difficult.
  - Duration: not indicated.
  3. - Design: quasi-experimental, posttest only.
  - Comparison conditions: reading a non-intercultural book chosen from a list and completing the same activities; no reading or activities at all.
  - Group sample sizes:  $n = 33$  vs.  $n = 33$  vs.  $n = 30$ .
  - Instruments: *Word Association Task*: ingroup and outgroup stereotypes (self-developed); *Intergroup Attitudes* (Liebkind & McAlister, 1999; adapted); *Hypothetical Contact Scenario*: behavioral intentions toward outgroups (Cameron & Rutland, 2006); *Desire for Future Contact* (Tropp & Bianchi, 2007; adapted); *Closeness to Outgroup* (Aron et al., 1992; adapted); *Ingroup Identification* (Capozza, Brown, Aharpour & Falvo, 2006; adapted).
  4. - Empirical support: full.
  - Experimental vs. other conditions combined: on all measures, higher scores in experimental condition ( $d$  ranging from .56 on Closeness to outgroup, to 1.16 on Intergroup attitudes).
  - Experimental vs. non-intercultural reading condition: on all measures, higher scores in experimental condition ( $d$  ranging from .52 on Closeness to outgroup, to 1.22 on Intergroup attitudes).
  - Lower ingroup identification scores in experimental condition, but unclear if this was due to the intervention.
- White (1995)
1. -  $N = 83$
  - Grade 9,  $M_{\text{age}} =$  not given
  - United States
  2. - Category: affect understanding of fictional others (understanding characters and their behavior).

- Texts: two short stories, *Indian Camp* and *The End of Something* (Hemingway, 1925). Thematical considerations and reflections on literariness of stories, in terms of writing style.
- Task: autobiographical writing before reading; writing tasks asked students to explore multiple perspectives and allowed for drawing from secondary sources (i.e., seen or heard instead of experienced).
- Duration: no information given.
- 3. - Design: quasi-experimental, switching replications, posttest only.
- Comparison condition: no autobiographical writing before reading
- Group sample sizes not given.
- Instrument: whole-class discussion after reading; analysis of students' responses, coded as off-task, contentless, describing or abstracting.
- 4. - Empirical support: full.
- After autobiographical writing as pre-reading activity, discussion responses showed more abstract character descriptions (moving beyond literal level), more engagement, and more elaboration. Effect sizes could not be calculated.

*Note.* <sup>a</sup> Effect size transformed from  $\eta^2$ . <sup>b</sup> Effect size calculated based on reported statistics and group sample sizes. <sup>c</sup> Stevahn et al. (1996) report twelve tables with results for four conditions on three measurement moments; only a few examples are summarized here.

## APPENDIX C. DESCRIPTION OF TDLT-1

### UNIT 1: PREPARATION

1. Activating prior knowledge and personal experiences: students were asked to write down previous knowledge about dialogues: what characterized a good and less good dialogue?
2. Observational learning: after watching a video of a non-constructive dialogue, the teacher led a classroom talk to evaluate the dialogue and discuss the video dialogue could have gone better. This led up to explicating dialogue guidelines. The guideline of using follow-up questions was explicated further: students observed their teacher who modeled the use of follow-up questions, in a dialogue about reading experiences with one of the students.  
*Intermezza.* students read a few sentences and a description of five stories, and indicated which two they would like to read best during lesson 3.
3. Practicing a small-group dialogue: students were asked to talk in groups of three or four about which considerations they take into account when choosing books to read, for school and at home. To practice the use of follow-up questions, they used cards on which such questions were printed (e.g., 'Can you give an example?' 'Could you explain that?' 'Could someone else have a different view?')
4. Exchange in class: the teacher asked the various group to share what they talked about in their groups and attended to different points of view that occurred.
5. Reflection: students were asked to write down what went well and what could have gone better in their small-group dialogues; a few students was called upon to share their reflection in class.

### UNIT 2: READING-AND-DIALOGUE

Story: *She was everywhere*, Ed van Eeden)

1. Activating prior knowledge about and personal experiences with a story theme: prior to reading, students were asked to write down their thoughts about how someone might react when a relationship ends, a theme that occurred in the story.
2. Internal dialogue with the story: students were asked to focus on the responses the story evoked in them during reading. They could take notes, but were not required to do so. Directly after reading, they were asked to indicate to which extent they had noticed experiences such as imagery, identification and sympathy. In this way, they determined what kind of reading experience was prominent to them, to prepare for external dialogues.
3. Dialogue in small groups: students were grouped according to their most prominent reading experience, to explore this experience more in-depth. For instance, those who had indicated that they felt sympathy for a character were asked to compare the moments in the story where each of them experienced this, and to share in their group what they thought and felt at those moments. Next, they were asked to brainstorm about what kind of help would be of avail to the protagonist. As a third step, they reached a conclusion about what kind of help they would offer the protagonist, by talking about issues like: how feasible would the ideas be? What

would be best for the protagonist? How would you take action? How would the protagonist respond? Students were asked to take notes of dialogic tasks; in this case, for instance, one student in the group would write down the ideas that emerged during the brainstorm.

4. Dialogue at classroom level: students shared their small-group conclusions in class and experienced that other groups had explored different experiences.
5. Reflection: students were asked to consider whether they could also have explored a different kind of reading experience, now that they heard the conclusions of other groups.

### UNIT 3: READING-AND-DIALOGUE

Stories to choose from: *A plate with spaghetti*, Adriaan van Dis; *The freshwater steak*, Hans Dorresteyn; *The right*, Annelies Verbeke; *Blood*, Gerard Reve; *Curious story*, Elke Geurts

1. Activating prior knowledge about and personal experiences with a story theme: based on students' indications of which stories they would like to read (see unit 1), the teacher assigned them to story groups. Each group was given a thematical statement prior to reading the story. Students were asked to write a short response to this statement and talk shortly about their responses in their group.
2. Internal dialogue with the story: reading instruction similar to unit 2; reflection task after reading was worded and organized slightly differently.
3. Dialogue in small groups: students were asked to talk about which moments and events they thought were most important in the story they read and to summarize these events by sketching a story board. This enabled them to exchange their interpretations of what happened in the complex social situations in the stories. They were then asked to formulate as a group a 'life lesson' or 'worldly wisdom' based on the story they just read.
4. Dialogue at classroom level: group representatives were asked to present their life lesson. After all groups had exchanged them, the class voted for the most inspiring one.
5. Reflection: short individual written reflection about whether students' original opinion about the story theme (see phase 1) had changed: if so, how, and if not, why?

### UNIT 4: READING-AND-DIALOGUE

Story: *Flight behavior*, Bertram Koeleman

1. Activating prior knowledge and personal experiences: students were asked to write down their ideas of a possible afterlife, a theme that occurred in the story.
2. Internal dialogue with the story: Students were instructed to focus during reading on the responses the story evoked in them, as they had practiced in the two previous units. Directly after reading, students were asked to write down as many short responses to the story as possible, next to the story.
3. Dialogue in pairs: in a speed date activity, students were asked to talk in three rounds of several minutes about their responses to the story and the meaning of the story end. Guiding questions were given on a PowerPoint slide and prompted students to talk about aspects of transformative reading and elements in the story that evoked their reading experiences (e.g., 'Talk about whether you could picture in your mind what happened in the story. Which story elements caused this?'). They took notes of the dialogues.
4. Reflection and dialogue at classroom level: the unit – and thereby the intervention – was finished by a classroom dialogue about what students felt they learned from the project, after they had written down their reflections individually.

## APPENDIX D. DESCRIPTION OF TDLT-2

Numbers in superscript correspond with phases mentioned in Table 6.2, see p. 162.

### UNIT 1: PREPARATION

Story: opening excerpt from *Flight behavior*, Bertram Koeleman

- Introduction: teacher introduces 'literature': how can we define it? Why it is attended to in Dutch class? Introduction of TDLT goals: students reflect on their starting level of four main objectives, using a rubric (see p. 280).
- <sup>[1]</sup> Internal dialogue (implicit): teacher reads opening excerpt from the story aloud, students read along on paper; no particular reading instruction.
- <sup>[2]</sup> Learning-by-observation: students watch two videos of peers modeling dialogues about the excerpt, write down strengths and weaknesses of observed dialogues, exchange these in class.
- <sup>[3]</sup> Explicit instruction: teacher introduces guidelines for external dialogues (i.e., dialogue strategy) and discusses the 'first aid card'.
- <sup>[4]</sup> External dialogue (practice): students talk in small groups about famous quotes about literature and reading, applying the dialogue strategy.
- Homework: students write about a situation characterized by 'injustice'.

### UNIT 2: READING-AND-DIALOGUE

Story: *Death*, Martin Bril

- Introduction: students reflect shortly, in writing, on what they learned in unit 1; they exchange this in pairs. The teacher introduces the goals of unit 2.
- <sup>[1]</sup> External dialogue applied to theme: students talk in pairs or small groups (chosen by the teacher) about their 'injustice' homework: what is unjust or unfair about the situation? They individually write down a definition of injustice.
- <sup>[2]</sup> Internal dialogue (implicit): the teacher reads the story aloud, students read along on paper; no particular reading instruction.
- <sup>[3]</sup> External dialogue applied to story: students talk in small groups about injustice in the story, about their opinion about the story, and support their opinion with references to literary devices. They create a small poster to summarize their results.
- Reflection: students reflect individually on their own and the groups' application of the dialogue guidelines.

### UNIT 3: READING-AND-DIALOGUE

Story: *Blood*, Gerard Reve

- Introduction: teacher introduces the goals of the unit; connects these to previous units.

- <sup>[1]</sup> Preparation for internal dialogue: students respond in writing to a moral statement (relevant to the theme of the story) and share ideas in pairs; the teacher calls upon a few students to share their ideas in class.
- <sup>[2]</sup> Explicit instruction about internal dialogue: the teacher explains how readers can focus on noticing their responses, how they can annotate these, and why that is important for sharing responses with others.
- <sup>[3]</sup> Learning-by-observation: students observe the teacher, who reads the first part of the story while thinking aloud; the teacher models 'noticing and annotating responses'.
- <sup>[4]</sup> Internal dialogue: students are instructed to notice and annotate their responses when reading the second part of the story; they reflect on their responses individually, using statements about transformative reading experiences; they indicate which reading experience was most prominent (e.g., imagery, experience-taking).
- <sup>[5]</sup> External dialogue (small group): students deepen a prominent transformative reading experience in a small-group dialogue, e.g., for imagery, they talk about what the characters and setting would look like if the story were transformed into a movie.
- <sup>[6]</sup> External dialogue (class): the groups present the outcomes of their dialogues; other students listen and write down at least one question for each group; students are randomly called upon by the teacher to ask their question to the presenting group.

#### UNIT 4: READING-AND-DIALOGUE

Story: *Following the rules*, Mirjam Bonting

- Introduction: teacher introduces the goals of the unit; connects these to previous units.
- <sup>[1]</sup> Internal dialogue: students are instructed to notice and write down their responses in their own way when reading the first part of the story; they reflect on their responses individually, using statements about transformative reading experiences; they indicate which reading experience was most prominent (e.g., imagery, experience-taking).
- <sup>[2]</sup> External dialogue (small groups): students imagine the perspective of the protagonist of the story and talk in a small group about how the story might end, coming up with as many options as possible, and about whether those ends are just or unjust, and for whom.
- <sup>[3]</sup> Internal dialogue: in response to the story, students individually write a story end and a short reflection on which literary devices they used and how (depending on scheduling on of the units, this is a homework task).
- <sup>[4]</sup> External dialogue (pairs): students exchange their story ends; they write down and talk about feedback on each other's story end.
- <sup>[5]</sup> Internal dialogue: the teacher reads aloud the end of the story, students read along on paper and are instructed to notice and annotate their responses.
- <sup>[6]</sup> External dialogue (classroom): teacher-led dialogue about justice and injustice of the different story ends.

#### UNIT 5: READING-AND-DIALOGUE

Stories from units 2, 3 and 4

- Introduction: teacher introduces the goals of the unit; connects these to previous units.
- <sup>[1]</sup> Internal dialogue (recall): students are asked to browse through the stories and their annotations from previous units.

- <sup>[2]</sup> External dialogue (small group): students compare the responses these stories evoked in them and how the theme of injustice was addressed; comparisons are visualized on worksheet. Based on the comparisons of the stories, students formulate a life lesson.
- Reflection: students reflect individually on their own and the groups' application of the dialogue guidelines.
- <sup>[3]</sup> External dialogue (classroom): the teacher selects some of the life lessons and guides a classroom dialogue about them.

## UNIT 6: READING-AND-DIALOGUE

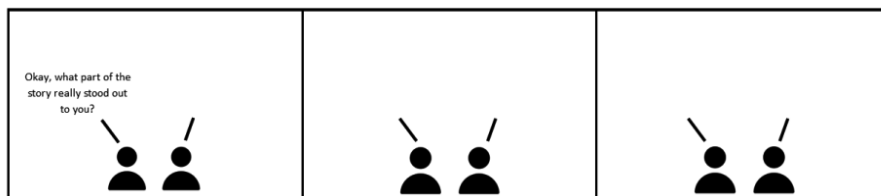
Story: *Count oneself lucky*, Marga Minco; stories to choose from for final task: *The freshwater steak*, Hans Dorresteyn; *An autumn day*, Thomas Heerma van Voss; *The right*, Annelies Verbeke; *Hula*, Cees Nooteboom

- Introduction: teacher introduces the goals of the unit; connects these to previous units.
- Explicit instruction: the teacher explains examples from research into how fiction (reading and tv shows) can influence empathy.
- <sup>[1]</sup> Internal dialogue: students are instructed to notice and write down their responses in their own way when reading the story.
- <sup>[2]</sup> External dialogue (pairs): students engage in a speed date activity: they enact a dialogue from the story (round 1), talk about how empathetic a character is (round 2), about literary devices and how they influence their reading experience (round 3), and about injustice in this story compared to the other stories (round 4).
- Reflection: students evaluated their progress during TDLT by again filling in the rubric (see p. 280).
- <sup>[3]</sup> Homework: internal dialogue: final TDLT task.

## FINAL TDLT TASK

Students selected one story to read; during reading, they annotated their responses. They were then asked to write a dialogue with an imaginary peer, in the form of a comic. The instruction read:

*Imagine you are having a dialogue about the story with a classmate. You talk, for example, about how you experienced the story, about its theme, the characters, things you found unclear... Write this dialogue on the next pages, as a comic (see below). You start with the sentence that is already given. Try to make it a real dialogue, not a question-and-answer interview. Use at least two pages.*



## SELF-EVALUATION RUBRIC

Students could select one of the three levels, or a point in between two levels. In unit 1, they indicated their starting levels; in unit 6, they evaluated their progress.

Noticing responses while reading	When I read a story, I mainly notice whether I understand the story (comprehension level) and what my opinion is about the story, such as fun, boring or exciting (evaluative level).	When I read a story, I notice my responses on the comprehension and evaluative level, but I also pay attention to whether the story evokes any feelings in me, such as compassion, horror, outrage or sadness (emotional level).	When I read a story, I notice my responses on the comprehension, evaluative and emotional level, but I also pay attention whether the story offers me new insights in myself, in others, in life or in what literature is (insight level).
Dealing with difficulties while reading	When I read a story, I am not actively focussing on whether I understand all of it: I just continue reading and think afterwards about what the story might mean. If I really don't understand it, I quit reading.	When I read a story, I notice during reading whether there are things that I find unclear, strange or difficult, but I continue reading. After reading, I consider whether I understood the story, or I discuss it with others.	When I read a story, I'm actively focusing on parts that I find unclear, strange or difficult. When I come across one, I stop reading for a moment and think about what it might mean. If I really can't come to a conclusion, I ask for help.
Gaining insights in reading experiences	I haven't read that much yet, so I cannot describe very well which literary characteristics I (don't) appreciate and what kind of reading experiences fit me.	I can describe, up to a certain extent, which literary characteristics I (don't) appreciate. For example: I like psychological tension; I don't like flash backs and flash forwards.	I can describe quite well which literary characteristics I (don't) appreciate and connect this to the kind of reading experiences that fit me. For example: I like it when focalization switches, because then I experience the story from the point of view of various characters.
Talking actively about stories and reading experiences	In dialogues about stories, I usually don't have that much to say about what I read. I mostly listen to what others have to say.	In dialogues about stories, I talk about what I read and how I experienced it. I listen to others and sometimes ask them a question about their ideas and experiences.	In dialogues about stories, I actively focus on the content of the dialogue. I ask others how they experienced reading the story and compare it to my own experience. During the dialogue, I consider things from multiple perspectives.

APPENDIX E. DESCRIPTIVE DATA AND CODING SCHEMES  
TO CHAPTER 5

CORRELATIONS BETWEEN TREQ SCALES

Measurement and scale		1	2	3	4	5	6	7	8
T1	1 Imagery	-	.40	.48	.39	.31	.29	.21	.19
	2 Identification		-	<b>.66</b>	.55	<b>.64</b>	.55	<b>.65</b>	<b>.60</b>
	3 Experience-taking			-	.40	<b>.61</b>	.38	.40	.39
	4 Character evaluation				-	.44	.38	.38	.41
	5 Sympathy					-	.44	.52	.49
	6 Aesthetic awareness						-	.56	.52
	7 Self-insights							-	<b>.74</b>
	8 Insights into others								-
T2	1 Imagery	-	.47	.56	.47	.44	.23	.26	.28
	2 Identification		-	<b>.68</b>	.58	<b>.69</b>	.55	<b>.70</b>	<b>.70</b>
	3 Experience-taking			-	.48	<b>.68</b>	.38	.44	.45
	4 Character evaluation				-	.51	.34	.42	.49
	5 Sympathy					-	.45	.54	.56
	6 Aesthetic awareness						-	.58	.52
	7 Self-insights							-	<b>.83</b>
	8 Insights into others								-
T3	1 Imagery	-	.33	.59	.42	.42	.23	.19	.20
	2 Identification		-	<b>.61</b>	.53	<b>.67</b>	.58	<b>.70</b>	<b>.68</b>
	3 Experience-taking			-	.44	<b>.66</b>	.33	.42	.39
	4 Character evaluation				-	.50	.38	.38	.47
	5 Sympathy					-	.46	.51	.50
	6 Aesthetic awareness						-	.58	.49
	7 Self-insights							-	<b>.80</b>
	8 Insights into others								-

*Note.* All correlations are significant ( $p < .01$ , \*  $p < .05$ ). Correlations  $\geq .60$  are printed in bold.

## CODING OF STORY RESPONSES

Variable	Type	Codes	Agreement
Extensiveness of story responses	Scale	0. No responses; 1. Non-extensive: one to four short responses of one or several words; 2. Medium extensive: five to ten short responses, or up to four phrases or sentences; 3. Extensive: more than ten short responses, or more than four phrases or sentences.	90.9% $\kappa = .88$
Story responses (content)	Categorical	<ul style="list-style-type: none"> <li>- Self-related responses: explicit connections between reader and story, such as personal associations, identification, or signs of imagery ('Makes me think about...'; 'I picture it in my mind').</li> <li>- Socially-related responses: references to characters' thoughts, behaviors or relations ('He really cares for her') or people in real world ('Why do people get married at all?').</li> <li>- Emotional-evaluative responses: references to emotions, either felt by the reader and/or ascribed to a character ('Sad'), or evaluations of the story or part of it ('Boring').</li> <li>- Story-related responses: descriptive references to literary or stylistic features ('Metaphor') or story events ('He is running away').</li> <li>- Questions or confusions: superficial or more substantive questions ('Why does this happen?'), or expressions of incomprehension or confusion ('Huh?'; 'What does this mean?').</li> </ul>	$\kappa = .74$ (multiple codes per response taken into account)*
Message (question 1)	Scale	1. No message: answer box left blank, and responses such as 'It has no message for me'. 2. Literal inferences, if responses focused in a literal way on story events ('You should not go swimming by yourself' in a story where someone drowns). 3. Basic life lessons, if story events are transformed into a guideline for students' own lives and lives of others, as a short statement ('Always be honest to people'). 4. Life lessons of a more sophisticated nature, for life lessons with meaningful elaborations such as an explanation, rationale or additional perspective on the issue ('You should always be honest to people, or it may have serious consequences').	83.3% $\kappa = .78$

Prediction (question 2)	Scale	<ol style="list-style-type: none"> <li>0. No prediction: answer box left blank, and responses such as 'No idea what'll happen to them'.</li> <li>1. Naive or conformist predictions, for 'happy endings' ('Everything will turn out fine for them'), and confirmations of a story event ('The father will drown' when the protagonist tries to tell his mother exactly that).</li> <li>2. Predictions about events and/or characters' behavior: if responses focus on what may happen in the story and/or what characters may do, without attending to their inner lives ('They will not talk to each other and the wife will leave her husband').</li> <li>3. Predictions also including characters' thoughts and/or emotions: if students make inferences about what characters may think and/or how they may feel because of predicted events and behaviors ('Even though it's hard and it hurts her a lot, she will leave him in the end to choose for herself').</li> </ol>	83.3% $\kappa = .76$
Values of stories (question 3)	Categorical  Single code per response	<ul style="list-style-type: none"> <li>- Insights into self and own life: how they are, think, behave, live their lives, position themselves in the world;</li> <li>- Insights into others and their lives: how other people are, think, behave, live their lives, and position themselves in the world;</li> <li>- Insights into experiences of characters: understand what characters 'go through', identify with them, and put themselves in their situation;</li> <li>- Insights into how stories evoke or influence emotions, actions or reactions: they make readers feel some-thing, help them to deal with, come to terms with, let go of, or decide some-thing;</li> <li>- Insights into the world and life in general: worldviews, ideas about life, worldly wisdoms, and life lessons;</li> <li>- Unspecified insights, for responses in which it is not specified what it is exactly that readers learn, think, understand, and gain insights in through story reading.**</li> </ul>	79.2% $\kappa = .76$

*Note.* \* Four steps: a) Overlap between codes for each response calculated (e.g., rater A coded response as 'story-related', 'self-related' and 'evaluation or emotion', rater B coded 'story-related' and 'self-related': agreement = 0.66); b) Agreement values summed for all responses (288.29); divided by number of responses (339): observed agreement of  $288.29 / 339 = .85$ ; c) Expected agreement by chance alone calculated: sum of agreement values divided by value that represented how often responses were coded (i.e., 339 responses coded by two raters = 678): expected agree-

ment of  $288.29 / 678 = 0.43$ ; d) Kappa = (observed agreement – expected agreement) / (1 – expected agreement) =  $(.85 - 0.43) / (1 - 0.43)$ , and therefore  $\kappa = .74$  (see Eccleston, Werneke, Armon, Stephenson, & MacFaul, 2001). \*\* In addition, we coded for 'other values', not related to gaining insights from story reading (e.g., enjoyment, relaxation), and 'no value'.

### RAW MEAN SCORES (SD) FOR TREQ SCALES

Scale	Measurement 1		Measurement 2		Measurement 3	
	Group A	Group B	Group A	Group B	Group A	Group B
Imagery	4.08 (.74)	4.09 (.86)	3.85 (.91)	3.94 (.82)	3.78 (.90)	4.00 (.88)
Identification	3.21 (.68)	3.25 (.78)	2.89 (.81)	3.14 (.85)	2.95 (.80)	2.95 (.84)
Experience-taking	3.38 (.79)	3.39 (.89)	3.23 (.85)	3.32 (.80)	3.25 (.85)	3.37 (.85)
Character evaluation	3.69 (.75)	3.76 (.80)	3.45 (.85)	3.69 (.84)	3.44 (.84)	3.59 (.91)
Sympathy	3.05 (.99)	3.21 (1.01)	2.94 (.97)	3.21 (.96)	2.98 (.94)	3.18 (.95)
Aesthetic awareness	2.62 (.91)	2.66 (.99)	2.64 (.93)	2.67 (.98)	2.54 (.96)	2.60 (.99)
Self-insights	2.41 (.76)	2.53 (.82)	2.35 (.81)	2.58 (.84)	2.39 (.80)	2.50 (.82)
Insights into others	2.72 (.84)	2.94 (.89)	2.60 (.83)	2.92 (.89)	2.65 (.87)	2.85 (.87)

*Note.* Group A = intervention-control; Group B = control-intervention

### RAW MEAN SCORES (SD) FOR STORY TASK DATA CODED WITH SCALE SCORES (0-3)

Variable	Measurement 1		Measurement 2		Measurement 3	
	Group A	Group B	Group A	Group B	Group A	Group B
Extensiveness of story responses	1.61 (1.01)	1.94 (.98)	1.46 (.92)	1.90 (.96)	1.66 (.99)	1.78 (.99)
Message	1.64 (.93)	1.74 (1.02)	1.88 (1.09)	1.85 (.97)	1.55 (.95)	1.78 (1.05)
Prediction	1.72 (.93)	1.62 (.91)	1.74 (1.03)	1.53 (.96)	1.74 (.94)	1.72 (.94)

*Note.* Group A = intervention-control; Group B = control-intervention

PERCENTAGE OF WRITING TASKS (N = 99 IN GROUP A AND B)  
IN WHICH RESPONSE TYPES OCCURRED AT LEAST ONCE

Response type	Measurement 1		Measurement 2		Measurement 3	
	Group A %	Group B %	Group A %	Group B %	Group A %	Group B %
Self-related	32.3	38.4	23.2	30.3	18.2	27.3
Socially-related	40.4	54.4	28.3	46.5	32.3	38.4
Emotional- evaluative	59.6	67.7	46.5	70.7	72.7	68.7
Story-related	52.2	51.1	44.4	55.6	51.5	56.6
Question or confusion	60.6	70.7	71.7	74.7	66.7	62.6

*Note.* Group A = intervention-control; Group B = control-intervention

PERCENTAGE OF RESPONSES TO QUESTION 3 (VALUES OF STORIES; N = 99  
IN GROUP A AND B) IN WHICH INSIGHT TYPE OCCURRED AT LEAST ONCE

Insight into...	Measurement 1		Measurement 2		Measurement 3	
	Group A %	Group B %	Group A %	Group B %	Group A %	Group B %
Self and own life	10.1	6.1	12.1	7.1	15.2	19.2
Others and their lives	8.1	18.2	12.1	20.2	9.1	13.1
Experiences of characters	14.1	19.2	11.1	22.2	14.1	17.2
Emotions, actions, reactions	16.2	14.1	18.2	12.1	11.1	2.0
The world and life in general	14.1	10.1	8.1	16.2	8.1	19.2
Unspecified	31.3	29.3	29.3	29.3	30.3	28.3

*Note.* Group A = intervention-control; Group B = control-intervention



APPENDIX F. DESCRIPTIVE DATA AND CODING SCHEMES  
TO CHAPTER 6

CORRELATIONS BETWEEN TREQ SCALES

Measurement and scale		1	2	3	4	5	6	7	8
T1	1 Imagery	-	.44	.49	.47	.42	.31	.17	.28
	2 Identification		-	.50	.47	<b>.67</b>	.57	<b>.64</b>	<b>.67</b>
	3 Experience-taking			-	.23	.48	.27	.28	.38
	4 Character evaluation				-	.43	.29	.28	.40
	5 Sympathy					-	.47	.46	.53
	6 Aesthetic awareness						-	.57	.54
	7 Self-insights							-	<b>.76</b>
	8 Insights into others								-
T2	1 Imagery	-	.38	<b>.60</b>	.40	.43	.28	.13*	.35
	2 Identification		-	.55	.37	.51	.56	<b>.67</b>	<b>.65</b>
	3 Experience-taking			-	.40	.47	.35	.34	.47
	4 Character evaluation				-	.40	.37	.26	.43
	5 Sympathy					-	.36	.34	.43
	6 Aesthetic awareness						-	.56	.53
	7 Self-insights							-	<b>.70</b>
	8 Insights into others								-
T3	1 Imagery	-	.34	.30	.28	.30	.22	.14*	.23
	2 Identification		-	.49	.43	.57	.49	<b>.61</b>	<b>.66</b>
	3 Experience-taking			-	.13*	.43	.27	.51	.40
	4 Character evaluation				-	.34	.30	.26	.35
	5 Sympathy					-	.45	.47	.45
	6 Aesthetic awareness						-	.50	.51
	7 Self-insights							-	<b>.73</b>
	8 Insights into others								-

*Note.* All correlations are significant ( $p < .01$ , \*  $p < .05$ ). Correlations  $\geq .60$  are printed in bold.

## RAW MEAN SCORES (SD) FOR TREQ SCALES

Scale	Measurement 1		Measurement 2		Measurement 3	
	TDLT	Control	TDLT	Control	TDLT	Control
Imagery	3.86 (.88)	3.84 (.90)	4.00 (.71)	3.70 (.90)	4.09 (.67)	3.84 (.86)
Identification	2.86 (.82)	2.81 (.77)	3.15 (.73)	2.83 (.79)	3.17 (.74)	3.04 (.74)
Experience-taking	3.17 (.70)	3.13 (.83)	3.54 (.67)	3.34 (.86)	3.29 (.79)	3.01 (.96)
Character evaluation	3.51 (.89)	3.46 (.80)	3.61 (.79)	3.43 (.82)	3.69 (.76)	3.59 (.81)
Sympathy	3.03 (1.07)	3.02 (1.00)	3.21 (.92)	3.00 (.96)	3.22 (.88)	3.03 (.86)
Aesthetic awareness	2.46 (.82)	2.52 (.95)	3.19 (.84)	2.62 (.91)	3.09 (.86)	2.87 (.97)
Self-insights	2.25 (.78)	2.33 (.80)	2.85 (.82)	2.35 (.76)	2.67 (.75)	2.45 (.82)
Insights into others	2.73 (.82)	2.73 (.91)	3.24 (.79)	2.86 (.84)	3.02 (.82)	2.91 (.85)

## RAW MEAN SCORES (SD) FOR OTHER VARIABLES

Scale	Measurement 1		Measurement 2		Measurement 3	
	TDLT	Control	TDLT	Control	TDLT	Control
Empathy	2.25 (.64)	2.25 (.70)	2.32 (.76)	2.30 (.83)	2.38 (.85)	2.35 (.78)
Eudaimonic reasons	3.91 (1.15)	3.95 (1.36)	4.42 (1.24)	3.76 (1.21)	4.35 (1.23)	3.95 (1.15)
Hedonic reasons	5.28 (.83)	5.08 (.85)	4.95 (.76)	4.95 (.92)	5.21 (.79)	5.23 (.82)
Autonomy	2.60 (.66)	2.69 (.73)	3.59 (.53)	3.03 (.63)	-	-
Competence	2.90 (.66)	2.82 (.79)	3.35 (.68)	2.94 (.76)	-	-
Relatedness	3.33 (.58)	3.26 (.60)	3.74 (.52)	3.39 (.58)	-	-
Strategy use	2.97 (.60)	3.00 (.59)	3.31 (.64)	3.09 (.56)	-	-
Moral competence	22.52 (15.66)	18.27 (13.80)	22.82 (15.34)	22.48 (15.86)	-	-

## EXAMPLES OF WRITTEN DIALOGUE WITH IMAGINARY PEER

Translation *MS*. Clarifications between brackets.

*Example 1 (female TDLT student)*

What part of the story really stood out to you?

The part where the man was given the death penalty.

How come?

Because he got the death penalty and not even begged for reduction of his sentence.

Maybe that was simply because he acknowledges that it was wrong what he did.

But still, you would beg for reduction of the sentence? Nobody wants to die.

I think he had so much regret that he thought the death penalty did him justice.

Yeah, that could be.

He was after all a kind man, according to his friends and colleagues.

Yeah, the only thing he said to his lawyer [*the main character; is actually the prison guard*] was that he lost his patience.

Yeah, if you work on a summer camp, sometimes you have annoying children, but still you can't lose your patience.

No, right.

But I think the lawyer also lost his patience at a certain point.

Why?

Because the prisoner never actually said something that might have helped him.

Yes, you're right. But did he actually murder children or were they only wounded?

That's not very clearly stated in the text.

I think he murdered someone because it says that a heavily wounded girl was able to escape, which means the other children couldn't escape.

But if he really murdered children, I think the death penalty is an easy way out.

How come?

Well, because if you're dead, you can't have any regrets anymore for what you've done.

Yes, that's why I think the death penalty should be abolished.

But on the other hand, it costs society lots of money to keep that man imprisoned for life.

Yes that's true.

But what does the title have to do with the text?

I think it's called 'An autumn day' because it's not just an ordinary autumn day for the lawyer and the prisoner.

Yeah, I think so too.

I did think, by the way, that there was little emotion in the story.

Yes I agree, because the prisoner also doesn't say anything.

*Example 2 (male TDLT student)*

What part of the story really stood out to you?

That the 50 year-old has slaughtered children.

Yeah man, for me as well.

[unreadable], Yes, that's bad, right.

In any case [slang] [unreadable]

But anyway, what do you think of the story?

I really like it, you?

I also think it's a good story but could be better.

What could be better then?

It should've been a bit longer.

Yeah true. It's a pity that is was so short.

Nothing to be done about it, right, hahaha.

What would you do if a 50 year-old would want to slaughter you?

In such a situation you can't do much, right?

True, as a child you can hardly do something against a 50 year-old with a machete.

Yeah man, but sad for all those children.

Hahaha, luckily it's just a story.

Right.

But a question, okay...

Yeah, what's up?

Do you think the title fits the story?

No, not really, no.

Why not?

Because the story is about something completely different than an autumn day.

Yes, I think so too.

I don't even understand where they got that title from.

Agree, I also don't really know why they chose that one.

Strange author, right?

Yeah man, hahaha.

But okay, not a problem, right?

Right, no problem, story is just nice.

Yeah, that's why.

## CODING SCHEME FOR WRITTEN DIALOGUES

Code	Code description	Example sentences from segments
<i>Code 1</i> Descriptive- evaluative statements	<p>Descriptions of what happens in the story or how it is told, or for basic evaluations:</p> <ul style="list-style-type: none"> <li>- Reiterations of story events.</li> <li>- Descriptions of story characteristics (setting, time, author, who is the protagonist, and so on).</li> <li>- Simple evaluations without substantiation: nice, boring, stupid, interesting, shocking, vague, and so on.</li> <li>- Incomprehension without attempts to solve it.</li> </ul>	<p>The part where the man died had a lot of tension.</p> <p>Yes, agree. It was very mysterious. [...] It was like something was missing.</p> <p>Was it difficult to read?</p> <p>Yes.</p> <p>What did Arthur like to do?</p> <p>Playing with cars.</p> <p>What did you think of the story?</p> <p>I thought it was gloomy and emotionless, you?</p> <p>I also thought it was emotionless.</p>
<i>Code 2A</i> Referential: real (social) world	<p>References to the world outside the text, without moral considerations:</p> <ul style="list-style-type: none"> <li>- Statements that story events might happen in reality.</li> <li>- Extension from characters to real-world people ('people', 'everyone').</li> <li>- Utterances designated by students themselves as 'life lessons' (see 1.1; unit 5 and 6).</li> </ul>	<p>Can't believe this still happens in the world.</p> <p>Agree.</p> <p>I hope that kids like Johan will get stronger eventually.</p> <p>Me too!</p> <p>Anyway, respect for people who do this job.</p> <p>Yes, agree. I think they have more stories to tell than just this one.</p>

<i>Code 2B</i> Referential: self	<p>References to oneself; not necessarily very 'deep' insights:</p> <ul style="list-style-type: none"> <li>- Questions like: what would you do? How would you react? What would you choose? (students refer to themselves when using 'you', because in fact they 'talk to themselves').</li> <li>- Comparisons of own behavior and responses to the story: I would never do that, I would feel the same way.</li> <li>- References to personal life outside the story (family, friends), as long as there is a direct relation with the story.</li> </ul>	<p>How would you respond if you were the prisoner? I think I would remain silent. Why? If I've done something really awful, I feel I don't have a right to speak. But do your parents also get angry so fast? Usually not, but sometimes. And how do you deal with that? Usually I just listen and think about what I can do better.</p>
<i>Code 3A</i> Reasoning: evaluative	<p>Evaluative statements are explored more in-depth or substantiated by argumentation, either 'spontaneously' or by using follow-up questions.</p>	<p>What part did have most tension for you? I thought the end had a lot of tension. Why? Because they were building tension until the boy was slapped. How then? <u>By slowing down the pace of the story.</u> I thought it was a bit scary. And you? Oh, no, for me not really. What was scary about it, then? Maybe more tension, by the threat that the prisoner might escape. Yes, I see what you mean. But I didn't have that at all.</p>
<i>Code 3B</i> Reasoning: hypothetical	<p>Reasoning about what 'might happen'. Examples:</p> <ul style="list-style-type: none"> <li>- Ideas about what could happen after the story ends.</li> <li>- Ideas about what could have gone differently in the story, what could also have happened.</li> <li>- Ideas about which consequences it would have had if something in the story had been different.</li> <li>- Ideas about what would have made the story better or worse.</li> </ul>	<p>I think the lawyer can come to terms with what happened, because he told it to his son. And he will have a happy life. Yes, I think so too. I thought he would say something that changed everything. Yeah, that he hadn't done it or something. Yes, something like that. Or that something completely different had happened. Yes, that his wife had been threatened or something.</p>

<p><i>Code 3C</i> Reasoning: understanding story events</p>	<p>Reasoning about what particular story events might mean. Examples:</p> <ul style="list-style-type: none"> <li>- Interpretation of particular story events: what does it mean that this happens?</li> <li>- Reasoning about why the author has made particular choices in the story.</li> <li>- Incomprehension followed by attempts to solve it.</li> </ul>	<p>The single thing that I don't understand are the last two lines. I don't understand what the sawmill has to do with it. And what would 'hula' mean? Maybe 'no way', in the sense of [<i>Dutch expression sounding like 'hula'</i>].</p> <hr/> <p>The characters were described as 'smells'.</p> <p>Why do you think the author did that? I think because the way family members [...] smell can be very typical. I think so too. And probably the author wanted to make the reader think because many people potentially recognize this.</p>
<p><i>Code 3D</i> Reasoning: understanding characters</p>	<p>Reasoning about character's actions, thoughts, feelings, motives, behavior, etc.; attempts to characterize or 'understand' the character.</p> <ul style="list-style-type: none"> <li>- Reasoning why a character does, says, thinks, feels something.</li> <li>- Reasoning what might be 'going on' with a character.</li> <li>- Reasoning why a character is the way (s)he is.</li> </ul>	<p>What surprises me is that this man of the prison does not forget this particular prisoner, but he doesn't remember others.</p> <p>I think that's because the prisoner reminds him of someone.</p> <hr/> <p>What else did you notice about the story?</p> <p>That the boy is really quiet and everything around him seemed to happen vaguely around him, he is there, but not with his mind. I noticed that too, yes, he also didn't seem to be too joyful. No, he didn't really seem happy.</p>
<p><i>Code 3E</i> Reasoning: moral, confined to story world</p>	<p>Moral reasoning that directly concerns and does not go beyond the story situation. Examples:</p> <ul style="list-style-type: none"> <li>- Reasoning whether something in the story is right, just, good, deserved, fair, etc.</li> <li>- Reasoning about guilt, blame, innocence, etc. in the story.</li> <li>- Reasoning that includes a moral/ethical judgment about a character: behavior, thoughts, etc. are not right, just, fair, should not be allowed, etc.</li> </ul>	<p>Did you think there was justice in this story?</p> <p>Yes and no. The suspect is sentenced, but I think the sentence isn't just. Why then?</p> <p>I think the death penalty is too severe.</p> <hr/> <p>What do you think is unjust about the story?</p> <p>That this man simply doesn't trust her. Yes, that is definitely unfair! I also think it is awful that he let himself go like that about the way she dressed. And on top of that, that he even pinched her breasts. That was bad, yeah.</p>

<i>Code 3F</i> Reasoning: moral, beyond story world	Moral reasoning that is related to a story theme, but goes beyond the story situation. Examples: - Reasoning whether certain real-world situations or human behaviors are right, just, good, deserved, fair. - Reasoning about guilt, blame, innocence etc. in real-world situations or related to human behavior. - Reasoning that includes a moral/ethical judgment about real-world human beings, extended from the story.	What do you think about the death penalty? For some people that's just the best option. Why then? Because some people just don't belong in this world anymore. I disagree. I think physical abuse should be abolished. Do you agree? Yes, that happens way too often these days. Something should be done about it.
<i>Code 4A</i> Transformative reading: imagery	Utterances about picturing the setting, story situation and/or characters vividly in your mind.	Can you also picture the story in your mind? Yes, I already pictured it when he caught that fish. I also noticed he said 'coke' tastes specifically, he really went into detail. Yes, then you can picture it even more.
<i>Code 4B</i> Transformative reading: identification	Utterances about recognizing something or someone in the story.	Did you recognize something? I recognized the part where the father was reading aloud. It is very recognizable. I agree! I think many people recognize this, because most boys play like that with toy cars. In particular calling 'toot, toot' gives the finishing touch!
<i>Code 4C</i> Transformative reading: experience- taking	Utterances about feeling as if the experience, position, or situation of a character is taken for a moment.	I think the story is written as if it happens now and you're really there. Yes, I totally agree. Could you put yourself in Johan's position? Yes, thereby I could imagine how he feels.

<i>Code 4D</i> Transformative reading: character evaluation	Utterance in which a character is evaluated: positively (e.g., kind, friendly), negatively (terrible, racist, violent), as 'vague', 'weird', etc., or neutral / descriptively (e.g., insensitive, absent-minded).	What do you think about the main character? I think he comes across as a kind person. It seemed as if he cared about the prisoner. I think Arthur's nephew is a rather insensitive person. I disagree, at the end you notice he <u>does</u> have feelings.
<i>Code 4E</i> Transformative reading: sympathy	Utterance about feeling sympathy or compassion for a character; also expressed as finding a character 'sad', 'pitiful', and so on.	How did you feel about the person who got the death penalty? I thought it was really bad and I felt a bit of pity after all. But it's sad for the boy, having nightmares every time. Yes, I feel sorry for him.
<i>Code 4F</i> Transformative reading: aesthetic awareness	Utterance that shows that language use and/or writing style are noticed; often qualifications like 'beautifully written', 'well-written', etc. (no narratological distinction with 'well-told' or 'beautifully told').	The atmosphere was really well-written. Yes, it was very apt that it was so cold and moist. What did you think of the story? It was special, sometimes difficult to understand, but I thought the writing style was beautiful. Me too, so we agree!
<i>Code 5A</i> Content-irrelevant: other responses	Utterances that are not related to the story. Examples: - Sentences to round off the dialogue. - Responses about school, reading etc., without a direct link with the story. - Irrelevant and nonsense sentences.	I don't want to disturb, but I need to go now. Bye! Bye! It was a nice talk. Now you're nagging, you know that? Ask me a question, then! Okay, let me think...
<i>Code 5B</i> Content-irrelevant: unclear	Utterances that are unreadable.	I think he [ <i>unreadable</i> ] because [ <i>unreadable</i> ].

*Note.* In calculating inter-rater agreement, multiple codings per segment were taken into account, in four steps: a) Overlap between codes for each segment calculated (e.g., rater A coded segment as 'Reasoning: understanding characters', 'Referential: self' and 'TR: character evaluation', rater B coded Referential: self' and 'TR: character evaluation': agreement = 0.66); b) Agreement values summed for all segments (83.87); divided by number of segments (100): observed agreement of  $83.87 / 100 = .84$ ; c) Expected agreement by chance alone calculated: sum of agreement values divided by value that represented how often segments were coded (i.e., 100 segments coded by two raters = 200): expected agreement of  $83.87 / 200 = 0.42$ ; d) Kappa = (observed agreement - expected agreement) / (1 - expected agreement) =  $(.84 - 0.42) / (1 - 0.42)$ , and therefore  $\kappa = .72$  (Eccleston, Werneke, Armon, Stephenson, & MacFaul, 2001).



## CONTRIBUTIONS OF AUTHORS

CRedit Author statements, see <http://casrai.org/credit/>

### CHAPTER 2

Based on:

Schrijvers, M., Janssen, T., Fialho, O. & Rijlaarsdam, G. (2016). The impact of literature education on students' perceptions of self and others: Exploring personal and social learning experiences in relation to teacher approach. *L1 – Educational Studies in Language and Literature*, 17, 1-37. doi:10.17239/L1ESLL-2016.16.04.01

*Marloes Schrijvers*: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Writing: original draft, Writing: review and editing. *Tanja Janssen*: Conceptualization, Funding acquisition, Methodology, Resources, Supervision, Writing: review and editing. *Olivia Fialho*: Conceptualization, Funding acquisition, Methodology, Resources, Supervision, Writing: review and editing. *Gert Rijlaarsdam*: Conceptualization, Funding acquisition, Methodology, Supervision, Writing: review and editing.

### CHAPTER 3

Based on:

Schrijvers, M., Janssen, T., Fialho, O. & Rijlaarsdam, G. (2018). Gaining insight into human nature: A review of literature classroom intervention studies. *Review of Educational Research*. Advance online publication. doi:10.3102/0034654318812914

*Marloes Schrijvers*: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Writing: original draft, Writing: review and editing. *Tanja Janssen*: Conceptualization, Data curation, Funding acquisition, Methodology, Supervision, Writing: review and editing. *Olivia Fialho*: Conceptualization, Funding acquisition, Methodology, Supervision, Writing: review and editing. *Gert Rijlaarsdam*: Conceptualization, Funding acquisition, Methodology, Supervision, Writing: review and editing.

## CHAPTER 4

Based on:

Schrijvers, M., Janssen, T., Fialho, O. & Rijlaarsdam, G. (2018). *Designing a literature classroom intervention to foster 10<sup>th</sup> grade students' insight into human nature*. Manuscript submitted.

*Marloes Schrijvers*: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Writing: original draft, Writing: review and editing. *Tanja Janssen*: Conceptualization, Funding acquisition, Methodology, Resources, Supervision, Writing: review and editing. *Olivia Fialho*: Conceptualization, Funding acquisition, Methodology, Supervision, Writing: review and editing. *Gert Rijlaarsdam*: Conceptualization, Funding acquisition, Methodology, Supervision, Writing: review and editing.

## CHAPTER 5

Based on:

Schrijvers, M., Janssen, T., Fialho, O. & Rijlaarsdam, G. (2018). *Effects of dialogic literary instruction on 10<sup>th</sup> grade students' insight into human nature*. Manuscript submitted.

*Marloes Schrijvers*: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Writing: original draft, Writing: review and editing. *Tanja Janssen*: Conceptualization, Funding acquisition, Methodology, Resources, Supervision, Writing: review and editing. *Olivia Fialho*: Conceptualization, Funding acquisition, Methodology, Supervision, Writing: review and editing. *Gert Rijlaarsdam*: Conceptualization, Funding acquisition, Methodology, Supervision, Writing: review and editing.

## CHAPTER 6

Based on:

Schrijvers, M., Janssen, T., Fialho, O. & Rijlaarsdam, G. (2018). *Transformative Dialogic Literature Teaching fosters adolescents' insight into human nature*. Manuscript under revision.

*Marloes Schrijvers*: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Writing: original draft, Writing: review and editing. *Tanja Janssen*: Conceptualization, Funding acquisition, Resources, Supervision, Writing: review and editing. *Olivia Fialho*: Conceptualization, Funding acquisition, Supervision, Writing: review and editing. *Gert Rijlaarsdam*: Conceptualization, Funding acquisition, Supervision, Writing: review and editing.

## RELATED PUBLICATIONS AND PRESENTATIONS

### *Practice-oriented publications*

- Schrijvers, M. (2018). Zelfinzicht en sociaal inzicht opdoen in literatuurlessen: ontwerp van een interventie voor havo 4. *Levende Talen Tijdschrift*, 19(3), 3-12.
- Schrijvers, M. (2018). *Integratie van fictie- en literatuuronderwijs met gespreksvaardigheden en schrijfvaardigheid: een verkenning van praktijkgerichte initiatieven*. Enschede: SLO.
- Schrijvers, M., Janssen, T. & Rijlaarsdam, G. (2017). Aandacht voor zelfinzicht en sociaal inzicht in dialogische literatuurlessen: hoe doe je dat? In A. Mottart & S. Vanhooren (Eds.), *Bundel 3<sup>e</sup> Conferentie Onderwijs Nederlands* (pp. 287-291). Gent: Skibris.
- Schrijvers, M., Janssen, T. & Rijlaarsdam, G. (2017). Literatuur leert je het leven? Over het stimuleren van persoonlijke en sociale inzichten via literatuuronderwijs. In R. van Steensel & E. Segers (Eds.), *Succesvol lezen in het onderwijs* (pp. 183-197). Amsterdam: Stichting Lezen.
- Schrijvers, M., Janssen, T. & Rijlaarsdam, G. (2016). Leert literatuur je het leven? Het stimuleren van persoonlijke en sociale inzichten via literatuuronderwijs. In A. Mottart & S. Vanhooren (Eds.), *Bundel 3<sup>e</sup> Conferentie Onderwijs Nederlands* (pp. 164-167). Gent: Skibris.
- Schrijvers, M., Janssen, T., & Rijlaarsdam, G. (2016). 'Dat een boek kan veranderen hoe je naar de wereld kijkt': De impact van literatuuronderwijs op zelfinzicht en sociaal inzicht van bovenbouwleerlingen in havo en vwo. *Levende Talen Tijdschrift*, 17(1), 3-13.

### *Contributions to international conferences*

- Schrijvers, M. & Janssen, T. (2018). What students learn about themselves and others in the literature classroom: An analysis of written learner reports. Paper presented as part of Symposium 'Professional competence in teaching and learning from a domain specific research perspective', organized by I. Pieper, GEBF Conference 2018, Basel, Switzerland.
- Schrijvers, M., Janssen, T., Fialho, O. & Rijlaarsdam, G. (2017). *Gaining personal and social insights in the literature classroom: Effects of dialogic learning*. Paper presented at JURE 2017, Tampere, Finland.
- Schrijvers, M., Janssen, T., Fialho, O. & Rijlaarsdam, G. (2017). *Fostering personal and social insights in the literature classroom: Review of intervention studies*. Paper presented at EARLI 2017, Tampere, Finland.
- Schrijvers, M., Janssen, T., Fialho, O. & Rijlaarsdam, G. (2017). *Talking about literature and life: Effects of dialogic literary instruction on students' personal and social insights*. Paper presented as part of the SIG Research on Literature Education Symposium 'Changing perspectives on self and others in the literature classroom', organized by M. Schrijvers, T. Janssen and I. Pieper, ARLE Conference 2017, Tallinn, Estonia.
- Schrijvers, M., Janssen, T., Fialho, O. & Rijlaarsdam, G. (2016). *The impact of literature education on students' perceptions of self and others*. Paper presented at IGEL 15<sup>th</sup> Biennial Meeting 2016, Chicago, Illinois, United States.
- Schrijvers, M., Janssen, T., Fialho, O. & Rijlaarsdam, G. (2015). *Literature education: Changing students' perceptions of self and others?* Poster presented at JURE 2015, Limasol, Cyprus.
- Schrijvers, M., Janssen, T., Fialho, O. & Rijlaarsdam, G. (2015). *Learning about oneself and others in the literature classroom: An analysis of students' written learner reports*. Paper presented at ARLE (IAIMTE) Conference 2015, Odense, Denmark.



## CURRICULUM VITAE

Marloes Schrijvers (1989) obtained her bachelor's degree in Text & Communication Studies (2011, cum laude), her master's degree in Children's Literature (2013, cum laude), and her master's teaching degree in Dutch Language and Literature (2014, cum laude) from Tilburg University. In the context of an internship at Cubiss (Tilburg) during the Children's Literature program, she developed an instructional approach for literature teaching to pre-vocational education students. She published an article about her master's thesis on autobiographical picture books in the European Journal of Life Writing. For her thesis, she received the 2014 IBBY Miep Diekmann Thesis Award for research in children's literature. She was also an editorial member of *Literatuur zonder leeftijd*. In 2013-2014, she worked as a pre-service teacher at De Nieuwste School (Tilburg).

In September 2014, she started her PhD research at the University of Amsterdam, under supervision of Gert Rijlaarsdam, Tanja Janssen, and Olivia Fialho, as part of the project *Uses of Literary Narrative Fiction in Social Contexts*, which was funded by a grant from the Dutch Organization for Scientific Research (NWO). She presented the results of her PhD research at international conferences, such as ARLE (2015; 2017), the ARLE SIG Research on Literature Education (2016; 2018), JURE/EARLI (2015; 2017), IGEL (2016), and GEBF (2018), as well as at national conferences such as Het Schoolvak Nederlands (2016; 2017) and those of Stichting Lezen (2016; 2018). She published several practice-oriented articles based on her research, for example, in *Levende Talen Tijdschrift*. In Spring 2018, she was a visiting scholar at Stanford University for two months, at the invitation of Dr. Sarah Levine.

During her PhD research, Marloes engaged in various ancillary activities. She has been a member of the Educational Committee of ICO, the graduate school for Educational Sciences (2015-2018), the Advisory Board of the Dutch Literature Foundation (2015-2019), and the Werkgroep Onderzoek & Didactiek Nederlands (SBN Levende Talen, from 2018 onwards). In 2016, she organized a PhD event for the ARLE SIG Research on Literature Education at the University of Amsterdam. In 2017, she developed an online course on literature teaching for E-Wise; in 2018, she conducted a practice-oriented review study on the integration of literature teaching, writing, and oracy, for the Dutch Institute for Curriculum Development (SLO). In addition, Marloes has been a member of the jury for the Jenny Smelik-IBBY Award (2014), the Miep Diekmann Thesis Award (2017), and the Stichting Lezen Award for research on reading promotion (2018). Since 2018, she has been an editor for *L1 – Educational Studies in Language and Literature*; for the same journal, she has set up a special issue on the design of literature classroom interventions, together with P. Karen Murphy and Gert Rijlaarsdam (to appear in 2019).



## DANKWOORD

*A book is a dream that you hold in your hand.*

~ Neil Gaiman

Als een boek een droom is, dan is dit proefschrift dat ook. Zeker in de laatste maanden van 2018 nam dat soms letterlijke vormen aan, wanneer mijn brein me 's nachts trakteerde op dromen over statistische analyses, revisies van papers, en deadlines die ik al dan niet zou halen. Maar het boek is er, en ik besteed graag nog enkele pagina's aan een woord van dank aan al diegenen die op de één of andere manier aan dit proefschrift hebben bijgedragen.

Gert – inspirator, wegwijzer, juiste-vragen-steller – dank voor je begeleiding, je vertrouwen, en je vermogen om te zien of ik even wat ondersteuning nodig had, of juist meer autonomie. Zeker in het begin zag ik heel wat beren op de weg, maar jij verwelkomde elke beer met een vriendelijk woord en stuurde 'm dan het bos in, zodat ik weer verder kon. Meer dan eens wees je me op een zijpad: door jou zag ik nieuwe kansen en mogelijkheden.

Tanja, dankjewel voor de sturing in het begin en de ruimte later in het proces, voor je ideeën waar mijn werk beter van werd, voor je scherpe blik op alle teksten die ik schreef. Jij en Gert hebben me niet alleen begeleid bij het schrijven van mijn dissertatie, maar me ook gestimuleerd en gesteund in het ondernemen van allerlei andere activiteiten (voor ARLE, voor L1, voor SLO). Het was heel fijn dat we telkens verder konden kijken dan het proefschrift.

Olivia, thank you for your help in making this dissertation an interdisciplinary work. It wasn't always easy to learn to speak each other's "language" and to connect different perspectives, but I learned so much during all that. Muito obrigado for your positive criticism on my papers (yay! 😊) and your continued understanding whenever I struggled in the analysis and writing processes.

Mijn dank gaat daarnaast uit naar de overige leden van de promotiecommissie: Carla van Boxtel, Monique Volman, Lia van Gemert, Irene Pieper en Helma van Lierop. Irene, vielen dank, not only for being in the committee, but also for welcoming me so warmly in the ARLE community and the SIG Research on Litera-

ture Education. Your support and encouragement means a lot to me. Helma, al in Tilburg stimuleerde jij me om verder te gaan met onderzoek, en mede dankzij jouw vertrouwen en aanmoediging durfde ik die droom na te jagen.

Deze dissertatie zou er bovendien niet zijn geweest zonder de NWO-groep, de Nussbaum-club, onder aanvoering van Frank Hakemulder. Frank, Inge, Matthijs: dank voor jullie belangstelling en ideeën!

Praktijkgericht onderwijsonderzoek is alleen mogelijk met medewerking van leerlingen en docenten. Ik zou iedereen wel willen noemen, maar ik laat het hier bij een bijzonder woord van dank aan Janetta de With, Corrine Verschoor en Els van Hall, die aan twee interventiestudies én het herontwerpproces meewerkten. Mijn dank gaat ook uit naar Hans Goosen, die afreisde naar Amsterdam om mee te denken over de lessenserie. Liselore en Sonja, dank voor jullie bereidheid om de pilots uit te voeren, en voor jullie waardevolle ideeën om de lessen te verbeteren.

Bij de dataverzamelingen kreeg ik gelukkig hulp van betrokken, flexibele onderzoeksassistenten: dank aan jullie allemaal! Een speciaal woord van dank aan Jelle, die telkens bereid was om op onzalige tijdstippen in de trein te stappen, en aan Angèle: jij hebt meermaals een organisatorische crisis bezworen. Jouw attente manier van werken en communiceren hielp mij erop vertrouwen dat elke dataverzameling uiteindelijk op z'n pootjes terecht zou komen. Ontzettend bedankt.

Ook anderen hebben op verschillende manieren aan deze dissertatie bijgedragen. Sander Bax gaf me in Tilburg mijn basis in de vakdidactiek mee en wees me op de vacature voor dit project. Janneke Staaks hielp me om de reviewstudie in het derde hoofdstuk van dit proefschrift van de grond te krijgen. Thomas Bijen ontwierp het prachtige omslag ervan.

En dan zijn er al die mensen – ARLE, L1, HSN, WODN – die warme belangstelling toonden en met wie ik heb mogen sparren en samenwerken. Stefanie, dank voor je codeerhulp, en hoera voor ons vervolgproject! Jimmy, inmiddels zijn we allebei editor bij L1, maar het was fijn om al voor die tijd herkenning te vinden, over reviewers te klagen, en elkaar eraan proberen te herinneren dat we niet altijd overal “ja” op zouden moeten zeggen (meestal tevergeefs ☺).

A big thank you as well to Sarah Levine, Karoline Trepper and Mary Hauser, who welcomed me so kindly at Stanford University in Spring 2018. Being there, in the sunny Californian bubble, and being challenged to look at my research from new perspectives, turned out to be exactly what I needed at that moment.

*We read to know that we are not alone.*

~ William Nicholson

Toen ik aan mijn onderzoek begon, klonk wel eens de waarschuwing dat het een solitaire bezigheid zou zijn. Ik prijs me gelukkig met alle mensen die het tegendeel bewezen.

Hoewel een zeker muurtje het einde inluidde van D8.12 oude stijl, bedank ik graag mijn (oud-)collega's in deze en andere kamers voor de fijne sfeer: Janneke, Alexander, Rianne, Sietske, Loes, Marjolein, Gerhard, Uddhava, Maartje, Sevinç, Johan, Joost, Marit, en wie ik maar mocht vergeten.

Degenen van de wekelijkse lunchlab-club die ik niet elders noem – Anouk, Chelsea, Marrit, Suzanne, Suzanne, Thao, Lieke, en alle anderen: het was heel fijn om jullie bijna wekelijks te zien en tijdens de lunch-uurtjes met elkaar mee te denken. Paula, Miriam, Bet, Juan Carlos, and in particular Marina, host of the wonderful crazy weekend in L'Escala: thanks for making the UvA an international place! I enjoyed the talks, ideas, lunches, dinners and general *gezelligheid* while I learned a few words of Spanish and Catalan.

Martijn, wat liepen we in het begin mooi in de pas met ons onderzoek en alle conferenties. De verdwenen laptop op Cyprus, roomservice op het balkon, de breeduit geclaimde rij in het vliegtuig naar Chicago en de Anaheim Scrambler (met "froo") zal ik niet snel vergeten. Jouw kritische blik, relativierungsvermogen, onderwijs- én levenservaring hebben veel voor me betekend.

Daphne en Geerte, als postdocs en later UD's zijn jullie een voorbeeld voor me geweest. Daph, het was zo fijn dat jij er altijd was met een luisterend oor en een arm om me heen als dat even nodig was. Er zijn maar weinig mensen die zo oprecht en attent zijn als jij. Dankjewel voor alle onderzoeks-, maar vooral alle levenswijsheden die je met me deelde. Geerte, op onze eerste dag aan de UvA gaf je me een waardevol advies: "Je vindt altijd dat je proefschrift nóg beter kan, maar probeer toch maar los te laten dat het een meesterwerk moet worden." Als loslaten even niet lukte, was er altijd jouw relativering, warmte en aanmoediging (en een huis in Blijdorp! ☺).

Collega's aan de bureaus om me heen werden dierbare vrienden. Lieve eilanders, you mean the world to me! Ik zat dan misschien op een eiland, maar wel mooi met jullie.

Saskia, het lijkt alweer lang geleden dat je geregeld aan het bureau naast me zat. Ik voelde me ontzettend welkom, en jouw vrolijkheid, je verhalen en je gemompel tegen je scherm brachten altijd iets extra's in D8.12. Met jou, Mark en Lucija op de eerste versie van het eiland had ik me geen betere start aan de UvA kunnen wensen!

Magdalena, I'm so happy that you and Carlos came to Amsterdam! I admire how you settled here in the Netherlands, even though winter was coming ;-). I very much enjoyed having you as my new neighbour on the Island. You added a big smile, many expressive gestures, and an occasionally spilled soup to it, as well as your friendship and understanding of how important certain things are in PhD life: the need to express enthusiasm as well as doubts, the "bubble", a fitness ball, a yoga position in the office every now and then, and a couch at crawling distance to crash on after Lucija's defense party.

Lucija, you truly are a unique person in your kindness, honesty, and encouragement on so many levels: research, relationships, rakia... ;-). You always focus on the bright side of things. When you were still at the other side of the island, ever so often you would try to convince us of something (*seriously, guys, listen to this!*), like the best place to buy a bike, even if none of us was in need of a bike at all. But seriously, dear, listen to this: I feel so lucky that we met when I started at UvA, that we became friends, and that you were there – again! – on my first day at the VU.

Monique, jij bewoonde Lucija's bureau na haar vertrek, maar eigenlijk was je natuurlijk al veel eerder een eilander. Onze UvA-expedities gingen vooral richting de chocomel-automaat, maar erbuiten gingen we van picknick naar museum, van Pizzabakkers naar Kriterion, van Tikibad naar schaatsbaan, en van Amsterdam naar Catalunya, alwaar we Catalaanse verkeersborden probeerden te stelen, een reddingsactie uitvoerden, en met weinig succes een veel te grote auto wilden fileparkeren. Naast al die gekkigheid werd je in de afgelopen jaren ook in meer serieuze aangelegenheden een dierbare vriendin: als ik er even doorheen zat, zag jij dat altijd meteen en was je er voor me. Dankjewel voor je vrolijkheid en knuffels (altijd op voorraad!) als ik die nodig had.

En Mark, mijn (wederzijdse) paranimf! Als er niet af en toe nog gepromoveerd werd, wat al regelmatig werd onderbroken door een schop onder het bureau, keken we naar wielrennen of zat ik bij jou achterop de fiets naar bor-

rels, films, feestjes of andere gezelligheid in Amsterdam of Utrecht. De stapel romans die je in mijn UvA-kast achterliet bleef maar groeien, net als het aantal rekeningen van de Pizzabakkers, de gedeelde ervaringen, en de vriendschap. De afgelopen jaren zouden niet hetzelfde zijn geweest zonder jou aan het bureau tegenover me.

Charlotte en Pauline: wat is het fijn dat we elkaar van tijd tot tijd blijven zien, al sinds het begin van onze studietijd in Tilburg. Pauline, dankjewel voor de lieve kaartjes en bloemen die me zo verrassen, de wandelingen en etentjes, en de levenswijsheden. Charlotte, dankjewel voor je hulp bij het coderen van studies, maar vooral voor je belangstelling en onze telefoongesprekken van drie uur, of ik nu in Utrecht, Rotterdam, Den Haag, Duitsland of de VS zat.

Lieve papa en mama, na drie diploma-uitreikingen waarbij jullie al zaten te stralen van trots, is er nu nog de verdediging van mijn proefschrift. Dankzij jullie vertrouwen heb ik steeds mijn eigen weg durven zoeken. Wat ik ook doe en waarheen ik ook ga: jullie hebben daarvoor de basis gelegd. Voor die basis en voor jullie belangstelling en steun ben ik jullie ongelooflijk dankbaar.

Liset, lief schwessie, toen ik aan dit onderzoek begon stond al vast dat jij een van mijn paranimfen zou zijn. Ik keek laatst nog eens naar die allereerste foto van ons samen: ik een paar uur oud, mijn grote zus nieuwsgierig naast me. Dat jij nu naast me staat als ik mijn proefschrift mag verdedigen, betekent zoveel voor me. Of we nu een appartement delen, burens zijn of ietsje verder uit elkaar wonen, jij bent er *áltijd* voor me. Iedereen zou een zus zoals jij moeten hebben.

Lieve Erwin, andere helft van mijn team. We leerden elkaar kennen toen ik net had gehoord dat ik aan dit promotieonderzoek mocht beginnen, en op de dag nadat ik mijn proefschrift verdedig zijn we vijf jaar samen. Met jou aan mijn zijde heb ik in de afgelopen jaren niet alleen kunnen groeien in mijn werk, maar ook in wie ik ben, en zijn we samen gegroeid tot waar we nu zijn. Soms denk ik aan de woorden uit het liedje dat jou herinnert aan onze begindagen: *If it wasn't for you...* Dankjewel, mijn lief, voor je steun, voor je optimisme, voor je liefde, voor alles.



## ICO DISSERTATION SERIES

295. Rijt, J.W.H. van der (11-9-2014). *Instilling a thirst for learning. Understanding the role of proactive feedback and help seeking in stimulating workplace learning*. Maastricht: Maastricht University.
296. Engelen, J. (11-09-2014). *Comprehending texts and pictures: Interactions between linguistic and visual processes in children and adults*. Rotterdam: Erasmus University Rotterdam.
297. Khaled, A.E. (7-10-2014). *Innovations in hands-on simulations for competence development. Authenticity and ownership of learning and their effects on student learning in secondary and higher vocational education*. Wageningen: Wageningen University.
298. Gaikhorst, L. (29-10-2014). *Supporting beginning teachers in urban environments*. Amsterdam: University of Amsterdam.
299. Wijnia, L. (14-11-2014). *Motivation and achievement in problem-based learning: The role of interest, tutors, and self-directed study*. Rotterdam: Erasmus University Rotterdam.
300. Gabelica, C. (4-12-2014). *Moving teams forward. Effects of feedback and team reflexivity on team performance*. Maastricht: Maastricht University.
301. Leenaars, F.A.J. (10-12-2014). *Drawing gears and chains of reasoning*. Enschede: University of Twente.
302. Huizinga, T. (12-12-2014). *Developing curriculum design expertise through teacher design teams*. Enschede: University of Twente.
303. Strien, J.L.H. van (19-12-2014). *Who to trust and what to believe? Effects of prior attitudes and epistemic beliefs on processing and justification of conflicting information from multiple sources*. Heerlen: Open University of the Netherlands.
304. Wijaya, A. (21-01-2015). *Context-based mathematics tasks in Indonesia: Towards better practice and achievement*. Utrecht: Utrecht University.
305. Goossens, N.A.M.C. (22-01-2015). *Distributed practice and retrieval practice in primary school vocabulary learning*. Rotterdam: Erasmus University.
306. Jupri, A. (28-01-2015). *The use of applets to improve Indonesian student performance in algebra*. Utrecht: Utrecht University.
307. Griethuisen, R.A.L.F. van (11-03-2015). *Relationships between students' interest in science, views of science and science teaching in upper primary and lower secondary education*. Eindhoven: Eindhoven University of Technology.
308. De Smet, C. (11-05-2015). *Using a learning management system in secondary education: Design and implementation characteristics of learning paths*. Ghent: Ghent University.
309. Aesaert, K. (19-05-2015). *Identification and assessment of digital competences in primary education*. Ghent: Ghent University.
310. Ardies, J. (22-05-2015). *Students' attitudes towards technology. A cross-sectional and longitudinal study in secondary education*. Antwerp: University of Antwerp
311. Donker, A. (11-06-2015). *Towards effective learning strategies*. Groningen: Groningen University
312. Veldhuis, M. (24-06-2015). *Improving classroom assessment in primary mathematics education*. Utrecht: Utrecht University.

313. Leeuwen, A. van (30-06-2015). *Teacher regulation of CSCL: Exploring the complexity of teacher regulation and the supporting role of learning analytics*. Utrecht: Utrecht University.
314. Boschman, F.B. (28-08-2015) *Collaborative design of ICT-rich early literacy learning material: Design talk in teacher teams* Enschede: Twente University.
315. Dijk, M.L. van (04-09-2015) *Physical activity, cognitive performance and academic achievement in adolescents* Heerlen: Open University of the Netherlands.
316. Jaarsma, T. (04-09-2015). *Expertise development under the microscope: Visual problem solving in clinical pathology*. Heerlen: Open University of the Netherlands.
317. Isac, M.M. (01-10-2015). *Effective civic and citizenship education. A cross-cultural perspective*. Groningen: Groningen University.
318. Spelt, E.J.H. (26-10-2015). *Teaching and learning of interdisciplinary thinking in higher education in engineering*. Wageningen: Wageningen University.
319. Ritzema, E.S. (05-11-2015). *Professional development in data use: The effects of primary school teacher training on teaching practices and students' mathematical proficiency*. Groningen: Groningen University.
320. Gijsselaers H.J.M. (06-11-2015). *Biological lifestyle factors in adult distance education: Predicting cognitive and learning performance*. Heerlen: Open University of the Netherlands.
321. Oude Groote Beverborg, A. (12-11-2015). *Fostering sustained teacher learning: Co-creating purposeful and empowering workplaces*. Enschede: Twente University.
322. Dijkstra, E.M. (13-11-2015). *Teaching high-ability pupils in early primary school*. Heerlen: Open University of the Netherlands.
323. Want, A.C. van der (17-11-2015). *Teachers' interpersonal role identity*. Eindhoven: Eindhoven University of Technology.
324. Wal, M.M van der (4-12-2015). *The role of computer models in social learning for participatory natural resource management*. Heerlen: Open University of the Netherlands.
325. Herten, M. van (11-12-2015). *Learning communities, informal learning and the humanities. An empirical study of book discussion groups*. Heerlen: Open University of the Netherlands.
326. Lucieer, S.M. (06-01-2016). *Selecting students for medical education: Exploring novel approaches*. Rotterdam: Erasmus University Rotterdam.
327. Ouwehand, K.H.R. (15-01-2016). *Effects of observing and producing deictic gestures on memory and learning in different age groups*. Rotterdam: Erasmus University Rotterdam.
328. Bogert, N.J. van der (21-01-2016). *On teachers' visual perception and interpretation of classroom events using eye tracking and collaborative tagging methodologies*. Eindhoven: Eindhoven University of Technology.
329. Wolff, C. (16-02-2016). *Revisiting 'witness': Differences in teachers' representations, perceptions, and interpretations of classroom management*. Heerlen: Open University of the Netherlands.
330. Kok, E.M. (01-04-2016). *Developing visual expertise; from shades of grey to diagnostic reasoning in radiology*. Maastricht: Maastricht University.
331. De Beer, H.T. (11-05-2016). *Exploring instantaneous speed in grade five: A design research*. Eindhoven: Eindhoven University of Technology.
332. Ebbeler, J. (12-05-2016). *Implementing data use in schools: effects on the professional development of educators and the role of school leaders in data teams*. Enschede: University of Twente.
333. Draaijer, S. (10-06-2016). *Supporting teachers in higher education in designing test items*. Amsterdam: Vrije Universiteit Amsterdam.
334. Bos, L.T. (15-06-2016). *Moving beyond words. Supporting text processing using a situation model approach*. Amsterdam: Vrije Universiteit Amsterdam.

335. Vrugte, J. ter (16-06-2016). *Serious support for serious gaming*. Enschede: University of Twente.
336. Kock, Z.D.Q.P. (23-06-2016). *Toward physics education in agreement with the nature of science: Grade 9 electricity as a case*. Eindhoven: Eindhoven University of Technology.
337. Trinh Ba, T. (28-6-2016). *Development of a course on integrating ICT into inquiry-based science education*. Amsterdam: Vrije Universiteit Amsterdam.
338. Gerken, M. (29-06-2016). *How do employees learn at work ? Understanding informal learning from others in different workplaces*. Maastricht: Maastricht University.
339. Louws, M.L. (06-07-2016). *Professional learning: what teachers want to learn*. Leiden: Leiden University.
340. Geel, M.J.M. van, & Keuning T. (08-07-2016). *Implementation and effects of a schoolwide data-based decision making intervention: a large-scale study*. Enschede: University of Twente.
341. Bouwer, I.R., & Koster, M.P. (02-09-2016). *Bringing writing research into the classroom: The effectiveness of Tekster, a newly developed writing program for elementary students*. Utrecht: Utrecht University.
342. Reijnders, P.B.G. (02-09-2016). *Retrieval as a cognitive and metacognitive study technique to learn from expository text*. Heerlen: Open University of the Netherlands.
343. Hubers, M.D. (08-09-2016). *Capacity building by data team members to sustain schools' data use*. Enschede: University of Twente.
344. Hsiao, Y.P. (23-09-2016). *Peer support to facilitate knowledge sharing on complex tasks*. Heerlen: Open University of the Netherlands.
345. Scheer, E.A. (23-09-2016). *Data-based decision making put to the test*. Enschede: University of Twente.
346. Bohle Carbonell, K. (28-9-2016). *May I ask you...? The influence of individual, dyadic, and network factors on the emergence of information in exchange teams*. Maastricht: Maastricht University.
347. Claessens, L.C.A. (30-09-2016). *Be on my side, I'll be on your side: Teachers' perceptions of teacher-student relationships*. Utrecht: Utrecht university.
348. Jansen in de Wal, J. (18-11-2016). *Secondary school teachers' motivation for professional learning*. Heerlen: Open University of the Netherlands.
349. Kock, W.D. de. (24-11-2016). *The effectiveness of hints during computer supported word problem solving*. Groningen: University of Groningen.
350. Oonk, C. (07-12-2016). *Learning and teaching in the regional learning environment: Enabling students and teachers to cross boundaries in multi-stakeholder practices*. Wageningen: Wageningen University.
351. Beckers, J. (09-12-2016). *With a little help from my e-portfolio: supporting students' self directed learning in senior vocational education*. Maastricht: Maastricht University.
352. Osagie, E.R. (14-12-2016). *Learning and corporate social responsibility. A study on the role of the learning organization, individual competencies, goal orientation and the learning climate in the CSR adaptation process*. Wageningen: Wageningen University.
353. Baggen, Y. (13-01-2017). *LLLIGHT 'in' Europe - Lifelong learning, innovation, growth and human capital tracks in Europe*. Wageningen: Wageningen University.
354. Wouters, A. (09-02-2017). *Effects of medical school selection. On the motivation of the student population and applicant pool*. Amsterdam: VU Medisch Centrum.
355. Baas, D.M. (01-05-2017). *Assessment for Learning: more than a tool*. Maastricht: Maastricht University.
356. Pennings, J.M. (04-05-2017). *Interpersonal dynamics in teacher-student interactions and relationships*. Utrecht: Utrecht University.

357. Lans, R.M. (18-05-2017). *Teacher evaluation through observation*. Groningen: University of Groningen.
358. Grohnert, T. (18-05-2017). *Judge/Fail/Learn; enabling auditors to make high-quality judgments by designing effective learning environments*. Maastricht: Maastricht University.
359. Brouwer, J. (22-05-2017). *Connecting, interacting and supporting. Social capital, peer network and cognitive perspectives on small group teaching*. Groningen: University of Groningen.
360. Van Lankveld, T.A.M. (20-06-2017). *Strengthening medical teachers' professional identity. Understanding identity development and the role of teacher communities and teaching courses*. Amsterdam: Vrije Universiteit Amsterdam.
361. Janssen, N. (23-06-2017). *Supporting teachers' technology integration in lesson plans*. Enschede: University of Twente.
362. Tuithof, J.I.G.M. (23-06-2017). *The characteristics of Dutch experienced history teachers' PCK in the context of a curriculum innovation*. Utrecht: Utrecht University.
363. Van Waes, S. (23-06-2017). *The ties that teach: Teaching networks in higher education*. Antwerp: University of Antwerp.
364. Evens, M. (30-06-2017). *Pedagogical content knowledge of French as a foreign language: Unraveling its development*. Leuven: KU Leuven.
365. Moses, I. (07-09-2017). *Student-teachers' commitment to teaching*. Leiden: Leiden University.
366. Wansink, B.G.J. (15-09-2017). *Between fact and interpretation. Teachers' beliefs and practices in interpretational history teaching*. Utrecht: Utrecht University.
367. Binkhorst, F. (20-10-2017). *Connecting the dots. Supporting the implementation of Teacher Design Teams*. Enschede: University of Twente.
368. Stoel, G.L. (14-11-2017). *Teaching towards historical expertise. Developing students' ability to reason causally in history*. Amsterdam: University of Amsterdam.
369. Van der Veen, M. (28-11-2017). *Dialogic classroom talk in early childhood education*. Amsterdam: Vrije Universiteit Amsterdam.
370. Frèrejean, J. (08-12-2017). *Instruction for information problem solving*. Heerlen: Open University of the Netherlands.
371. Rezende Da Cunha Junior, F. (19-12-2017). *Online groups in secondary education*. Amsterdam: Vrije Universiteit Amsterdam.
372. Van Dijk, A.M. (22-12-2017). *Learning together in mixed-ability elementary classrooms*. Enschede: University of Twente.



tb