The story, the self, the other

Developing insight into human nature in the literature classroom

Schrijvers, M.S.T.

Publication date
2019
Document Version
Other 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].

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 10th 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	- Four units - 200 minutes of classroom work - Exact same structure in all units	- Six units - 300 minutes of classroom work - About 45 minutes of homework - More variety in structure of units
Text selection	- Reading three stories (literary fiction) - Each story addresses a different social-moral theme	- Reading six stories (literary fiction) - All stories address the same social-moral theme ("justice and injustice")
Internal dialogues	- Applied in three units (2-4) - No explicit attention for importance - No explicit strategy instruction	- Applied in six units (1-6) - Explicit attention for importance - Explicit strategy instruction in unit 3 - Monitoring progress with rubric
External dialogues	- Applied in four units (1-4) - Some attention for importance in unit 1 - Strategy instruction with dialogue guidelines in unit 1	- Applied in six units (1-6) - More explicit attention for importance in units 1 and 3 - Strategy instruction with dialogue guidelines and additional prompts in unit 1, available on "first aid card" - Monitoring progress with rubric
Support for teachers	- Teacher guidelines and lesson plans - Face-to-face preparatory meeting with each teacher individually	- Teacher guidelines and lesson plans, including prompts for guiding talk and giving feedback - Workshop for new teachers, including walk-through of materials and feedback practice

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 10th 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 diffi-
culties during literary reading, and their rather low motivation for literature ed-
ucation. 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 mo-
tivation. 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 au-
tonomy, 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 broad-
ly than in Study 4: we defined additional indicators for it (i.e., students' empa-
thy 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 moti-
     vation for literature education, in comparison to students in a control con-
     dition 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 eudai-
     monic 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 condi-
tion 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 read-ing experiences were fostered as a result of participating in TDLT-2.

In our view, thus, this dissertation is innovative in two ways. First, the find-ings 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 empiri-cal 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 ena-ble us to actually validate the model of transformative reading for this popula-tion 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 trans-formative 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.
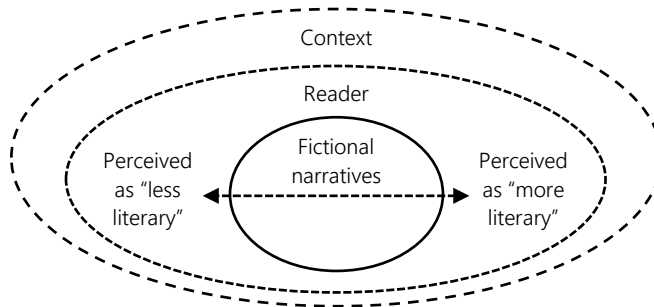
2.3 *Literariness*

TDLT was situated in the context of *literature* education in Dutch 10th 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 per-spective and focalization, theme and motifs, gaps that must be filled in by the reader (Iser, 1980), and stylistic features that deviate from conventional lan-guage use (e.g., Van Peer, Hakemulder, & Zyngier, 2007). In the Dutch litera-ture 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 10th grade students. Usually, teachers expect their 10th 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 flash-back? 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 sub-stantiating 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, & Alexan-der, 2009, for research into discussion-based reading programs). This disserta-tion contributes to the existing body of research on dialogic teaching by apply-ing 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 dia-
logue), 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. Ros-
enblatt 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 imple-
mented 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, how-
ever, demonstrated that internal dialogues did not come naturally in our sam-
ple of 10th 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 dia-
logues 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, reason-
ing, and asking questions stimulate thinking and the development of know-
ledge. 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 think-
ing 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 TDT-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, dela 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 10th 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 10th 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 nation-wide 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 10th 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. Morover, 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, Handelzalts, 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 10th 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.