TOWARD THE VALIDATION OF A LITERATURE CLASSROOM INTERVENTION TO FOSTER ADOLESCENTS' INSIGHT INTO HUMAN NATURE

An iterative design process

MARLOES SCHRIJVERS*, TANJA JANSSEN*, OLIVIA FIALHO** & GERT RIJLAARSDAM*

**University of Amsterdam ** Utrecht University*

Abstract

We report on the design process 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. Starting from a model of transformative reading, an exploration of the educational context, and a review of previous intervention studies, we designed an intervention in an iterative process. We evaluated the validity and practicality of two versions of the intervention. From teacher and student data, we concluded that the validity and practicality of the first version were suboptimal and identified various suggestions for improvement. In a second iteration, the initial design principles were reoperationalized. Based on these reoperationalized principles, we designed a second version of the intervention, which was found to be sufficiently valid and practical. In addition, the second iteration led to specifiying the initial design principles, by formulating subprinciples for operationalization in the classroom. All in all, this study demonstrates that an iterative design process is needed to arrive at a valid and practical intervention, and that this process may have the potential to further specify initial design principles.

Keywords: literature education, design study, design principles, validity, practicality

Schrijvers, M., Janssen, T., Fialho, O. & Rijlaarsdam, G. (2019). Toward the validation of a literature classroom intervention to foster adolescents' insight into human nature: An iterative design process. Contribution to a special issue Systematically Designed Literature Classroom Interventions: Design Principles, Development and Implementation, edited by Marloes Schrijvers, Karen Murphy, and Gert Rijlaarsdam. L1-Educational Studies in Language and Literature, 19, 1-47. https://doi.org/10.17239/L1ESLL-2019.19.04.02

Corresponding author: Marloes Schrijvers, University of Amsterdam, email: m.s.t.schrijvers@hva.nl

1. INTRODUCTION

Intervention studies in literature classrooms empirically evaluate whether a particular instructional approach helps students to achieve particular 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 also 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 et al. called for more comprehensive descriptions of interventions in order to strengthen their *replicability*. Interventions often tend to remain 'black boxes': it is unclear what happened in the classroom, why it happened, and how it was developed. This threatens the validity and replicability of interventions and hampers detailed insights into domain-specific instructional activities that are designed to achieve particular aims (also see Schrijvers, Janssen, Fialho, & Rijlaarsdam, 2018).

In this paper, we aim to respond to the call for more comprehensive and replicable intervention descriptions, by describing in detail the iterative design process of an intervention for 10th grade literature classrooms that aims to foster students' insight into human nature—insight into themselves, fictional others, and real-world others. Designing and validating an intervention is not so different from designing a measurement instrument. Both start by reviewing relevant literature, conducting exploratory work, and developing a theoretical model, which is, in case of an intervention, identifying design principles. In a first iteration, an initial construct is designed: a first version of an intervention, or a set of items in an instrument. In practice tests, data are collected to assess validity and practicality (of interventions) or reliability and feasibility (of instruments). If these are unsatisfactory, a second iteration will

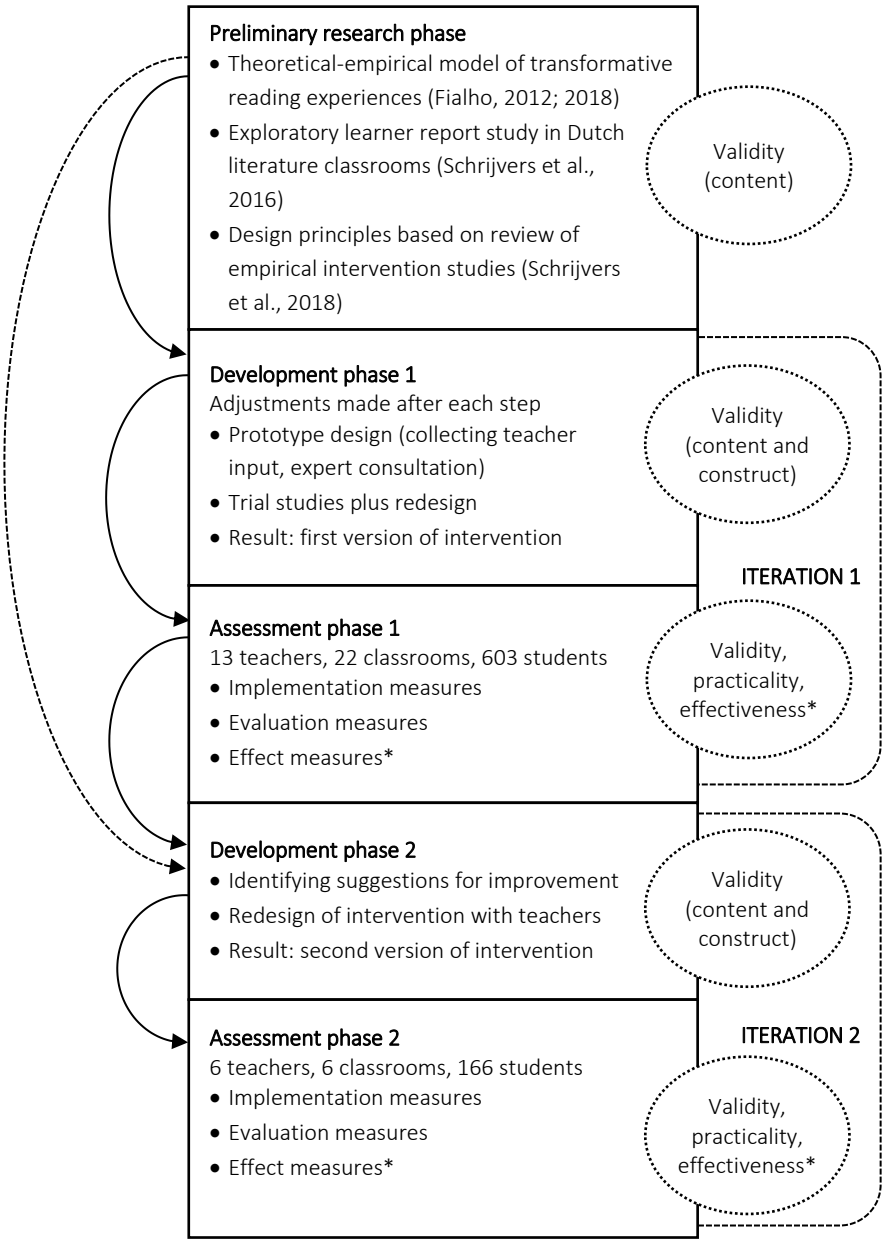
follow: the construct and/or the operationalization is adjusted and tested again. Iterations continue until a valid, practical intervention or measurement instrument is constructed. Just like instrument design may contribute to theory about the measured variable, intervention design may further specify the initial design principles in terms of their operationalization in practice.

Therefore, 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 1 shows, the design process 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 10th 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. 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. There are only three objectives that students work toward: 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 one's own life, or considering how thoughts, feelings and behaviors of characters in fictional situations represent human responses to similar real-life situations.

This paper is organized along the lines of Figure 1. We first outline the preliminary research results, including three initial design principles, that we based on a systematic review study of previous (quasi-)experimental intervention studies in literature classrooms. In the Method section, we describe the design process, the teachers and students involved, and the instruments used to assess the implementation and evaluation of two subsequent versions of the intervention, which yielded information about their validity and practicality. In the Results section, we subsequently address: a) the development phase of iteration 1, resulting in the first version of the intervention; b) the assessment phase of iteration 1, indicating how valid and practical the first version has been; c) the development phase of iteration 2, that concerns the reoperationalization of the initial design principles and results in a second version of the intervention; and d) the assessment phase of iteration 2, which illuminated the validity and practicality of the second version. We end this section with an overview of how the design process affected the initial design principles, after which we turn to the Discussion.

Figure 1. Iterations and phases in design process, attending to validity and practicality.



Note. * Effectiveness is not discussed in this paper (see Schrijvers, Janssen, Fialho, De Maeyer, & Rijlaarsdam, 2019).

2. OUTCOMES OF THE PRELIMINARY RESEARCH PHASE

The preliminary research phase consisted of the development of a theoretical-empirical model of transformative reading (Fialho, 2012; 2018), an exploration of students' learning experiences about themselves and others 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).

2.1 A 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 in 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 (2012; 2018) distinguished two outcomes of transformative reading—insights into oneself and into others—and identified six underlying components. The model was based on in-depth interviews with adult readers who talked 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 (*evaluation of protagonist*), 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*).

The transformative reading model (Fialho, 2012; 2018) has not yet been validated for adolescent readers. However, various 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 focused on non-literary reading. 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 'transformative' modes of reading.

2.2 Context: Do students learn about self and others in the literature classroom?

We asked students in Dutch upper secondary literature classrooms to complete a written learner report and found that learning about self and others was among their learning experiences (Schrijvers et al., 2016). 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 in classes of teachers who reported that they allowed for more student autonomy and interaction in the classroom.

Literature education thus appears to 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 development, and helps them to think about people's choices, about themselves, others, and the world (Curriculum.nu, 2018). Moreover, teachers consider fostering students' personal growth or personal development to be an important aim of literature teaching (Janssen, 1998; Oberon, 2016).

Fostering personal growth and social development in education is not new, as is illustrated by numerous widely-implemented approaches to teaching and learning such as social and emotional learning (Elias et al., 1997), citizenship education (Derricott, 2014), and moral and character education (Nucci, Krettenauer, & Narváez, 2014). Philosophers, too, have argued that education plays a vital role in guiding young people toward reflection on what it means to be human (e.g., Biesta, 2007; Nussbaum, 2010). Curriculum designers and researchers suggest that literature teaching, specifically, may contribute to students' personal and social development, as is indicated by handbooks for teachers (e.g., Beach, Appleman, Hynds, & Wilhelm, 2011; Langer, 2011; Wilhelm, 2016) and small-scale qualitative studies (e.g., Banks, 2009; Bender-Slack, 2002). However, experimental and quasi-experimental studies in this area are relatively scarce. Such studies shed light on potential causal effects of instructional approaches on students' learning about themselves and others in the literature classroom. In the next section, we therefore summarize the findings of a systematic review that resulted in a set of empirically grounded design principles.

2.3 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). We selected studies that aimed at fostering some form of insight into self, fictional others, or real-world others, were conducted in regular, first-language sec-

ondary 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). These studies elicited three underlying design principles.

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 these would impact themselves and others (see Mar & Oatley, 2008). As researchers hardly considered whether the fictional texts used were *literary* texts, the review remained inconclusive as for how to conceptualize 'literariness' and whether it might successfully foster students' insight into human nature.

Empirical literary studies into the effects of reading on with adult readers' self-other insights, however, suggest that literary texts in particular may evoke such insights, for various reasons: because characters in literary texts are more complex than in non-literary texts (Mar & Oatley, 2008), because striking and unconventional language may make the content of the text more vivid, which may cause the reader to look at the world differently (Miall & Kuiken, 1994), and because ambiguity and indeterminacy in literary texts may evoke deeper (self-)reflection than less ambiguous non-literary texts (Koopman & Hakemulder, 2015). Likewise, the model of transformative reading is a model of *literary* response, taking the perspective that 'responses to literary texts combine verbal, emotional and cognitive elements that may account for the distinctiveness of the literary experience. In this sense, then, the concept of literariness [is] revisited and shaped by evidence-based observations' (Fialho, 2012, p. 22; also see Fialho, 2007). In other words, literariness comes about in the interaction between text and reader (cf. Rosenblatt, 1938/1968), in such a way that transformations in self-other concepts may occur. All in all, even though the studies included in the review did not position themselves in terms of literariness, we hold the position that reading and responding to *literary* texts may be key for developing 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 reflect on responses 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, 1938/1968), because students may become aware of their responses to 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 guided internal and external dialogues with and about texts and to focus on the transformative reading experiences outlined in section 2.1.

2.4 Aims and research questions

Our aims and research questions were twofold. First, we aimed to design, in an iterative process, a valid and practical construct of an intervention aimed at stimulating students' insight into human nature in the literature classroom. Our first research question is: *What is the eventual construct of a valid, practical intervention that is based on the model of transformative reading and the design principles concerning text choice, internal dialogues and external dialogues?* Second, we aimed to reflect on how the initial design principles are affected by their operationalization in the classroom, which resulted in the second research question: *Which changes to the initial principles—e.g., removing, reformulating, adding principles—can be made as a result of the design process?*

3. METHOD

We first address the design process, which we combine with presenting the samples of teachers and students involved (3.1); we follow the phases as visualized in Figure 1. We then indicate which instruments were used to collect data and how these data were analyzed (3.2).

3.1 Design process and participants

3.1.1 Development phase 1

When designing a prototype, we first published a call for input on social media. Six teachers of Dutch language and literature responded to this call and shared their ideas with us via email. We asked for practical suggestions, for instance, concerning the number of units, as well as for examples of suitable texts and tasks. We also discussed how transformative reading experiences could be integrated in the units, and designed a prototype.

We discussed the prototype with four experienced teachers in an expert consultation session and made several adjustments. Subsequently, two female teachers tested the adjusted prototype in their classrooms. The first author observed all lessons. After adjusting the prototype, we named the intervention *Transformative Dialogic Literature Teaching,* or TDLT.

3.1.2 Assessment phase 1

This 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 the units to either one or two classes (22 classes in total), in two to four weeks. The teachers had on average 14.4 years of teaching experience in upper secondary education (*SD* = 10.9); two of them were male. Student participants (*N* = 603) 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.

In this phase, we assessed the implementation and evaluation of the first version of the intervention (TDLT-1). We focused on three aspects:

1) Content validity, by collecting evaluation data and looking for endorsements and skepticism regarding TDLT-1 and its relevance.
2) Construct validity, by collecting data concerning coherence and clarity of TDLT-1.
3) Practicality, by analyzing whether units were implemented as intended, and by focusing on data that concerned feasibility.

Whereas validity and practicality 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.

3.1.3 Development phase 2

Data from assessment phase 1 yielded suggestions for improvement. Based on these suggestions, the first author redesigned TDLT-1 together with three teachers who had been involved in the first assessment phase. Two of them had five years of teaching experience; the third had been a teacher for 28 years. In the redesign the initial design principles were maintained, but operationalized differently. For example, a suggestion for improvement was that students needed more time to learn to engage in internal and external dialogues; thus, we maintained a dialogic approach, but decided to devote six units to TDLT-2.

3.1.4 Assessment phase 2

Six teachers taught TDLT-2 to a single class: those involved in development phase 2, and three new teachers. They had on average 11.7 years of experience in upper secondary education (*SD* = 9.7); 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. The intervention 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. We evaluated validity and practicality, after which a third design iteration was not deemed necessary.

3.2 Instruments and data-analysis

We used five instruments to collect data of teachers and students: (a) teacher logs, (b) time on task observations, (c) teacher interviews as well as (d) student evaluation forms, and (e) student evaluation tasks. Table 1 indicates in which indicators of validity and practicality (appreciation and relevance, structure and coherence, clarity and feasibility) the instruments provided insight, in which assessment phase instruments were used (1 or 2, or both), and how many responses and observations we collected. Below, we provide details of instruments and analysis.

3.2.1 Teacher logs

Teachers completed one online log per unit. Each unit 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 were asked to indicate why (e.g., 'not enough time', 'forgot about it') and to add elaborations. We also asked teachers to evaluate fully or partly completed phases. They indicated on 5-point scales:

A. how interested and engaged students seemed to be;
B. how clear the phase seemed for students;
C. how attainable it was to teach the phase;
D. how much order and discipline there was in the classroom.

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

Table 1. Overview of instruments and derived quality indicators.

Instrument	Quality indicators	Assessment phase	Response rate or *N* observations
Teacher logs	*Content validity*: evaluation statement A, appreciation, relevance for students *Construct validity*: evaluation statement B, clarity for students *Practicality*: - Percentage of fully, partly, not completed phases - Evaluation statement C, feasibility of teaching a phase - Evaluation statement D, feasibility of maintaining order in class	1 2	96.6% (for 21 phases from 4 units, in 22 classes) 94.5% (for 31 phases from 6 units, in 6 classes)
Time on task observations	*Practicality*: percentages time on task (overall, and for intended activities)	1 2	*N* = 1690 (in 22 visits; 1 unit in each class) *N* = 877 (in 12 visits; 2 units in each class)
Teacher interviews	*Practicality, content and construct validity*: indications of appreciation and relevance, structure and coherence, clarity and feasibility	1	100% (all 13 teachers)
Student evaluation form	*Content validity*: - 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 *Construct validity*: - Overall clarity and comprehensibility - Clarity and comprehensibility of dialogue tasks *Additional support for indicators*: strengths, suggestions for improvement	1	85.4%
Student evaluation task	*Content validity*: relevance of intervention elements, story appreciation *Construct validity*: coherence between intervention elements and objectives *Practicality*: story difficulty	2	90.4% (story items) 70.5% (intervention elements)

3.2.2 Time on task observations

To assess the proportion of available learning time that students were engaged in intervention tasks, which is an indication of teaching quality, trained research assistants conducted time on task observations (Karweit, 1984). Six students in the classroom were randomly selected for observation. Each student was observed for about one minute, during which the observer scored twice (after 20 seconds) whether the student had been mainly on task or off task. A code for off task was assigned if the student was clearly not engaged in the lesson content (e.g. chatting, looking at cell phones). 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 see whether a student was on task or off task (e.g., if another student moved into the line of sight) we used 'unclear'. For each observation moment, we also coded the intended learning activity: 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. After all six students had been observed twice, the same students were observed again, in the same order. These rounds continued until the end of the lesson. All units of TDLT-1 were observed at least once. Students were told that the observer was present in the classroom 'to see what happens in the unit'. They were not aware that task behavior was observed, nor did they know which six students were observed.

3.2.3 Teacher interviews

The first author held semi-structured interviews with the teachers involved in assessment phase 1, within maximum three weeks after the intervention. To stimulate recall, she encouraged them to browse through TDLT materials if needed. The interview guideline consisted of three parts:

A. *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.
B. *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?

C. *Tips for new teachers.* Teachers were asked about tips and tricks for new teachers who would teach TDLT. Finally, they could indicate whether there was anything else left to discuss and received a gift card as a token of appreciation.

As the guideline indicates, the teacher interviews were largely based upon teachers' observations and interpretations of what happened in their own classrooms. Because the interview data represented only a small share of a much larger data set, we chose not to engage in detailed transcriptions and inductive coding. Rather, we were interested in teacher experiences and observations that either substantiated or contradicted findings from other data sources. Therefore, we analyzed the interviews directly from notes and audio recordings and looked selectively for information that helped us in interpreting the data from teacher logs, evaluation forms, and time on task observations. In particular, we focused on observations and experiences shared by multiple teachers.

3.2.4 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' '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.% 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 score (1 = *did not like it at all*; 10 = *liked it very much*). 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). Presumably, 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 (e.g., 'When talking in small groups, I felt sufficiently at ease to share my thoughts and experiences') formed an internally consistent scale (Cronbach's α = .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 the intervention 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.

3.2.5 Student evaluation task

In assessment phase 2, we assessed story appreciation and difficulty selectively (1 = *did not like it at all / not at all difficult*; 5 = *liked it very much / very difficult*), by asking students to evaluate these aspects for one story only: the one they read for the final TDLT-2 task, which they had chosen from four options.

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 in 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 indicated 1784 times that an element had been helpful (on average 15 indications per student).

4. RESULTS

4.1 First development phase

This phase resulted in TDLT-1. Teachers indicated that they preferred a four-lesson project (of 50 minutes each, the conventional length at Dutch schools), in view of their regular teaching program. We decided that each TDLT unit should fit within a single lesson; TDLT-1 thus was a four-unit intervention.

Its overall objective was, as described in the teacher guideline, 'to help students identify connections between literary short 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 a preparatory unit, aimed at learning strategies for engaging in peer-led dialogues about stories, and three 'reading-and-dialogue' units. In students' work books, we presented for each unit the learning objectives, instructions, and tasks including writing space. Stories were bundled in a separate booklet. TDLT-1 is described in detail in Appendix A; here, we address the operationalization of the initial design principles.

4.1.1 Operationalization of text choice principle

The text choice principle indicated that thematically relevant fictional texts should be selected, but did not offer any directions in terms of 'literary' texts. Yet, other studies suggest that literary texts in particular may offer readers the opportunity to gain insight into themselves and others (see section 2.3). We therefore decided to select literary fictional texts and asked teachers in the first development phase about the kinds of literary texts they usually read in their classrooms. Their suggestions led us to choose a particular genre: the literary short story. This genre is often used in 10th grade literature classrooms. The choice for short stories with a reading time of 10-15 minutes, rather than a novel or play, allowed us to offer students a variety of texts. It also enabled to organize dialogic activities prior to and after reading, in a single unit.

We originally planned to offer students freedom of choice in reading materials to increase engagement (e.g., Lenters, 2006). However, consulted experts suggested that reading the same—thus preselected—stories would enable 'deeper' talk about themes and related insights into themselves and others. Moreover, 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; see Vezzali et al., 2012). We decided to preselect literary short stories for students to read. These were written in Dutch by acclaimed Dutch and Flemish literary authors who are represented by renowned publishers of literary fiction in the Netherlands and Belgium. In line with the design principle, all stories addressed peculiar, presumptuous or painful social interactions between characters (see Table 2 for descriptions). We selected both canonical short stories often used in 10th grade (e.g., *A plate with spaghetti* and *Blood*) and more recently published stories (e.g., *She was everywhere* and *Flight behavior*). Students read a story in units 2, 3 and 4; for unit 3, they chose from five options.

Table 2. Stories used in TDLT-1.

Story and author	Description	Unit
Ze was overal ('She was everywhere'), 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
Een bord met spaghetti ('A plate with spaghetti'), Adriaan van Dis	White man in restaurant thinks a black man stole his plate; he passive-aggressively confronts him, but finds out he misjudged.	3*
De biefstuk van het zoete water ('The freshwater steak'), Hans Dorresteijn	Boy is nervous during a fishing trip, because his father always physically punishes him several days after he has misbehaved.	3*
Het recht ('The right'), Annelies Verbeke	White man distrusts his black cleaning lady; he tries to trick her into stealing, but she gets the better of him.	3*
Bloed ('Blood'), Gerard Reve	Child is physically abused by guardian and takes revenge by causing him to take a deathly fall.	3*
Merkwaardig verhaal ('Curious story'), 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*
Vluchtgedrag ('Flight behaviour'), 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

4.1.2 Operationalization of the internal dialogue principle

To stimulate internal dialogues, we designed writing tasks focused on activating previous personal experiences before reading, noticing experiences during reading, and/or reflecting on evoked experiences directly after reading. We therefore used prereading tasks, reading instructions, and individual reflection tasks that focused on story themes and transformative reading experiences. For example, in unit 4, students wrote about their ideas of an afterlife, prior to reading about someone in a plane crash who has, to his own disbelief, a religious experience. 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, an example of reflection directly after reading was to indicate which part of the story stood out most (stimulating aesthetic awareness; cf. Fialho, 2012) and to which extent the story evoked, for example, imagery and sympathy.

4.1.3 Operationalization of the external dialogue principle

Expert consultation already indicated that constructive, meaningful peer-led dialogues might be difficult to achieve: teachers felt that students needed explicit strat-

egies to do so. The teachers involved in the trial study observed that students struggled to deepen their talk about stories and reading experiences and offered suggestions for improvement. We adjusted the prototype accordingly. In the preparatory unit of TDLT-1, students first considered what characterizes good dialogues and were given 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'). They observed their teacher modeling how to ask follow-up questions and talked in small groups about a reading-related topic, using questioning cards (e.g., 'Can you give an example?', 'Who has a different opinion?'). Small-group experiences were shared in class and students reflected individually, in their work books, on strengths and points for improvement of their dialogues. In the reading-and-dialogue units, dialogues in pairs or small groups always preceded classroom talk. For example, 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, which were presented by group representatives to the class. Following a suggestion from one of the consulted experts, we included in the teacher guidelines concrete suggestions for teachers' roles and their interaction with students; these were further specified after the trial study.

4.2 First assessment phase

In Table 3, we summarize the findings regarding the validity and practicality of TDLT-1, for the intervention as a whole, selected stories, internal dialogue tasks, and external dialogue tasks. This table includes quantitative results from teacher logs, time on task observations, and evaluation forms, as well as summarized findings from teacher interviews; Table 4 contains example responses from these interviews.

4.2.1 TDLT-1 as a whole

Content validity. Teachers indicated that students were interested and engaged in TDLT-1. Students themselves neither highly valued nor highly disliked it, but agreed to feel safe during the lessons. Students' mixed responses were reflected in strengths and suggestions for improvement they listed (for all strengths and suggestions listed by students, see Figure 2 and 3). For example, 23.7% of the strengths referred to valuable, relevant learning outcomes, whereas 21.7% of the suggestions for improvement concerned TDLT-1 not being fun or engaging. Teachers agreed in interviews that students needed clearer goals and more insight into the steps to work toward those goals.

Table 3. Overview of validity and practicality indices for TDLT-1.

Topic	Content validity	Construct validity	Practicality
TDLT as a whole	*Student evaluation forms* - Overall appreciation and relevance: $M = 3.0$; $SD = .8$; $M < 3.0$ and $M > 3.0$ evenly distributed - Sense of safety: $M = 3.9$, $SD = .7$ - Strenghts: valuable, relevant learning outcomes (23.7%); general strengths, e.g. 'important' or 'fun' (7.5%); safe social atmosphere in class (2.3%) - Improvements: TDLT-1 not fun or engaging (21.7%); purpose and relevance questioned (6%). *Teacher logs* - Appreciation and relevance for students: $M = 3.8$, $SD = .8$ *Teacher interviews* [1] - Students in need of clearer goals and more insight into the steps to work toward those goals, e.g., with a rubric.	*Student evaluation forms* - Overall clarity, comprehensibility: $M = 3.6$, $SD = .8$ - Strengths: clarity of units (8.5%); pace and structure of units (3.7%) - Improvement: more variety in units (13.1%) *Teacher logs* [2] - Units clear for students: $M = 4.1$, $SD = .5$	*Student evaluation forms* - Improvement: feasibility (12.2%), either not feasible or not challenging enough. *Teacher logs* - 88% of phases fully completed, 9.4% partly completed, 2.7% not completed. - Teaching attainable: $M = 4.2$, $SD = .5$ - Order and discipline: $M = 4.0$, $SD = .6$. *Students' time on task* On task 72.5%, off task 24.9%, with large variety among: - Units: $\chi^2(6) = 15.09$, $p = .020$ - Teachers: $\chi^2(24) = 124.21$, $p < .001$; range 54.1 - 84.8% - Activities: $\chi^2(8) = 56.54$, $p < .001$ *Teacher interviews* [3] - Active pedagogy feasible for some teachers, but not for others. - Suggestions: develop an introductory workshop.

Selected stories	*Student evaluation forms* - Stories read by all students: *She was everywhere, M* = 6.3, *SD* = 1.3; *Flight behavior, M* = 6.7, *SD* = 1.3; other stories 6.1 - 6.9, similar *SD*'s. - Story selection mentioned as 9.1% of strengths, but as 17% of suggestions for improvement. *Teacher interviews* [4] - Overall positive, some reservations about one story; however, no teacher would replace a story.	*Teacher interviews* [5] - Suggestions to select stories centering around a single theme.	*Students' time on task* - 91% on task for story reading *Teacher interviews* [6] - Students mostly finished reading in time. - Some teachers read aloud to increase practicality. - Short stories more practical than fragments from novels.
Internal dialogues	*Student evaluation forms* - Appreciation and relevance: *M* = 2.7; *SD* = .7. - Strenghts and suggestions for improvement: internal dialogues not mentioned *Teacher interviews* [7] - Tasks rather open; purpose not explicit enough for students. - Tasks nonetheless evaluated positively. - Students nonetheless generally engaged.	*Student evaluation forms* - Clarity and comprehensibility: *M* = 2.8; *SD* = 1.0 (applies to internal and external dialogues). *Teacher interviews* [8] - Internal dialogue tasks clear and well-structured. - Internal dialogue tasks sometimes created coherent starting points for external dialogues. - Suggestions: teacher think-alouds of responses during reading; offering step-by-step strategies to deal with difficulties in stories.	*Students' time on task* - 77% on task for individual activities *Teacher interviews* [9] - Tasks were practical and well-organized. - Annotating or highlighting the text when noticing a response and guided reflection after reading worked well for students.

External dialogues	*Student evaluation forms* - Appreciation and relevance of external dialogues: M = 2.8; SD = .8 - Talking in groups mentioned as 26.6% of strengths and as 7% of suggestions for improvement *Teacher interviews* [10] - Mixed evaluations, e.g.: positively surprised, observing that relevance for students was not always clear; observing a lack of follow-up questions, superficial talk, and disinterest in literature. - Suggestions: to offer a set of follow-up questions to use anytime, and to clarify why talking about literature can be beneficial.	*Student evaluation forms* - Clarity and comprehensibility: *M* = 2.8; *SD* = 1.0 (applies to internal and external dialogues). *Teacher interviews* [11] - Dialogue guidelines from unit 1 provided coherence, helped to engage in dialogues. Guidelines often returned in all lessons.	*Students' time on task* - 74.4% on task for whole-class activities - 65.5% on task for dialogues in pairs or small groups *Teacher logs* - Most non-completed phases occurred toward end of unit: not enough time left for classroom dialogues. *Teacher interviews* [12] - Suggestions to leave out whole-class dialogues or shift them to next lesson; units may cross boundaries of lessons. - Suggestions to spend more time on TDLT; students get used to talking about reading experiences. - Suggestions to include questions and prompts in teacher guidelines to stimulate students to deepen dialogues.

Note. [1-12] Example responses from teacher interviews; see Table 4.

Table 4. Example responses from teacher interviews about TDLT-1, corresponding with Table 3.

Topic	Quality indicator	Example responses
TDLT-1 as a whole	1. Content validity	'A rubric for 'noticing reading experiences' and talking about it might be helpful.'
	2. Construct validity	'The build-up and structure of the tasks offered students something to hold on to, because they returned in all the next lessons. They knew what to expect.'
	3. Practicality	'It's something else'; 'It challenges students'; 'You can organize so much more in a lesson than I expected'; 'It's too much fuss'; 'It's strange for students, they are too passive for this.'
Selected stories	4. Content validity	'*Flight behavior* was really imaginable for them and concrete enough to talk about, they can really put themselves in the [plane crash] situation'; '*She was everywhere* has a recognizable setting and comprehensible style, but maybe it also is, you know, intangible, too open.'
	5. Construct validity	'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.'
	6. Practicality	'Short stories are convenient. They allow you to offer the sense of a beginning and end in a single lesson.'
Internal dialogues	7. Content validity	'What students find strange is that these writing tasks are very open. For them it feels […] as if anything goes'; 'Thinking about a theme prior to reading is valuable and safe. We should do it more often'; 'My students completed the tasks rather seriously and felt they were heard and taken seriously.'
	8. Construct validity	'The tasks had a clear buildup. I think that helped my students focus'; 'This went really well. The students understood the steps and had enough time to complete the tasks'; 'Things like 'Huh?' or 'I don't get it' are genuine responses that occur frequently among 10th graders, but they often just don't know what to do when it happens.'
	9. Practicality	'Highlighting where you notice something in particular [a certain experience, *MS*] helps, because it narrows down what you are asked to reflect upon.'
External dialogues	10. Content validity	'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'; '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,' 'The dialogues were… well, mediocre. Students are too easily satisfied. They don't ask follow-up questions, it just doesn't interest them. The talk remains artificial.'
	11. Construct validity	'Using these dialogue guidelines, it doesn't come naturally. I constantly reminded them.'
	12. Practicality	'Yes, prompting. 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. But it can be quite difficult to prompt, without being too directive.'

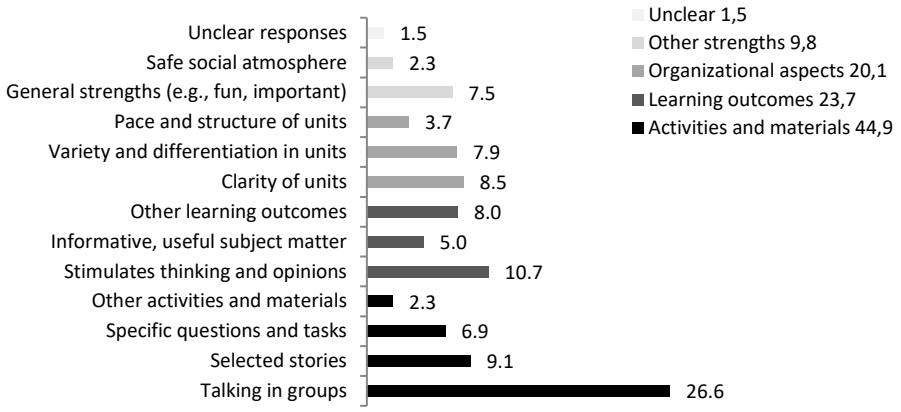


Figure 2. Categories of strengths listed by students (% of 683 strengths).

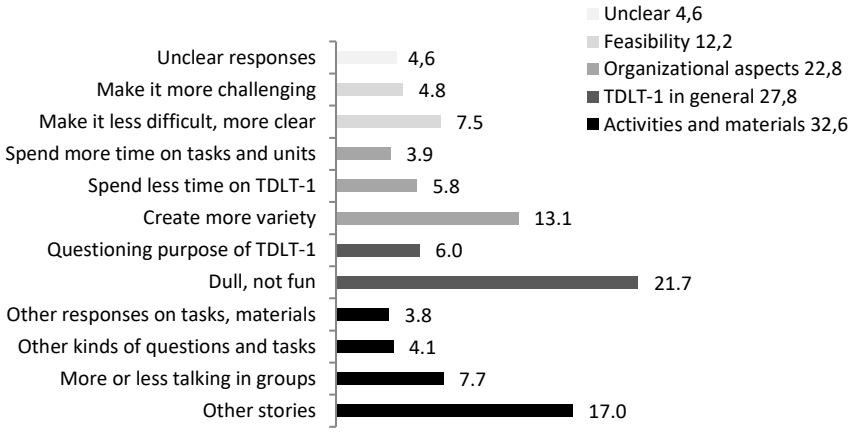


Figure 3. Categories of suggestions for improvement listed by students (% of 482 suggestions).

Construct validity. Students were neutral to positive about the overall clarity and comprehensibility of TDLT-1. Some of the strengths they listed concerned the clarity, pace and structure of the units. As 13.1% of the suggestions for improvement concerned calls for more variety in the activities, we hypothesized that TDLT-1 might even have been *too* coherently structured. Teachers indicated that the units were clear for students.

Practicality. Although some students commented on the feasibility of TDLT-1 (e.g., 'It was hard to complete everything in detail in the time that we had'), teachers indicated that the large majority of the unit phases was completed as intended. Teaching the phases was attainable for them, and there was order and discipline in the classrooms. However, the overall on-task percentage was below the standard of 80% that has been suggested in studies on effective teaching (e.g., Kauchak & Eggen, 1993; Muijs & Reynolds, 2010). It varied significantly among units, due to unit 3 (67.7% versus 73.1%, 73.9%, 74.2%), as well as among teachers: only two of them scored above the 80% norm. Moreover, the on-task percentage varied among intended activities, as percentages for story reading, internal and external dialogue tasks in Table 3 indicate. Teachers related practicality to the active pedagogy of TDLT-1 (e.g., short phases in a high pace, organizing group work and short presentations), which they either appreciated or not; some would have appreciated a workshop to become familiar with TDLT.

4.2.2 Selected stories

Content validity. Table 3 shows that students evaluated the stories neutrally to positively, although *SD*'s indicated considerable variety. On the other hand, 17% of their 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*. Some teachers had reservations about *She was everywhere*. Yet, when asked if they would replace a story, none of them indicated that they would.

Construct validity. In student data, we found no responses about the coherence of the stories. Three teachers suggested to increase coherence by selecting stories that center around a single theme.

Practicality. The high on-task percentage for story reading indicated that teachers successfully engaged their students in this activity. Teachers indicated the stories were practical in terms of reading time, although some chose to read aloud so that students finished and started the next task simultaneously. Due to the relatively short reading time and, consequently, the opportunity to organize activities prior to and after reading, short stories were considered a practical genre to use in class.

4.2.3 Internal dialogues

Content validity. In terms of appreciation and relevance, students did not value internal dialogue tasks highly (see Table 3). They did not refer to these tasks in their strengths and suggestions for improvement. Possibly, internal dialogues were not prominent enough for students to reflect upon and, potentially, value them. Several teachers endorsed that the purpose of internal dialogues could be more explicit.

Apart from this suggestion, teachers evaluated prereading tasks, tasks to notice responses, and reflection tasks after reading positively; students appeared to be generally engaged in the tasks.

Construct validity. Whereas students evaluated the clarity and comprehensibility of internal dialogue tasks negatively to neutrally, teachers felt the tasks were clear and well-structured. In terms of overall coherence, the tasks sometimes created coherent starting points for external dialogues, which were then more profound than the teachers usually observed; however, other teachers said this happened 'not as much as you would expect'. Students seemed to struggle to notice responses during reading and to actively deal with difficulties in stories: they would often ignore difficulties and simply continue reading. Therefore, teachers suggested to include teacher think-alouds and 'step-by-step strategies to deal with difficulties', to enhance the clarity and comprehensibility of internal dialogue tasks.

Practicality. Despite being difficult for students, teachers mostly felt that internal dialogue tasks were practical and well-organized. Teachers further indicated that annotating or highlighting the text when noticing a response and guided reflection after reading worked well for students. The on-task percentage for individual student activities was 77%.

4.2.4 External dialogues

Content validity. Similar to internal dialogues, students did not value external dialogues highly, even though 'talking in groups' was the most frequently mentioned strength. Teachers considered talking about stories to be valuable, but their evaluations were mixed. Some teachers were positive, but most 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.

Construct validity. As indicated, students evaluated the clarity and comprehensibility of dialogue tasks negatively to neutrally. In terms of coherence, teachers mostly felt that the guidelines, introduced in unit 1, helped students to engage in dialogues: therefore, they actively reminded students of the guidelines.

Practicality. Concerns were raised about the practicality of external dialogue tasks. During whole-class activities, students were on task in 74.4% of the time. In contrast, they 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. Moreover, logs indicated that there was not always enough time for whole-class dialogues at the end of a lesson. Three teachers suggested to shift these dialogues to a next lesson. Units, thus, could potentially cross the boundaries of 50-minute lessons. Finally, teachers suggested to list example questions and prompts in the teacher guidelines, to offer new teachers more guidance in scaffolding students' dialogues.

4.3 Second development phase

The second development phase resulted in a six-unit intervention: TDLT-2. Again, it started with a preparatory unit in which the main objective was to understand and apply strategies for engaging in and deepening external dialogues about stories and reading experiences. The subsequent units were 'reading-and-dialogue' units. To 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 B; here, we present specific reoperationalizations of the initial design principles.

4.3.1 Reoperationalization of text choice principle

Several teachers suggested that more thematical coherence would help students to identify relations between stories and, consequently, to consider more deeply how a story theme might be connected to themselves and their perceptions of others. For TDLT-2, we selected stories centering around 'justice and injustice'. We maintained some stories from TDLT-1, but substituted others (see Table 5). The most appreciated story from TDLT-1 (*Flight behavior*) did not fit the theme, but was given another role: we recorded two short videos of students modeling a good and weak example of a dialogue about this story, which were observed and discussed in the preparatory unit of TDLT-2.

Table 5. Stories used in TDLT-2.

Story and author	Description	Unit
Excerpt from *Vluchtgedrag,* Bertram Koeleman	See Table 2.	1
Dood ('Death'), Martin Bril	A girl realizes her love interest has died because of senseless violence, and thinks about all that will never happen anymore.	2 and 5*
Bloed, Gerard Reve	See Table 2.	3 and 5*
Volgens de regels ('Following the rules'), Mirjam Bonting	A father has always set strict rules for his daughter; when he falls during mountain climbing, she follows the rules and leaves him behind.	4 and 5*
Van geluk spreken ('Count oneself lucky'), 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
De biefstuk van het zoete water, Hans Dorresteijn	See Table 2.	6**
Een najaarsdag ('An autumn day'), 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**
Het recht, Annelies Verbeke	See Table 2.	6**
Hoela ('Hula'), Cees Nooteboom	From behind a window, a boy at a family birthday party watches his little nephew drown in a pond in the garden, 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.

4.3.2 Reoperationalization of internal dialogue principle

The most important change regarding internal dialogue tasks was to explicate for students *why* and *how* they learn to engage in internal dialogues. We designed a self-evaluation 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 in reading experiences, for which three levels were described (see Appendix C). Students indicated their starting level in unit 1 and evaluated their progress at the end of the intervention.

In addition, we postponed explicit attention to 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.

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 engaged in a think-aloud during reading, while being observed by students. Only after these

explicit preparations, students were asked to focus on their 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.

4.3.3 Reoperationalization of external dialogue principle

Several major adjustments were made to the operationalization of the external dialogue principle. First, we extended TDLT 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 being in the protagonist's position, and talked from that point of view about possible story ends. 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. On 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 the units were taught in four to eight weeks, instead of 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 the purpose and relevance of sharing reading experiences in unit 3. 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 prompts and questions teachers could use in stimulating students to deepen a dialogue. During the workshop for new teachers, such feedback situations were imagined, discussed and rehearsed.

4.4 Second assessment phase

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

4.4.1 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 or not completed (9.4% and 2.7%). 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, teaching was attainable (*M* = 4.1, *SD* = .7) and there was order and discipline in the classrooms (*M* = 3.9, *SD* = .8). These scores did not differ significantly from TDLT-1.

The overall on task percentage was 85.2%. The on-task percentage was well above the standard of 80% (Kauchak & Eggen, 1993; Muijs & 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 lessons, nor across teachers. All teachers seemed to be equally able to engage their students in efficiently spent learning time.

4.4.2 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; range *M* = 2.9 for *Hula* to *M* = 3.8 for *An autumn day*). Students thus evaluated them neutrally to positively. Even though we used a different scale than in assessment phase 1, story appreciation was rather similar. 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 6). We collected no data that specifically indicated the construct validity of selected stories.

Table 6. Distribution intervention elements perceived as helpful (N = 1784)

Intervention element	% of total (*N* = 1784)	% of total per objective* 1	2	3	4
Observing peer modeling videos: good and less good dialogue about reading experiences	5.8	2.0	0.9	0.8	2.1
Dialogue guidelines on first aid card, e.g., listening carefully, asking follow-up questions	9.5	2.5	1.6	1.8	**3.6**
Other suggestions on card, e.g., strategies to deal with incomprehension, follow-up questions	9.6	2.4	2.1	2.3	2.9
Reading the selected stories	15.2	3.6	**4.1**	**4.1**	3.4
Observing the teacher modeling to notice responses during reading (teacher think-aloud)	14.5	**4.9**	3.1	**3.4**	3.1
Focus on noticing own responses through reading instructions (e.g., annotating)	13.7	**4.5**	3.1	**3.4**	2.7
External dialogue tasks in pairs, groups or class	14.7	3.1	**3.4**	3.1	**5.1**
Teacher's explanations about noticing responses, sharing responses, and other topics	16.9	**4.0**	**4.3**	**4.1**	**4.4**
Totals	100	27.0	22.6	23.0	27.4

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 lower than in TDLT-1 (91%).

4.4.3 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 6). Thus, at least for some 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 6 indicates, the latter was most often considered helpful for learning to notice responses. Thus, there appeared to be coherence between the objective (noticing responses) and the reoperationalization (i.e., more explanations and modeling by the teacher).

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

4.4.4 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 6). In particular, they felt these tasks helped them to learn to actively contribute to dialogues about stories and reading experiences. Support for external dialogues was less often seen as helpful (peer modeling videos: 5.8%; dialogue guidelines: 9.5%). Teachers' explanations, including those about small-group dialogues, were most helpful to students, as indicated above (16.9%). Clearly, the role of the teacher in TDLT-2 could not 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%), which was in stark contrast with TDLT-1 (65.5%).

4.5 From initial design principles to sub-principles for operationalization

The design process described in this paper started from three rather general design principles. Here, we complement them with sub-principles 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 thematical 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 10th grade students, literary fictional texts can be selected.* Following teachers' suggestions, we used literary short stories. Observations indicated that on-task percentages were high during reading these stories. 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 that 10th grade students and their teachers considered reading literary texts to be appropriate and helpful for the objective 'gaining insight into reading experiences', even though the effects on students' insight into human nature requires further analysis.

2) *To stimulate an internal dialogue between reader and text as preparation for external dialogues, 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 reflect on responses directly after reading.*

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 10th grade are not necessarily used to paying attention to the responses evoked by a literary text. TDLT-2 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 TDLT-2: 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 transformative reading experiences.

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 fruitful dialogue.
   b. *Students should receive guidelines and other suggestions for external dialogues*. The 'first aid card' in TDLT-2 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 receive guidelines for providing students with feedback and guiding their dialogic processes*. For teachers who are new to dialogic approaches, it may be challenging to assist students in conducting a 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 of 50-minutes seemed 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.

5. DISCUSSION

5.1 Conclusions

In this paper, we aimed to answer two research questions. Our first question concerned the construct of a valid, practical literature classroom intervention that aims to foster students' insight into human nature. The iterative process of design and assessment led to the development TDLT-2, which is characterized by 1) reading stories that address an overarching social-moral theme; 2) an emphasis on learning to notice responses to stories in internal dialogues, including 3) responses related to transformative reading; and 4) to share these responses in external dialogues with peers. The eventual construct, thus, illustrates that the initial design principles (Schrijvers et al., 2018) and the theoretical-empirical model of transformative reading (Fialho, 2012; 2018) can be implemented in literature teaching practices in upper secondary schools.

Simultaneously, concerning our second research question, this study has shown that the initial design principles have been affected by operationalizations in the classroom. We have demonstrated that an iterative design process was needed to arrive at a valid and practical intervention, which resulted in further *specifying* the initial principles. In other words, due to the design process, sub-principles for operationalization could be added to the initial principles, which we believe offer practical suggestions for teachers and other curriculum designers who engage in developing a literature classroom intervention that aims to foster students' insight into human nature.

The results of this study show that TDLT-2 can be considered a more valid and practical intervention than TDLT-1, and was thus of a better quality (Nieveen, 1999). The study has also shed light on the importance of triangulation of data sources (McKenny, Nieveen, & Van den Akker, 2006), which explicated a wide range of experiences of teachers and students: from these data, we derived valuable information for redesigning TDLT. The most prominent changes from TDLT-1 to TDLT-2 include the single-theme approach, extended lesson time, and increased strategy instruction about internal and external dialogues. Yet, changes also took place on a more general level. With TDLT-2, a more fruitful learning culture appeared to arise, potentially

because TDLT-2 had a broader scope than TDLT-1. More attention was paid to integrating the development of insight into human nature in the regular literature curriculum, for example, by learning strategies to approach literary texts in case of comprehension difficulties and by more often practicing external dialogues about stories and reading experiences. Learning to engage in external dialogues was more explicitly connected to working toward the literature exam, which in the Netherlands usually is an oral exam. Such adjustments may have positively affected the relevance of the project for students: in TDLT-2, learning about self and others was framed as an integral part of becoming a more experienced literary reader.

The transition from TDLT-1 to TDLT-2 had consequences for teachers as well. In TDLT-2, the teacher were more involved in the instruction process, for example, in functioning as a model to acquaint students with internal dialogues. We aimed to prepare new teachers for their roles by organizing a preparatory workshop. However, as we will point out in section 6.2, TDLT may require fundamental discourse shifts in the literature classroom: teachers must reflect on what kind of discourse currently exists in their classroom (e.g., mainly monologic, already moving toward dialogic) and their own abilities to foster dialogic discourse in the classroom. Therefore, a workshop may in fact not suffice: implementing all aspects of TDLT, in negotiation with current practices, students' needs and expectations, and curricular demands, may require a more extensive professionalization trajectory.

5.2 Dialogic learning and the role of the teacher

As the name TDLT—Transformative Dialogic Literature Teaching—constitutes, dialogic teaching and learning was pivotal in this instructional approach. We suggest in this study that dialogues in the literature classroom may take place in two phases: between the individual reader and the text (the internal dialogue), and among several readers in response to the text (the external dialogue), in which the former functions as a prerequisite for the latter.

5.2.1 Internal dialogues

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, including 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).

Indeed, task that prompted internal dialogues in TDLT-2 stimulated students to focus on their initial, highly personal reactions and responses to literary texts. Yet, according to the general impressions of their teachers, internal dialogues did not come naturally to these students: for example, students would simply ignore difficult parts of a story and the responses they shared afterwards remained rather superficial. Therefore, in TDLT-2, we included more strategy instruction, via which students learned how to engage in internal dialogues with the text. Considering that an internal dialogue is an invisible cognitive and affective process, the teacher needs to function as a model to open up this process to students, by thinking and responding aloud while reading a text.

5.2.2 External dialogues

In TDLT, internal dialogues generated the responses to discuss in external dialogues. External dialogues enable learners to not only communicate and elaborate their own ideas, responses and interpretations, but also to take the perspective of others into account, engage in co-construction, and higher-order thinking activities (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. Researchers have debated the pros and cons of peer-led as well as teacher-led dialogues. Peer-led talk about literature may bring students to challenging and negotiating positions of power, related to 'competing identities students must address within themselves and others' (Lewis, 1997, 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, 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.

Chinn, Anderson, and Waggoner (2001) showed that moving from teacher-led toward student-led talk about literature may result in changed patterns of discourse: students may gain 'greater control over when to speak, how long to speak, and what

to speak about' (p. 403), the stance in the classroom may shift from efferent to critical-analytical and aesthetic (also see Murphy, Wilkinson, Soter, Hennessey, & Alexander, 2009), and there may be greater engagement and intellectual productivity in the classroom. 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. TDLT-2 adhered to these insights, as teachers were asked to move away from a monologic initiation-response-evaluation pattern when guiding whole-class and small-group talk. 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), which was discussed during the redesign meeting and practiced in the workshop for new teachers. 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).

5.3 Literary texts

TDLT was situated in the context of literature education in Dutch 10th grade classrooms. Although one of the initial design principles referred only to fictional texts, we chose to select literary texts for TDLT. In doing so, we not only adhered to insights from empirical literary studies (e.g., see Koopman & Hakemulder, 2015; Mar & Oatley, 2008; Miall & Kuiken, 1994) and phenomenological work (Fialho, 2012; 2018), but we also 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. After evaluating students' and teachers' experiences with these texts, we further specified the initial text choice principle that had emerged from the review study, by adding the notion of literariness to it.

5.4 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, it yielded a large variety of perspectives on and experiences with TDLT, for example, during expert consultation, trial studies and interviews. On the other hand, some teachers

mentioned that it was challenging to teach TDLT precisely because they had not been involved in developing it.

Second, the instruments used in the assessment phases yielded valuable information regarding the validity and practicality of TDLT, both from teachers and students. Digital teacher logs and time on task observations proved to be practical and time-efficient instruments for collecting implementation data; in particular, it was valuable to discuss log data in the interviews, as teachers appeared to appreciate that their logged experiences were taken up and discussed more extensively. Importantly, inspection of the data in assessment phase 1 suggested that teachers, despite being critical, evaluated TDLT-1 more positively than their students. However, we did not consider this a problem: on the contrary, it indicated that we built up a nuanced picture of validity and practicality.

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.

Fourth, in the iterative design process of TDLT, the role of teachers became increasingly important, but their roles in the process—for instance, of the three teachers who were involved in the redesign—were not outlined from the start. In hindsight, we would have preferred to collect more information about all teachers involved, as it would have been valuable to know more about their regular literature teaching practices, their motivations to participate in the TDLT studies, their expectations of designing and/or teaching, and so forth. In future studies, researchers might develop teacher portraits, or vignettes of their practices, to more strongly relate teacher information to the design process, implementation, and/or effects of a new instructional approach.

Furthermore, 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 compared teacher log data and time on task results of TDLT-1 to TDLT-2, but these findings must be considered cautiously as the groups of teachers and students may not have been fully comparable.

5.5 Implications

In this research project, we assessed not only the validity and practicality of both TDLT versions, but also their effectivity. Therefore, the lessons were embedded in quasi-experimental studies. Our next step is to examine the effects of TDLT-1 and 2,

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. As an implication of integrating intervention development with quasi-experimental effect studies, tension may arise between the paradigms of literary instruction (e.g., tailoring instruction to students' needs, providing them with freedom of choice) and methodological requirements. During the implementation of TDLT-1 and 2, already we noticed that teachers were sometimes concerned about 'sticking to the guidelines' and the time frame of effect measures. In contrast, students 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 their own teachers during regular hours of literature class, its ecological validity may have been affected. In future design projects, rather than integrating, 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 focus on learning outcomes.

Finally, this paper may have implications for educational practices in two ways. First, as all teachers who successfully implemented TDLT-2 were involved in either its design or a preparatory workshop, we believe there might be a need for setting up a professional development course for pre- and in-service teachers that emphasizes students' interaction with and about literary texts. For future design studies, 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.

ACKNOWLEDGEMENTS

This study is part of the project 'Uses of Literary Narrative Fiction in Social Contexts', which is supported by a grant from the Dutch Organization for Scientific Research (NWO; 360 30 240). We would like to thank all teachers, students and research assistants who contributed to this study.

REFERENCES

Adler, E.S., & Foster, P. (1997). A literature-based approach to teaching values to adolescents: Does it work? *Adolescence, 32*(126), 275-286. https://www.ncbi.nlm.nih.gov/pubmed/9179324

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. https://doi.org/10.3102/00028312040003685

Banks, W. P. (2009). Literacy, sexuality, and the value(s) of queer young adult literatures. *English Journal, 98*(4), 33-36.

Beach, R., Appleman, D., Hynds, S., & Wilhelm, J. D. (2011). *Teaching literature to adolescents* (2nd ed.). New York, NY: Routledge.

Bender-Slack, D. (2002). Using literature to teach global education: A humanist approach. *English Journal, 91*, 70-75. https://doi.org/10.2307/821401

Biesta, G. (2007). The education-socialisation conundrum, or 'Who is afraid of education?' *Utbildning & Demokrati, 16*(3), 25-36.

Booth, W. C. (1983). *The retoric of fiction* (2nd Ed.). Chicago, IL: University of Chicago Press.

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*(2), 241-263. https://doi.org/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. https://doi.org/10.1598/rrq.36.4.3

Curriculum.nu (2018). *Tweede tussenproduct Nederlands* [Second intermediate report Dutch]. Retrieved from https://curriculum.nu/ontwikkelteam/nederlands/

Derricott, R. (2014). *Citizenship for the 21st century: An international perspective on education.* London, UK: Routledge.

Dutch Institute for Curriculum Development (2012). *Handreiking schoolexamen Nederlands havo/vwo.* [Guide school exam Dutch higher general/pre-university education.] Enschede, The Netherlands: SLO.

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.

Fialho, O. (2007). Foregrounding and refamiliarization: understanding readers' response to literary texts. *Language and Literature, 16*, 105-123. https://doi.org/10.1177/0963947007075979

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. IGEL Conference paper, July 2018, University of Stavanger, Norway.

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, The Netherlands: John Benjamins.

Iser, W. (1974). *The implied reader. Patterns of communication in prose fiction from Bunyan to Beckett.* Baltimore, MD: Johns Hopkins University Press.

Janssen, T. (1998). *Literatuuronderwijs bij benadering* [Approaches of literature education] (Doctoral dissertation). Retrieved from http://dare.uva.nl/home

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*(1), 91-116. https://doi.org/10.17239/l1esll-2009.09.01.05

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*(1), 117-137. https://doi.org/10.17239/L1ESLL-2009.09.01.06

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.

Koopman, E.M., & Hakemulder, F. (2015). Effects of literature on empathy and self-reflection: A theoretical-empirical framework. *Journal of Literary Theory, 9*(1), 79-111. https://doi.org/10.1515/jlt-2015-0005

Langer, J. A. (2011). *Envisioning literature. Literary understanding and literature instruction.* New York, NY: Teachers College.

Lenters, K. (2006). Resistance, struggle, and the adolescent reader. *Journal of Adolescent & Adult Literacy, 50*(2), 136-146. https://doi.org/10.1598/jaal.50.2.6

Levine, S., & Horton, W.S. (2013). Using affective appraisal to help readers construct literary interpretations. *Scientific Study of Literature, 3*(1), 105-136. https://doi.org/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.

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. http://www.ncte.org/journals/rte/issues/v48-4

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), 1-34. https://doi.org/10.15763/issn.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. https://doi.org/10.1111/j.1745-6924.2008.00073.x

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.

Miall, D. S., & Kuiken, D. (1994). Foregrounding, defamiliarization, and affect: Response to literary stories. *Poetics, 22*, 389-407. https://doi.org/10.1016/0304-422x(94)00011-5

Muijs, D., & Reynolds, D. (2010). *Effective teaching: Research and practice*. London, UK: Paul Chapman.

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. https://doi.org/10.1037/a0015576

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, the Netherlands: Kluwer.

Nucci, L., Krettenauer, T., & Narváez, D. (Eds.). (2014). *Handbook of moral and character education.* New York, NY: Routledge.

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.

Oatley, K. (1999). Meetings of minds: Dialogue, sympathy, and identification, in reading fiction. *Poetics, 26*, 439-454. https://doi.org/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, The 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*(1), 33-84. https://doi.org/10.3102/0034654307313793

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, The Netherlands: SLO.

Probst, R. E. (1988). Dialogue with a text. *English Journal, 77*, 32-38. https://doi.org/10.2307/818021

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.

Richardson, P.W., & Eccles, J.S. (2007). Rewards of reading: Toward the development of possible selves and identities. *International Journal of Educational Research, 46*(6), 341-356. https://doi.org/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*(2), 173-225. https://doi.org/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, The Netherlands: Brill.

Rosenblatt, L.M. (1968). *Literature as exploration.* Champaign, IL: NCTE. (Original work published 1938)

Rothbauer, P. (2011). Rural teens on the role of reading in their lives. *The 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

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.

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. Contribution to a special issue on The Role of Writing in Literature Education, edited by Tanja Janssen and Irene Pieper. *L1-Educational Studies in Language and Literature, 16*, p. 1-37. https://doi.org/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, 89*(1), 3-45. https://doi.org/10.3102/0034654318812914

Schrijvers, M., Janssen, T., Fialho, O., De Maeyer, S., & Rijlaarsdam, G. (2019). Transformative Dialogic Literature Teaching fosters' adolescents insight into human nature and motivation. *Learning & Instruction, 63*. (advance online publication) https://doi.org/10.1016/j.learninstruc.2019.101216

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, The Netherlands: SLO.

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*(2), 148-162. https://doi.org/10.1002/pits.20621

White, B.F. (1995). Effects of autobiographical writing before reading on students' responses to short stories. *The Journal of Educational Research, 88*(3), 173-184. https://doi.org/10.1080/00220671.1995.9941296

Wilhelm, J. D. (2016). *You gotta BE the book. Teaching engaged and reflective reading with adolescents* (3rd Ed.). New York, NY: Teachers College.

APPENDIX A. 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.
3) *Intermezzo*: students read a few sentences and a description of five stories, and indicated which two they would like to read best during lesson 3.
4) 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?')
5) 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.

6) 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 Dorresteijn; *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 'worldy 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 B. DESCRIPTION OF TDLT-2

Unit 1: preparation

Story: opening excerpt from *Flight behavior*, Bertram Koeleman

1) 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 Appendix C).
2) Internal dialogue (implicit): teacher reads opening excerpt from the story aloud, students read along on paper; no particular reading instruction.
3) 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.
4) Explicit instruction: teacher introduces guidelines for external dialogues (i.e., dialogue strategy) and discusses the 'first aid card'.
5) External dialogue (practice): students talk in small groups about famous quotes about literature and reading, applying the dialogue strategy.
6) Homework: students write about a situation characterized by 'injustice'.

Unit 2: reading-and-dialogue

Story: *Death*, Martin Bril

1) 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.
2) 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.
3) Internal dialogue (implicit): the teacher reads the story aloud, students read along on paper; no particular reading instruction.
4) 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.
5) Reflection: students reflect individually on their own and the groups' application of the dialogue guidelines.

Unit 3: reading-and-dialogue

Story: *Blood*, Gerard Reve

1) Introduction: teacher introduces the goals of the unit; connects these to previous units.

2) 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.
3) 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.
4) 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'.
5) 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).
6) 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.
7) 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

1) Introduction: teacher introduces the goals of the unit; connects these to previous units.
2) 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).
3) 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.
4) 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).
5) External dialogue (pairs): students exchange their story ends; they write down and talk about feedback on each other's story end.

6) 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.
7) 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

1) Introduction: teacher introduces the goals of the unit; connects these to previous units.
2) Internal dialogue (recall): students are asked to browse through the stories and their annotations from previous units.
3) 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.
4) Reflection: students reflect individually on their own and the groups' application of the dialogue guidelines.
5) External dialogue (classroom): the teacher selects some of the life lessons and guides a classroom dialogue about them.

Unit 6: reading-and-dialogue

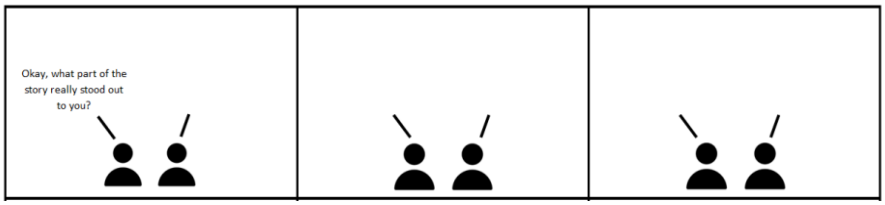
Story: *Count oneself lucky*, Marga Minco; stories to choose from for final task: *The freshwater steak,* Hans Dorresteijn; *An autumn day*, Thomas Heerma van Voss; *The right*, Annelies Verbeke; *Hula*, Cees Nooteboom

1) Introduction: teacher introduces the goals of the unit; connects these to previous units.
2) Explicit instruction: the teacher explains examples from research into how fiction (reading and tv shows) can influence empathy.
3) Internal dialogue: students are instructed to notice and write down their responses in their own way when reading the story.
4) 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).
5) Reflection: students evaluated their progress during TDLT by again filling in the rubric (see Appendix C).
6) Homework: internal dialogue: final TDLT task (see below).

Final task of TDLT-2

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.



APPENDIX C. SELF-EVALUATION RUBRIC

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 focusing 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 expe-riences	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.