Gaining insight into human nature: A review of literature classroom intervention studies

Schrijvers, M.; Janssen, T.; Fialho, O.; Rijlaarsdam, G.

DOI
10.3102/0034654318812914

Publication date
2019

Document Version
Final published version

Published in
Review of Educational Research

Review of Educational Research
February 2019, Vol. 89, No. 1, pp. 3–45
DOI: 10.3102/0034654318812914
Article reuse guidelines: sagepub.com/journals-permissions

Gaining Insight Into Human Nature: A Review of Literature Classroom Intervention Studies

Marloes Schrijvers and Tanja Janssen
University of Amsterdam

Olivia Fialho
Utrecht University

Gert Rijlaarsdam
University of Amsterdam

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

Keywords: literature education, insight into human nature, instructional approaches, dialogic learning, adolescents

Contemporary society finds itself in turbulent times. In an era of globalization, migration, and polarization, there appears to be a need for people to be able to reflect on their own nature as well as on the nature of others. This may include their

own position in the world, their views of themselves, and their perceptions of and relationships with other people. Numerous approaches to teaching and learning indicate that education may play a pivotal role in helping young people gain insight into human nature, such as social and emotional learning (Elias et al., 1997), citizenship education (Derricott, 2014), moral and character education (Nucci, Krettenauer, & Narváez, 2014), and values education (Halstead & Taylor, 1996). Despite having their own backgrounds, frames of reference, and terminology, these approaches all indicate that teachers of any subject may attend to "human nature."

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

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

Fictional and Literary Texts

The literature classroom includes all sorts of written fictional texts, for instance, stories, novels, poetry, drama, song texts, and so forth. The term *fictional text* refers to texts in which characters are "not presented as existing in the real world" (Koopman, 2016, p. 106). The world these characters inhabit does not exist in reality but may function as a safe abstraction and simulation of the real world

(Mar & Oatley, 2008). The borders between fictional and *nonfictional* texts are blurred, for example, when fictionalized stories are based on true events or refer to real-world places. In addition, although fictional texts and stories are often used synonymously, a fictional text is not necessarily a narrative. Poems are considered nonnarrative; however, they are oftentimes fictional texts.

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

Insight Into Human Nature

In this section, we relate reading fictional and literary texts to gaining insight into human nature and attempt to characterize this kind of learning in the literature classroom. Research in developmental psychology indicates that learning about human nature is pivotal during adolescence—a stage of life during which humans develop their sense of self and their social and moral identity (Hart, Atkins, & Ford, 1998). Adolescents become increasingly aware of their inner self and realize that others have an inner self as well, thereby acknowledging the relativity of their own perceptions, thoughts, and feelings (Kohlberg & Gilligan, 1971). Thus, adolescents come to understand "the mutuality of perspectives [which] includes a view of both self and other as complex psychological systems of values, beliefs, attitudes etcetera" (Selman, 1975, p. 40). As such, adolescents' insights, attitudes, beliefs, responses, and behavior related to themselves and others are constantly evolving. Reading fiction and literature, it seems, may play a role in fostering adolescents' insight into human nature.

The Role of Fictional and Literary Reading

Insight into human nature may come about during and after reading fictional and literary texts, as Koopman and Hakemulder (2015) indicate in a synthesis of research. They distinguish between empathy (i.e., insight into the nature of others) and reflection (i.e., insight into one's own nature).

The various definitions of empathy to which Koopman and Hakemulder (2015) refer all relate to the metaphor of putting oneself in the shoes of others, either cognitively or emotionally. Research included in Koopman and Hakemulder's synthesis indicated, for example, that reading fictional texts was found to enhance adult readers' scores on various empathy measures and the accuracy of their perceptions of social interactions (e.g., Bal & Veltkamp, 2013; Mar, Oatley, Hirsh, dela Paz, & Peterson, 2006). Moreover, reading fictional and literary texts may positively affect readers' out-group perceptions (i.e., people's attitudes toward

groups of human beings other than the group with which they identify) and are closely related to feeling empathy for others (Hakemulder, 2000; Johnson, 2013; Kaufman & Libby, 2012).

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

Other studies have investigated which concepts and processes may underlie these effects on insight into human nature and which relationships may exist among them (Fialho, 2012; Fialho, Zyngier, & Burke, 2016; Hakemulder, Fialho, & Bal, 2016). These studies suggest, for instance, that experiencing changes in notions of self and others evolve in particular when readers respond to passages that are highly metaphoric or stylistically deviant from conventional language use (Miall & Kuiken, 1994). They further indicate close relationships between readers' perceptions of themselves and their perceptions of others, which supports the view that there is, conceptually, no other without a self (Zahavi, 2014).

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

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

& Oatley, 2008). For example, readers may compare characters, events, concepts, and themes in stories with real-life situations.

Instructional Approaches to Literature Teaching

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

Text Selection

In literature curricula, numerous choices must be made in terms of what kind of texts are read in the classroom (genre), what these texts are about (theme), and to which extent they can be characterized as fictional or literary texts (literariness).

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

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

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

Tasks

Literature teachers have many different types of tasks at their disposal to ask their students to respond to fictional and literary texts. These may include, for example, dialogues and discussions, formal and creative writing, performing drama, and creating visual and audiovisual arts (e.g., Beach, Appleman, Hynds, &

Wilhelm, 2011; Galda & Beach, 2001; Soter, Wilkinson, Connors, Murphy, & Shen, 2010; Wilhelm, 2016). Such tasks are intended to prompt students' learning activities (Rijlaarsdam, Janssen, Rietdijk, & Van Weijen, 2017). Learning activities are cognitive or affective activities in relation to, in the case of this article, fictional or literary texts. For example, a particular kind of writing task may prompt various learning activities, such as evaluating a story and explaining the evaluation—Did it appeal to students, why or why not?

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

Teachers' Roles

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

Dialogic discourse in the literature classroom, on the other hand, allows students and the teacher to explore and share ideas among each other, which is what we would expect if the intention were to foster students' insight into human nature. The focus would be on stimulating students' divergent thinking and on developing and deepening ideas and experiences. The authenticity of questions and responses, expressed by both students and the teacher, is key for learning in dialogic literature classrooms. Students must be seen as capable partners in open conversations, response writing, and creative performances, which may be achieved by working in small groups. The teacher's task is to guide and to support students in their responses. They may do so, for example, by offering prompts for exploratory talk (Mercer & Dawes, 2008), by thinking aloud during reading to model their own authentic responses to texts, or by making explicit their own difficulties in response writing (Wilhelm, 2016). Specifically for facilitating small-group talk in response to texts, Wei, Murphy, and Firetto (2018) identified a taxonomy of subtle discourse moves that teachers may use, such as backchanneling (indicating that they are

listening to their students), clarifying (inviting a student to provide a clearer response), prompting (helping students construct an elaborate response, by asking for reasons and evidence from a text), and summarizing (giving an overview of what has been said during the talk). In our analysis, we will examine to which extent teachers' roles are addressed and explicated—as Wei et al. (2018) note, information about how teachers interact with their students may not always be given, or may remain implicit.

Stance Toward Texts

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

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

We expect that taking an expressive and a critical–analytical stance may be apt for fostering students' insight into human nature in the literature classroom. Research with adult participants suggests that insight into oneself and others is preceded by various kinds of spontaneous responses and emotional connections to a literary text, such as imagery of setting and characters, identification with characters, and feeling sympathy for characters (Fialho, 2018)—experiences that may be addressed in particular if an expressive stance toward texts is taken. Furthermore, insight into human nature may include students' understanding of complex social situations. Because fictional and literary texts may function as simulations of the real social world (Mar & Oatley, 2008), they may evoke numerous relevant questions, for example: Why do characters in this text think or behave in a particular way? What in the text may explain their thoughts and behavior? Can we classify their behavior as being "right" or "wrong"? Would people in real life behave similarly? Addressing such issues requires students to investigate and reason about worldviews, beliefs, assumptions, and so forth, that are represented in a fictional or literary text. Therefore, we assume that students' insight into human nature may be fostered by taking a critical–analytical stance in the literature classroom.

Aims and Research Questions

Literature education is a promising domain for fostering students' insight into human nature. However, an overview of research that investigates whether

literature teaching—in particular, what kind of instructional approach—is effective in doing so is not yet available. Therefore, we systematically reviewed empirical intervention studies in the literature classroom that deliberately focus on fostering adolescents' insight into human nature.

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

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

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

Method

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

Search Terms

We combined four clusters of search terms. The first cluster contained terms regarding the intended student population, that is, adolescents in upper secondary education. Examples of search terms included *secondary education, high school, secondary school students*, and *grade* (9 to 12). The second cluster focused on texts used in literature teaching. We understood the literature classroom as a first-language context in secondary schools. We broadly defined the term *texts* to include fictional texts, such as novels, short stories, plays, and poems written by published authors. Search terms included *literature, novels, poetry, fiction,*

literature education, literature lessons, and *literature class*. The third and fourth clusters included terms related to human nature. We distinguished between insight into one's own nature and into the nature of others and social relationships. Examples of search terms in the third cluster included *personality, self-understanding, possible selves, emotional experience, identity*, and *self-concept*. In the fourth cluster, search terms included *cultural awareness, empathy, social attitudes, social experience, prejudice, equality*, and *out-group*. Appendix A (available in the online version of the journal) contains all terms in search syntaxes and shows how we applied variations of search terms and adjacent terms.

Inclusion and Exclusion Criteria

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

We further selected intervention studies based on five inclusion and exclusion criteria regarding type of intervention, hypotheses and measures, classroom context, research design, and publication language (see Table 1). For the first criterion, we broadly defined "intervention types." We also included intervention studies in which activities were initiated in the literature classroom, but ultimately performed outside of it (e.g., school-assigned book readings at home), and studies that focused on reading various types of texts with particular reading instructions given in the classroom. If researchers used self-written stories or manipulated text features (e.g., Andringa, 1996), the study was excluded because we sought to identify design principles based on published texts.

Regarding the second criterion, we focused on intervention studies' central hypotheses and the measures used to assess them. If researchers expected that an intervention would in some way affect readers' insight into human nature (e.g., their perceptions of self, attitudes toward others, understandings of others, ways of handling particular social situations, and so forth), their studies were included—provided that these expected effects were measured. When relevant variables were presented as side effects to other variables (e.g., studies that examine both text comprehension and empathy), we included the study but focused on the variables relevant for this review.

For the third criterion, we deliberately only included intervention studies conducted in first-language classrooms because a foreign language–teaching context introduces comprehension challenges that potentially interfere with gaining insight into human nature. We were particularly interested in intervention studies with adolescent participants in the upper grades of secondary education. However, we decided to also include studies conducted in lower grades of secondary school. If they occurred in the records, intervention studies in Grade 7 or 8, for example,

TABLE 1

Inclusion and exclusion criteria

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

can perhaps inform interventions in upper grades. Thus, we decided to include intervention studies with adolescent participants ranging in age from 13 to 18 years. We excluded intervention studies conducted at the college or university level because these often involve older students who voluntarily enroll in a particular program, such as sociology, medical school, or literary studies, which hampers the generalizability of outcomes to secondary school students.

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

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

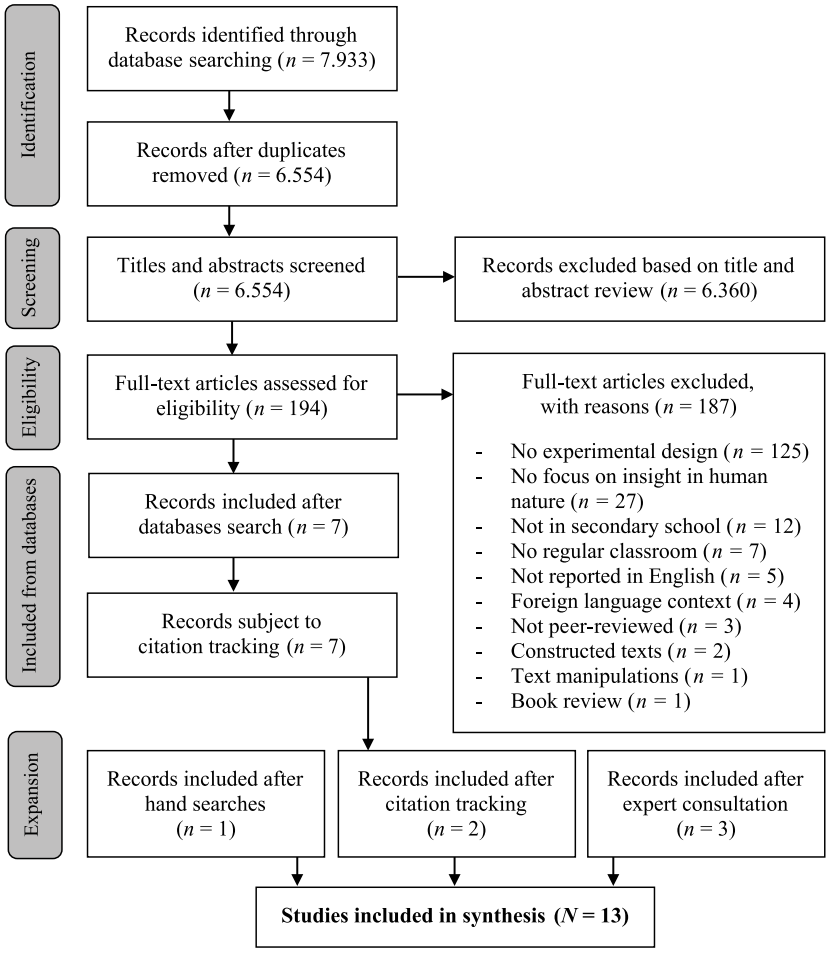

FIGURE 1. *Flow chart of study selection procedure.*

Search Procedure

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

Screening Database Records

The first and second author screened titles and abstracts. The second author screened results from Scopus, and the first author screened records from the other databases. Figure 1 shows an overview of the selection procedure. The large

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

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

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

Search Expansion

We expanded the search by applying citation tracking, by conducting hand searches, and by consulting experts in the field. First, for the seven intervention studies included from the database search, the first author screened whether they referred to relevant studies or whether these seven studies themselves were cited in other relevant studies. We included two additional intervention studies (Darragh, 2015; Malo-Juvera, 2016), both of which cited a study from the database search (Malo-Juvera, 2014). For these two new intervention studies, we also performed citation tracking, which yielded no new studies to be included.

Furthermore, we hand-searched two journals that were not in the databases. First, we screened abstracts from *Pedagogy: Critical Approaches to Teaching Literature, Language, Composition, and Culture* from 2011 to 2017. Because no

empirical intervention studies were published in this period, we waived further screening. Second, we screened *Study and Scrutiny: Research on Young Adult Literature* from its start in 2015. Apart from one study (Malo-Juvera, 2016), which was already included via citation tracking, this journal yielded no other studies to be included. We further consulted the *Annotated Bibliography of Research in the Teaching of English* (National Council of Teachers of English, 2003–2014), one online bibliography (Runge, 2012), and seven reviews and meta-analyses (De Leon, 2017; Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011; Galda & Beach, 2001; Koopman & Hakemulder, 2015; Murphy et al., 2009; Nystrand, 2006; Sigvardsson, 2016). From Koopman and Hakemulder (2015), we included an intervention study by Adler and Foster (1997). Citation tracking of this study yielded no additional intervention studies to be included.

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

As Figure 1 shows, the number of intervention studies included in the expansion phase was rather large compared with the number retrieved from the database search. This may be explained by the fact that most studies found in the expansion phase did not include keywords. Only Adler and Foster (1997) included one keyword, "bibliotherapy." This particular keyword was not included in our search syntax; however, because the study tested a literature classroom intervention rather than small-group therapeutic sessions, we included it. Another reason for appearance in the expansion phase rather than the database search may be that studies were published in small electronic journals that were not included in large databases (Darragh, 2015; Malo-Juvera, 2016). Finally, our search syntax may not have complied fully with descriptions in abstracts. For example, the relevance of the work by Stevahn et al. (1996) and Stevahn et al. (1997) lies in its descriptions of "conflict resolution," a term not included in the syntax.

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

Quality Assessment

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

Methodological Quality

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

Analysis. Table 2 shows the coding scheme, which contains 15 indicators distributed over five categories: (a) description and rationale of comparison condition(s); (b) reliability of measures; (c) instructors and implementation; (d) data, results, and conclusions; and (e) attrition rate. Indicators that were not applicable (e.g., reliability of qualitative measurements if a study only included quantitative measures) were coded accordingly. If a study compared multiple experimental conditions, without using a control condition, indicators for the comparison condition were coded as "not applicable." Of 195 scores (13 studies * 15 indicators), 144 were applicable. The two attrition rate indicators were nominally scored: 0 (*not reported*) or 1 (*reported*). The other 13 indicators were scored from 0 (*not at all* or *very poor*) to 4 (*completely* or *excellent*). We used scale scores because intervention characteristics could be described in both more or less detail. For example, Henschel, Meier, and Roick (2016) specifically reported the duration of the comparison condition (135 minutes) and scored a 4 on this aspect, whereas Eva-Wood (2004) reported that the comparison group followed a 4-week program, without specifying the duration of these lessons, resulting in a score of 2. Scale scores were also used to score reliability of measures because most researchers used multiple instruments. If the reliability of all measures was sufficient, a score of 4 was assigned. If the reliability of one or more measures was insufficient, a lower score was assigned. In case researchers reported only percentage agreement to report interrater reliability of coding schemes, without taking chance agreement into account, we assigned a score of 0 (Adler & Foster, 1997).

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

Outcomes. As Table 2 shows, some researchers adequately described what happened in the comparison conditions (Henschel et al., 2016; Vezzali, Stathi, & Giovannini, 2012; White, 1995). This was not the case in other studies. For example, Garrod (1989) only reported that "the comparison group curriculum had some features in

TABLE 2

Assigned quality scores (0–4) to methodological quality indicators, per category

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

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

common with the treatment" (p. 68) followed by two examples of selected texts. Other researchers merely mentioned that comparison group students read an unrelated text (Hakemulder, 2008) or were part of a wait-list (Malo-Juvera, 2016). Rationales for the choice of texts and tasks (A3, A5) were seldom reported. For instance, Adler and Foster (1997) only mentioned that texts in the comparison condition were part of the regular curriculum. Malo-Juvera (2014) reported that students in the comparison group read a text by Shakespeare—instead of the young adult novel that was related to the intervention theme, which was read by the experimental group—but in both groups, "similar instructional methods" (p. 416) were used. However, the purpose of designing the comparison condition in this particular way was not discussed.

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

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

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

Finally, attrition rate was reported in only 4 of the 13 studies. In two of these, more than 10% of the students dropped out between the pretest and posttest (Adler & Foster, 1997; Malo-Juvera, 2014). If sample sizes differ at the pretest and

TABLE 3

Assigned quality scores (0–4) to characteristics of intervention descriptions

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

posttest, internal validity may be affected, which may result in incorrect statistical analyses. Such issues, however, were not considered in the included studies.

Quality of Intervention Descriptions

In Table 3, the coding scheme and outcomes of the assessment of the quality of intervention descriptions are presented.

Analysis. We assessed to which extent the interventions were described in detail, by focusing on five basic characteristics (see Table 3): information about (a) the duration of the intervention, (b) the selected texts, (c) the tasks students were asked to complete, (d) rationales for selecting these texts, and (e) rationales for designing these tasks. Rationales are important because they indicate *why* the selected texts and tasks would be suitable to achieve a particular purpose, thereby helping educational designers make well-informed choices (Rijlaarsdam et al., 2017). The quality of descriptions for these five indicators was scored on a scale from 0 (*not at all* or *very poor*) to 4 (*completely* or *excellent*).

To establish interrater reliability, an independent second rater scored the 5 indicators for all 13 studies. The intraclass correlation between the first and second rater was excellent, ICC = 0.78, $p < .001$ (Cicchetti, 1994). Consequently, the scores of the first author of this article were used for interpretation.

Outcomes. First, Table 3 indicates a rather large variety across studies regarding information about the duration of interventions. For example, Malo-Juvera (2014) reported in detail that the instructional unit "lasted 5 weeks and consisted of 12 classes (each lasting 1 hour and 45 minutes)" (p. 415). In contrast, Garrod (1989)

mentioned neither the exact number of lessons nor their duration; the only information given was that the intervention was part of a year-long curriculum (which received a score of 1). Second, most researchers reported which texts were used in interventions by providing titles, authors, and sometimes summaries (Adler & Foster, 1997). A low score on this criterion is exemplified by Eva-Wood (2004) who reported author's names but not how many and which poems were selected. Third, most studies contained good or excellent descriptions of intervention tasks.

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

Quality Assessment Conclusions

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

Data Analysis

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

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

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

Results

Expected Effects and Empirical Support

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

As Table 4 shows, we distinguished three categories of expected effects on students' insight into human nature. Researchers expected to affect (a) students' insight into themselves, (b) their understanding of fictional others, and/or (c) their understanding of, views on, or intended behavior toward real-world others. These categories emerged from researchers' use of measures that focused either on insight into oneself (e.g., a scale with items such as "Reading literature makes me sensitive to aspects of my life that I usually ignore"; Miall & Kuiken, 1994, p. 55, in Eva-Wood, 2004), into fictional others (e.g., "I can easily empathize with one of the characters from the text"; Henschel et al., 2016, p. 16), or into real-world others (e.g., "When girls wear low cut tops and short skirts they're just asking for trouble"; Malo-Juvera, 2014, p. 419). One researcher expected effects on both students' self-insight and their understanding of fictional characters (Eva-Wood, 2004).

Table 4 also indicates to which extent there was empirical support for the expected effects. There was empirical support for 9 out of 13 interventions, although two of these could only provide partial support (Adler & Foster, 1997; Eva-Wood, 2004). Effect sizes are included when they were reported in the study or could be calculated based on the data and annotated if they were not applicable or could not be calculated.

Insight Into Oneself

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

Halász (1991) asked students to write down memories and associations that were evoked during reading a text or to do so in response to salient words from the text. First, he expected that students would rely more on personal experiences when responding to a literary text than to an expository text or an essay. Second, he expected the same result if, after reading, students' writings responded to high frequent, salient words from these texts. Third, he expected that responding to salient words after reading would result in more personal responses than responding to the same words without reading the texts. All three hypotheses were confirmed. In all cases, reading the literary text evoked more personal, affective, and detailed responses, which most often consisted of personal references and indications of emotion, than reading the other texts or reading no text. Halász also observed that students, in their personal responses, predominantly referred to secondary sources such as fiction, music, art, or experiences they garnered from others. This finding indicates that drawing on such sources may help students interact with a literary text with respect to their own lives. Because analyses were based

TABLE 4

Expected and demonstrated intervention effects

Study	Insight into oneself	Understanding of fictional others	Understanding of, views on, or intended behavior toward real-world others	Empirical support
Halász (1991)	Insight in personal memories			Full
Malo-Juvera (2014)			Views on sexual harassment	Full
Malo-Juvera (2016)			Views on others' sexual orientation	Full
Stevahn et al. (1996)			Understanding of, intended behavior in, and views on conflict situations	Full
Stevahn et al. (1997)				Full
Vezzali et al. (2012)			Views on and intended behavior toward immigrants	Full
White (1995)		Understanding characters and their behavior		Full
Adler and Foster (1997)			Views on caring for others	Partial[a]
Eva-Wood (2004)	Insight in own qualities and world[b]	Empathy for characters and understanding poems' speakers[b]		Partial[c]
Darragh (2015)			Views on and intended behavior toward disabled people	No
Garrod (1989)			Understanding people's moral dilemmas	No
Hakemulder (2008)			Views on immigrants	No
Henschel et al. (2016)		Empathy for characters[b]		No

[a]Empirical support based one out of three measures that were used. [b]Ancillary effect rather than primary aim of intervention. [c]Empirical support only provided for understanding poems' speakers.

on merged categories (e.g., "personal references" consisted of four categories) and means and standard deviations were only reported for individual categories, effect sizes of merged categories could not be calculated.

Eva-Wood (2004) implemented a think-and-feel-aloud pedagogy in response to poetry. The pedagogy was primarily expected to enhance students' transactions with poems, in terms of higher levels of engagement in analysis and more sophisticated responses, but Eva-Wood also assessed whether it enhanced students' insight into previously unrecognized qualities in themselves and in their world. To assess this, she used the *Insights* scale of the Literary Response Questionnaire (Miall & Kuiken, 1994). On this scale, no differences were found between the experimental condition and a comparison condition that focused on structural analysis of the poems.

Understanding of Fictional Others

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

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

In addition to their self-insight, Eva-Wood (2004) expected students' empathy for fictional characters to be affected (measured using the *Empathy* scale of the Literary Response Questionnaire; Miall & Kuiken, 1994) and expected students to personally engage and identify with poems' speakers as measured by contributions to classroom discussions. No differences on empathy were found between the experimental and the comparison condition. However, students in the experimental condition showed greater personal engagement and identification with poems' speakers, for example, by expressing their understanding of a speaker who had lost a loved one. Thus, Eva-Wood's intervention study showed partial empirical support for fostering students' understanding of fictional others; effect sizes for personal engagement and identification could not be calculated because standard deviations were not reported.

Henschel et al. (2016) expected that completing reader-oriented tasks, which focused on personal emotional engagement and creative responses, would increase students' empathy for fictional characters—as opposed to completing text-based tasks that stimulated text analysis via cognitive activities. The hypothesis was tested by using three items adapted from the *Fantasy Empathy* scale of the Interpersonal Reactivity Index (Davis, 1983). As indicated in the Method section, posttest mean scores did not differ between conditions, and pretest scores were not taken into account in the statistical analysis. Based on the reported results, we could not infer that the intervention study yielded empirical support for an expected effect on students' empathy for characters.

Understanding of, Views on, or Intended Behavior Toward Real-World Others

In 9 out of 13 studies, researchers expected to affect students' understanding of, views on, or intended behavior toward other human beings in the real world. The measures they used did not apply specifically to reading (as did the scales used by Eva-Wood, 2004, and Henschel et al., 2016) but to the world beyond students' experiences with the text. Of nine interventions in this category, five provided full empirical support (Malo-Juvera, 2014; 2016; Stevahn et al., 1996; Stevahn et al., 1997; Vezzali et al., 2012); one provided partial empirical support (Adler & Foster, 1997), and three provided no empirical support (Darragh, 2015; Garrod, 1989; Hakemulder, 2008).

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

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

Stevahn et al. (1996) and Stevahn et al. (1997) expected that conflict resolution training, in which conflicts from fictional novels were used, would result in improved understandings of how to solve conflicts, better application of this knowledge in conflict scenarios (intended behavior), and more constructive, positive views on conflicts. In the comparison condition, students read the same novel but completed tasks that did not focus on conflicts. Measures included writing down steps to solve a conflict (understanding), writing short essays about how a conflict scenario could be solved (intended behavior), and writing words associated with conflict (views on conflict)—coded as negative/destructive, neutral, or positive/constructive. Students in the experimental condition scored higher on understanding how to solve conflicts and on intended behavior in conflict situations than students in the control condition, both at the posttest and the delayed posttest (ds > 1.00). Students in the experimental condition also listed more positive associations than students in the control condition. More specifically, Stevahn et al. (1996) compared two versions of the intervention: a cooperative and an individualistic version. Results indicated an interaction effect of condition and version. The cooperative version of the intervention most effectively fostered understandings of conflict resolution, intended behavior to solve conflicts, and positive views on conflict.

Finally, Vezzali et al. (2012) asked students to read a novel featuring immigrant characters and to complete a writing and evaluation task afterward. The authors expected that the intervention would result in more positive views on

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

Adler and Foster (1997) developed an intervention which they expected to increase students' support for the value "caring for others." Students read three novels in which this theme was prominent and participated in classroom discussions and exercises that were designed to reinforce the theme of the books. In the comparison condition, students read novels from the regular curriculum. All students completed three essays as pretests and the same three essays as posttests, which included topics about caring for a family member, caring for strangers who lost their home to a fire, and about friends as stand-ins for family. In the essays about friends, more students in the experimental group showed positive change in valuing "caring for others" than in the control group, a difference that was statistically significant. For the other two pre- and posttest essays, no differences were found. Thus, there was partial empirical support for the expectation that the intervention would foster students' support for the value of caring for others. An effect size could not be calculated because group sample sizes were not given.

Three studies remained in which no empirical support was found for the expected effects. Darragh (2015) asked students to read a young adult novel featuring a disabled character and to respond to it in writing tasks. She expected that students would develop more positive views on and intended behavior toward disabled people, as compared with students who read novels that did not feature disabled characters. One measure assessed views on disabled people via a word association task; the other asked students whether they would undertake particular activities with a disabled person. No differences between conditions were found.

Hakemulder (2008) expected that reading a fragment from an intercultural novel while following a role-taking reading instruction would evoke a positive view of immigrants. The intervention was compared with three other conditions: (a) reading the same text but focusing on its structure, (b) reading an essay about the same theme without instruction, and (c) reading an unrelated text without instruction. Five items were used to assess intervention effects. On two of these, a statistically significant difference between the experimental and a comparison condition was found; however, these effects were in favor of the comparison conditions. Thus, the intervention effect contradicted the hypothesis; however, it

should be noted that this was determined using a single item rather than a validated attitude scale.

Finally, Garrod (1989) developed an intervention in which students engaged in Socratic discussions about particular moral dilemmas in literary texts. He expected that this would result in students developing their moral reasoning abilities compared with a condition in which Socratic discussions focused on other texts. Moral reasoning development was assessed using the written version of the Kohlberg Moral Judgment Interview. Of all students in both conditions ($N = 44$), 17 students formed a subgroup with the lowest pretest scores. In this subgroup, the intervention had the expected effect: Students in the experimental condition ($n = 5$) achieved more growth than those in the comparison condition ($n = 12$), a difference that was statistically significant. However, the small condition sample sizes in this subgroup may have affected the statistical results. In addition, there was no overall difference between the experimental and the comparison condition.

Interventions With Empirical Support

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

Instructional Approaches

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

Texts Used

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

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

were used were characterized as a complex metaphorical story (*The Vulture* by Kafka, in Halász, 1991) and two literary stories (*Indian Camp* and *The End of Something* by Hemingway, in White, 1995).

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

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

Only Eva-Wood (2004) considered literariness from a theoretical perspective, stating that reading poems defamiliarizes readers when they encounter stylistic devices that are specific to literary texts, such as metaphors and similes. Literary texts thus deviate from the conventional understandings of words and the relationships among them (Eva-Wood, 2004; Miall & Kuiken, 1994). All in all, literariness seemed of little concern in the studies included in this review.

Tasks

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

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

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

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

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

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

Reflective response writing. Reflective response writing occurred most often in the included interventions. This happened either after a full text had been read or after students had finished part of the text. Vezzali et al. (2012) asked students to

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

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

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

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

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

In several interventions, writing prepared students to engage in small-group or whole-class dialogues. Dialogues, then, added a layer of responses, perspectives, and interpretations to the kind that individual students explored in writing. After

writing and reading, White (1995) engaged students in classroom dialogues about the characters, their actions, and the consequences of these actions. The teacher guided the dialogues by following a protocol. The protocol contained 10 questions. Seven questions initiated describing characters or explaining their actions. These were followed by a prediction question about what might happen next, a question about which message for the real world might be embedded in the text, and one question that addressed why the author made particular choices to include or to describe a character. The study indicated that autobiographical writing prior to reading helped students explore in these dialogues what fictional characters are like and why they behave in a certain way.

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

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

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

Conclusions. All in all, writing and dialogues were the most salient types of tasks in interventions with full or partial empirical support. Notably, some interventions featured minimal instructions, such as a single writing task (Halász, 1991) or a written summary and encircled evaluations (Vezzali et al., 2012). Some tasks functioned as intervention activities—in the sense that they were presented to students as response tasks—and research instruments simultaneously. For example, Vezzali et al. used students' evaluations to control for appreciation and difficulty in statistical analyses, Halász used students' written responses for analysis, and White (1995) analyzed students' responses in classroom dialogues.

The Role of the Teacher

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

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

In short, Malo-Juvera (2014) and Eva-Wood (2004) appeared to envision somewhat different roles for teachers. Nonetheless, both roles allow for dialogic discourse in the classroom, in which the teacher acts as a facilitator rather than an authoritative figure. No details were provided on specific, subtle discourse moves that teachers might make to facilitate students' small-group talk—even though "modeling" was included in the Teacher Move Taxonomy by Wei et al. (2018) as well as in Eva-Wood's work, the latter did not address it as a discourse move, but as an instructional strategy for a particular mode of reading.

Other interventions did not shed light on teachers' roles, neither in terms of guiding students' talk, nor in terms of facilitating or providing feedback on their writing or role-playing. Researchers often dispensed the interventions (Halász, 1991; Vezzali et al., 2012) or provided no information on how teachers interacted

with students during group talk, writing tasks or role-playing (Adler & Foster, 1997; Stevahn et al., 1996; Stevahn et al., 1997). White (1995) described the questions that teachers asked during classroom dialogues, but did not describe how teachers guided students' autobiographical writing (e.g., if students encountered difficulties or were hesitant to complete the task). All in all, the data in this review were not fit to draw firm conclusions about teachers' roles. Therefore, our analysis remains inconclusive with regard to this aspect of instructional approaches.

Stance Toward Texts

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

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

A critical–analytical stance toward the texts was taken in four interventions (Malo-Juvera, 2014; 2016; Stevahn et al., 1996; Stevahn et al., 1997). Malo-Juvera asked students to critically interrogate texts in terms of ideas, assumptions, and worldviews regarding sexual harassment and sexual orientation they presented. The tasks prompted students, for instance, to evaluate whether a protagonist was telling the truth and to compare their evaluations with other students. According to Malo-Juvera, students were invited to explore "moral reasoning about sensitive topics" (2014, p. 421). Stevahn et al. asked students to search the text for conflicts and to reason about how these conflicts could be solved by characters; as such, students used the text rather instrumentally to deepen their understanding of conflicts and to enhance their skills at solving them.

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

All in all, the existing categorization suggests that taking both an expressive and a critical–analytical stance toward texts seems most promising for fostering

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

Discussion

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

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

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

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

From Instructional Approaches to Design Principles

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

> If we want to increase the likelihood that adolescent students gain insight into human nature in the literature classroom, we are best advised to: (1) select fictional texts such as novels, short stories, passages, or poems, that are thematically relevant for the intended outcomes of the intervention; (2) design writing tasks related to fictional texts and text themes that prompt students to (a) activate previous personal experiences before reading, (b) notice and annotate their experiences during the reading process, and/or (c) reflect on evoked experiences directly after reading; and (3) design exploratory dialogic activities that stimulate students to verbally share their personal experiences related to fictional texts and text themes.

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

The Principle of Text Selection

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

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

Notably, researchers did not attend to the possibility that social or moral themes in fiction may be sensitive issues in students' lives, such as family relationships (White,

1995), sexual harassment (Malo-Juvera, 2014), sexual orientation (Malo-Juvera, 2016), and immigration (Vezzali et al., 2012). Creating a safe learning environment to talk about such themes was not addressed in the reviewed intervention studies. Students' sense of safety may be considered a default, that is, not always critically reflected upon (Boostrom, 1998).

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

The Principle of Writing About Personal Experiences

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

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

Research on learning processes has demonstrated the importance of establishing a meaningful basis of prior knowledge in which new information can be embedded (Dochy, Segers, & Buehl, 1999; Merrill, 2002; Pressley et al., 1992). This basis of prior knowledge, we argue, may well include previous life experiences. From a more domain-specific perspective, writing prior to reading may enhance students' emotional involvement in a text (Janssen & Braaksma, 2016). In addition, writing tasks assigned during and directly after reading may help students engage in internal dialogues with the texts. This line of reasoning traces back to Rosenblatt's (1938/1968) transactional theory of reading, which outlines how activities like engaging, constructing, and imagining are part of people's reading experience. Beach (1993) built on Rosenblatt's work and argued for experiential approaches to teaching literature that stimulate these aspects of the

reading experience, which appear to be crucial for reading experiences to affect oneself and oneself in relation to others (Fialho, 2018). All in all, completing writing tasks may function as "writing-to-learn" (Klein, Boscolo, Kirkpatrick, & Gelati, 2014). If it allows students to draw on personal experiences in particular, writing-to-learn is considered a valuable activity in literature classrooms (Newell, 1996).

The Principle of Verbally Sharing Personal Experiences

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

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

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

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

Interventions Without Empirical Support

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

One might argue that all interventions, either with or without empirical support, were rather alike in their instructional approaches. Therefore, it may seem invalid to consider these instructional approaches as being informative for designing future literature classroom interventions. However, numerous methodological or contextual reasons may explain why 4 out of 13 studies found no empirical support for their interventions. For example, the instruments used may not have been apt for capturing these effects, an intervention may not have been implemented as originally intended, or the contrast between the experimental and comparison condition may not have been large enough. Such reasons cannot be determined with certainty in this review because sufficient information about instruments, coding schemes, implementation fidelity, and comparison conditions was not always provided in the studies without empirical support. These shortcomings repeatedly resulted in rather low scores on methodological quality indicators.

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

Limitations

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

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

Implementation fidelity was a major issue in the included intervention studies. Although it is crucial to know whether interventions are implemented as intended (O'Donnell, 2008), few studies sufficiently accounted for it. In addition, our review remains inconclusive regarding the role of the teacher. In the majority of the included studies, descriptions of teachers' roles were insufficient. The two researchers who addressed teachers' roles (Eva-Wood, 2004; Malo-Juvera, 2014) both seemed to suggest that the teacher should allow dialogic discourse in the classroom, but operationalized this in different ways. Moreover, in neither of these studies, teachers' subtle discourse moves were addresses, even though such moves may be "influential in promoting or hindering students' learning outcomes" (Wei et al., 2018, p. 579). All in all, our review yielded too little information to formulate a design principle about teacher–student interactions.

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

Future Research

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

Conclusion

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

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

Note

This research 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 Janneke Staaks and Charlotte Huigh for their valuable assistance during the preparation of this article.

References

**Studies included in the review are denoted with an asterisk.*

*Adler, E. S., & Foster, P. (1997). A literature-based approach to teaching values to adolescents: Does it work? *Adolescence*, *32*, 275–286.

Andringa, E. (1996). Effects of "narrative distance" on readers' emotional involvement and response. *Poetics*, *23*, 431–452. doi:10.1016/0304-422x(95)00009-9

Applebee, A. N., Langer, J. A., Nystrand, M., & Gamoran, A. (2003). Discussion-based approaches to developing understanding: Classroom instruction and student performance in middle and high school English. *American Educational Research Journal*, *40*, 685–730. doi:10.3102/00028312040003685

Bal, M., & Veltkamp, M. (2013). How does fiction reading influence empathy? An experimental investigation on the role of emotional transportation. *PLoS One*, *8*, 1–12. doi:10.1371/journal.pone.0055341

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

Barnes, D. (1976). *From communication to curriculum*. Harmondsworth, England: Penguin.

Beach, R. (1993). *Reader response theories*. Urbana, IL: National Council for Teacher Education.

Beach, R., Appleman, D., Hynds, S., & Wilhelm, J. (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*(5), 70–75.

Blackie, M., & Wear, D. (2015). Three things to do with stories: Using literature in medical, health professions, and interprofessional education. *Academic Medicine*, *90*, 1309–1313. doi:10.1097/acm.0000000000000761

Boostrom, R. (1998). "Safe Spaces": Reflections on an educational metaphor. *Journal of Curriculum Studies*, *30*, 397–408. doi:10.1080/002202798183549

Bourdieu, P. (1996). *The rules of art: Genesis and structure of the literary field*. Stanford, CA: Stanford University Press.

Buitrago, C. B. (2017). EFL students' social identities construction through gender-based short stories. *Colombian Applied Linguistics Journal*, *19*, 11–21. doi:10.14483/calj.v19n1.10641

Charlton, M., Pette, C., & Burbaum, C. (2004). Reading strategies in everyday life: Different ways of reading a novel which make a distinction. *Poetics Today*, *25*, 241–263. doi:10.1215/03335372-25-2-241

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, *6*, 284–290. doi:10.1037//1040-3590.6.4.284

Cummins, A. (2013). Border crossings: Undocumented migration between Mexico and the United States in contemporary young adult literature. *Children's Literature in Education*, *44*, 57–73. doi:10.1007/s10583-012-9176-1

Curriculum. (2017). *Nederlands—Vakgebonden Eindtermen* [Dutch—Domain specific objectives]. Retrieved from http://eindtermen.vlaanderen.be/secundair-onderwijs/derde-graad/aso/vakgebonden/nederlands/eindtermen.htm

Curwood, J. S. (2013). Redefining normal: A critical analysis of (dis)ability in young adult literature. *Children's Literature in Education*, *44*, 15–28. doi:10.1007/s10583-012-9177-0

*Darragh, J. J. (2015). Exploring the effects of reading young adult literature that portrays people with disabilities in the inclusion classroom. *Electronic Journal for Inclusive Education*, *3*(4), 1–48.

Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology*, *44*, 113–126. doi:10.1037/0022-3514.44.1.113

De Leon, C. (2017). Engaging adolescents in mental health discussions through young adult literature: A literature review. *Queen's University Graduate Student Symposium*. Retrieved from http://qspace.library.queensu.ca/handle/1974/15468

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

Dochy, F., Segers, M., & Buehl, M. M. (1999). The relation between assessment practices and outcomes of studies: The case of research on prior knowledge. *Review of Educational Research*, *69*, 145–186. doi:10.3102/00346543069002145

Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development*, *82*, 405–432. doi:10.1111/j.1467-8624.2010.01564.x

Dutch Institute for Curriculum Development. (2015). *Curriculumspiegel Deel B: Vakspecifieke Trendanalyse* [Curricular Description Part B: Domain-specific trend analysis]. Enschede, Netherlands: SLO. Retrieved from http://downloads.slo.nl/Repository/curriculumspiegel-2015-deel-b.pdf

Elias, M. J., Zins, J. E., Weissberg, R. P., Frey, K. S., Greenberg, M. T., Haynes, N. M., . . . Shriver, T. P. (1997). *Promoting social and emotional learning: Guidelines for educators*. Alexandria, VA: Association for Supervision and Curriculum Development.

Ellis, J. M. (1974). *The theory of literary criticism: A logical analysis*. Berkeley: University of California Press.

Engberg, M. E. (2004). Improving intergroup relations in higher education: A critical examination of the influence of educational interventions on racial bias. *Review of Educational Research*, *74*, 473–524. doi:10.3102/00346543074004473

*Eva-Wood, A. L. (2004). How think-and-feel-aloud instruction influences poetry readers. *Discourse Processes*, *38*, 173–192. doi:10.1207/s15326950dp3802_2

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, July). *Deepening readers' perceptions of self and others* (IGEL Conference paper), University of Stavanger, Norway.

Fialho, O., Zyngier, S., & Burke, M. (2016). Empirical approaches to the study of literature in learning environments. In M. Burke, O. Fialho, & S. Zyngier (Eds.), *Scientific approaches to literature in learning environments* (pp. 1–16). Amsterdam, Netherlands: John Benjamins.

Fialho, O., Zyngier, S., & Miall, D. (2011). Interpretation and experience: Two pedagogical interventions observed. *English in Education*, *45*, 236–253. doi:10.1111/j.1754-8845.2011.01103.x

Fredricks, L. (2012). The benefits and challenges of culturally responsive EFL critical literature circles. *Journal of Adolescent & Adult Literacy*, *55*, 494–504. doi:10.1002/jaal.00059

Galda, L., & Beach, R. (2001). Response to literature as a cultural activity. *Reading Research Quarterly*, *36*(1), 64–73. doi:10.1598/rrq.36.1.4

*Garrod, A. (1989). Promoting moral development through a high school English curriculum. *Alberta Journal of Educational Research*, *35*(1), 61–79.

Gersten, R., Fuchs, L. S., Compton, D., Coyne, M., Greenwood, C., & Innocenti, M. S. (2005). Quality indicators for group experimental and quasi-experimental research in special education. *Exceptional Children*, *71*, 149–164. doi:10.1177/001440290507100202

Hakemulder, F. (2000). *The moral laboratory: Experiments examining the effects of reading literature on social perception and moral self-knowledge*. Amsterdam, Netherlands: John Benjamins.

*Hakemulder, F. (2008). Imagining what could happen: Effects of taking the role of a character on social cognition. In S. Zyngier, M. Bortolussi, A. Chesnokova, & J. Auracher (Eds.), *Directions in empirical literary studies: In honor of Willie van Peer* (pp. 139–160). Amsterdam, Netherlands: John Benjamins.

Hakemulder, F., Fialho, O., & Bal, P. M. (2016). Learning from literature: Empirical research on readers in schools and at the workplace. In M. Burke, O. Fialho, & S. Zyngier (Eds.), *Scientific approaches to literature in learning environments* (pp. 19–38). Amsterdam, Netherlands: John Benjamins.

*Halász, L. (1991). Emotional effect and reminding in literary processing. *Poetics*, *20*, 247–272. doi:10.1016/0304-422x(91)90009-e

Halstead, J. M., & Taylor, M. J. (Eds.). (1996). *Values in education and education in values*. London, England: Falmer Press.

Hart, D., Atkins, R., & Ford, D. (1998). Urban America as a context for the development of moral identity in adolescence. *Journal of Social Issues*, *54*, 513–530. doi:10.1111/0022-4537.801998080

Hebert, M., Simpson, A., & Graham, S. (2013). Comparing effects of different writing activities on reading comprehension: A meta-analysis. *Reading and Writing*, *26*, 111–138. doi:10.1007/s11145-012-9386-3

*Henschel, S., Meier, C., & Roick, T. (2016). Effects of two types of task instructions on literary text comprehension and motivational and affective factors. *Learning and Instruction*, *44*, 11–21. doi:10.1016/j.learninstruc.2016.02.005

Iser, W. (1980). Interaction between text and reader. In S. R. Suleiman & I. Crosman (Eds.), *The reader in the text* (pp. 106–120). Princeton, NJ: Princeton University Press.

Janssen, T., & Braaksma, M. (2016). Effects of creative writing on adolescent students' literary response. In M. Burke, O. Fialho, & S. Zyngier (Eds.), *Scientific approaches to literature in learning environments* (pp. 193–211). Amsterdam, Netherlands: John Benjamins.

Johnson, D. R. (2013). Transportation into literary fiction reduces prejudice against and increases empathy for Arab-Muslims. *Scientific Study of Literature*, *3*, 77–92. doi:10.1075/ssol.3.1.08joh

Kaufman, G. F., & Libby, L. K. (2012). Changing beliefs and behavior through experience-taking. *Journal of Personality and Social Psychology*, *103*, 1–19. doi:10.1037/a0027525

Kaywell, J. F. (1993). *Adolescents at risk: A guide to fiction and nonfiction for young adults, parents, and professionals*. Westport, CT: Greenwood Press.

Kidd, D. C., & Castano, E. (2013). Reading literary fiction improves theory of mind. *Science*, *342*, 377–380. doi:10.1126/science.1239918

Klein, P. D., Boscolo, P., Kirkpatrick, L. C., & Gelati, C. (Eds.) (2014). *Writing as a learning activity*. Leiden, Netherlands: Brill.

Kohlberg, L., & Gilligan, C. (1971). The adolescent as a philosopher: The discovery of the self in a postconventional world. *Daedalus*, *100*, 1051–1086.

Koopman, E. M. (2016). *Reading suffering: An empirical inquiry into empathic and reflective responses to literary narratives* (Doctoral dissertation). Erasmus Research Center for Media, Communication and Culture, Rotterdam, Netherlands.

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

*Malo-Juvera, V. (2014). Speak: The effect of literary instruction on adolescents' rape myth acceptance. *Research in the Teaching of English*, *48*(4), 407–427.

*Malo-Juvera, V. (2016). The effect of an LGBTQ themed literary instructional unit on adolescents' homophobia. *Study and Scrutiny*, *2*, 1–34. doi: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. doi:10.1111/j.1745-6924.2008.00073.x

Mar, R. A., Oatley, K., Hirsh, J., dela Paz, J., & Peterson, J. B. (2006). Bookworms versus Nerds: Exposure to fiction versus non-fiction, divergent associations with social ability, and the simulation of fictional social worlds. *Journal of Research in Personality*, *40*, 694–712. doi:10.1016/j.jrp.2005.08.002

McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods*, *1*, 30–46. doi:10.1037/1082-989X.1.1.30

Mercer, N., & Dawes, L. (2008). The value of exploratory talk. In N. Mercer & S. Hodgkinson (Eds.), *Exploring talk in school* (pp. 55–71). London, England: Sage.

Merrill, M. D. (2002). First principles of instruction. *Educational Technology Research and Development*, *50*, 43–59. doi:10.1007/bf02505024

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

Miall, D. S., & Kuiken, D. (1999). What is literariness? Three components of literary reading. *Discourse Processes*, *28*, 121–138. doi:10.1080/01638539909545076

Muřakovský, J. (1976). *On poetic language* (J. Burbank & P. Steiner, Trans.). Lisse, Netherlands: Peter de Ridder Press.

Murphy, P. K., Wilkinson, I., Soter, A., Hennessey, M. N., & Alexander, J. F. (2009). Examining the effects of classroom discussion on students' comprehension of text:

A meta-analysis. *Journal of Educational Psychology*, *101*, 740–764. doi:10.1037/a0015576

National Council of Teachers of English. (2003–2014). *Annotated bibliography of research in the teaching of English*. Retrieved from http://www.ncte.org/journals/rte/biblios

National Council of Teachers of English. (2012). *NCTE/IRA standards for the English language arts*. Retrieved from http://www.ncte.org/standards/ncte-ira

Newell, G. E. (1996). Reader-based and teacher-centered instructional asks: Writing and learning about a short story in middle-track classrooms. *Journal of Literacy Research*, *28*, 147–172. doi:10.1080/10862969609547914

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

Nussbaum, M. C. (1995). *Poetic justice: The literary imagination and public life*. Boston, MA: Beacon Press.

Nystrand, M. (1997). *Opening dialogue: Understanding the dynamics of language and learning in the English classroom*. New York, NY: Teachers College Press.

Nystrand, M. (2006). Research on the role of classroom discourse as it affects reading comprehension. *Research in the Teaching of English*, *40*(4), 392–412.

O'Donnell, C. L. (2008). Defining, Conceptualizing, and measuring fidelity of implementation and its relationship to outcomes in K-12 curriculum intervention research. *Review of Educational Research*, *78*, 33–84. doi:10.3102/0034654307313793

Pressley, M., Wood, E., Woloshyn, V. E., Martin, V., King, A., & Menke, D. (1992). Encouraging mindful use of prior knowledge: Attempting to construct explanatory answers facilitates learning. *Educational Psychologist*, *27*, 91–109. doi:10.1207/s15326985ep2701_7

Pyle, D., Pyle, N., Lignugaris/Kraft, B., Duran, L., & Akers, J. (2017). Academic effects of peer-mediated interventions with English language learners: A research synthesis. *Review of Educational Research*, *87*, 103–133. doi:10.3102/0034654316653663

Reigeluth, C. M. (1999). What is instructional-design theory and how is it changing? In C. M. Reigeluth (Ed.), *Instructional-design theories and models: A new paradigm of instructional theory* (Vol. 2, pp. 5–29). Mahwah, NJ: Lawrence Erlbaum.

Richardson, P. W., & Eccles, J. S. (2007). Rewards of reading: Toward the development of possible selves and identities. *International Journal of Educational Research*, *46*, 341–356. doi:10.1016/j.ijer.2007.06.002

Rijlaarsdam, G., Janssen, T., Rietdijk, S., & Van Weijen, D. (2017). Reporting design principles for effective instruction of writing: Interventions as constructs. In R. Fidalgo & T. Olive (Series Eds.), & R. Fidalgo, K. Harris, & M. Braaksma (Vol. Eds.), *Studies in writing series: Vol. 34. Design principles for teaching effective writing* (pp. 280–313). Leiden, Netherlands: Brill.

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

Rosenblatt, L. M. (1994). *The reader, the text, the poem*. Carbondale: Southern Illinois University Press. (Original work published 1978)

Rothbauer, P. (2011). Rural teens on the role of reading in their lives. *Journal of Research on Libraries and Young Adults*, *1*(2). Retrieved from www.yalsa.ala.org/jrlya/2011/02/rural-teens-on-the-role-of-reading-in-their-lives

Runge, L. L. (2012). *Annotated bibliography on literature pedagogy*. Retrieved from http://chuma.cas.usf.edu/~runge/Class_Bib.htm

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

Selman, R. L. (1975). Level of social perspective taking and the development of empathy in children: Speculations from a social-cognitive viewpoint. *Journal of Moral Education*, *5*, 35–43. doi:10.1080/0305724750050105

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton, Mifflin.

Shen, Y. (2008). Two levels of foregrounding in literary narratives. In S. Zyngier, M. Bortolussi, A. Chesnokova, & J. Auracher (Eds.), *Directions in empirical literary studies: In honor of Willie van Peer* (pp. 103–111). Amsterdam, Netherlands: John Benjamins.

Sigvardsson, A. (2016). Teaching poetry reading in secondary education: Findings from a systematic literature review. *Scandinavian Journal of Educational Research*, *61*, 1–16. doi:10.1080/00313831.2016.1172503

Sikora, S., Kuiken, D., & Miall, D. S. (2011). Expressive reading: A phenomenological study of readers' experience of Coleridge's The rime of the ancient mariner. *Psychology of Aesthetics, Creativity, and the Arts, 5*, 258–268. doi:10.1037/a0021999

Slone, M., Tarrasch, R., & Hallis, D. (2000). Ethnic stereotypic attitudes among Israeli children: Two intervention programs. *Merrill-Palmer Quarterly*, *46*(2), 370–389.

Soter, A., Wilkinson, I., Connors, S., Murphy, P., & Shen, V. (2010). Deconstructing "aesthetic response" in small-group discussions about literature: A possible solution to the "aesthetic response" dilemma. *English Education*, *42*(2), 204–225.

Soter, A., Wilkinson, I., Murphy, P. K., Rudge, L., Reninger, K., & Edwards, M. (2008). What the discourse tells us: Talk and indicators of high-level comprehension. *International Journal of Educational Research*, *47*, 372–391. doi:10.1016/j.ijer.2009.01.001

Steen, G., & Schram, D. (2001). The empirical study of literature: Psychology, sociology and other disciplines. In D. Schram & G. Steen (Eds.), *The psychology and sociology of literature* (pp. 1–16). Amsterdam, Netherlands: John Benjamins.

*Stevahn, L., Johnson, D. W., Johnson, R. T., Green, K., & Laginski, A. M. (1997). Effects on high school students of conflict resolution training integrated into English literature. *Journal of Social Psychology*, *137*, 302–315. doi:10.1080/00224549709595442

*Stevahn, L., Johnson, D. W., Johnson, R. T., & Real, D. (1996). The impact of a cooperative or individualistic context on the effectiveness of conflict resolution training. *American Educational Research Journal*, *33*, 801–823. doi:10.3102/00028312033004801

Turner, R. N., Hewstone, M., Voci, A., Paolini, S., & Christ, O. (2007). Reducing prejudice via direct and extended crossgroup friendship. *European Review of Social Psychology*, *18*, 212–255. doi:10.1080/10463280701680297

Turner, R. N., Hewstone, M., Voci, A., & Vonofakou, C. (2008). A test of the extended contact hypothesis: The mediating role of intergroup anxiety, perceived ingroup and outgroup norms, and inclusion of the outgroup in the self. *Journal of Personality and Social Psychology*, *95*, 843–860. doi:10.1037/a0011434

UK Literary Association. (2016). *Curriculum and assessment in English 3 to 19: A better plan. The essentials of English*. Retrieved from https://ukla.org/resources/details/curriculum-and-assessment-in-english-3-to-19-a-better-plan

Van den Akker, J. (1999). Principles and methods of development research. In J. van den Akker, R. M. Branch, K. Gustafson, N. Nieveen, & T. Plomp (Eds.), *Design approaches and tools in education and training* (pp. 1–14). Dordrecht, Netherlands: Springer.

Van den Akker, J., Gravemeijer, K., McKenney, S., & Nieveen, N. (Eds.). (2006). *Educational design research*. New York, NY: Routledge.

Van Peer, W., Hakemulder, J., & Zyngier, S. (2007). Lines on feeling: Foregrounding, aesthetics and meaning. *Language and Literature*, *16*, 197–213. doi:10.1177/0963947007075985

*Vezzali, L., Stathi, S., & Giovannini, D. (2012). Indirect contact through book reading: Improving Adolescents' attitudes and behavioral intentions toward immigrants. *Psychology in the Schools*, *49*, 148–162. doi:10.1002/pits.20621

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Weber, C. D. (2010). Literary fiction as a tool for teaching social theory and critical consciousness. *Teaching Sociology*, *38*, 350–361. doi:10.1177/0092055x10378816

Wei, L., Murphy, P. K., & Firetto, C. M. (2018). How can teachers facilitate productive small-group talk? An integrated taxonomy of teacher discourse moves. *Elementary School Journal*, *118*, 578–609. doi:10.1086/697531

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

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

Zahavi, D. (2014). *Self and other: Exploring subjectivity, empathy, and shame*. New York, NY: Oxford University Press.

Authors

MARLOES SCHRIJVERS is a PhD candidate at the Research Institute of Child Development and Education, University of Amsterdam, Netherlands; email: *m.s.t.schrijvers@uva.nl*. She holds a master's degree in teaching Dutch language and literature, as well as a master's degree in children's literature. She specializes in domain-specific intervention research in literature classrooms.

TANJA JANSSEN a senior researcher at the Research Institute of Child Development and Education, University of Amsterdam, Netherlands; email: *t.m.janssen@uva.nl*. She specializes in research in literature education and writing instruction, and supervises research of students, PhD candidates, and teachers.

OLIVIA FIALHO is a postdoctoral researcher at the Department of Media and Culture Studies, Utrecht University, Netherlands; email: *o.fialho@uu.nl*. She holds a master's degree in applied linguistics, as well as a PhD in comparative literature from the University of Alberta. She specializes in phenomenological research into reading experiences that transform readers' self and social concepts.

GERT RIJLAARSDAM is a professor at the Research Institute of Child Development and Education, University of Amsterdam, Netherlands, and the Department of Linguistics at Antwerp University, Belgium; email: *g.c.w.rijlaarsdam@uva.nl*. His focus is on domain-specific research in language, arts, and literature. He specializes in both process studies and intervention studies of teaching and learning in elementary and secondary school classrooms.