SELF-QUESTIONING IN THE LITERATURE CLASSROOM: EFFECTS ON STUDENTS' INTERPRETATION AND APPRECIATION OF SHORT STORIES

TANJA JANSSEN, MARTINE BRAAKSMA
& MICHEL COUZIJN

Graduate School for Teaching & Learning, University of Amsterdam, the Netherlands

Abstract. In this study we examined the effects of self-questioning on students' interpretation and appreciation of complex short stories. Two experiments were carried out, in which tenth grade students from different secondary schools participated. In Experiment 1 self-questioning instruction was compared to instructor-made questions about stories. In Experiment 2 two forms of self-questioning instruction were compared: an unguided and a guided form. Literature discussions in peer groups formed a substantial part of all conditions.
Results showed that (unguided) self-questioning had a positive effect on students' appreciation of literary stories, compared to instructor-prepared questions and to guided self-questioning. The results for quality of interpretation were more diffuse. In Experiment 1 effects on students' story interpretation could not be established. In Experiment 2 a main effect on story interpretation was found for both the guided and unguided form of self-questioning instruction. In addition, students' reading experience appeared to be important for the effectiveness of the unguided self-questioning condition: avid readers tended to benefit more from this condition than infrequent readers. We conclude that an open literature approach, based on 'authentic' student-generated questions in response to short stories, can be beneficial for students' story interpretation and appreciation.
Key words: self-questioning, literary interpretation, story appreciation, literature discussions.

Chinese
[Translation Shek Kam Tse]
文學課上的自我提問———學生對短故事的闡釋與欣賞效果

摘要：本研究調查了自我提問在學生對複雜的短故事的闡釋與欣賞時所起的效果。來自不同中學的十年級學生參加了兩項實驗。實驗2中比較了兩種形式的自我提問教學：一個無指導，一個有指導。每種情況都以同伴群體對文學作品的討論為主體。

Janssen, T., Braaksma, M., & Couzijn, M. (2009). Self-questioning in the literature classroom: Effects on students' interpretation and appreciation of short stories.
L1 – Educational Studies in Language and Literature, 9*(1), 91-116.*
© International Association for the Improvement of Mother Tongue Education

Correspondence concerning this article should be directed to Tanja Janssen, Graduate school of Teaching and Learning, University of Amsterdam, Spinozastraat 55, 1017 HJ, Amsterdam, the Netherlans. Electronic mail: T.M. Janssen@uva.nl

結果顯示和老師準備問題以及有指導的自我提問相比，無指導的自我提問對於學生對文學故事的鑒賞有積極作用。理解力的結果更加發散。實驗1中，學生對故事的理解效應不存在，實驗2指導和無指導的自我提問教學都存在故事理解的主效應。此外學生閱讀經驗對於無指導的自我提問顯得比較重要：經常閱讀的熱切閱讀者在該組中比不常閱讀者收益更大。我們總結出：一個讓學生提出可信問題的開放性的文學學習方法有助於學生對於故事的理解和欣賞。

關鍵詞： 自我提問， 文學闡釋，故事欣賞，文學作品討論

Dutch

[Translation Tanja Janssen]

TITEL: Zelf vragen stellen door leerlingen in de literatuurles; effecten op de interpretatie en waardering van korte verhalen.

SAMENVATTING. In dit onderzoek bestudeerden we de effecten van het zelf vragen stellen op de interpretatie en waardering van leerlingen voor korte verhalen. Twee experimenten werden uitgevoerd waaraan leerlingen van verschillende scholen deelnamen. In Experiment 1 vergeleken we twee condities: instructie in zelf vragen stellen door leerlingen versus docentvragen over verhalen. In Experiment 2 werden twee vormen van zelf vragen stellen instructie met elkaar vergeleken: een met en een zonder begeleiding bij het vragen stellen. Literatuurgesprekken in kleine groepjes vormden een belangrijk bestanddeel van alle condities.

De resultaten laten zien dat (onbegeleid) zelf vragen stellen een positief effect had op de verhaalwaardering van de leerlingen, vergeleken met docentvragen en begeleid zelf vragen stellen. De resultaten voor kwaliteit van interpretatie waren minder duidelijk. In Experiment 1 werden geen effecten op verhaalinterpretatie van de leerlingen gevonden. In Experiment 2, daarentegen, werd een hoofdeffect gevonden op verhaalinterpretatie, zowel van de begeleide als onbegeleide vorm van zelf vragen stellen. Bovendien bleek de leeservaring van de leerlingen van belang voor de effectiviteit van onbegeleid zelf vragen stellen: leerlingen die in hun vrije tijd veel lazen profiteerden meer van deze conditie dan weinig-lezers. We concluderen dat een open literatuurbenadering, die uitgaat van 'echte' leerlingvragen in reactie op gelezen verhalen positief kan uitwerken op de verhaalinterpretatie en –appreciatie van leerlingen.

TREFWOORDEN: zelf vragen stellen, literaire interpretatie, verhaalwaardering, literatuurgesprekken.

Finnish

[Translation Katri Sarmavuori]

TITTELI : Itsekysely kirjallisuusluokassa: Novellien vaikutukset oppilaiden tulkintoihin ja ymmärrykseen

ABSTRAKTI: Tässä tutkimuksessa selvitimme itsekyselyn vaikutuksia oppilaiden tulkintoihin ja mutkikkaiden novellien ymmärtämiseen. Tehtiin kaksi koetta, joissa 10. luokan oppilaat eri yläkouluista osallistuivat. Kokeessa 1 itsekyselyopetusta verrattiin ohjaajan kertomuksesta tekemiin kysymyksiin. Kokeessa 2 verrattiin kahta itsekyselyopetusmuotoa; ei-ohjattu ja ohjattu muoto. Kirjallisuuskeskustelut ikäryhmissä muodostivat olennaisen osan järjestelyistä.

Tulokset näyttävät, että (ei-johdetuilla) itsekyselyillä oli positiivinen vaikutus oppilaiden kertomusten ymmärtämiseen verrattuna ohjaajan valmistamiin kysymyksiin ja ohjattuihin itsekyselyihin. Tulokset tulkintojen laatuun nähden olivat hajanaisemmat. Kokeessa 1 ei voitu havaita vaikutuksia oppilaiden kertomustulkintoihin. Kokeessa 2 päävaikutus kertomuksen tulkintaan löydettiin sekä ohjattuun että ohjaamattomaan itsekyselyopetuksen muotoon. Lisäksi oppilaiden lukukokemus osoittautui tärkeäksi ohjaamattoman itsekyselyn tehokkuudelle: himolukijat hyötyivät tästä tilanteesta enemmän kuin vähän lukevat. Päättelemme, että avoin kirjallisuuden lähestymistapa, joka perustuu 'autenttisiin' novelleista tehtyihin oppilasjohtoisiin kysymyksiin, voi olla hyödyllinen oppilaiden kertomustulkintoihin ja niiden ymmärtämiseen.

AVAINSANAT: itsekysely, kirjallinen tulkinta, kertomuksen ymmärtäminen, kirjallisuuskeskustelut.

French

[Translation Laurence Pasa]

TITRE : L'auto-questionnement en cours de literature: Les effets sur l'interpretation et l' appreciation de nouvelles par les élèves .

RÉSUMÉ : Dans cette étude nous examinons les effets de l'auto-questionnement sur l'interprétation et l'appréciation de nouvelles complexes par des élèves. Deux expériences ont été réalisées, auxquelles des élèves de seconde issus de différents établissements secondaires ont participé. Dans l'expérience 1, la pratique de l'auto-questionnement a été comparée à la pratique d'un questionnement initié par l'enseignant. Dans l'expérience 2, deux types d'auto-questionnement ont été comparés : une forme non-guidée et une autre guidée. Des débats littéraires entre pairs ont été réalisés dans toutes les conditions expérimentales. Les résultats montrent que l'auto-questionnement non-guidé a un impact positif sur l'appréciation des nouvelles littéraires par les élèves, davantage que les questions préparées par l'enseignant et que l'auto-questionnement guidé. En ce qui concerne la qualité de l'interprétation, les résultats sont plus nuancés. Dans l'expérience 1, les effets sur l'interprétation de l'histoire par les élèves n'ont pas pu être établis. Dans l'expérience 2, l'impact sur l'interprétation de l'histoire a été confirmé, à la fois pour l'auto-questionnement guidé et non-guidé. En outre, l'expérience de la lecture qu'ont les élèves s'est avérée importante pour l'efficacité de l'auto-questionnement non-guidé : les plus fervents lecteurs ont eu tendance à en bénéficier davantage que les lecteurs occasionnels. Nous concluons qu'une approche ouverte de la littérature, basée sur un questionnement authentique et spontané des nouvelles par les élèves, peut être bénéfique pour leur interprétation et leur appréciation de l'histoire.
MOTS-CLÉS : auto-questionnement, interprétation littéraire, appréciation d'histoire, débats littéraires.

Greek

[Translation by Panatoya Papoulia Tzelepi]

Τίτλος: Αυτοερώτηση στην τάξη της λογοτεχνίας: Αποτελέσματα στην επεξεργασία και εκτίμηση διηγημάτων από τους μαθητές

Περίληψη: Στη μελέτη αυτή εξετάσαμε τα αποτελέσματα της αυτοερώτησης στην επεξεργασία και εκτίμηση σύνθετων μικρών ιστοριών (διηγημάτων). Δύο πειράματα έλαβαν χώρα στα οποία μαθητές της 10ης τάξης (πρώτη Λυκείου) από διάφορα σχολεία έλαβαν μέρος. Στο Πείραμα 1 η αυτοερώτηση συγκρίθηκε με ερωτήσεις κατασκευασμένες από το δάσκαλο. Στο Πείραμα 2 δύο τύποι διδασκαλίας αυτοερώτησης συγκρίθηκαν· μια ακαθοδήγητη και μια καθοδηγημένη. Οι συζητήσεις για τη Λογοτεχνία σε ομάδες μαθητών αποτελούσαν ένα ουσιαστικό κομμάτι σε όλες τις συνθήκες πειραματισμού. Τα αποτελέσματα έδειξαν ότι (ακαθοδήγητη) αυτοερώτηση είχε θετικό αποτέλεσμα στην εκτίμηση των διηγημάτων εκ μέρους των μαθητών, συγκρινόμενες με τις κατασκευασμένες ερωτήσεις από το δάσκαλο, ή με τις καθοδηγούμενες αυτοερωτήσεις. Τα αποτελέσματα σχετικά με την ποιοτική επεξεργασία υπήρξαν πιο συγκεχυμένα. Στο Πείραμα 1, αποτελέσματα στην επεξεργασία της ιστορίας από τους μαθητές, δεν είναι δυνατόν να πιστοποιηθούν. Στο Πείραμα 2, ένα κύριο αποτέλεσμα για την επεξεργασία της ιστορίας ανευρέθη στην ακαθοδήγητη και στην καθοδηγημένη μορφή διδασκαλίας αυτοερώτησης. Επιπρόσθετα, η αναγνωστική εμπειρία των μαθητών φαίνεται ότι είναι σημαντική για την αποτελεσματικότητα της συνθήκης της ακαθοδήγητης αυτοερώτησης: οι βιβλιοφάγοι τείνουν να ωφελούνται περισσότερο από αυτή τη συνθήκη από ό,τι οι περιστασιακοί αναγνώστες.Καταλήγουμε ότι μια ανοιχτή προσέγγιση στη λογοτεχνία, βασιζόμενη σε «αυθεντικές» ερωτήσεις των μαθητών στη μελέτη διηγημάτων, μπορεί να είναι ωφέλιμη για την από μέρους των μαθητών επεξεργασία κει εκτίμηση των αφηγήσεων.
Λέξεις κλειδιά: Αυτοερώτηση, λογοτεχνική επεξεργασία, εκτίμηση διηγήματος, λογοτεχνικές συζητήσεις

Italian

[Translation Manuela Delfino, Francesco Caviglia]

TITOLO: Porsi delle domande nelle ore di letteratura: gli effetti sull'interpretazione e sull'apprezzamento di racconti brevi da parte degli studenti

SINTESI: In questo contributo abbiamo studiato gli effetti che il porsi delle domande (*self-questioning*) ha sull'interpretazione e sull'apprezzamento di racconti brevi complessi. Sono stati condotti due esperimenti a cui hanno partecipato studenti del secondo anno della scuola secondaria. Nell'esperimento 1 le domande poste autonomamente sono state confrontate con le domande sui racconti poste dal docente. Nell'esperimento 2 sono state confrontate due forme di domande poste autonomamente; una guidata, l'altra no. In tutte le configurazioni degli esperimenti, un ruolo sostanziale è stato comunque riservato alla discussione letteraria all'interno del gruppo dei pari.

I risultati mostrano che il porsi delle domande in modo non guidato - se confrontato con le domande predisposte dall'insegnante e con quelle guidate - ha avuto un effetto positivo sull'apprezzamento dei racconti brevi da parte degli studenti. I risultati sulla qualità dell'interpretazione sono più articolati. Nell'esperimento 1 non è stato possibile stabilire gli effetti sull'interpretazione del racconto da parte degli studenti. Nell'esperimento 2 si è trovato un chiaro influsso sull'interpretazione del racconto da parte delle

domande poste autonomamente, sia guidate che non guidate. Inoltre, sembra che l'esperienza di lettura degli studenti sia importante per l'efficacia della modalità non guidata di porsi domande (self-questioning): i grandi lettori tendono a ricevere un maggior beneficio da questa condizione rispetto ai lettori saltuari. Per concludere, un approccio aperto alla letteratura, basato su domande 'autentiche' formulate dagli studenti in merito ai racconti brevi può avere dei benefici sull'interpretazione e sull'apprezzamento dei racconti da parte degli studenti.
PAROLE CHIAVE: porsi domande, interpretazione letteraria, apprezzamento della narrativa, discussioni letterarie

Polish

[Translation Elżbieta Awramiuk]

TITUŁ: SAMODZIELNE STAWIANIE PYTAŃ DO TEKSTU NA LEKCJI LITERATURY: WPŁYW NA UCZNIOWSKIE INTERPRETACJE I ROZUMIENIE KRÓTKICH OPOWIADAŃ

STRESZCZENIE: W niniejszym artykule relacjonujemy wpływ samodzielnie stawianych pytań do tekstu na uczniowskie interpretacje i rozumienie skomplikowanych krótkich opowiadań. Przeprowadzono dwa eksperymenty, których uczestnikami byli uczniowie klas dziesiątych z różnych szkół średnich. W eksperymencie 1. nauczanie oparte na pytaniach budowanych przez uczniów zostało porównane z nauczaniem opartym na pytaniach do tekstu zbudowanych przez nauczyciela. W eksperymencie 2. porównano dwie formy kształcenia za pomocą samodzielnie budowanych pytań do tekstu: ze wskazówkami i bez. Podstawę wszystkich eksperymentów stanowiła dyskusja nad literaturą w grupie rówieśniczej.

Rezultaty pokazały, że samodzielne budowanie pytań do tekstu nie opierające się na jakichkolwiek wskazówkach miało – w porównaniu do pytań zbudowanych przez nauczyciela i pytań stawianych samodzielnie, ale w oparciu o wskazówki – pozytywny wpływ na rozumienie przez uczniów literackich opowiadań. Rezultaty jakości interpretacji były bardziej rozmyte. Na podstawie eksperymentu 1. nie można było określić wpływu na uczniowską interpretację opowiadania. W eksperymencie 2. główny wpływ na interpretację opowiadania miały samodzielnie stawiane pytania do tekstu, zarówno te budowane ze wskazówkami, jak i bez. Dodatkowo znaczące dla efektywności pytań bez wskazówek okazały się wcześniejsze doświadczenia uczniów z czytaniem: lubiący czytać starali się uzyskać w tych warunkach więcej niż uczniowie, którzy czytają rzadko. We wnioskach podkreślamy, że otwarte podejście do literatury bazujące na autentycznych pytaniach generowanych przez uczniów w wypadku krótkich opowiadań pozytywnie wpływa na uczniowskie interpretacje i rozumienie..

SŁOWA-KLUCZE: samodzielnie stawianych pytań do tekstu, interpretacja literatury, rozumienie opowiadania, dyskusja o literaturze

Spanish

[Translation Ingrid Marquez]

TÍTULO: PREGUNTAS GENERADAS POR EL ESTUDIANTE EN EL SALÓN DE LITERATURA: EFECTOS EN LA INTERPRETACIÓN ESTUDIANTIL Y SU APRECIO DE CUENTOS CORTOS

RESUMEN: En este estudio, examinamos el efecto de las preguntas generadas por el estudiante en su interpretación y aprecio de cuentos cortos complejos. Se hicieron dos experimentos con la participación de estudiantes de tercer semester de bachillerato provinientes de diferentes escuelas. En Experimento 1, la instrucción de hacer preguntas propias se comparó con el uso de preguntas planteadas por el maestro acerca de los cuentos. En Experimento 2, dos formas de instrucción de generación de preguntas se compararon: una forma guiada y no guiada. Pláticas sobre la literatura en grupo formaron una parte importante de los dos experimentos.

Los resultados mostraron que la generación de preguntas no guiada tuvo un efecto positivo al aumentar el aprecio estudiantil de los cuentos literarios, comparado con preguntas preparadas por el profesor y con la generación guiada de preguntas. Los resultados en cuanto a la calidad de la interpretación fueron más difusos. En Experimento 1, los efectos en la interpretación estudiantil del cuento no se pudieron establecer. En Experimento 2, un efecto sustancial se notó en la interpretación del cuento tanto para la forma guiada y no guiada de generar preguntas. Adicionalmente, la experiencia estudiantil de la lectura pareció ser más importante para la eficacia de la condición de generación de preguntas no guiada: los lectores ávidos solían sacar más provecho de esta condición que los que leían poco. Concluimos que un enfoque abierto para la literatura, basado en "auténticas" preguntas generadas por los estudiantes como respuesta a los cuentos cortos, pueden ayudar a que los estudiantes interpreten y aprecien un cuento.

PALABRAS CLAVE: preguntas generadas por el estudiante, interpretación literaria, aprecio de un cuento, pláticas sobre literatura

> "On the day they were going to kill him, Santiago Nasar got up at five-thirty in the morning to wait for the boat the bishop was coming on. He'd dreamed he was going through a grove of timber trees where a gentle drizzle was falling, and for an instant he was happy in his dream, but when he awoke he felt completely spattered with bird shit."

The opening sentences of *Chronicle of a Death Foretold*, a novella written by Gabriel García Márquez, raise many questions; Who is Santiago Nasar? Who are 'they'? Why are 'they' going to kill him? Does Santiago know he is going to be killed that day? Does his dream have a meaning?

As the story progresses, we – the readers – receive answers to some of these questions. We learn that Santiago is suspected of having deflowered the beautiful Angela. Her brothers have sworn to kill him to restore the family's honor. The entire town knows of the brothers' plan, but no one tries to stop them, or to warn Santiago. Eventually, we are wondering whether Santiago is guilty at all. The answer to this question remains a mystery.

As this case illustrates, literature reading is - to a certain extent - question-driven. Since literary texts are often ambiguous and open, readers are stimulated to ask themselves questions during reading, and to seek explanations for the who, what, why, when, and how in a story (Andringa, 1995; Trabasso & Magliano, 1996). However, not *all* readers do so. Think aloud studies indicate that an open, questioning way of reading is characteristic of more experienced readers of literature. Less experienced readers tend to read in a closed, passive manner; they ask themselves fewer questions, and less often generate alternative explanations (Andringa, 1995; Earthman, 1992).

To examine how Dutch tenth grade students interacted with literature, we analysed students' think aloud responses to short literary stories (Janssen, Braaksma & Rijlaarsdam, 2006). The participating students were carefully selected; 9 were 'weak' literature readers according to their teachers, while 10 were known to be 'strong' literature readers. Each student read five complex short stories under think aloud conditions. Results showed that weak and strong adolescent readers differed in the way they attempted to make sense of stories during reading. Weak readers engaged less often in problem detecting and questioning during reading, among other things, than strong readers of literature. Instead, weak readers tended to focus on the story events, taking these events at face value.

If questioning is an important component of literature reading and literary interpretation, it seems wise to teach students to ask themselves and others questions during and after their reading of literary texts. Weak literature readers in particular might benefit from learning to generate questions about what puzzles them in literary texts, and to discuss their questions with peers. This idea inspired us to develop a series of literature lessons in which self-questioning and peer discussion of self-generated questions played a central part. To test the lesson series for its effectiveness, we carried out two experiments. Before we report on these experiments, we will briefly discuss what we know about the effects of self-questioning instruction on reading outcomes from previous research. (For a more extensive discussion see; Janssen, 2002.)

1. RESEARCH ON EFFECTS OF SELF-QUESTIONING IN READING; A BRIEF OVERVIEW

There is a vast body of research on questioning and reading (Janssen, 2002). Here, we will focus on experimental training studies. Descriptive studies of classroom discourse, that analyze the kinds and patterning of students' questions and responses, will be left out of consideration.

The first training studies on student-generated questions date from the 1960s. Since then, much research has been done on the benefits of stimulating students to generate questions during and after reading difficult texts. Over 95 training studies on self-questioning during reading have been published between 1965 and 2000 (see reviews of Huang, 1992; Janssen, 2002; Rosenshine, Meister & Chapman, 1996; Taboada, 2003; Wong, 1985).

According to Rosenshine et al. (1996), three theoretical frameworks can be discerned in these studies; active processing theory, schema theory, and metacognitive theory. According to *active processing theory*, asking yourself higher-level questions during reading leads to active engagement with the text, which in turn results in increased text comprehension and recall. *Schema theory* emphasizes the role of readers' prior knowledge in text comprehension. In this theory it is assumed that comprehension during reading can be hampered not only by a lack of prior knowledge, but also by failing to *activate* prior knowledge. By asking questions, students can be stimulated to activate their relevant prior knowledge, resulting in better text comprehension. Finally, *metacognitive theory* focuses on self-regulation processes during reading. By asking questions students may monitor their reading process, become aware of their reading problems or knowledge deficits, and thus reach higher levels of textual understanding. According to Veenman (2004), self-questioning is highly interwoven with metacognition, if not entirely being part of it.

The effects of self-questioning has been most widely studied for the reading of expository texts. Most often self-questioning is presented as a reading or study strategy, to be used by students individually when reading texts about specific content matter (biology or history for instance). Reviewers of this research converge in their conclusion that students' text comprehension and learning performance are improved by training in self-questioning skills and that self-questioning training can to be more effective than responding to teacher-made questions (Rosenshine et al., 1996; Wong, 1985). Moreover, different instructional approaches to self-questioning were found to be equally effective, be it reciprocal or conventional teaching, peer-assisted or teacher-assisted procedures. Two factors were associated with higher effect sizes; program duration and the use of particular prompts, such as signal words (who, what, when, where, why, how) (Janssen, 2002).

Effects of self-questioning approaches have been studied less often for literature reading. A well-known study is that of Singer and Donlan (1982), who trained students in generating and answering questions based on story-grammar categories (e.g. setting, main character, character's goal, theme). Students were taught to apply a set of general questions to specific short literary stories. Singer and Donlan's main focus was on teaching story grammar; no attention was paid to more spontaneous, reader-based questions.

During the 1990s reader response approaches entered the literature classroom. In these approaches the importance of developing an open, questioning frame of mind in response to literature, was well recognized (e.g., Chambers, 1993; Langer, 1994; Rosenblatt, 1938/1999). However, instead of teaching students to ask particular (meta)cognitive questions (as in the Singer & Donlan study), the emphasis was on students' personal responses to literature and on eliciting genuine, reader-based questions; students should be invited to ask questions about anything that captured their interest or puzzled them during their reading of a literary text. Special importance was placed on fostering a sense of 'personal relevance' and 'ownership' in response to literature.

In reader response approaches, questioning is not so much presented as an individual reading strategy, but rather as a starting point for meaningful conversations with other readers. The assumption is that, by exchanging and discussing their questions in peer discussions and classroom conversations, students may become aware of multiple perspectives on a literary text. Social interaction, especially with peers, may also contribute to students' reading engagement, more than when an individual processing of a literary text is required. Increased reading engagement in turn may lead to higher levels of response and a deeper understanding of the text (Applebee, Langer, Nystrand & Gamoran, 2003; Eeds & Wells, 1989; Nystrand, 2006).

A few case studies were published on reader-based questions in the literature classroom (Commeyras & Sumner, 1998; Kooy, 1992). In these studies, students were stimulated to express their uncertainties, wonderings, and hunches about stories, in the form of questions. Questions were first written down individually in personal reading logs, and then discussed in explorative conversations, aimed at sharing and challenging each other's ideas. According to the teachers and researchers, students were motivated to ask meaningful questions, and to think more deeply about stories. Both students and teachers responded positively toward the literature approach (Commeyras & Sumner, 1998; Kooy, 1992). Thus, these case studies provide indications for the motivating effect of self-questioning in the literature classroom.

However, the assumptions about the effects of 'authentic', student-generated questions on literary response have not been tested experimentally as yet. It remains uncertain, for instance, whether a reader response approach based on 'authentic' student-questions is more effective than a regular approach to literature reading, in which teacher-prepared questions dominate (Janssen, 1996). Also, it is unclear how much and what type of guidance students need - in particular the weak, inexperienced readers among them – in order to be able to generate meaningful questions in response to literary texts and develop a 'questioning frame of mind'.

2. RESEARCH QUESTIONS AND EXPERIMENTS

The purpose of this study was to design and test an instructional approach to literature reading that is based on 'authentic' student-generated questions. We carried out two experiments, involving 15- to 16-year old students. Our main research questions were:

1) Which instructional approach is more effective; self-questioning or responding to instructor-prepared questions? (Experiment 1);
2) Which type of self-questioning instruction is more effective; with or without guidance in self-questioning? (Experiment 2).

The relevance of Experiment 2 depends on the outcome of Experiment 1. If the first experiment would yield a result strongly in favour of instructor-prepared questions, there would be no reason to conduct the second experiment. This is why we conducted the experiments independently and consecutively. This is reflected in the way the experiments are presented in this article.

We examined the effects on two variables: the quality of students' interpretations and their appreciation of the complex short stories that they read. Our first hypothesis was that students who are stimulated to generate and discuss their own questions in response to short stories will outperform students who are required to respond to given, instructor-prepared questions, both in the quality of their interpretation and in their appreciation of stories. This hypothesis was tested in Experiment 1. Furthermore, we hypothesized that the less-experienced readers (students who read relatively little fiction in their spare time) will benefit more from guidance in self-questioning than their more experienced peers. In this study, the guidance consisted of providing students with weak and good examples ('models') of student-questions and other responses to short stories. This hypothesis was tested in Experiment 2.

3. THE EXPERIMENTAL LESSON SERIES

In this section, we describe the experimental program, as it was implemented (with variations) in both experiments.

The program was intended for students in 10th grade of higher general secondary education and pre-university education in the Netherlands, and general secondary education in Belgium (Dutch speaking part). Students at these educational levels are 15- to 16-year olds who are receiving formal literature education for the first time, since formal literature education only starts in 10th grade, at least in the Netherlands. Most students, then, are novices in the field of literature, they have little experience with literature written for adult readers.

It should be noted that in the Netherlands, there is no formal or national curriculum for literary education. Schools and teachers are free to spend as much time on literary education as the total instruction time for mother-tongue education and the other learning goals (reading, writing, oral skills) allow them to. They are also free to choose the literary texts, subject matter, pedagogy and learning activities of their own liking. However, there are formal goals for literary education that teachers are obliged to assess at the end of the last form (the final examination), with students being 17-18 years of age:

- *Literary development*: the candidate is able to report and elaborate on his reading experiences with 8 to 12 Dutch literary works, 3 of which are published before 1880.
- *Literary terminology*: the candidate is able to recognize and distinguish literary genres, and to use literary terminology in the interpretation of literary texts.

- *Literary history*: the candidate is able to present an outline of Dutch literary history, and to place the literary works he/she read within this historical perspective.

The experimental lesson series can be seen as a preparation for students' independent reading of literary texts, and as part of stimulating students' literary development.

The lesson series consisted of 5 lessons of 60 minutes (in Experiment 1) or 6 lessons of 50 minutes (in Experiment 2). The lessons centered on one literary genre; the short story. In each lesson one or two short stories were read and discussed. The stories had to be appealing for beginning readers of literature, and contain a clear story-line. At the same time, they also must raise questions and invite multiple interpretations.

We used short stories, written by recognized authors of modern literary fiction. With one exception ('Poison' by Roald Dahl), the stories were originally written in Dutch (by René Appel, Manon Uphoff, Bob den Uyl, and others). For the Roald Dahl story we used an authorized translation. Most stories were short; between 1 and 5 pages in print. Each student received a booklet containing all stories, and worksheets containing all assignments.

In the lessons, students were taught the following strategy for the reading and interpreting of short stories:

'Read the story, and

- ask yourself questions while reading;
- after reading, choose one essential question;
- discuss your question with peers; decide on one or more answers or solutions to your question;
- substantiate your answer(s).'

Each lesson focused on one of these strategic steps, but previous steps were continually repeated. In the final lesson we asked students to apply all the steps, working in pairs, while reading and responding to two stories of their own choice (chosen out of five different short stories).

We strove for variation in student activities, that is, we required the students to read, to listen to read-aloud stories, to write different types of responses, to make posters, and to discuss. Students worked individually, as well as in pairs and small groups. Each lesson started with an individual reading of the story; during reading students were asked to write down their questions (about 20 minutes, depending on the story). This was followed by small group discussions or exchanges in pairs (about 15 minutes). The group results were then discussed in a whole-class conversation, led by the teacher (about 10 minutes). Generating and discussing questions formed the principal part of each lesson.

In principle, students were allowed to pursue any question that interested them; no explicit instruction in question types was given (e.g. questions concerning particular story elements). The main requirement was that a question should be genuine or 'authentic'; students were discouraged to ask rhetorical questions or test questions for which they already knew the answer (e.g., 'what's the title?' 'what's the name of the dog?') (Van der Meij, 1993; Kooy, 1992). Furthermore, thought-provoking, dis-

cussion-type questions were preferred to questions for which the answers could be looked up in a dictionary or encyclopaedia. We called the preferred type of questions 'essential' questions; questions that appear to go to the heart of the story, that do not have one obvious 'right' answer, and keep lingering in your mind after having read the story.

The teacher's role was foremost that of a facilitator. We instructed teachers not to 'teach' the meaning of the stories, by evaluating student questions in terms of 'good' or 'bad' questions, or by providing 'right' answers. Their task was to assist students in presenting their own questions and in developing their own responses in explorative discussions. During discussions the teachers were instructed to support students' developing understandings as much as possible, for instance by summarizing, by asking for clarifications and elaborations ('Why do you ask this question?' 'How did you come to this idea?'), by relating responses to each other ('Does this mean that you agree with ….?'), and by giving positive feedback ('The point you raise is very interesting') (Chambers, 1993).

The experimental lesson series was first tested in two 9th grade classes, and then revised. Based on the experimental program two comparison programs were designed; one in which self-questioning was replaced by instructor-prepared questions, and one which contained question examples or 'models' (see below; Method).

4. METHOD

In this section we will successively present the methodology for the two experiments that together comprise our study (see section 'Research questions and experiments'). Experiment 1 was conducted prior to experiment 2. We present the data of both experiments in the 'Results' section.

4.1 Experiment 1

4.1.1 Design

In Experiment 1, two conditions were compared; an experimental and a comparison condition. In the experimental condition students read short stories and were stimulated to generate and discuss their own questions during and after reading a story (as described in the previous section). In the comparison condition students read and discussed the same stories, but they were not stimulated to generate questions themselves. Instead, questions were provided to them on worksheets. These questions were open-ended interpretation questions, previously generated by 9th grade students in the pilot study (e.g., 'Do you think this story really happened?' 'How does the title relate to the story?' 'Why does she let him treat her so badly?').

Thus, in both conditions student-questions were used; the main difference was whether the questions were self-generated by the participating students, or not. In all other respects, the conditions were kept as similar as possible; the same stories, in the same order, the same activities, the same amount of discussion, etc..

4.1.2 Participants

Participants were 67 students from four different secondary schools in Amsterdam (10th grade). The students volunteered to follow the lessons after school time; they received a small financial reward. Their mean age was 15 years. There were as many students from higher general secondary education (32), as students from pre-university education (35). Girls were overrepresented (50 girls versus 17 boys). Not all students were avid readers of literature. One third of the students (23) reported to read at least one book of fiction per month; the other students reported to seldomly (23) or never (21) read fiction in their spare time. Most students (61 %) considered themselves to be average or below average in literature as a school subject.

The students were assigned at random to one of the two conditions: the experimental ($N = 35$), or the control condition ($N = 32$). Two experimental and two comparison groups were formed, each consisting of 16 to 18 students. No differences were found between the experimental and comparison groups on a range of background variables that appeared to be relevant, such as gender, mean age, ethnic background, self reported frequency of reading fiction, or achievement in literature at school. The groups were also comparable with regard to students' general orientation toward literature, as assessed with the Literary Response Questionnaire of Miall and Kuiken (1995).

Teachers

The lessons were led by two experienced teachers of Dutch language and literature. Both were teacher-trainers at the Graduate School of Teaching and Learning, and involved in the research project. To avoid teacher-effects, each teacher coached one experimental and one comparison group.

4.1.3 Procedure

The students came to the university once a week, during a period of 7 weeks, in November and December. Their total time investment was about 8 hours, including a pretest session (90 minutes), 5 lessons (60 minutes each), and a posttest session (90 minutes). All lessons were observed by one of the researchers, and discussed with the teacher.

4.1.4 Measures

Several measures were used to assess students' story understanding and appreciation. Before and after the lesson series, each student read a short story under think aloud conditions. Five different short stories were used, written by Kader Abdolah, Primo Levi, Marianne de Nooyer, Cees Nooteboom, and Jeanette Winterson. The stories were distributed at random to students, in such a way that each student read a different story at the two moments of testing. The stories and think aloud procedure had been tested in a previous study (Janssen, Braaksma & Rijlaarsdam, 2006).

After reading and thinking aloud, students wrote a short response to the story they had just read. Students were asked to critically review the story for an imagi-

nary peer audience. In their review, students had to describe the story they had just read, to give a personal opinion of the story, and provide arguments in support of their opinion. Their text should contain at least 200 words. Students wrote their text directly on the computer. Afterwards, all reviews were printed, made anonymous and put in a random order.

Students' reviews were rated in terms of the overall quality of their story interpretations. Two levels of interpretation were distinguished; a local level (local inferences, e.g., concerning setting or story characters) and a global level (interpretive statements concerning the story as a whole, such as main idea, theme of the story, or author's intentions). The overall quality of the local and global interpretations in the reviews was rated separately, using a three-point scale (1 = weak, 2 = fair, 3 = good), by three independent raters. The overall reliability score or Cronbach's alpha for local interpretations was .83; for global interpretations .72. See the Appendix for examples of students' story interpretations and ratings.

Students' appreciation of the stories was assessed by way of a questionnaire administered at the end of the pre- and posttest sessions. In the questionnaire students were asked: 'How much did you like the story you read previously, in the think aloud session?' Students were asked to rate the story on a scale from 0 (strong dislike) to 10 (strong liking of the story).

4.1.5 Implementation

The two teachers delivered the lessons in constant consultation with the researchers. All lessons were observed by one of the researchers in order to monitor the implementation of the lessons. On the whole, lessons were carried out according to plan, and students were engaged in the intended learning activities most of the time. The instructions and assignments, provided on worksheets, presented few problems for students or teachers.

4.2 Experiment 2

4.2.1 Design

In Experiment 2 again two conditions were compared (condition 3 and 4). In condition 3, students were stimulated to generate and discuss questions during and after reading a story, in much the same way as in Experiment 1. (See section The experimental program). In condition 4, students were also stimulated to generate and discuss their own questions, but – in addition – received good and weak examples of student questions and responses to stories. The examples were derived from 'authentic' responses, generated by students in Experiment 1. Students were asked to reflect on these examples, and to make judgments on their quality and interestingness. Findings were discussed.

We decided to conduct this experiment in real classrooms, as part of the regular literature curriculum (and thus not outside the school, at the university, as in Experiment 1). This would enable us to examine the effects of the self-questioning approach under more or less ordinary circumstances.

4.2.2 Participants

Ten whole classes (10th grade) participated, of five different schools in the Netherlands and Belgium. We assigned the classes randomly to one of the conditions; five classes to condition 3 ($N = 127$), and five classes to condition 4 ($N = 118$).

The students' mean age was 15.5 years. The classes assigned to condition 3 were found to contain relatively more boys than the classes assigned to condition 4 (53 versus 36 %). In other respects the experimental groups appeared to be comparable; no significant differences were found in self reported achievement in Dutch language and literature as a school subject, nor in self reported frequency of fiction reading.

Teachers

Nine experienced teachers of Dutch language and literature volunteered to participate in the research project by delivering the lessons. Eight teachers participated with one of their classes; one teacher participated with two classes (one in each condition).

We discussed the lesson series with them beforehand, and provided each teacher with a manual describing each lesson, the lesson goals and a time schedule for student activities.

4.2.3 Procedure

The experiment took place between January and April 2005. All students received a collection of the short stories and worksheets. In most classes, the teachers dedicated one hour a week to the experimental lessons, over a period of about 6 weeks. The pretest and posttest each took one lesson, so that the entire experiment extended over a period of 8 weeks.

4.2.4 Measures

Students' experience with reading fiction was assessed (along with some background variables) by a questionnaire at pretest. In the questionnaire we asked the students to indicate on a 5-point scale how many books of fiction they had read during the last three years, ranging from 'not one book' to 'one or more books per week'. There were three items forming a reliable scale (Cronbach's alpha .92).

Before and after the lesson series, each student read two short stories. Two parallel test versions were constructed; version A (containing a story by Vonne van der Meer and one by Jeanette Winterson) and version B (containing a story by Primo Levi and Kader Abdolah). To avoid story-effects, the versions were counterbalanced; half of the students in each class received version A as pretest, and version B as posttest, whereas the other half received the versions in reversed order.

Students were asked to read the two stories and write down their initial responses in the margins. Next, they received open-ended interpretation questions concerning different aspects of meaning (e.g., 'In the story, the weeping willow and handkerchiefs are often mentioned. What role do they play in the story?' 'The story ends

with a question: 'Who are they with fish and starfish in their hair?' How would you answer that question?'). The questions targeted both local and global aspects of meaning. Students were asked to write down their answers in their own words.

Student answers were rated for 'depth' of interpretation on a 3-point scale (0 = answer is missing; 1 = superficial answer; 2 = answer reflects some depth of interpretation and/or reflect some awareness of multiple perspectives on the story). The minimum score was 0, the maximum score 16 (2 points for each of the 8 questions). Reliability over 8 items was just sufficient (Cronbach's alpha .70).

The ratings were performed by one of the researchers. To examine the reliability of ratings, 320 student answers were rated by the researcher again after some time, and also by a second independent rater. The agreement between the researcher's first and second rating was 91 % (Cohen's kappa .85), the agreement between the two independent raters was 88 % (kappa .80). See the Appendix for examples of student answers and ratings.

To measure students' story appreciation, students were asked how much they liked each story. In answer to this question, students rated each story on a scale from 0 (strong dislike) to 10 (strong liking of the story).

4.2.5 Implementation

To monitor the implementation of the lessons, we asked the teachers to keep a log in which they reported on their own and students' experiences. In these logs some deviations of the original lesson plans were reported. One teacher, for instance, had decided to read all stories aloud in order to accommodate students with reading problems. Another teacher had replaced one of the stories, because his students already had read that particular story.

Yet, on the whole the teachers faithfully carried out the lessons. Teachers unanimously responded positively toward the self-questioning approach. The lesson material was clear, well-structured, and instructive for students according to the teachers. Some noted that lively literature discussions had taken place in the classroom. Others noted, however, that students needed time getting used to their active thinking-role, especially in the beginning. Asking questions and participating in explorative discussions was new and appeared to be difficult for some students.

To assess students' time-on-task, nine lessons (one lesson per teacher) were observed by one of the researchers or a research-assistant. We observed a number of individual students systematically during 2-minutes-intervals. Results indicated that, in general, students showed relatively little off-task behavior (in minutes: M =14, SD = 5.9). No significant difference was found in the amount of off-task behavior between conditions (F = .261, df 1,9, p = .623). In one respect, students' activities in the guided self-questioning condition differed from those in the unguided self-questioning condition; in the latter condition students spend relatively more time on reading (28 versus 17 per cent of the time; F = 5.144, df 1,9, p = .053). Most likely, this can be explained by the examples in the learning material, which required some extra reading time.

5. RESULTS

5.1.1 Story appreciation

We found significant correlations between students' pretest and posttest scores for story appreciation in all four conditions (r = .29 - .57, p < .05). Moreover, we found no significant differences in pretest story appreciation between the groups. This means that the experimental groups were comparable with respect to pretest story appreciation.

To examine treatment effects on story appreciation we performed a one-way univariate analysis of covariance (ANCOVA). ANCOVA allows testing of treatment effects while controlling for the influence of relevant other variables. In our analyses, we included story appreciation at pretest as the covariate in order to control for pretest differences (Table 1).

Table 1. Story appreciation at pre- and posttest in Experiment 1 and 2 (mean scores at pretest and posttest on a 10-point scale, standard deviations) and results of analyses of covariance.

Condition	Pretest M (SD)	Posttest M (SD)	Analyses of covariance			
			R^2	F	df	p
Experiment 1						
Condition 1 Self-questioning	5.7 (1.7)	6.5 (1.3)*				
Condition 2 Given questions	5.8 (1.8)	5.3 (1.8)				
			.205	8.26	2	.001
Experiment 2						
Condition 3 Unguided self-questioning	5.7 (1.6)	6.1 (1.2)*				
Condition 4 Guided self-questioning	5.7 (1.3)	5.5 (1.4)				
			.273	39.48	2	.000

* Difference between pre-and posttest: p < .05

Pretest scores for story appreciation were found to significantly influence posttest scores (p < .05 in Experiment 1, p < .001 in Experiment 2).

Table 1 shows that in both experiments, significant effects of condition on story appreciation were found, after controlling for pretest story appreciation. In Experiment 1, students who had received instruction in self-questioning showed a significantly higher story appreciation at posttest than students in the comparison group, who had not engaged in self-questioning but had responded to given, instructor-made questions. In Experiment 2, students who had received unguided self-questioning instruction showed a significantly higher story appreciation at posttest than students who had received some guidance in the form of question and response examples.

5.2 Story interpretation

In Experiment 1 students' written reviews were found to contain hardly any 'global' interpretations, at pretest as well as at posttest. Only four students provided global interpretations at both moments of testing. The rarity of global interpretations in students' reviews prevents us from drawing any conclusions concerning changes in their quality as a result of the lessons.

Students' reviews did contain a large number of 'local' interpretations (on average 4 per review) which were rated for quality by the three raters. Only one student did not provide any local interpretations at pretest and posttest. In Table 2, we present the mean scores for local interpretations, within the two conditions (self-questioning and responding to given questions).

Table 2. Quality of students' local interpretations in Experiment 1 (mean scores at pretest and posttest, on a three point scale; 1 = weak, 2 = fair, 3 = good).

Condition	Pretest M (SD)	Posttest M (SD)
Condition 1 Self-questioning	2.1 (.5)	1.9 (.5)
Condition 2 Given questions	2.1 (.5)	2.1 (.5)

The mean scores in Table 2 suggest a general lack of improvement between pretest and posttest in both conditions. However, no significant correlation was found between the quality of local interpretations at pretest, and their quality at posttest (r = .11, p = .40). This means that, in point of fact, we cannot compare the pretest- and posttest scores. As a consequence, for Experiment 1 we cannot ascertain whether students improved their story interpretation skill as a result of the lessons, or not.

Apparently, the written review used in Experiment 1 was not a valid measure for story interpretation. Therefore, we decided to use a different measure in Experiment 2; instead of writing a review, students were asked to answer open-ended interpretation questions about two different stories. This time, we found a significant medium sized correlation (r = .40, p < .001) between students' pretest and posttest interpretation scores.

To test the effects of guided versus unguided instruction in self-questioning in Experiment 2, we performed a one-way univariate analysis (ANCOVA), with interpretation quality at posttest as the dependent variable, and interpretation quality at pretest as the covariate to control for pretest differences in interpretation quality. Results are presented in Table 3.

As expected, the covariate significantly affected the posttest interpretation scores (p < .001). After adjusting for pretest quality of interpretation, students who had received unguided self-questioning instruction showed significantly higher scores for quality of interpretation than students in the guided self-questioning condition.

Table 3. Quality of students' story interpretations in Experiment 2 (mean scores at pretest and posttest; on a 16-point scale, standard deviations) and result of the analysis of covariance.

Condition	Pretest M (SD)	Posttest M (SD)	Analysis of covariance R^2	F	df	p
Condition 3 Unguided self-questioning	9.1 (2.8)	10.5 (2.7)*				
Condition 4 Guided self-questioning	8.5 (2.7)	9.2 (3.1)				
			.196	25.29	2	.000

* Difference between pre-and posttest: $p < .001$

5.3 Reading frequency

We expected that inexperienced readers of literature would benefit more from guidance in self-questioning than more experienced readers among the students. In other words, we expected an interaction effect of condition and reading experience (or frequency) on students' posttest scores for story appreciation and interpretation. To examine this, we performed two regression analyses; one with posttest story appreciation and one with posttest story interpretation as the dependent variable.

For story appreciation, no interaction effect of condition and reading frequency was found. This indicates that the beneficial effect of the unguided self-questioning condition did not depend on the level of students' reading experience. For story interpretation, an interaction effect was found of condition and reading frequency on students' posttest scores (after controlling for pretest differences). The interaction is visualized in Figure 1.

The horizontal axis in Figure 1 represents levels of reading frequency; varying from non-reading (-2 *SD*) to very frequent reading of fiction in spare time (2 *SD*). On the vertical axis, students' posttest interpretation scores are presented in standard scores (or Z-scores) (adjusted for differences in pretest interpretation scores). The two lines represent readers in the two conditions.

The figure shows that the effect of the guided self-questioning condition did not vary, depending on students' reading frequency (the straight line). In the unguided condition, however, we found an interaction with reading frequency. The effect of this condition on students' interpretation scores depended on their level of reading frequency; students who read relatively much fiction in their spare time, the so called avid readers, profited more from unguided questioning than students who had little experience in reading fiction.

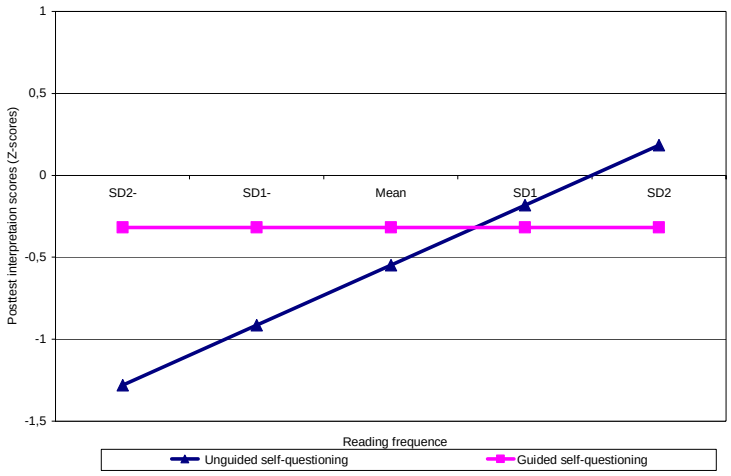

Figure 1. Effect of the interaction between condition and students' reading frequency on their posttest interpretation scores (in standard scores)

5.4 Overview of results

In Table 4, we present an overview of the treatment effects we found. We used Cohen's effect size index *d* as indication of the magnitude of an effect. Cohen (1988) defined an index of .20 as a small effect, an effect of .50 as medium-sized, and an effect of .80 as a large effect.

Table 4. Overview of treatment effects and effect sizes (ES) in Experiment 1 and 2

Condition	Effect on story appreciation ES (*d*)	Effect on story interpretation ES (*d*)
Experiment 1 Condition 1 Self-questioning Condition 2 Given questions	.68	
Experiment 2 Condition 3 Unguided self-questioning Condition 4 Guided self-questioning	.43	.42

As Table 4 shows, Experiment 1 resulted in one main effect; students who received self-questioning instruction responded more favourably to stories than students who had received instruction based on instructor-prepared questions about short stories. According to Cohen's (1988) definition, the effect is medium-sized (d = .68). Whether the self-questioning instruction also influenced students' interpretation skill more than instructor-made questions, remains undecided due to measurement problems.

In Experiment 2, two main effects were found; students who received self-questioning instruction without any guidance outperformed students who had reflected on examples of student questions and responses, both in story appreciation and in quality of story interpretations. The effect sizes are small (d = .42 - .43). The main effect for story interpretation was undermined by an interaction effect in the unguided self-questioning condition. In this condition, students benefitted more depending on their level of reading experience. For the guided self-questioning condition, no interaction effect with reading experience was found.

6. DISCUSSION

The results of this study suggest that a literature approach based on 'authentic' student-generated questions in response to complex literary stories has a positive impact on students' appreciation of such stories. We found that self-generated questions were more beneficial in this respect than instructor-prepared questions about stories. In addition, we found that an open, unguided approach to questioning was more successful in increasing students' story appreciation than an approach in which students first responded to examples of (student) questions. These findings suggest that 'personal ownership' of questions and/or the process of generating questions yourself, are important motivating factors. This finding is in line with reader response theories, and with case studies of student-questioning in the literature classroom (Commeyreas & Sumner, 1998; Kooy, 1992).

Thus, our first hypothesis was confirmed, at least partially. We were unable to determine whether self-questioning instruction also leads to a deeper understanding of stories compared to the use of instructor-made questions. The reason for this will be discussed below.

Furthermore, we had expected that the less-experienced readers (students who read little fiction in their spare time) would benefit more from guidance in self-questioning than their more experienced peers. This hypothesis was not confirmed. We found that the effects of guided self-questioning did not depend on the level of students' reading experience. However, in the open, unguided form of self-questioning instruction, students' reading experience appeared to matter for their story interpretation scores at posttest. In this condition, students who read relatively much fiction in their spare time scored better on the interpretation posttest than students who read relatively little fiction.

A strong point of our study is that we tested the experimental program twice, in two experiments. The experiments were carried out in different settings, with different students, and different teachers, but with similar instruments. With respect to

story appreciation, the findings converge; an open, unguided approach to student-questioning in response to literary stories was found to be most beneficial. The replication enhances the validity of this finding.

In Experiment 1 two similar groups were created by randomly assigning students to conditions. Shadish, Cook and Campbell (2002, p. 252) strongly recommend random assignment, because it facilitates causal inference in many ways, making alternative explanations implausible. Moreover, so called mortality did not occur in Experiment 1; the groups were identical at post- and pretest. Furthermore, teacher-effects were ruled out by counterbalancing; two well-informed, experienced teachers taught both conditions. They implemented the treatments with great precision in constant consultation with the researchers. An analysis of variance, performed as a check, indicated that there was no teacher-effect on students' appreciation or interpretation scores at posttest.

It should be noted that the two conditions in Experiment 1 were very similar. The same short stories were discussed, in the same setting. Even the questions about the stories that were discussed were comparable, since the instructor-made questions in the comparison group were based on authentic student-questions in response to the stories. The conditions differed in one respect: whether questions were generated and posed by the students themselves, or posed by the instructor. The close similarity of the conditions diminished the chance of finding differences in posttest performance.

However, Experiment 1 had limitations as well, especially regarding its ecological validity. The experiment took place outside the school, at the university. The students were volunteers who participated in the project after school time. Students did not know each other, for the most part, and the groups were rather small compared to normal classes. Moreover, girls were in the majority, since few boys volunteered. Although we have indications that most participating students were 'average' students, not just highly motivated students or high achievers in the subject of literature, the whole setting was clearly different from a regular literature classroom.

In Experiment 2, we have attempted to counter some of these threats to validity by conducting a quasi-experiment in real literature classrooms. Different schools were involved, with different classes, taught by nine different teachers. A larger number of students partipated (N = 245) than in the first experiment (N = 67). The experimental program was carried out as part of the regular literature curriculum.

Conducting an experiment in real classrooms, however, brings along other threats to validity and reliability. First of all, the reliability of treatment can be questioned. Obviously, the nine teachers each held their own views on literature teaching and they each had their own teaching style. Almost certainly this has caused differences in the way the self-questioning instruction was delivered. For instance, teachers may have differed in the way they responded to student-questions (supportive or less supportive) during whole-class discussions. We have attempted to control the variability between and within conditions in several ways; (1) by discussing the key points and pitfalls of the lesson series with the teachers before and during the experiment; (2) by providing a manual for the teachers, including lesson plans and a time schedule; (3) by providing worksheets for the students, which contained all stories and tasks, and (4) by monitoring the lessons via teachers' logs and our own

observations. The available data indicate that, although there were variations between teachers, most lessons were carried out as intended.

Second, although whole classes were randomly assigned to conditions, students within these classes were not. Possibly, the students in the two conditions already differed from each other at the start of the experiment. So, we may have confounded treatment effects with population differences (Shadish, Cook & Campbell, 2002, p. 56). For a number of personal variables we have examined whether our experimental and comparison groups were comparable. The groups were found to be similar, except for gender; boys were overrepresented in the unguided self-questioning condition. As was to be expected, girls performed significantly better than boys, in all conditions. However, we did not find any differences in treatment effects between boys and girls; the open, unguided self-questioning approach was found to be most effective, irrespective of students' gender.

Third, a crucial question (in both experiments) is whether students did indeed learn to use a self-questioning strategy in response to short stories, and whether they applied that strategy more frequently at posttest than at pretest. If not, we cannot ascribe our results to the self-questioning instruction students' received. In Experiment 1, we collected think aloud responses to stories at pre- and posttest, which were analysed for the reading activities the students used during reading (for the procedure see; Janssen et al., 2006). An explorative analysis showed that students more often engaged in questioning during reading at the posttest than at pretest. In the experimental condition, the majority of students (94 %) increased their number of questions during reading between pretest and posttest. In the comparison condition, about half of the students (46 %) showed an increase in questions at posttest.
In Experiment 2, students did not think aloud, but wrote down their initial responses to stories in the margins. We analysed these responses for the questions they contained. In both conditions, students asked more questions at posttest than at pretest. About 58 % of the students (in both conditions) increased the number of their questions in initial response to stories. These findings suggest that most students had indeed learned to use the self-questioning strategy in response to short stories. In future studies, the relationship between self-questioning during literature reading and reading outcomes (appreciation, interpretation) should be examined more closely. Not just the frequency but also the content and quality of student-questions during reading should be analysed.

Finally, questions may arise with regards to our measurement of 'story appreciation' and 'quality of story interpretation'. Both are complex concepts. Obviously, the measurement of 'story appreciation' can be improved, since we limited ourselves to a simple questionnaire (in response to just one or two stories). More and more open instruments have to be used in future studies. Finding a valid way of measuring 'quality of interpretation' proved to be difficult. In Experiment 1, we used a written review of a story as instrument. Although the reliability of the ratings was sufficient, we have serious doubts about the validity of the instrument. No correlations were found between students' pretest and posttest scores, and no general improvement was observed. Apparently, we measured something different at the two moments of testing. As a result we cannot draw any conclusions about students' improvement in story interpretation in Experiment 1. A different instrument was used in Experiment

2, consisting of open-ended interpretation questions about two different stories. Although not quite consistent with the nature of self-questioning instruction and being far from perfect, this instrument proved to be useful. Constructing reliable instruments to measure literary understanding must have high priority in future research.

Students often lack motivation for reading and studying literature at school, especially in the higher grades of secondary education (Van Schooten, 2005). One of the reasons could be that students - novices in the field of literature - are confronted with complex and ambiguous literary texts, written for adults, which are explained to them by the teacher. Often, students are not stimulated to think for themselves; they are expected to find 'right' answers to the teacher's questions and/or to search for deeper layers of meaning the teacher has in mind. In Dutch literature education, this approach appears to be the default (Janssen, 1996). Results of this study indicate that a more open learning environment, in which students are encouraged to think for themselves about what they read and in which their personal questions are valued and discussed, may contribute not only to students' appreciation but also to their understanding of complex literary texts.

ACKNOWLEDGEMENTS

The experiments were part of the research project "Learning to interpret literary texts in the 'studiehuis'" (2001-2006), carried out at the Graduate School of Teaching and Learning, University of Amsterdam. The project was funded by NWO, the Dutch Society of Scientific Research (grant 411-21-008).
Our special thanks goes to the teachers and students who participated in our study.
We would also like to thank Gert Rijlaarsdam and Mary Kooy for their support and valuable advice during different stages of the project.

REFERENCES

Andringa, E. (1995). Strategieën bij het lezen van literatuur.[Strategies in literature reading]. *Spiegel*, *13*(3), 7-33.

Applebee, A.N., Langer, J., 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*(3), 685-730.

Chambers, A. (1993). *Tell me: Children, reading and talk.* South Woodchester: The Thimble Press.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* (2nd ed.). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Commeyras, M., & Sumner, G. (1998). Literature questions children want to discuss: What teachers and students learned in a second-grade classroom. *Elementary School Journal*, *99*(2), 129-152.

Earthman, E.A. (1992). Creating the virtual work: readers' processes in understanding literary texts. *Research in the Teaching of English*, *26*(4), 351-384.

Eeds, M., & Wells, D. (1989). Grand conversations: An exploration of meaning construction in literature study groups. *Research in the Teaching of English, 23*, 4-29.

Huang, Z. (1992). *A meta-analysis of student self-questioning strategies*. Dissertation, Hofstra University. (UMI Dissertation Services No. 9211456).

Janssen, T. (1996). Asking for trouble. Teacher questions and assignments in the literature classroom. *SPIEL: Siegener Periodicum zur Internationalen Empirischen Literaturwissenschaft*, *15*(1), 8-23.

Janssen, T. (2002). Instruction in self-questioning as a literary reading strategy: An exploration of empirical research. *L1, Educational Studies in Language and Literature, 2*(2), 95-120.

Janssen, T., Braaksma, M. & Rijlaarsdam, G. (2006). Literary reading activities of good and weak students; a think aloud study. *European Journal of Psychology of Education, 21*(1), 35-52.

Kooy, M. (1992). Questioning classroom questioning. *Journal of Learning about Learning*, *5*(1), 14-26.

Langer, J. A. (1994). A response-based approach to reading literature. *Language Arts*, *71*(3), 203-211.

Miall, D. S., & Kuiken, D. (1995). Aspects of literary response: A new questionnaire. *Research in the Teaching of English*, *29*, 37-58.

Rosenblatt, L.M. (1938). *Literature as exploration*. New York: D. Appleton Century Company Inc.

Rosenshine, B., Meister, C., & Chapman, S. (1996). Teaching students to generate questions: A review of intervention studies. *Review of Educational Research*, *66*(2), 181-221.

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

Singer, H., & Donlan, D. (1982). Active comprehension: Problem-solving schema with question generation for comprehension of complex short stories. *Reading Research Quarterly*, *17*(2), 166-186.

Taboada, A.M. (2003). *The association of student questioning with reading comprehension*. Dissertation. University of Maryland.

Trabasso, T., & Magliano, J.P. (1996). Conscious understanding during comprehension. *Discourse Processes*, *21*(2), 255-288.

Van Schooten, E. (2005). *Literary response and attitude toward reading fiction*. Dissertation. Groningen: Rijksuniversiteit Groningen.

Van der Meij, H. (1993). What's the title? A case study of questioning in reading. *Journal of Research in Reading*, *16*(1), 46-56.

Veenman, M. (2004). *Self-questioning as a metacognitive skill.* Unpublished paper. Amsterdam: Graduate School of Teaching and Learning, University of Amsterdam.

Winterson, J. (1998). The three friends. In J. Winterson, *The world and other places* (pp. 47-52). London: Vintage Books.

Wong, B.Y.L. (1985). Self-questioning instructional research: A review. *Review of Educational Research*, *55*, 227-268.

APPENDIX

QUALITY OF STORY INTERPRETATIONS: EXAMPLES OF RATINGS

Here, we will present some examples of student responses rated on quality of story interpretations. We will present these examples separately for experiment 1 and 2 because story interpretation was assessed differently in both experiments. The examples are taken from students' responses to one of the stories that we used in both experiments: The three friends by Jeanette Winterson (1998). This short story (two and a half pages long, less than 500 words) begins as a simple fairy tale ("Once upon a time there were three friends who found a third"), but then develops into an absurd, post-modern story. Three friends decide to go on a quest. They decide to look not for gold or wives, but for "that which cannot be found". On their quest they first find gold, then they find women, and finally they manage to find what they were looking for, or rather it finds them. It is suggested, not told explicitly, that the three friends meet their death by drowning in the sea. The story ends with a question: "Who are they with fish and starfish in their hair?"

Experiment 1: Ratings of written reviews

As described previously in this article, students wrote a short response to the story. They were asked to critically review the story for an imaginary peer audience; they had to describe the story, give a personal opinion of the story, and provide arguments in support of their opinion. The students' reviews were rated in terms of the overall quality of their story interpretations. Two levels of interpretation were distinguished; a local level (local inferences, e.g. concerning setting or story characters) and a global level (interpretive statements concerning the story as a whole, such as main idea, theme of the story, or author's intentions). The overall quality of the local and global interpretations in the reviews was rated separately, using a three-point scale (1 = weak, 2 = fair, 3 = good).

Now, we will present for each scale an example of the interpretations at *local level*, starting with weak interpretations, defined as incorrect interpretations. For instance, the student statement "In the end the three friends probably saw mermaids, because the final sentence says: "Who are they with fish and starfish in their hair?"". This statement was rated as a weak interpretation, since the student failed to make the inference that "they" probably refers to the three friends.

Fair interpretations are acceptable, and could be correct, but are for instance superficial, or show small misconceptions that do not affect the understanding of the text as a whole. For instance, "The friends should be very rich, because of that palace", or "Probably, they were bored and therefore they decided to seek "that which cannot be found"".

Good interpretations are correct, acceptable and can be quite elaborate. Also nice, striking observations of the student were rated as 'good interpretations'. For

instance, "The three friends remind me of the three little pigs. They are comic figures, like Huey, Dewey and Louie."

Also for the interpretations at *global level* we will present for each scale (weak, fair, good) examples of student responses. Weak interpretations are interpretations that are senseless and are not supported by the text at all, for instance, "The three friends is about how miserly people could be". Fair interpretations are interpretations that are not implausible, some support can be found in the text, but they are not very satisfying (e.g., too superficial, not elaborated). Examples of fair interpretations are: "The three friends is all about greed", or "During reading, you discover that there is more to life than money. You need an ideal to live for". Good interpretations show that the essence of the story is well struck; the student gives a satisfying, quite elaborate global interpretation of the story. Also surprising and original interpretations that are supported by the text are rated as good interpretations, for instance, "You cannot find death, but in the end it will find you. Therefore, death is "that which cannot be found", you'd better not look for it" or "The story is quite ironic. The friends find all the things they are not looking for: gold and women. And that what they are looking for, finds them."

Experiment 2: Ratings of answers to open-ended questions

In Experiment 2, students did not write a review but received open-ended interpretation questions concerning different aspects of meaning. The questions targeted both local and global aspects of meaning. The questions were based on interpretation difficulties shown by the written reviews in experiment 1 (e.g., "The story ends with a question: "Who are they with fish and starfish in their hair?" How would you answer that question?"). Students were asked to write down their answers in their own words. Their answers were rated for depth of interpretation on a 3-point scale (0 = 'no idea', the answer is missing, or the answer is circular, repeats the question; 1 = superficial answer; 2 = answer reflects some depth of interpretation and/or reflects some awareness of multiple perspectives on the story). The minimum score was 0, the maximum score 16 (2 points for each of the 8 questions).

An example of a superficial answer to the question who "they" could be in the story's final sentence is: "They are mermaids or people who live in the sea". An example of a student response rated as reflecting some depth is: "I guess 'they' in the final sentence refers to the three friends who have met the Grim Reaper, and thus found "that which cannot be found"; they have drowned and are now lying at the bottom of the sea."

TANJA JANSSEN

Graduate Institute of Teaching and Learning
University of Amsterdam
Spinozastraat 55
1018 HJ Amsterdam
the Netherlands
E-mail: T.M.Janssen@uva.nl
MARTINE BRAAKSMA

Graduate Institute of Teaching and Learning
University of Amsterdam
Spinozastraat 55
1018 HJ Amsterdam
the Netherlands
E-mail: M.A.H.Braaksma@uva.nl

MICHEL COUZIJN

Graduate Institute of Teaching and Learning
University of Amsterdam
Spinozastraat 55
1018 HJ Amsterdam
the Netherlands
E-mail: M.J.Couzijn@uva.nl

Filename: Janssen Braaksma Couzijn published version
Directory: C:\Data\Mijn documenten\Text\L1 journal\Vol Janssen & Pieper\Published Versions
Template: C:\DOCUME~1\Owner\LOCALS~1\Temp\fctemp\Kluwer Vrevised.dot
Title: A
Subject:
Author: dana colarusso
Keywords:
Comments:
Creation Date: 3/15/2009 8:31:00 PM
Change Number: 5
Last Saved On: 3/15/2009 8:38:00 PM
Last Saved By: grijlaa1
Total Editing Time: 12 Minutes
Last Printed On: 3/15/2009 8:38:00 PM
As of Last Complete Printing
Number of Pages: 26
Number of Words: 11.478 (approx.)
Number of Characters: 65.425 (approx.)