Observational learning for narrative writing in elementary school. Text quality and self-efficacy in students with normal hearing and students with hearing loss.

Grenner, Emily

2021

Document Version:
Publisher's PDF, also known as Version of record

Observational learning for narrative writing in elementary school

Text quality and self-efficacy in students with normal hearing and students with hearing loss

EMILY GRENNER
FACULTY OF MEDICINE | LUND UNIVERSITY

Printed by Media-Tryck, Lund 2021 NORDIC SWAN ECOLABEL 3041 0903

In this dissertation I describe the development and evaluation of an observational learning intervention for narrative writing and a corresponding self-efficacy scale. The intervention was implemented in Swedish elementary school classes for students with normal hearing in Grade 5 (n=55) and for students with hearing loss in Grades 5–8 (n=11). The self-efficacy scale was implemented in classes for students with normal hearing in Grade 5.

The students wrote personal narratives and the students with normal hearing filled out a self-efficacy scale. Verbal working memory capacity, language comprehension and reading comprehension were measured. For students with hearing loss, audiological data was collected.

The dissertation demonstrates that observational learning may complement other structured writing interventions supporting young writers.

LUND UNIVERSITY

FACULTY OF MEDICINE

Department of Clinical Sciences, Lund
Logopedics, Phoniatrics and Audiology

Lund University, Faculty of Medicine
Doctoral Dissertation Series 2021:78
ISBN 978-91-8021-084-3
ISSN 1652-8220

Observational learning for narrative writing in elementary school

Text quality and self-efficacy in students with normal hearing and students with hearing loss

Observational learning for narrative writing in elementary school

Text quality and self-efficacy in students with normal hearing and students with hearing loss

Emily Grenner

LUND
UNIVERSITY

DOCTORAL DISSERTATION

by due permission of the Faculty of Medicine, Lund University, Sweden.
To be defended at Belfragesalen, BMC D, September 9th, 2021, 13:15.

Faculty opponent
Associate Professor Jakob Åsberg Johnels, Gothenburg University

Organization LUND UNIVERSITY	**Document name** DOCTORAL DISSERTATION
Faculty of Medicine, Department of Clinical Sciences, Lund, Logopedics, Phoniatrics and Audiology	**Date of issue** September 9th, 2021
Author EMILY GRENNER	Sponsoring organization Marcus and Amalia Wallenberg Foundation

Title and subtitle

Observational learning for narrative writing in elementary school

Text quality and self-efficacy in students with normal hearing and students with hearing loss

ABSTRACT

Background: Writing skills are crucial in our society. They should therefore be trained early and developed throughout school with effective methods.

Aim: This dissertation aimed at developing and evaluating an observational learning intervention and a self-efficacy scale for supporting narrative writing in elementary school students with normal hearing (NH) and with hearing loss (HL). The intervention and the scale were developed for possible use in a Swedish school context.

Method: A five-lesson observational learning intervention and a self-efficacy scale were developed. Grade 5 students with NH (n=55) and students with HL in Grades 5–8 (n=11) wrote four narrative texts which were rated holistically. Students with NH filled out a self-efficacy scale twice. Verbal working memory capacity (WM), language (sentence) comprehension and reading comprehension were measured. Audiological data was collected for students with HL.

Results: An observational learning intervention for writing and a corresponding self-efficacy scale could be developed and used in the classroom, in line with the Swedish curriculum. Mixed effects regression analysis showed a positive intervention effect on text quality ratings for students with NH. The effect was not sustained at follow-up. Higher WM scores were associated with higher text quality ratings (NH). Higher scores on the reading comprehension test were associated with stronger intervention effects (NH). For students with HL, there were no significant text quality changes. Cognitive and linguistic factors had no overall effects (HL). The students with HL with the highest text quality ratings were amplified earlier than students with the lowest text quality ratings. Female gender was associated with higher text quality ratings (NH and HL). Repeated measures ANOVA showed that self-efficacy for writing increased significantly after the intervention (NH). Students with high text quality ratings had strong self-efficacy (NH). Boys and girls had similar self-efficacy (NH).

Conclusion: Observational learning could be implemented in Swedish schools and may complement other structured writing interventions. The reinforcement of new skills may be important for sustained effects, especially for students with poorer language skills. Despite girls' higher text quality ratings, boys with NH had similar self-efficacy to girls with NH. Writing motivation may increase by observing and identifying with models. The higher text quality in students who were diagnosed with HL earlier underline the importance of early diagnosis of HL for optimal audiological, linguistic and pedagogical support. Adapting, evidence-basing and implementing new teaching practices as well as evidence-basing current methods is important in improving the teaching of writing and strengthening language outcomes. The continuous support of students' writing development is a question of facilitating their academic achievement in school and their participation in and contribution to a sustainable society.

Keywords modeling; pediatric hearing loss; personal narratives; writing intervention

Classification system and/or index terms (if any)

Supplementary bibliographical information: Lund University, Faculty of Medicine Doctoral Dissertation Series 2021:78	**Language** English
ISSN and key title 1652-8220 Observational learning for narrative writing in elementary school	**ISBN** 978-91-8021-084-3

Recipient's notes	**Number of pages** 91	Price 0
	Security classification	

Signature  Date 2021-07-05

Observational learning for narrative writing in elementary school

Text quality and self-efficacy in students with normal hearing and students with hearing loss

Emily Grenner

LUND
UNIVERSITY

Cover illustration by Signe Grenner

Cover photo by Emily Grenner

Faculty of Medicine
Department of Clinical Sciences, Lund, Logopedics, Phoniatrics and Audiology

ISBN 978-91-8021-084-3
ISSN 1652-8220
Lund university, Faculty of Medicine Doctoral Dissertation Series 2021:78

Printed in Sweden by Media-Tryck, Lund University
Lund 2021

This is how you do it: you sit down at the keyboard and you put one word after another until its done. It's that easy, and that hard. *(Neil Gaiman)*

Table of Contents

List of publications

Paper I

Grenner, E., Åkerlund, V., Asker-Árnason, L., van de Weijer, J., Johansson, V., & Sahlén, B. (2020). Improving narrative writing skills through observational learning: A study of Swedish 5th-grade students. *Educational Review 72*(6) 691-710.

(Advance online publication, 2018-11-05).

DOI:10.1080/00131911.2018.1536035

Paper II

Grenner, E., van de Weijer, J. Johansson, V., & Sahlén, B. Predictors of narrative text quality in students with hearing loss. *Logopedics, Phoniatrics Vocology.*

(Advance online publication, 2021-02-10)

DOI: 10.1080/14015439.2021.1881613

Paper III

Grenner, E., Johansson, V., van de Weijer, J., & Sahlén, B. (2021). Effects of observational learning intervention on self-efficacy for narrative writing in elementary school. *Logopedics, Phoniatrics Vocology* 46(1) 1-10.

(Advance online publication, 2020-01-08).

DOI:10.1080/14015439.2019.1709539

Abbreviations

ANOVA	analysis of variance
BEHL	best ear hearing level
CI	cochlear implant
DLD	developmental language disorder
HA	hearing aid
HL	hearing loss
LD	language disorder
LME	linear mixed effects regression model
NH	normal hearing
PISA	Programme for International Student Assessment (of OECD)
SD	standard deviation
SES	socio-economic status
SLP	speech-language pathologist
VAS	Visual-Analogue Scale
WHO	World health organisation
WM	working memory

Papers at a glance

Table 1. Main findings

	Paper I	Paper II	Paper III
	Improving narrative writing skills through observational learning: A study of Swedish 5th-grade students	Predictors of narrative text quality in students with hearing loss	Effects of observational learning intervention on self-efficacy for narrative writing in elementary school
Aim	To develop an observational learning intervention To evaluate intervention effects on narrative text quality ratings To relate narrative text quality data to cognitive and linguistic data.	To perform the observational learning intervention in classes for students with HL To relate narrative text quality ratings to cognitive and linguistic data, school class and audiological data	To explore self-efficacy for narrative writing before and after an observational learning intervention
Data	Text quality data, WM, language comprehension and reading comprehension test scores	Text quality data, WM, language comprehension and reading comprehension test scores and audiological data	Text quality data, WM, language comprehension and reading comprehension test scores. Self-efficacy scales pre- and post-intervention
Participants	55 students with normal hearing	11 students with hearing loss	55 students with normal hearing (as in Paper I)
Design	Effects of predictors and intervention estimated in a mixed effects model Four measuring points, intervention between T1 and T2 or between T2 and T3	Effects of predictors estimated in a mixed effects model Four measuring points; intervention between T1 and T2	Associations between text quality and self-efficacy measured before and after intervention with repeated measures ANOVA
Results	Mild but significant intervention effects modulated by reading comprehension Intervention effects were not sustained Effects of WM on text quality ratings Boys' texts received significantly lower text quality ratings	Text quality did not improve Later age at amplification was associated with lower text quality ratings Boys' texts received significantly lower text quality ratings	Strong internal consistency and strong self-efficacy for narrative writing Higher self-efficacy ratings post-intervention Similar self-efficacy in boys and girls
Conclusion	The intervention was effective in the short term but effects were not sustained over time.	Early diagnosis and early rehabilitation are important for language development. Intervention should be further developed for students with HL before evaluation and implementation.	Increased self-efficacy for writing was a possible positive effect of the observational learning intervention. Results support previous findings of strong self-efficacy at this age.

Preface

Once upon a time, before joining this research project, I was working as a school speech-language pathologist (SLP), meeting students with language disorders (LD), which sometimes co-exist with hearing loss or other factors which can affect language development. Working with those students and collaborating with their teachers and other school personnel has given me a valuable perspective on this research. Since I started out as a school SLP over ten years ago, school SLPs have become somewhat more common. Hopefully that trend will continue, developing effective collaboration between teachers, SLPs and other professionals in the schools, as students with LD need considerable support and time for their language, reading and writing development.

In this dissertation I describe the development and evaluation of a narrative writing intervention by observational learning in two settings: mainstream school classes with a full range of students in Grade 5 and school classes exclusively for students with hearing loss in Grades 5–8. This dissertation may be relevant for school SLPs and teachers, both those who are working in settings with students with hearing loss and those who are working in full range/mainstream school classes. I hope I will be able to continue working with writing intervention in different age groups and settings.

The dissertation is part of a project investigating observational learning and keystroke logging in school settings. It was financed by the Marcus and Amalia Wallenberg Foundation, Grant 2012.0038, and by the Faculty of Medicine, Lund University, Sweden.

Introduction

Writing skills are crucial in our society. We write in our social life, at work, almost regardless of line of work, and at school, regardless of school subject (Rijlaarsdam et al., 2005; Skolverket, 2019a). All students need support in developing writing throughout their education, as writing requires a complex interaction of language skills (Beard et al., 2009). Written communication is expected for expressing and exploring knowledge and views throughout school, not only in the traditional language subjects (Skolverket, 2019a). Students also often write to communicate with classmates and teachers. The acquisition of adequate writing skills is thus a democratic right, meaning that writing skills should be trained early and developed throughout school with effective methods. Writing instruction is, however, often too scarce to give students the skills they need (e.g. Brindle et al., 2016; Gilbert & Graham, 2010). In the age group of the students in this dissertation, 10–16 years, writing development has not been studied as extensively as young children's writing and writing in older students (Myhill, 2009b).

Students who are at risk of language and listening difficulties, for example students with hearing loss (HL), need particular support in developing adequate writing skills. Students with HL represent a heterogeneous group as for language skills and many meet the linguistic criteria for developmental language disorder (DLD) (Geers et al., 2016; Hansson et al., 2017; Sahlén et al., 2018). Language difficulties with no known biomedical aetiology, which create obstacles to communication in everyday life and are expected to continue over time are diagnosed as DLD (Bishop et al., 2017). An LD is diagnosed when the communicative criteria for DLD are met and there is a possible aetiologic cause of the disorder, for example 'LD associated with HL' (Bishop et al., 2017). Many students with HL fall behind peers with normal hearing (NH) in basic language domains such as phonology (Sundström et al., 2018), grammar and lexicon. The development of the complex language skills essential in school, for example narration, conversation, listening and reading comprehension and writing is therefore often delayed (Geers et al., 2016; Sandgren et al., 2015). The evidence base for effective writing instruction for students with HL is weak (Roos & Allard, 2016; Strassman & Schirmer, 2012).

Narration or storytelling is at the heart of humankind (Brown, 2013). It requires the cognitive abilities to grasp the concept of past, present and the future. We use spoken or written narratives or stories to carry and create our beliefs and history, our knowledge of the current world, and what we think the future holds. The (written)

narrative genre is considered a prerequisite for other genres, for example the expository and argumentative genres. Lastly, narrative skills have been found to predict academic achievement (Catts et al., 2003; Feagans & Appelbaum, 1986; Griffin et al., 2004). In the present dissertation, the writing instruction method *observational learning* and its effects on writing performance are in focus. During observational learning, the students observe other writers, in the present case peers, during their writing activities. The students observe, reflect on and learn about text creation separated in time from writing, setting the method apart from traditional writing exercises where an instruction on some aspect of writing is directly followed by writing according to the instruction. This may alleviate the cognitive load (Braaksma et al., 2002). Evidence of positive effects of observational learning on writing performance and on self-efficacy for writing is accumulating, often in studies with participants in high school or at undergraduate levels who write in expository or argumentative genres (Bouwer et al., 2017; Braaksma et al., 2002, 2018; Braaksma et al., 2004; Raedts et al., 2007; Rijlaarsdam et al., 2008). Observational learning studies in similar age-groups to the participants with NH in the present dissertation are few and studies targeting students with HL are fewer (but see van de Weijer et al., 2018, for a study on university students with HL). To the knowledge of the author, there were no published studies of observational learning for narrative writing in Swedish schools prior to the studies in this dissertation, nor were there any Swedish studies of observational learning for 12–16-year-old students with HL.

Writing is cognitively demanding and for example working memory (WM) plays a crucial role in it (Hayes & Berninger, 2014; Kellogg et al., 2013; McCutchen, 2000). As Graham and Harris write, "having to switch attention while writing to a mechanical demand such as figuring out how to spell a word, may lead a writer to forget ideas or plans held in working memory, influencing sentence construction, how much they write, and the quality of their text" (Graham & Harris, 2014, p. 94). Dealing simultaneously with low-level processes such as transcription and spelling and high-level processes such as planning and reading through a text, is taxing. Many students, particularly those with weak language skills, must spend considerable resources on low-level processes of writing, which leaves less resources for higher order processes like organization of the text (Berninger & Swanson, 1994). The separation of learning from writing in observational learning may thus be beneficial. Another possible benefit of the observational learning paradigm is that observing peers performing a task is considered highly motivating (Raedts et al., 2017). Self-efficacy is defined as a person's beliefs about their ability to perform a task (Bandura, 1997). It is linked to motivation and may thus influence how well the task is performed, as a slight overestimation of capability will lead to trying harder on a particular task, according to Bandura (1997).

The following paragraphs introduce an overview of language and early literacy development, writing development, writing instruction and the Swedish school

setting and curriculum. Thereafter, children with HL and their specific educational challenges and needs in language and writing are described. Lastly, the paradigms of observational learning and self-efficacy are described.

Language and literacy development in children with normal hearing

Language development is highly dependent on social context – social, perceptual, motor and cognitive skills being essential (Paul et al., 2018). Hearing is developed in the 23rd week of pregnancy and sensitivity to sound improves the following weeks (J. Grenner, 2017) . Infants react more to speech (especially child-directed speech) than to other sound signals (Lagercrantz, 2017). Typically children's first words are interjected in their babbling at around 12-15 months and at this time they can already understand the meaning of many words as well as prosodic and syntactic elements of the language (languages) they hear (Paul et al., 2018). By age two, the spoken vocabulary may contain several hundred words (E. Clark, 2016, p. 87). At three years old, basic syntax is established in expressive language, and at four years old, phonological development has usually reached a level where the child's speech can be understood without contextual cues (Paul et al., 2018). Phonological awareness, for example the ability to identify phonemes in words and rhymes, continue to develop during the preschool years alongside early literacy development, for example recognizing the initial letter of one's name, understanding phoneme–grapheme mapping and reading direction (Runnion & Gray, 2019). At age six, when Swedish children start compulsory school, a child often understands around 14, 000 words (E. Clark, 2016, p. 87). Socio-economic factors, for example school district, impact the amount of words and complexity of language which children hear (Neuman et al., 2018).

Language comprehension is crucial for continuing language development, reading comprehension (Gough & Tunmer, 1986) and understanding and developing narratives (Blom & Boerma, 2016; Reuterskiöld et al., 2011). Poor language comprehension on the other hand is a predictor of DLD (e.g. Bishop et al., 2016), and children with language delay at 2.5 years often have oral narrative difficulties at 7-8 years (Miniscalco et al., 2007).

Young children's retelling of narratives occur first in "heaps" of information without clear structure, then with increasing cohesion of sequences and settings (Applebee, 1978). Already at six years old, children's retelling of narratives has developed into the schema of true narratives containing a beginning, an action and a resolution (Karmiloff-Smith, 1981). By the end of preschool, the spoken narrative genre has thus become well established in spoken language (Berman & Slobin, 1994; Nordqvist Palviainen, 2001; Westby & Culatta, 2016). Children's spoken narrative

skills have turned out to predict reading during school as well as academic achievement (Applebee, 1978; Catts et al., 2002; Feagans & Appelbaum, 1986). Further, reading comprehension has repeatedly been linked to working memory (Carretti et al., 2009) and to academic achievement (Swanson & Alloway, 2012; Titz & Karbach, 2014).

The Swedish preschool curriculum stresses children's language and literacy development (Skolverket, 2018), underscoring the importance of supporting narrative development from an early age. Preschool curriculum goals include letting children listen to narratives and other types of texts, discussing and interpreting them, supporting children's lexical development, and supporting children's interest in different modes of communication (Skolverket, 2018).

Writing development during elementary school

Writing requires many simultaneous processes which rely on basic cognitive and linguistic prerequisites (Beard et al., 2009; Hayes & Berninger, 2014). The most cited cognitive model of writing was originally proposed by Hayes and Flower (1980) and has been developed further several times (e.g. Hayes, 1996; Hayes & Berninger, 2014). In the model, writing is described as consisting of different cognitive processes during *planning* (generating ideas, organizing them and setting goals for the text), *translating* (putting the ideas into words and transcribing those words) and *reviewing* (by reading, evaluating and revising). During a writing task, the writer must keep the task itself, the reader and the topic in mind, recall earlier writing experiences and writing plans, while staying motivated. Focusing and shifting attention between tasks taxes limited cognitive resources. A combination of Hayes and Flower's cognitive model of writing and a model of WM was proposed by Kellogg (1996).

In early elementary school or before, children become aware that a written text is not speech written down, but this awareness of the difference between spoken and written language may not yet be reflected in their own texts (Myhill, 2009a). The writer is preoccupied with low-level transcription skills, which include forming letters or finding them on the keyboard, orthography, spelling and punctuation (Beard et al., 2009; Myhill, 2009a). Low-level transcription skills may not be established until 11–12 years (Berninger & Swanson, 1994; Fayol et al., 1999). As long as lower level writing skills, for example transcription and spelling, are not automatized, there is less capacity for more complex writing activities influencing text quality such as planning, sentence generation and reviewing (Arfé et al., 2014; Drijbooms et al., 2015; McCutchen, 2000). With advancing age, children's writing takes on more of the lexical and grammatical features which set it apart from spoken language (Myhill, 2009b). Written narratives develop further during elementary

school. During these years, writing is characterized by a linear structure, meaning that the writer seldom changes earlier parts of the text after writing later parts of it, and that revision generally focuses on local editing and spelling (Johansson, 2009; Limpo et al., 2014). Berninger and colleagues showed that students up to age 9 seldom revised their texts except for word for word, but from 9–11, more planning and revising occurred even after writing a portion of a text (Berninger et al., 1994; Berninger et al., 1996; Berninger et al., 1992). These skills were not fully developed until 12–15 years (Berninger et al., 1994; Berninger et al., 1996; Berninger et al., 1992). At 14–16 years, many students have reader awareness and thus revise their texts to suit the reader (Myhill & Jones, 2007). Lexical and syntactic development during these school years are influenced by children's reading skills, social context and general development (Myhill, 2009b). As the writer becomes more proficient, more aspects of the developing text are added to the task. Thus, an older student may make the same effort as a younger one, while producing a text of higher quality, as the freed cognitive capacities due to automatized lower level processing are replaced with, for example, genre awareness and reader awareness (McCutchen, 2000). In a study of Swedish groups of students who were 10, 13, 17 years old, and of adults at university, 10- and 13- year-olds showed linearity in their narrative writing, while the older age groups would also revise texts not only directly after writing a word (revision on a local level) but by going back further (global revision), to a greater extent (Johansson, 2009). Narrative writing was found to be highly influenced by spoken narrative skills during elementary school, while the written narratives began to influence spoken language in older students (Johansson, 2009).

Some studies on writing and narrative development in elementary school report gender differences (Kanaris, 1999; Myhill, 2008). Girls have been found to write longer, more elaborated and more complex narratives (Kanaris, 1999). Other studies find no gender differences (Johansson, 2009). In the curriculum for the compulsory school, writing and reading narrative texts are two of the central objectives for Grades 1–3 (ages 7–9) (Skolverket, 2019a).

Effective writing instruction

Writing instruction for elementary school students has been evaluated in several meta-analyses, for example (Graham & Harris, 2018; Graham, Liu, et al., 2018; Graham et al., 2012; Koster et al., 2015). Instruction methods which have been found effective include modeling, peer assistance/collaboration and strategy instruction (Graham & Harris, 2018; Graham et al., 2012; Koster et al., 2015). Increasing students' motivation, for example by improving self-efficacy for writing, has also been found effective for improving text quality (Graham & Harris, 2018). In strategy instruction, students may for example explicitly be taught how to plan and write a text in a specific genre, what structure certain texts have and how to

write and revise texts (Graham & Harris, 2018). Another type of strategy instruction, included in one of the mentioned meta-analyses, is observation of a model who is using effective writing strategies (Fidalgo et al., 2015). See under Observational learning and Method development below for a more detailed description. Strategy instruction has been found especially effective for increased text quality and strengthened self-efficacy and motivation, when combined with self-regulating strategies of writing such as goal-setting and self-monitoring (Graham et al., 2012; Koster et al., 2015).

The teaching of writing in Sweden

The syllabi of Swedish teacher training programs have changed over time according to different school curricula (1980, 1994 and 2011) and have local differences. Although prospective teachers are taught different theories on learning during teacher training, a socio-cultural perspective on learning has been influential in Swedish schools for a long time (Lundgren et al., 2017; Yassin Falk, 2017). From this perspective, learning develops through interaction with others, such as peers and teachers, within a social context as a social activity (Vygotsky, 1978; Yassin Falk, 2017). This affects the didactic methods which teachers are familiar with. Neither teacher training nor the Swedish curriculum stipulates which methods should be used for teaching and developing writing (Skolverket, 2019a). Instead, personal and local experience make a considerable influence on what students are taught. Around a decade ago, there was little focus in the curriculum on teaching writing (Skolverket, 2020b). PISA comparisons showed that Swedish students' literacy was declining during the first decade of this millennium, especially reading comprehension (OECD, 2013). To remedy this, a project for continued professional development (CPD) for teaching literacy (Läslyftet) led by the Swedish National Agency for Education was introduced in 2015 (Skolverket, 2020a). Some modules of the CPD included writing but the main focus was on reading (Roe & Tengberg, 2016). Since then, students' results on reading comprehension tasks have improved somewhat (OECD, 2016, 2019). Equity between students of different socio-economic backgrounds has, however, declined (SOU, 2020) and considerable differences in language and literacy skills in different schools districts are reported (Andersson et al., 2019). While measures are taken for improving Swedish students' literacy, effective methods for teaching writing suited to the Swedish school setting are still called for.

While there is freedom of didactic choices and teaching methods, the goals stated in the Swedish curriculum are more explicit. The curriculum addresses three school years at a time (Grades 1–3, 4–6, 7–9) and states the goals that students should reach in the subject of Swedish at the end of Grade 3 (10 years old), Grade 6 (13 years old), and Grade 9 (16 years old). For Grade 3, writing narrative texts is one central

objective of the curriculum (Skolverket, 2019a). Students at the end of Grade 3 are expected to know about the organization of a narrative (introduction, sequence of events and ending) and descriptions of literary figures. Students at the end of Grade 6 should know about the structure and linguistic features of different genres (e.g. descriptions, instructions, and argumentative texts), and should have developed their narrative writing, understanding "language features, structure and narrative perspectives in fiction for youth and adults" as well as "parallel action, flashbacks, descriptions of settings and persons, internal and external dialogues" (Skolverket, 2019a). Further, they should master syntactic and morphologic features including main and subordinate clauses, parts of speech, morphology, spelling rules, punctuation and cohesion (Skolverket, 2019a). They should know how to write by hand and on the computer. Further, they are expected to know how to organize and revise their own texts and how to revise others' texts and how to give and receive feedback.

Some methods used in Swedish schools include elements that are found in observational learning, for example strategy instruction, giving feedback to peers' texts and collaborative learning (Skolverket, 2020a).

Children with hearing loss

There are around 5,000 children with HL who need audiological services in the Swedish compulsory school for ages 6–16 (SOU, 2011:30). Most students with HL are "mainstreamed" in classes where most students have NH. In 2016, 85 percent were mainstreamed and the estimate for the coming years was 95 percent (SOU 2016:46). Some students with HL who follow the general Swedish curriculum are enrolled in classes exclusively for students with HL (Skolverket, 2019a). The students with HL in this dissertation belong to this group. Further, a small number of students with profound HL attend special schools which mainly use sign language (HL special school). Finally, there are special needs schools for students with intellectual disability and other disabilities, including HL. Students with HL have poorer academic achievement than their NH peers (Rydberg et al., 2009; SOU, 2016:46). Eighty-four percent of Swedish students with NH graduate from compulsory school eligible for high school (Skolverket, 2019b). In comparison, 78 percent of mainstreamed students with HL, 65 percent of students in classes exclusively for students with HL, and 38 percent of students in HL special school graduate with grades that make them eligible for high school (SPSM, 2008). According to the Swedish National agency for special needs education and schools (SPSM), the differences between school results for mainstreamed students with HL and students in classes exclusively for students with HL may in part depend on the fact that struggling students change schools, from mainstream classes to the exclusive classes (SPSM, 2008). Recent numbers have not been found, but

according to data from 2005, only ten to fifteen percent of students with HL continued to tertiary education (HRF, 2007), while around sixty percent of all Swedish students do (SCB, 2015). More recent American data show a similar pattern. Eleven percent of adults with NH were enrolled in tertiary education, compared to five percent of adults with HL (Garberoglio et al., 2019).

In Swedish classrooms for classes exclusively for students with HL there are often sufficient acoustic accommodations to improve listening conditions. The number of students is small, and classrooms are fitted with assistive listening devices. Aside from the time spent in the classroom however, much time is spent in environments where the acoustic environment is not adapted to persons with HL, for example in extracurricular activities, commuting to school and the school-age educare (fritidshem). In mainstream classrooms, acoustics vary considerably (Karjalainen et al., 2019; SOU, 2016:46). Noise impairs perception, even with well-fit amplification (Eisenberg et al., 2016). Thus, regardless of school placement and classroom accommodations, students with HL spend much time in contexts with adverse listening conditions. Children with HL have more listening difficulties in noise compared to peers with NH (Lewis et al., 2016; McFadden & Pittman, 2008; Torkildsen et al., 2019). Amplification and other acoustic accommodations are thus very important even for children with a HL classified as minimal or mild (Tomblin et al., 2015). The degraded speech signal caused by impaired perception of auditive input requires that students with HL must use more cognitive resources for listening. Brännström et al. (2018) found that nine-year-old students with NH listening in noise had poorer narrative comprehension than controls listening in quiet, which shows that students with NH also use more cognitive resources in non-optimal listening conditions. What is lacking in auditive input, the listener must fill in by, for example, contextual and visual cues (Rönnberg et al., 2013). Thus, in adverse listening conditions, fewer cognitive resources are available for the task at hand (Mattys et al., 2012).

Hearing loss is often measured with pure tone audiometry, i.e. the ability to detect tones in a quiet environment. The best ear hearing level (BEHL) is the audiometric average of hearing thresholds (in dB) at 500, 1,000, 2,000 and 4,000 Hz. According to WHO, BEHL can be used as "a useful summary of an individual's hearing thresholds" but does not determine how a person is affected by HL (WHO). For children, HL greater than 30 dB in the better hearing ear is considered disabling, while the cutoff for disabling HL for adults is 40 dB (WHO). Children are thus considered more sensitive than adults to HL. See Table 2 for an overview of degree of HL and corresponding BEHL. The degree of HL in the students in Paper II is classified according to J.G. Clark (1981), as this was the common classification when the students were diagnosed. Neither of these classification systems take unilateral HL into account, although a unilateral HL often makes listening difficult, especially in noise (Olusanya et al., 2019). Regardless of classification of HL,

hearing level and speech intelligibility with amplification in an ecologically valid setting is important for describing function, but it is not always measured.

Table 2. Degrees of hearing loss (HL) with corresponding better ear hearing level in dB (BEHL) according to the current classification of WHO, 2019, and according to J.G. Clark, 1981.

Degree of HL	WHO, 2019	Clark, 1981
Normal or minimal	< 26 dB	< 20 dB
Mild/Slight	26–40[1] dB	20–40 dB
Moderate	41–60 dB	41–70 dB
Severe	61–80 dB	71–90 dB
Profound	≥ 81 dB	>90 dB

[1] For children, a mild HL is considered disabling if >30 dB.

Neonatal hearing screening has become implemented throughout Sweden during the last two decades, but it was not in place when the participants of the study in the present dissertation were born. Late detection of HL still occurs. Children with mild and moderate HL seldom receive language assessment and intervention in Sweden (Rosén et al., 2019). Children with severe or profound HL usually receive more attention and intervention from audiological services. Those who receive a cochlear implant (CI) receive considerably more support than students with hearing aids (HA) (Rosén et al., 2019) in Sweden. Hearing aids amplify sound in an ear with HL, while cochlear implants bypass the ear, delivering sound signals directly to the auditory nerve (Gelfand, 2016). Those who have severe or profound HL may be eligible for CI surgery (Gelfand, 2016). There are no clear national guidelines in Sweden on how, when and by whom the follow-up on language development in children diagnosed with HL should be carried out.

Language and literacy development in children with hearing loss

Many children with HL fall behind peers with NH, in basic language domains e.g. phonology, grammar and lexicon as well as in complex language activities for example spoken and written narration, conversation, listening, reading and writing (Geers et al., 2016; Sandgren et al., 2015). Between twenty and fifty percent of pre-school children with HL meet criteria for language disorder (LD) (Briscoe et al., 2001; Geers et al., 2009). The degree of HL is often not proportional to language difficulties (Yoshinaga-Itano et al., 1998). Some studies show that even a minimal HL may lead to a considerable lag in language skills (Marschark & Knoors, 2012, 2018a; Tharpe, 2008). However, a recent study found that most students with mild or moderate HL caught up in language measures for example novel word reading and reading comprehension at 8 years, despite lagging behind in early literacy skills

at 5 years (Tomblin et al., 2020). The origin of language outcomes is multifactorial. The heterogeneity depends on the interaction of child-internal as well as external factors. Factors pertaining to the HL include degree of HL, etiology, age at identification and diagnosis, age at amplification, type of amplification and usage of it. Language skills such as grammar and vocabulary are better in children who use their HA consistently (Walker et al., 2015). Early exposure to sign language may also influence language skills including writing. For example, Gärdenfors and colleagues (2019) found that students with HL with early exposure to sign language made fewer spelling mistakes than has been found in other students with HL. Cognitive factors, for example WM, affect language skills (Arfé et al., 2015; Sahlén et al., 2018). For elementary school students with severe or profound HL, verbal WM capacity was found to contribute more to text quality of picture-elicited, retold written narratives, than age or reading comprehension (Arfé et al., 2015). Further, social and pedagogical factors, for example audiological and pedagogical support, and SLP services for the family and in the preschool or school also affect language skills. In sum, there is an interplay of risk and resilience factors, where no single factor can explain language outcomes in students with HL.

To end on a hopeful note, the published data on academic achievement in the previous section and the data on language outcomes may not correctly mirror the current situation. Neonatal hearing screening increases the chance of early diagnosis and amplification. Earlier diagnosis and amplification improve chances of language development on a par with hearing peers (Geers et al., 2016).

Writing instruction for students with hearing loss

Few well-controlled studies exist on effective writing instruction for students with HL. Implications for teaching are therefore inconclusive. According to a review of writing intervention for students with HL, “the evidence for practice is at best promising” (Strassman & Schirmer, 2012, p.176). Recent studies by Dostal, Wolbers and colleagues have shown positive results from strategic and interactive writing intervention writing instruction (SIWI) (Dostal & Wolbers, 2016; Wolbers et al., 2018; Wolbers et al., 2015). It includes activation of previous knowledge, teacher modeling and thinking aloud, writing practice while the teacher gradually decreases support, and feedback (Dostal & Wolbers, 2016; Wolbers et al., 2018; Wolbers et al., 2015). In a Swedish overview of reading and writing instruction for students with HL (Roos & Allard, 2016), the authors listed some possible success factors, based on the described problems and difficulties which these students have.

Observational learning

A successful strategy in writing instruction is the observation of a model (Bouwer et al., 2017; Braaksma et al., 2002; Fidalgo et al., 2008; Fidalgo et al., 2015; Harris et al., 2006; Rijlaarsdam et al., 2008). The model may be a teacher, or other "mastery model", showing students how to address an aspect of writing (Harris et al., 2006) or a peer with similar characteristics to the student of for example similar age and skill (Fidalgo et al., 2008; Raedts et al., 2017). Observation of a peer may be especially motivating and may affect students' self-efficacy for writing (Raedts et al., 2017). When students are 'modeling' they try to emulate their actions and strategies and even thoughts and beliefs, to the model (Schunk, 2003). Students may identify more with a peer model than with a mastery model, which motivates them to try to perform the task in the same way as the peer does (Schunk, 2003). Observation should be seen as an active process; the observing student does not passively absorb the modeled behavior and imitate it (Bandura, 1997). The observer must make a series of decisions, i.e. focusing attention, retention for later use, gauging the effectiveness of the behavior, for example by identifying similarities to the model and then using the same techniques if deemed useful (Bandura, 1997).

Observational learning is thought to relieve cognitive load by separating the learning activities from the writing (Braaksma et al., 2012). Writing requires much cognitive effort and may make following new instructions difficult while simultaneously juggling the many tasks which writing requires (Baker et al., 2003; McCutchen, 2000). Observing and listening to a peer who verbalizes how they complete a task while thinking aloud may make the observing student aware of what the model is doing, the observer thereby learning how to do the same thing next time they get a similar task (Braaksma et al., 2004). Thus, observation and reflection on a model makes students re-evaluate how they write their texts, which may give insights in what to do – or not do – in the next writing task, i.e. they have an opportunity of gaining metacognitive knowledge.

Observation of peers for improving writing has been evaluated in several studies with positive effects on text quality in expository and argumentative texts in high school and college students (Bouwer et al., 2017; Braaksma et al., 2018; Couzijn, 1999; Raedts et al., 2017; Rijlaarsdam et al., 2008; Zimmerman & Kitsantas, 2002). Some data show that similarity in skill between (peer) model and observer may affect learning and self-efficacy – a coping model, i.e. overcoming difficulties during the observed task may lead to better skill and higher self-efficacy (Schunk et al., 1987).

In a study on argumentative writing in 120 15-year-old students, Couzijn demonstrated that students who observed peers who were either writing or both writing and reading had a better writing performance than students who were doing writing exercises (1999). For a student to observe a peer reading and reacting to the

student's own text was also more effective than doing writing exercises. A study on college students similarly showed that the observation of a model improved writing skills (in this case syntax) more than learning without a model (Zimmerman & Kitsantas, 2002). Further, the authors showed that observation of a "coping" model, i.e. one who struggled and learned gradually, was more effective than observing a mastery model (Zimmerman & Kitsantas, 2002).

Self-efficacy

The concept of self-efficacy, or a person's beliefs about their capabilities to perform a task, was introduced by Albert Bandura within the framework of social cognitive theory (Bandura, 1977, 1986, 1997). When a person's self-efficacy is high, a difficult task will be viewed as an achievable challenge, motivating more effort. Slightly higher self-efficacy than ability increases motivation, whereas low self-efficacy on the other hand leads to investing less in a task, increase the risk of giving up, and thus reinforce lack of success (Bandura, 1997; Schunk, 2003).

Bandura identified four sources influencing self-efficacy (1986). The first is *enactive experience* – remembering a prior, similar activity which was successful may affect self-efficacy positively, whereas a prior perceived failure may affect it negatively. The second source is *vicarious experience* – watching somebody perform a task may positively or negatively influence self-efficacy depending on how the performer does. Observation of a peer is one form of vicarious experience. The third is *social persuasion* – encouragement or discouragement from a believable source influences self-efficacy. The fourth source is a person's *physiological and emotional state* – feeling at ease may increase a person's self-efficacy while they may interpret anxiety as a lack of capability or skill, (Bandura, 1986). Self-efficacy affects motivation and learning in widely different fields for example health, sales and academia (see Schunk 1991 for an overview). It is affected by task demands and may vary widely between different domains and skills and even between activities within a domain (Bandura, 1986, 1997). A valid self-efficacy scale should therefore measure self-efficacy in a domain-specific way (Bandura, 2006).

In the academic setting, self-efficacy has been found an important predictor of motivation and learning in children (Pajares, 1996; Schunk, 1994) as well as in adults (Richardson et al., 2012; Robbins et al., 2004). As an example, when comparing students from different math groups, strong self-efficacy was linked to better performance within each group, whether that group consisted of students with strong or weak performance (Coutinho & Neuman, 2008). Self-efficacy for writing has also been linked to writing performance (Bulut, 2017; Callinan et al., 2018; Pajares, 2003; Raedts et al., 2007; Schunk & Swartz, 1993). Some studies on writing self-efficacy and gender show that girls and boys report similar self-efficacy,

although girls *and* boys think that girls' writing is better (Pajares & Valiante, 1999, 2001).

Raedts and colleagues studied self-efficacy for writing in the observational learning context (2007, 2017). One study showed that observing models could increase self-efficacy directly (Raedts et al., 2007). Another showed that observation increased students' perceived value of the writing task, which in turn increased their self-efficacy for writing (Raedts et al., 2017). Students' self-efficacy for narrative writing has not been widely studied, especially in connection with writing intervention. Aspects of motivation such as self-efficacy may be especially important for struggling students, due to the difficulties they experience during writing (Schunk, 2003).

Self-efficacy changes over time. Young children often display strong and holistic self-efficacy which is not yet differentiated between domains or between aspects of a skill. It may comprise a general sense of skill, the perceived effort on a task, and self-assessment (Stipek, 2002). During childhood and adolescence, self-efficacy decreases and becomes more differentiated (Muenks et al., 2018; Pajares, 2007). This decrease may have several reasons. One is transition between schools, creating an unfamiliar environment (Schunk & Meece, 2006). Another is improved calibration to skill from a previous overestimation or, in other words, a deeper understanding of the complexity of a skill (Muenks et al., 2018; Stipek, 2002). A third reason is that self-efficacy is linked to self-esteem, which may decrease during adolescence. Further, motivation is not only affected by perceived *capability* to perform a task, but also by the perceived *result* of the effort and by the importance of the result to an individual (Geiger & Cooper, 1995).

Aim

The overarching aim of the research presented in this dissertation is to develop and evaluate an intervention for supporting narrative writing in elementary school students with NH and with HL and to develop and evaluate a self-efficacy scale for writing, which may be used in a Swedish school context.

To meet this aim, four research questions are posed:

1) How can an observational learning intervention for writing and a corresponding self-efficacy scale be developed and implemented in Swedish elementary school classes for students with NH and HL, in line with the Swedish curriculum, students' school grade and age?

2) How does a five-lesson observational learning intervention affect holistic text quality after intervention and at follow-up in students with NH and in students with HL?

3) How are cognitive (WM capacity) and linguistic (language comprehension and reading comprehension) factors and demographic factors (gender and, in the case of students with HL, school grade and audiological factors) associated with text quality before and after intervention?

4) How does self-efficacy change after intervention in students with NH and how is it associated with text quality and gender?

To address these questions, an intervention using observational learning was developed and implemented in typical school classes and classes for students with HL. A self-efficacy scale was developed to measure self-efficacy in relation to intervention effects. Narrative writing skills were measured by text quality and intervention effects were studied in relation to self-efficacy, gender, cognitive and linguistic factors and, in the case of students with HL, school grade and audiological factors. The three studies (Papers I–III) each had more specific aims and questions, which are briefly described under Results.

The first research question is mainly addressed in the chapter Method development. The second and third questions are addressed in Paper I for students with NH, and Paper II for students with HL. The fourth research question is addressed in Paper III.

Method development

This chapter describes the adaptation and development of observational learning intervention and self-efficacy measurement from the methods which inspired the present research. It further describes the preparation of the films used in the intervention and lesson development. The last paragraph of the chapter briefly describes a small parallel study, not reported in the papers in this dissertation.

Intervention design: Observation, reflection, and learning

The research in the present dissertation was inspired by several Dutch studies on observational learning, for example Couzijn's comparative study of different observational learning conditions (1999), and the "Yummy yummy case", reported in Rijlaarsdam et al. (2008). Adapting the observational learning paradigm for the purpose of this dissertation led to three major differences in design. Firstly, the intervention was adapted to suit the Swedish curriculum and school setting. An intervention series of five lessons would replace five other Swedish lessons for the students, not interrupting the rest of the syllabus for the semester. Secondly, the intervention would target younger students. Thirdly, to suit the younger students, the genre was changed to the narrative rather than the argumentative genre which has been studied more thoroughly in observational learning (Rijlaarsdam et al., 2008).

A series of five 40-minute lessons was developed, where observation was based on peers filmed working with writing activities. An intervention in that comparatively small scale would be possible to implement on a larger scale in Swedish schools if successful. It also made the data collection (aside from the follow-up) possible to fit within one semester. As the narrative genre would be well-known to the students, they would build their new knowledge on an already familiar genre, consolidating and developing their knowledge.

Lesson content

A classroom intervention in three steps was developed. To make the lessons familiar to the students, each lesson followed the same pattern. The lesson theme, for

example *"reader awareness"* was introduced. The five themes are listed below under Lesson content. The first step of the intervention was *Observation*, where the students would observe and listen to peers ('film peers') filmed working with writing activities. The second step was *Reflection*, where students would reflect individually and with classroom peers on what the film peers were doing. Several examples were shown where film peers were performing a writing activity. After each observation of one or a few film clips, a structured reflection with predefined questions would follow. The third and final step was *Learning*, where students would consolidate what they had learned by formulating it as advice for peers, i.e. "for next year's students in this grade". They would thereby reinforce what they found to be relevant to themselves and their peers. See Table 5 for examples of the advice which students wrote.

Based on the curriculum for Grades 4–6 (Skolverket, 2019a) and on writing activities suitable for observation, the intervention lessons were designed targeting the following areas: reader awareness, organization of events, conclusion and coda, revision of a written text, and online revision. Thus, high-level writing processes such as planning and organizing, as well as low-level writing skills such as spelling and punctuation were targeted. The following lessons were developed:

(1) The reader's perception: *What does the reader find important in a story?*
(2) Structure: *Different ways to start a story – in what order should the events unfold?*
(3) Conclusion: *How does one finish a story?*
(4) Revising (someone else's) text: *How can a text be improved?*
(5) Revising during writing: *What changes do writers make while they write?*

Preparation of observation material

Writing activities for film peers

To prepare the observation material, 10–13-year-old 'film peers' (unknown to the participants) were recruited and filmed in the Lund University Humanities Lab and in the department of Logopedics, Phoniatrics and Audiology at Lund University. The film peers came to the lab or department either on their own or two by two with a friend they already knew. The environment was set up to resemble a writing workshop or a small classroom and was without visual or auditive disturbances. The writing workshop took place after school and the whole session took around two hours. The film peers were introduced to the purpose of the films – that other students may be helped by observing them. The film peers were further asked to do their best but there was no right or wrong way to do the writing activities. The activities were prepared to resemble age-relevant classroom activities, so that the

resulting films did not have to be manipulated with regard to what the students were saying or doing. The film peers were introduced to the writing activities and asked to work on them without receiving verbal prompts from the researchers during filming. Further, they were prompted to speak loudly and clearly enough to be heard in the resulting films. During activities with only one film peer, the film peer was instructed to think aloud about what they were doing. Care was taken to make the films as ecologically valid as possible.

As expected, the different film peers carried out the writing activities in different ways because of differences in age, maturity, and writing skills. This yielded a rich material of film to choose clips from. In some cases, the film peers focused on other aspects of the activities than those the research team had anticipated when preparing the writing activities. If the film peers focused on something which was not expected by the researchers but valid to the curriculum, their focus was chosen, rather than trying to prompt other answers.

Selection of film material

The resulting films featured twelve film peers, five girls and seven boys. Ten of these were working in pairs for some of the activities. See table 3 for examples from each lesson of what writing activity the film peers were doing, the film outcomes and lesson outcomes. From the several hours of film material, clips which fit each lesson theme were selected from the different writing activities. Writing processes and text features on text level as well as word level were included in the lessons. As the twelve film peers had different writing proficiency and maturity, the students would thus observe a variety of approaches to the writing activities. The clips where selected to illustrate different levels of writing skill and to generate reflection in the participating students on different ways of taking on the writing activities. Four to seven different clips were selected for each lesson. A professional film editor edited the film clips and improved the sound by decreasing noise to provide the observing students with optimal conditions for understanding, especially important for the students with HL. Subtitles of what the film peers were saying were added, for those whose listening would be supported by reading.

Lesson preparation

Each lesson was carefully scripted around the film clips. Scripts for the lessons were prepared. These included information about timing of film clips and reflection as well as manuscripts of what to say. They could be followed verbatim, to ensure reliable intervention in all classes. The scripts included introduction of the lesson, prompts for reflection after film clips, and instructions for writing down advice. The prompts, one or two prompts for each film clip, were prepared, to elicit more answers if the students were reluctant to start sharing their thoughts. All lessons

were rehearsed in a "dry run" within the research group to further ensure reliability and to check the time schedule.

To help the participating students to formulate their advice succinctly, small sheets of paper (size A5) were prepared for them to write on. The students were told that their advice would be collected for "next year's students in this grade", so that they would think of important advice to themselves or their peers, rather than catering to a teacher or researcher. In the fifth and final lesson, the learning section of the lesson was to write down advice concerning that lesson theme (revising while writing), and any advice they wished to point out from the earlier lessons. Slides introducing each lesson theme were prepared to remind the students of the themes they had worked with. The students thus reflected on all five lessons and on what another learner would need to know when writing a narrative. The students were told that their advice would be collected and given back to them.

Table 3. Examples of writing activities, corresponding film outcomes and lesson outcomes from each lesson.

Lesson	Writing activity	Film outcome	Lesson outcome
1 The reader's perception	"Read and reflect on a personal narrative written by another student"	Some film peers react mostly on form aspects, others on content. The importance of making the reader want to continue to read becomes apparent.	Students reflect on the reader's perspective.
2 Structure	"Using this story board, try to think of different ways to tell the story. Can one start somewhere else than in the beginning?"	Some film peers play around with starting with the coda, or 'in medias res'. Some film peers stress the importance of following the chronological order.	Students reflect on alternatives to a linear narrative.
3 Conclusion	"Write about an occasion when you saved somebody from a jam" (written on a computer with keystroke logging)	Audio of film peer thinking aloud + screencast of computer screen during the writing process.	Students reflect on the importance of a coda.
4 Revising someone else's text	"Read this story, try to improve it by changing some of the words/wordings"	Film peers discuss how to improve a text (some successfully, some not).	Prompting students to reflect on others' writing processes, that text is not necessarily finished after the last full stop.
5 Online revision	"Write about an occasion when you saved somebody from a jam" (written on a computer with keystroke logging)	Screencast of the text making the writing process visible. Examples of editing of typos and of global revision, i.e. adding crucial information.	Making visible the writing process with its online editing, revising and thought processes during writing.

The written advice from the students was collected and printed in a booklet, "Tell me more!" which was given to all students after the data collection had finished. Thereby, the students thus had the content of the lessons summed up in the form of

their own advice. See Table 5 in the Method section for some examples of the students' advice reflecting each lesson.

Self-efficacy scale

Self-efficacy was measured in several of the observational learning studies which inspired the intervention in the present dissertation (Braaksma et al., 2002; Raedts et al., 2007). In the study by Raedts and colleagues, the participants were university students who wrote a literary review. Self-efficacy and text quality were measured. Only a third of the students were well "calibrated", i.e. had self-efficacy which was close to their actual performance, but after an observational learning intervention, fewer students overestimated their abilities (2007). Pajares and colleagues showed that although both girls and boys agreed that girls were the better writers (in the studies), girls and boys often reported similar self-efficacy (Pajares, 1996; Pajares et al., 1999; Pajares & Valiante, 2001). The authors suggested that girls may answer the self-efficacy scale differently from boys (Pajares, 1996; Pajares et al., 1999; Pajares & Valiante, 2001).

According to Bandura, self-efficacy should be measured in a domain-specific way (2006). Statements must be constructed for a specific genre, but not be task-specific, i.e. only for a specific type of assignment. If statements begin with "I can..." they are considered more specific to self-efficacy than for example questions (2006). Bandura further states that there must be many intermediate steps when marking the scale and suggests that one should answer with a number from 0 to 100, stating to what extent one agrees with that statement. This position is supported by Pajares and colleagues, who found that 0–100 scales were psychometrically sounder than 1–5 point Likert scales (2001). With these precautions, the person filling out the scale will state their self-efficacy, or self-perceived ability, rather than for example rating an earlier task. A Swedish scale developed to measure self-efficacy for reading and writing was available (Wolff, 2010). The scale is answered by choosing one of four options and not all statements start with "I can...". It would therefore not be valid for the purposes in this research project. After development of the scale in the present dissertation had begun, SEWS, a self-efficacy scale for writing was published (Bruning et al., 2013b). As scales and performance assessment should correspond, SEWS may not have been suitable for the purposes of this dissertation. Differences in self-efficacy may be based on students' age, as described earlier. However, reporting self-efficacy may be done differently in different age groups. Young children may not have as varied self-efficacy between domains or tasks. Interests and self-assessment of earlier work may contribute to self-efficacy, and valid scales may be difficult to make. Strong self-efficacy contributes to motivation and performance, and calibration (accuracy) of self-efficacy is difficult to assess.

The scales by Braaksma and colleagues (2002) were developed to measure self-efficacy for expository and argumentative writing in older students than those in the present dissertation. Their scale had eighteen statements in all, as did the scale developed for this dissertation. A third of the statements only had to be translated from Dutch (by bilingual members of the research team), as they were not genre-dependent or age-dependent but pertaining to general writing skills, for example spelling and revision. The developed scale (as Braaksma's scale), followed Bandura's recommendations for self-efficacy scales (Bandura, 2006). This means that it was constructed to be domain-specific but not task-specific and all statements started with "I can...". To make sure that the students would fill out the scale according to what they thought of their abilities, rather than assessing their performance in previous tasks, the following instructions were written on the first page of the booklet and would also be read aloud before they started filling out the scale:

> Imagine the following scenario: In school, you get a writing assignment where you are to write a narrative story about something you have experienced. For example, it could be writing a story about the most exciting time you had during the summer vacation. It should be written so that somebody in your class can understand it, and the text should be about one page long. You are not going to write this text yourself, but please answer some questions about what writing such an assignment would be like. Answer each question by marking the horizontal line beneath each question with a vertical line. The further to the right your mark is, the more you agree with the statement.

In the scale developed by Braaksma and colleagues and in Bandura's recommendations, students were asked to write a number between 0 and 100 indicating to what extent they agree with the statements (Bandura, 2006; Braaksma et al., 2002). As the students in the present dissertation were young, making a mark with a pen on a continuous scale was expected to be easier to grasp than writing down a number 0–100.

Cognitive and linguistic tests

Alleviation of the cognitive load has been suggested as an explanation of effects of observational learning (Braaksma et al., 2002). Despite this, cognitive measures such as WM capacity have not been measured in the studies described above. The measurement of WM capacity was therefore included in the studies presented here. Since language skills are crucial for academic achievement, and especially since students with HL often meet criteria for LD, language comprehension and reading comprehension tests were included in data collection as possible predictors for text quality.

Parallel study of intervention led by teachers

In Papers I-III, the intervention was carried out by researchers. In a separate study, described in a master's thesis by two SLP students (Hammarstedt & Karan, 2016), teachers used the intervention in their own Grade 4 and 5 classes after training, led by the research group. The intervention in the present dissertation was adapted to be used by teachers in the school district. Based on the filmed material and the scripted lessons, a teacher's manual for the intervention lessons was developed. Teachers in the school district of the students in Papers I and III were offered a full day of training by the research group and provided with the manual and the film clips. They could thus carry out the observational learning intervention in their classes. The teachers were asked to follow the manual closely. Some teachers performed the intervention in their classes. Narrative texts written before and after intervention were collected by the research group. The teachers split one intervention lesson in two and allowed students extended time for the *learning* section of the lessons. Results showed improved text quality after completion of the five intervention lessons (Hammarstedt & Karan, 2016).

Method

In the following section, the study design and procedure will first be described, followed by descriptions of the participating students and tests of cognitive and linguistic ability. Further, text quality measurement, intervention lessons and the self-efficacy scale will be presented. Lastly, considerations of statistical analyses and ethical issues will be reported.

Design and procedure

The full data collection (writing narratives, filling out self-efficacy scales and performing cognitive and linguistic tests) and intervention took thirteen lessons of 40 minutes each. Data collection and intervention were performed in the classrooms of the students, during their Swedish lessons, which were replaced with data collection or intervention lessons. An overview of the data collection and intervention is found in Figure 1. Written personal narratives were collected at four points in time (Texts 1–4). Self-efficacy scales were filled out twice, and tests of WM, language (sentence) comprehension and reading comprehension were administered. The intervention took place between Texts 1 and 2 (intervention first) or between Texts 2 and 3 (wait first). All data collection (aside from Text 4) was designed to fit within one semester to avoid general maturation effects. From the first data collection and until Text 3 was written, the teachers were asked not to focus on written narratives during their classes. Four classes of students with NH and two classes of students with HL participated.

Two classes of NH students in one school had intervention first, while two classes in another school had regular Swedish lessons first. Thus, students would constitute their own controls. In the classes for students with HL, the groups were small, and all students received intervention at the same time. In Grade 6 there are extensive national assessments of the subjects of Swedish, English and Math which affect the schedules of several weeks of the semester, as well as the focus of the students. Further, teachers recommended that the intervention should not be implemented in Grade 4 as they thought those students would be too young. This ultimately led to implementation in Grade 5 among students with NH. For practical reasons (foremost, the national assessments in Grade 6) all students with HL (who were in Grades 5–8) received intervention after Text 1 and had regular Swedish lessons after

Text 2. The self-efficacy scale and the cognitive and linguistic tests were piloted on students of similar age before data collection.

To make the intervention and the data collection ecologically valid, all interaction with the students took place in the students' regular classrooms during Swedish lessons. All lessons were led by two persons in the research group. The present author was responsible for data collection and led all intervention lessons, for optimal reliability (Ebbels, 2017; Ebbels et al., 2019). A research assistant (Viktoria Åkerlund) ran films and slides during intervention lessons and assisted during data collection. The present author gave all lessons except those missed due to parental leave – for the collection of Texts 3 and 4 in the classes of students with HL and Text 4 in the classes of students with NH. The research assistant was present for all lessons (assisted by another member of the research group for the collection of texts 3 and 4). The present author and the research assistant thus became well known to all students.

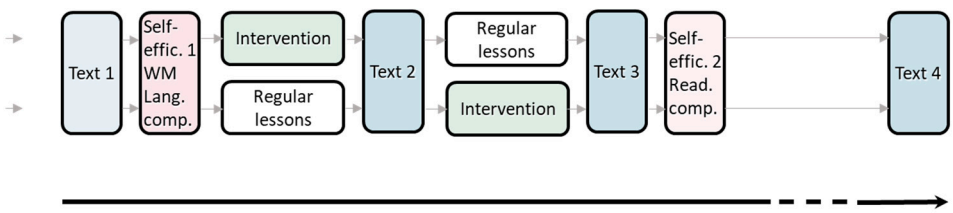

Figure 1. Overview of the design. All students with HL and one school with students with NH followed the upper path (intervention first). Students with NH in the other school followed the lower path (wait first). Personal narrative texts were written at four times (Texts 1-4). The self-efficacy scale (Self-effic.) was administered twice, after Texts 1 and 3. The working memory (WM) and language comprehension (Lang. comp.) tests were administered after Text 1 and the reading comprehension test (Read. comp.) was administered after Text 3.

Participants

The 55 students with NH (see Table 4) in Papers I and III were in Grade 5. Normal hearing is here defined as not having any suspected or diagnosed HL. The eleven students with HL in Paper II were in Grades 5–8. The students in papers I and III (all with NH) were in two schools in the same school district. The schools were located in similar areas, which were socio–economically relatively strong. For example, around 80 percent of parents in the area had received tertiary education, compared to 50-60 percent nationwide (publicly available data). The students in the classes were unselected, i.e. all students who would follow typical lessons could participate in the study. The majority of students had Swedish as their first language. There were few bilingual students.

The students in Paper II (all with HL) represented a very heterogeneous group. Their school classes were exclusively for students with HL and they followed the general Swedish curriculum (Skolverket, 2019a). The school received students from the whole geographical region. All had a spoken first language. Born before the introduction of neonatal hearing screening, some students' HL was not detected until they were several years old. Age at diagnosis with HL varied from a few months up to ten years. In this dissertation, degree of HL is classified according to J.G. Clark (1981) as this was the classification at the time of diagnosis. Some of the students' HL would thus be classified as more severe now, following the classification by WHO (see Table 2) (WHO). Parents or caregivers were asked to provide audiological information and/or access to audiological medical records. The data which was provided was somewhat limited, from caregivers as well as from medical records. A few caregivers did not give their consent for access to the medical records. Best ear hearing level (BEHL) could be found for all students with HA whose information we had access to. Hearing level with amplification had been measured in only one case, whereas in all other participants it was measured without amplification. Data from students who chose not to participate in the studies was discarded, while they were still included in the lessons. This is described under Ethical considerations and further details are found in Papers I–III.

Cognitive and linguistic tests

The cognitive and linguistic tests which were used in this dissertation were possible to administer in a whole class setting. Screenings and other tasks which are common in schools are often administered as paper-and-pen-tasks like the chosen tests. The tests are standardized or norm-referenced for students with NH. The tests are widely used in schools or by SLPs. More details on the tests can be found in the three papers. Table 4 gives an overview of the students' test results.

Verbal WM was measured with a subtest from the Swedish dyslexia screening test Lilla Duvan (Wolff, 2010). It has a process component and a recall component and the maximum score is 36. As an example, the students heard a letter "B" and were then asked "Is France larger than Denmark?". They responded to the question by raising a YES or a NO sign (process component). Then they heard another letter "J" and were asked "Is a bird a mammal?". Again they raised a YES or NO sign, and then they wrote down the two letters, in the right order (recall component).

Language comprehension was measured with a sentence comprehension test developed to assess comprehension of grammar, (the Swedish version of) TROG–2 (Bishop, 2009). With permission, the test was adapted slightly so that it could be given in the classroom. During the test, students listen to 80 sentences concerning 20 grammatical constructions. The sentences are short and grammatically complex

and the test is considered sensitive for language disorder. The students listened to each sentence, chose the corresponding picture in a booklet and marked it. The adaptation may have made the test somewhat more difficult. Each grammatical construction answered correctly in four cases yields one point. The maximum score is 20. This language comprehension test thus measured listening comprehension on sentence level.

Reading comprehension was measured with SL40, a Swedish norm-referenced test (Magnusson & Nauclér, 2010). The test consists of 40 sentences, each with several pictures to choose from, where one depicts the content of the sentence. It is norm-referenced only up to Grade 5. The reason for using this with older students with HL was that students with HL often fall behind, which made the risk of ceiling effects lower. However, good results on the test are difficult to interpret.

Table 4. Overview of the age, gender and cognitive and linguistic predispositions of the students

	Papers I, III	Paper II
Participants	55 students with NH	11 students with HL
Number of groups for intervention	4	2
F/M	29/26	5/6
Mean age (range)	11:2 (10:9–11:9)	12:11 (12:3–13:8) (class 5–6) 14:1 (13:4–15:3) (class 7–8)
Working memory test from Lilla Duvan, mean (max. 36)	31.7	31.6 (class 5–6) 28.8 (class 7–8)
Language comprehension test TROG–2, mean (max. 20)	16.5	13.2 (class 5–6) 15.0 (class 7–8)
Reading comprehension test SL40, mean (max. 40)	37.5	33.8 (class 5–6) 35.2 (class 7–8)

Age refers to the age when writing Text 1. The test Lilla Duvan is a screening test for reading and writing difficulties (Wolff, 2010), TROG-2 is a test of grammatical sentence comprehension (Bishop, 2009), and SL40 is a reading comprehension test (Magnusson & Nauclér, 2010).

Narrative texts and text quality

The personal narratives were written on the computer. In the observation films, the film peers wrote on the computer, and there were examples of revision within texts. A curricular goal is that students learn to write by hand and on the computer, and thus the students were used to this mode of writing (Skolverket, 2019a). Further, writing on the computer made using a keystroke logging program possible, so that further studies may explore writing processes before and after intervention. The topics of the narratives were chosen so that they would be comparable. Two of the topics had been used in earlier studies, generating texts similarly (Johansson, 2009;

Strömqvist, 1996). The other two were constructed to generate similar texts, in length and elaboration.

There are many ways of measuring text quality. Holistic or global ratings by trained raters have been found valid (Tillema et al., 2012). This method has also been used in earlier observational learning studies (Raedts et al., 2017). In the present project, a panel of six raters was assembled and trained to rate the text quality. They were trained to give a holistic rating based on aspects of genre, structure, organisation, punctuation, grammar, spelling, text length and content. The raters were university students in the social sciences, i.e. with less focus on language than students of Swedish or SLP students. The raters were trained by members of the research group. For reliability, they practiced on benchmark texts. The benchmark texts were texts of the same subjects, written by peers of similar age in a pilot before data collection. These texts were introduced to help the raters answer similarly. Text quality was assessed by marking a VAS scale. Markings were measured and text quality given a number from 0 to 100. Out of the six raters, three rated each text. The texts were distributed among the raters evenly so that all raters would rate texts from the four points in time and so that the same raters would rate as few of the same texts as possible. Apart from this, the texts were distributed at random. The raters did not know that intervention was involved or that some of the students had HL. Descriptive data in the papers show mean values of the text quality, but in the regression analyses, all data points were measured. The interrater reliability of the ratings was strong (Cronbach's $\alpha = .90$).

There was some data loss which was not considered vital. For three students with NH, there were three texts and for all others there were four. One text received only one rating, eight received two ratings, and one received four ratings. One student with HL only wrote three texts and one other text was not rated. Closer descriptions are found in Papers I and II, see also Statistical methods and analyses below.

Intervention lessons

The lesson plan for the intervention lessons could be followed as planned and scripted in all but one case. In that lesson, a slight adjustment from the lesson plan was needed as it took more time than was planned. The same adjustment was made the following times the lesson was given. The group sizes for the reflection and learning steps varied from pairs to groups of up to five students, to gain advantages of both small and large groups. Before each lesson, the students were divided into groups of two to five and the classroom was furnished accordingly. The researchers showed each student to their place to sit directly when they entered the classroom. Thus, time and focus were not lost when reassigning seats. Table 5 shows some examples of advice which the students wrote "for next year's students". Note that

the students had a focus on form (for example spelling and punctuation) but were also displaying reader awareness and text structure on a more global text level ("start with the exciting bit", "a bit funny") after the intervention.

The learning stage, in which students wrote down advice for "next year's students in this class", was done individually or in the groups working together that lesson. The five intervention lessons were 40 minutes each and followed the prepared written scripts. The lessons would be given over three weeks, in place of the regular Swedish lessons.

Table 5. Examples of students' advice for next year's students (Translation and Swedish original text)

Lesson theme	Advice	Swedish original text
1 The reader's perception	Readers like it when stories are spelled correctly and when they are a bit funny.	En läsare vill att en berättelse ska vara rättstavad och lite rolig.
2 Structure	I think you should start with the exciting bit. Otherwise the reader might not want to continue.	Jag tycker att man ska börja med det lite spännande. Annars kanske läsaren inte vill fortsätta.
3 Conclusion	You should think about having a good ending. Not like "...then he died. The end. "But like "I saw him fall towards the ground and I didn't know that was the end... " If you want a sad ending that's good.	Man ska tänka på att det ska vara ett bra slut. Inte till exempel: "...så dog han. Slut." Utan till exempel: "Jag såg honom falla till marken och jag visste inte att det var slut..." Om man vill ha ett sorgligt slut är det bra.
4 Revising someone else's text	Rephrase if needed. You can change said to shouted, if someone is shouting.	Omformulera om det behövs. Du kan till exempel ändra sa till skrek om någon skriker.
5 Online revision	If you forgot something. Like spaces, misspellings, full stop, commas or better sentences.	Om man glömt något. Till exempel mellanrum, stavfel, punkt, kommatecken eller bättre meningar.

Self-efficacy scale

The students filled out the 18 statements of the self-efficacy scale on paper by marking to what extent they agreed with each statement on a 100 mm VAS scale. See Table 6 for a list of the statements. Scoring was done by measuring the 100 mm scale and thus the self-efficacy score could vary between 0 and 100. The statements were adapted for narratives. They concerned different aspects of writing: form as well as content aspects; process as well as product aspects. Examples of form aspects are structure of a text, or spelling, while content aspects concern for example what is relevant to write about the setting. An example of a process aspect is finding letters on the keyboard, while product concerns for example text cohesion and overall structure.

Table 6. Self-efficacy statements

	Statement
1	I can quickly write a text on the computer.
2	I can find all the letters on the keyboard.
3	I can write a text without spelling errors.
4	I can use periods, question marks and exclamation marks in the right places.
5	I can use commas in the right places.
6	I can divide my text into paragraphs.
7	I can write a heading that fits with the content.
8	I can write a story in such a way that the reader understands where the story took place and who was in it.
9	I can divide my text into beginning, middle and end.
10	I can write a story with an unexpected ending.
11	I can write a story in such a way that the reader understands what happened.
12	I can begin the story in an exciting way, so that the reader wants to continue reading.
13	I can change one or several words in my text if I am not satisfied with what I had written initially.
14	I can write both long and short sentences.
15	I can decide which parts of my story need to be explained to the reader.
16	I can read through my text and correct spelling mistakes.
17	I can read through my text and make changes to improve it.
18	I can write a text which is one page long.

Statistical methods and analyses

In the present dissertation, linear mixed effects models (LME) and ANOVA were performed, aside from descriptive measurements. In LME, continuous variables (text quality, results on cognitive and linguistic tests) as well as categorical variables (gender; intervention Y/N, follow-up Y/N; class, school grade) can be used as predictors. LME is considered a robust method of measuring intervention even when there is some missing data, which is suitable in a school context where students are sometimes absent and where data may be lost (Quené & van den Bergh, 2004). These advantages have made LME very common in repeated measure designs as well as in other study designs with hierarchically structured data (Baayen et al., 2008; Casals et al., 2014; Meteyard & Davies, 2020). In paper III, the difference in self-efficacy before and after intervention was measured with repeated measures ANOVA. Aside from statistical analyses, a qualitative approach was taken to describe features of the participants in Paper II.

Inter-rater reliability was measured for all texts together (from students with NH and HL) with Cronbach's α.

Ethical considerations

The research in this dissertation was carried out in accordance with the recommendations of The Swedish Ethical Review Authority and the protocol was approved by the Regional Ethical Review Board in Lund (which preceded the present Swedish Ethical Review Authority), Dnr. 2013/270. The participating students and the film peers and their caregivers gave informed consent in accordance with the Declaration of Helsinki before inclusion. The film peers and their caregivers gave consent to use of the films in the intervention lessons during the research project.

Before any data collection, the project was discussed with head teachers and school board officials of potential schools. Further, the research group visited parent–teacher meetings, before initiating data collection or intervention, to describe the project and answer any questions. As mentioned in the Introduction, in Sweden the individual teacher has the mandate to choose teaching methods in accordance with the policy of that particular school. The head teachers and teachers thus had the authority to allow the research group to implement the intervention and perform data collection in the classes, replacing other lessons in the subject of Swedish. Originally, data collection from schools with no intervention was planned, but school officials did not agree to data collection without intervention. Thus, the students in the participating schools all received intervention, with different start times in relation to the intervention.

The Ethical Review Board recommended that the whole class should be included in data collection and intervention, in order not to single out any student. All students who ordinarily took part in classroom activities were thus present for the data collection and intervention and could not opt out of the lessons. Data from non-participating students were not saved or analyzed.

The small number of students with HL has been considered. The students in question have been described case by case in some respects, and some audiological data have been excluded to preserve their anonymity. Further, neither locations of schools nor the year of data collection have been revealed.

Results

The following paragraphs briefly describe the main results of each paper. The four main research questions of this dissertation are addressed in the summary of results.

Paper I

In Paper I, the aims were to develop the observational learning intervention and to evaluate its effects on text quality, considering the students' individual predispositions. The participants were fifty-five Grade 5 students with NH.

Research questions:

(a) Do narrative skills as measured by text quality improve as a consequence of the intervention?

(b) Do any effects last?

(c) Can WM capacity, language comprehension, reading comprehension, or gender predict text quality or intervention effects?

There was a mild, but statistically significant positive intervention effect on text quality. This effect had diminished at follow-up after four months and was no longer significant. Higher WM scores were linked to higher text quality ratings but there were no effects of language comprehension or reading comprehension. There was a clear effect of gender, showing considerably higher text quality ratings for girls' texts. While gender and WM predicted text quality in the regression, neither interacted significantly with intervention. Students with better reading comprehension did however derive significantly more benefit from the intervention. This constituted the only significant interaction in the mixed effects regression analysis.

Paper II

In Paper II, the aims were to investigate whether the observational learning intervention was feasible to use in classes with students with HL and to study the students' cognitive, linguistic, audiological, and demographic factors in relation to narrative text quality. Eleven students with HL, in Grades 5–6 and 7–8, participated in the study.

Research questions:

> (d) Is observational learning suitable for the training of narrative text writing in students with HL?
>
> (e) What are the effects of WM capacity, language comprehension, reading comprehension, school grade, and gender on narrative text quality in students with HL?
>
> (f) What associations are there between degree of HL, age at amplification and text quality?

The observational learning intervention was performed in two classes for students with HL, in school years 5–6 and 7–8, respectively, and found to work well in the classroom. Text quality did not change significantly over the course of the four texts. Some students, mainly girls, received slightly higher text quality ratings over time and some, mainly boys, slightly lower. Regression analysis did not show significant effects of WM capacity, language comprehension, or reading comprehension on text quality. Text quality varied considerably between students. Grouping the students by school class and gender showed significant differences in text quality. Girls in Grade 7–8 had the highest text quality, followed by girls in Grade 5–6. Boys in Grade 5–6 followed and boys in Grade 7–8 had the lowest text quality results. There was considerable individual variability between students regarding cognitive and linguistic factors, hearing sensitivity and time factors. Inspection of individual results showed that the three students with the highest text quality had lower age at amplification (at age 1–4) than the three students with the lowest text quality (amplification at age 6–10). The great heterogeneity of students with HL suggests additional individual adaptation in the further development of the observational learning intervention.

Paper III

In paper III, the aims were to explore self-efficacy for narrative writing before and after an intervention which may affect motivation. While Paper I showed differences in narrative text quality for boys and girls, but similar intervention effects, this paper explored associations between gender, text quality and self-efficacy. The participants were the same 55 students in Grade 5 with NH as in Paper I.

Research questions:

(g) Does self-efficacy for narrative writing change after the intervention, and if it does, in what way?

(h) Are there any gender differences?

(i) Is self-efficacy related to text quality, and if so, in what way?

Self-efficacy increased significantly after the intervention. Boys and girls had similar self-efficacy, despite the significantly higher text quality ratings of girls' texts. There was an association between self-efficacy and text quality pre- as well as post-intervention. Internal consistency within the self-efficacy scale was strong.

In conclusion, students' strong self-efficacy for narrative writing supports previous findings. The observational learning intervention may have reinforced students' self-efficacy. Gender may influence self-efficacy and how statements in self-efficacy scales are answered. Text quality ratings and self-efficacy had positive, moderately strong correlations before as well as after intervention. Students may not be able to differentiate between self-efficacy, general writing skills, task performance, and self-regulation at this age. Self-efficacy scales should thus be carefully constructed with respect to students' age, genre, instruction, and general educational context.

Summary of results

The five-lesson observational learning intervention and the self-efficacy scale was developed and implemented in classes for Grade 5 students with NH as well as in classes for students with HL in Grades 5–8. The observational learning paradigm was new to the students. The lessons were well received by the students. The intervention had a significant, positive effect on text quality for the group of students with NH. At a follow-up after four months, the effect was no longer significant. For the group of students with HL, there were no significant changes in text quality over time. In the students with NH, higher WM scores were associated with higher text

quality ratings. Further, students with NH with higher scores on the reading comprehension test had stronger intervention effects. In students with HL, there were no overall effects of cognitive and linguistic factors. The students with HL with the highest text quality were amplified earlier than students with low text quality. For all students, whether with NH or HL, female gender was associated with higher text quality ratings. Self-efficacy for writing increased significantly after the intervention for students with NH. Students with high text quality ratings had strong self-efficacy for narrative writing. Boys and girls had similar self-efficacy, despite the significantly higher text quality ratings of girls' texts.

Discussion

The four research questions are discussed in the General discussion. The first section addresses development and implementation of the intervention and the self-efficacy scale. The second section concerns text quality and text quality changes. In the third section, cognitive and linguistic factors, gender, and, for participants with HL, school grade and audiological factors are discussed. In the fourth section, self-efficacy and changes in self-efficacy are addressed. Under Methodological considerations, some strengths and weaknesses of the methodological choices in this dissertation are discussed. Finally, some future directions for method development and research are suggested.

General discussion

Development and implementation

Prior to the studies in this dissertation, observational learning intervention had not been implemented in the Swedish school context, to the knowledge of the author. Self-efficacy scales for writing are not commonly used in the Swedish school context. The self-efficacy scale was developed for these studies and was thus another new lesson feature for the participating students. Teaching students why and how to give and receive peer feedback is a curricular goal (Skolverket, 2019a, 2020a). Students may for example exchange texts, read and give feedback on a peer's text before revision. Swedish students are thus somewhat familiar with reflection on texts and revision. The structured group reflection on writing processes which was an integral part of the intervention, was however new to the students.

The first research question was whether it was feasible to perform the intervention during regular Swedish lessons in the age groups. The manual was easy to follow and students participated willingly in the intervention. Teachers and head teachers were interested and engaged in the process. As mentioned in the Method development chapter, some teachers who took part in the observational learning intervention training chose to use it in their own classes the subsequent school year, as a method of developing students' writing. Self-efficacy scales may be valuable tools to use in Swedish schools since self-efficacy and other aspects of motivation

are important predictors for goal achievement (Callinan et al., 2018; De Smedt et al., 2018; Harris et al., 2006; Limpo & Alves, 2017; Raedts et al., 2007; Schunk, 1991).

Several adaptations were made in order to make the intervention and the self-efficacy scale appropriate for the Swedish school setting and the chosen age group. Adaptations to the intervention were made to suit younger students, as earlier studies have often focused on high school or undergraduate levels (Braaksma et al., 2002; Couzijn, 1999; Raedts et al., 2017; Rijlaarsdam et al., 2008). The genre was thus changed from expository and argumentative writing, which earlier studies have focused on, to narrative writing, which the younger students in the present project were more used to and which suited the curriculum. Lastly, to include measures of language comprehension and reading comprehension in an observational learning study is also a novelty. So is the inclusion of a measure of WM capacity. Although alleviation of cognitive resources is seen as a possible reason for the efficacy of observational learning, WM was, for example, not studied in the papers which inspired the design of the present studies (Braaksma et al., 2002).

Text quality before and after intervention and at follow-up

Paper I showed mild, but significant intervention effects for students with NH. Paper II was not designed as an effect study, but descriptive data of text quality before and after intervention showed no differences in text quality in students with HL from pre- to post-intervention or to follow-up. The intervention effects for students with NH were no longer significant at follow-up. Positive intervention effects on text quality of observational learning were found in comparable earlier observational learning studies but they did not include a follow-up, and long-term effects are thus not clear (Couzijn, 1999; Raedts et al., 2007; Rijlaarsdam et al., 2008). Further, the students in the earlier mentioned studies were older than the students with NH in the present dissertation (Couzijn, 1999; Raedts et al., 2007; Rijlaarsdam et al., 2008). This may have affected their ability to generalize from observations of what peers were thinking and doing to their own writing.

Between data collection of Texts 3 and 4, the researchers did not have any contact with the students, and their teachers did not have specific information on the intervention. There was thus no reinforcement of knowledge related to the intervention which could be important for sustained skills. The follow-up texts (Text 4) were written in the semester after the intervention. For students with NH Text 4 was written during the spring semester, and for students with HL, it was written after the summer vacation. Several studies show that many students experience "summer loss" of knowledge and skills in various subjects after the summer vacation (Cooper et al., 1996; Rosqvist et al., 2019).

The lack of improved text quality after the observational learning lessons in students with HL in Paper II indicates that further development is needed before evaluation and implementation of observational learning intervention for students with HL. Students with HL receive degraded source signals not only due to the HL per se but also due to the linguistic limitations they often demonstrate. If the acoustic environment is non-optimal, speech perception and language comprehension are further challenged (Mattys et al., 2012).

Students with vulnerable listening skills or other processing limitations, for example students with HL and/or LD have even more pronounced needs for effective writing instruction, as fewer teaching methods have been found effective in clinical groups. Structured intervention where students may focus on one thing at a time has been found effective for students at risk for LD (Wolbers et al., 2018; Wolbers et al., 2015). Observational learning should therefore be further explored and developed. Teaching writing is an important element of teacher training and continued professional development. SLPs in school settings as well as teachers may use observational learning as one tool to strengthen students' writing skills. Implementation requires time and continued collaboration with and support from researchers (Ebbels et al., 2019; Markussen-Brown et al., 2017), or between teachers (van Schaik et al., 2019).

Despite amplification and acoustic accommodations in the classroom, the students with HL spend time in adverse listening conditions outside classrooms. They experience listening effort and need a structured education context. A possible improvement to the observational learning lessons which may increase effects could be more explicit instructions about the film peers' activities. For example, the teacher could offer linguistic and metalinguistic explicit teaching and scaffolding by asking students to point out when film peers set goals or instruct themselves. Such explicit teaching strategies have rendered effects in students with HL (Wolbers et al., 2018).

Influencing factors and predictors of text quality

For all students (NH and HL), gender was a significant predictor of text quality, girls receiving higher text quality ratings than boys. The effect of gender was greater than any intervention effects or other possible predicting factors in the mixed effects regressions of students with NH and HL in Papers I and II. This is not completely unexpected as girls generally do better in school in Sweden (Skolverket, 2019b). Studies of writing have sometimes shown gender effects favoring girls (Kanaris, 1999). Other studies have shown no gender differences, but this may depend on the included participants. For example, Johansson (2009) found no gender differences but the students with reading and writing difficulties, mostly boys, were excluded in that study.

Students with NH with higher scores on the reading comprehension test had somewhat stronger intervention effects. A possible explanation is the Matthew effect, i.e. the cumulative advantage of early reading skills which lead to continued reading success. This is recurrent in literacy studies. Students who read more may have more experience with narrative structure and written language.

In students with HL, school grade turned out as another important factor for text quality. Girls in Grade 7–8 received higher text quality ratings than girls in Grade 5–6, followed by boys in Grade 5–6, and lastly by boys in Grade 7–8. The students with HL represented a more diverse group as to age, audiological, linguistic and cognitive factors compared to the students with NH. Socio-economic status (SES) was not investigated in this study. Since the classes for students with HL received students from an entire region, SES may have been more diverse among the students with HL than among the students with NH, whose schools were located in areas with high SES. Results from PISA show that the gap in reading comprehension, mathematics and science knowledge widens between Swedish students from homes with high and low SES (OECD, 2016, 2019).

The finding that boys in Grade 7–8 had lower text quality than boys in Grade 5–6 is unexpected. A possible explanation is school placement changes. Students with HL who struggle in mainstream classes sometimes move to classes which are exclusive for students with HL as they get older (SPSM, 2008). With such a small number of participants, individual differences may render high impact in analyses.

The students with HL with the highest text quality ratings were two bimodally aided girls (using one HA and one CI). It may be worth noting that children with CI in Sweden receive considerably more rehabilitative support than children with HA. As an example, all children with unilateral or bilateral CI are followed regularly with checkups and have access to an SLP. This is sometimes not the case for children with HA, at least those with mild/moderate HL who, in many cases, never meet an SLP (Rosén et al., 2019), despite a high risk for LD in this group. The support from the healthcare system for students with CI may positively affect school motivation and performance, when comparing to students with HA. None of the other predictors explained differences in text quality in students with HL. The small group of eleven students in Grades 5–8 indicates that all results must be interpreted with great care. A lack of significant differences is not unexpected, as students with HL represent a very heterogeneous group (Marschark & Knoors, 2018b).

Regarding predicting factors, WM capacity turned out to predict text quality in students with NH but not in students with HL. It has been suggested that observational learning alleviates cognitive load compared to traditional writing exercises, since learning is separated from the writing task load (Rijlaarsdam et al., 2008). Writing is considered highly dependent on WM (McCutchen, 2011). Possibly, the students with NH were still working on automatizing transcription and other lower-level writing skills, thus spending WM capacity (Drijbooms et al.,

2015) and leaving less resources for higher-level skills such as finding the right words, organizing the story and making revisions. The somewhat older students with HL may have progressed further in this automatization, which could decrease their WM load during writing. WM capacity results did not modulate intervention effects, which may mean that students with different WM capacity received the intervention similarly. While WM may be alleviated during the observational learning intervention, freeing cognitive capacity for learning, WM will still be taxed during writing. More automatized low-level transcription skills predict text quality in students with NH as well as with HL, possibly because WM capacity can be freed for higher-level writing skills as planning and organization of the text (Arfé et al., 2014; Drijbooms et al., 2015; McCutchen, 2000). Verbal WM capacity was found to contribute more to text quality than age or reading comprehension in written narratives in elementary school students with severe or profound HL (Arfé et al., 2015).

The test scores of reading comprehension (Magnusson & Nauclér, 2010) and language (listening) comprehension (Bishop, 2009) were not significantly associated with text quality in this study. When examining the results of students with HL, the students with the lowest text quality ratings did however have low results on the language comprehension test.

Self-efficacy change, text quality and gender

As shown in Paper III, students' previously strong self-efficacy increased post-intervention. By observing peers and reflecting on their own and together with classmates, students may have increased their belief in their capabilities to perform the writing task described in the instructions to the scale (see Method development). The described task was similar to the personal narrative tasks which they were given. Returning to the sources of self-efficacy posited by Bandura (1986), the intervention lessons may have affected three of them (*vicarious experience, social persuasion,* and *physiological and emotional state*). First and foremost, students gained *vicarious experience* by observation of peers in the film clips. Additionally, the structured reflection on the film clips and writing down advice for another peer can be seen as a form of *social persuasion*. To discuss film clips and to write down advice may have "persuaded" the students that they had important, useful knowledge of narrative writing. Observing peers who are working both on their own and on others' texts may make students feel more at ease about the possibility of revising a text and how to do it. This may contribute to an improved *physiological and emotional state*. Observation of peers may thus be a valuable way of increasing students' motivation for writing.

The students were asked to indicate with marks on the VAS scales the extent to which they agreed with each self-efficacy statement, and most often marked the scales between "average" and "completely". Their self-efficacy was thus 'strong'

rather than ‘weak’. Strong self-efficacy at this age is in concurrence with earlier research (Bruning et al., 2013a; Pajares, 2003). At the age of the students (around 12 years) however, self-efficacy starts decreasing and becoming more diversified between domains. This decrease may be due to calibration to achievement (Muenks et al., 2018; Stipek, 2002). However, performance should improve with age, even if demands also increase. One explanation may be that self-efficacy is linked to self-esteem which decreases during adolescence.

Self-efficacy was significantly correlated with text quality before and after intervention. Further, internal consistency within the scale was high, meaning that no statement was judged very dissimilar to the others and suggesting that the students had a holistic sense of self-efficacy for the domain of narrative writing. Despite the obvious differences in text quality ratings of boys and girls, where girls’ texts were assessed as having better quality, boys and girls reported similar self-efficacy. This did not change after intervention. Similar results were found by Pajares and colleagues in a series of studies (Pajares et al., 1999; Pajares & Valiante, 1999, 2001). Although girls reported similar self-efficacy to boys, they often found that they were better writers than boys when comparing texts (Pajares et al., 1999; Pajares & Valiante, 1999). The authors interpreted this not as differences in students’ true self-efficacy, but as differences in how self-efficacy is reported, depending on gender or possibly gender orientation.

Methodological considerations

In the following, some aspects of the method and method development are discussed.

In Paper I, the two *wait first* classes were in one school and the two *intervention first* classes were in the other. An alternative would have been one class in each school following the *intervention first* path, to control for possible differences between the schools. This was decided against, as students would perhaps have discussed intervention between classes, which would have confounded the *wait* condition. In Paper II all eleven students followed the *intervention first* path. As the participants were few and the group heterogenous in age and as to audiological factors, the results must be interpreted with caution.

It was possible to give all lessons according to the same script in all classes. Similar reflections after observation of the films were made in all classes, whether NH or HL students. This is interpreted as sufficient reliability of the scripts of the intervention lessons. An alternative to researchers performing the intervention would have been training the students’ own teachers to perform the intervention, which could pose reliability issues. Some teachers would know their students better than other teachers or have more knowledge of students which may affect

prompting. Different teachers could unknowingly add or subtract from the method due to their previous knowledge or believe more or less in the efficacy of the intervention. These possible confounding factors may have resulted in the stronger intervention results in the parallel study of teachers using the intervention method the following school year.

The film peers were comparable in age (10–13 years) to the majority of the participating students, and thus probably in language and writing skills too. Seeing models with whom one can identify is an important motivational factor (Raedts et al., 2007; Raedts et al., 2017) but the extent to which the participants identified with film peers was not studied. Some students with HL were older than all film peers. It is possible that older students found film peers childish and not motivating as peers.

The two schools for students with NH were in one school district. Several head teachers in the school district were approached, and the head teacher at each school decided whether the intervention could be carried out in their school. The two schools were located in similar, socio-economically strong areas and a smaller portion of students were bilingual compared to many Swedish schools. The students in this dissertation did therefore not represent a cross-section of Swedish students regarding linguistic background and SES. There are associations between linguistic abilities and multilingual background, where results of formal tests measuring language skills favor monolingual students, but the school district contributes considerably more to differences between mono- and bilingual students than bilingualism itself (Rosqvist et al., 2019).

The intervention was developed for Grades 4–6. Grade 5 was chosen for practical reasons in discussion with school management and teachers. Many of the Swedish lessons in Grade 6 are used for national assessments as mentioned before and there would thus be less time for implementing a new intervention. Some teachers advised against implementation in Grade 4 as the students could be too young to have the full benefit of the method. Considering that observational learning has been developed for older students before, the intervention could be better suited for students in Grade 6. The reflection phase may be more beneficial to students with more developed metacognitive abilities, which could influence the whole group.

Performing researcher-led intervention in schools is challenging. The close collaboration with school management and personnel is time-consuming but crucial for optimal effects (Ebbels et al., 2019). However, even with careful planning and information to all parties, schedules are vulnerable, for example to sports days and national assessments, which may influence timing. To meet these challenges, the research group prepared meetings with school management, headmasters, and teachers to explain the theory, method, and design and took part in parent–teacher meetings.

Verbal WM capacity was measured with a subtest from a screening test for reading disabilities ‘Lilla Duvan’ (Wolff, 2010) developed for classroom screening. It was

carried out in the classroom, according to the test manual. As it is an auditory task, students' hearing sensitivity may have affected results. Language comprehension was measured on sentence level with a test of receptive grammar (Bishop, 2009). As in the WM task, results may have been confounded by poor hearing sensitivity. Hearing was not measured in students with NH. The results on the language comprehension test reflect the ability to understand short, grammatically dense sentences. To measure full language comprehension, one should ideally measure comprehension of words, clauses, longer utterances and inferencing skills. The reading comprehension test was norm-referenced for the students with NH. The results from students with HL could be compared to the reference group, but high scores were expected. Thus, the test may not have identified all students with reading difficulties.

Intervention effects can be measured in many ways. Here, they were measured by text quality, which may in turn be measured in different ways. Holistic text quality ratings based on benchmarking texts have been used in earlier observational learning studies and have been validated with sufficient inter-rater reliability (Tillema et al., 2012). This was also found in the present studies.

When the research resulting in this dissertation started there was no appropriate existing self-efficacy scale for narrative writing fitting the criteria suggested by Bandura (2006). A scale by Braaksma and colleagues was used as a point of departure when developing the scale. When adapting the scale, one change was to use a VAS scale instead of writing down a number. This was to make the scale easier to understand for a younger age group. A Swedish self-efficacy scale for reading and writing was available (Wolff, 2010). The statements in Wolff's scale were answered by choosing one of four options and they did not begin with "I can..." (as the statements in the present dissertation do). It was thus not an appropriate scale according to Bandura's guidelines (2006). It may have lower predictive value than answering "I can"-statements by marking a VAS scale or writing a number 0–100. The self-efficacy scale for writing (SEWS) was not published when the scale in this project was being developed but might have been an alternative (Bruning et al., 2013b). However, self-efficacy should be measured in relation to the skill in question, and comparisons between studies may be difficult to do. A possible drawback of how the scale was presented, i.e. by the researchers performing the intervention is that students may have been prone to marking the scales more positively than if the scale had been distributed differently.

Finally, the linear mixed effects model (LME) is commonly used in studies with repeated measures including studies in education (Meteyard & Davies, 2020), speech-language pathology (Quené & van den Bergh, 2004) and medicine (Casals et al., 2014). Despite its usefulness, there are concerns about how to report analysis for replicability (Meteyard & Davies, 2020). ANOVA was used in Paper III as there was less data to consider. The number of participants in Paper II was small, and the group heterogeneous. This is a recurring issue in clinical groups. Looking closer at

the characteristics of those who do well and those who struggle may be one way of studying effects qualitatively in a heterogeneous group (Marschark & Hauser, 2008). Therefore, Paper III described the students with the highest and lowest text quality ratings.

Future directions

The results in the present dissertation raise some questions about the role of observational learning for writing in elementary school. Further studies are needed to explore a range of issues such as the optimal length of an intervention, whether more explicit instruction and feedback on texts from teachers and peers could improve effects and whether intervention should be led by the teacher or someone else. For students with HL, controlled effect studies of observational learning are called for.

Students' development of adequate writing skills takes time, effort and motivation. Teaching methods must therefore be effective. The fact that positive effects of this structured five lesson intervention were found is important from an educational point of view, but only if students remember what they learned over time. An explanation for the lack of long-term effects may be that the intervention had too few lessons for students to recall what they learned without revisiting it later during the semester. Recurring elements of observation and reflection, or reinforcement by the teachers on what the students have learned, may make the students more prone to reflect on what other students do. An even closer collaboration with the schools during method development and preparations, for example focus groups with teachers may improve intervention lesson content and acceptance in the schools even more. Working closely with the school has been found an effective approach for improving outcomes but it requires considerable time from the teachers and support from school management (Ebbels et al., 2019; van Schaik et al., 2019).

The observational learning paradigm may in future studies be combined with other structured instruction methods. Feedback is one recurring effective element of writing intervention (Graham & Harris, 2018) and may be seen partly as *social persuasion* which also may increase self-efficacy (Bandura, 1997). Receiving feedback by observing the effects on a reader is another way of improving text quality, as shown by Couzijn (1999) and Rijlaarsdam et al. (2008). The students could observe peers who read and react to the students' texts.

Self-efficacy for writing and other aspects of motivation may be further explored for research and as an educational tool. Observational learning intervention may affect motivation differently depending on whether and how students identify with film peers. Self-efficacy has sometimes been found to decrease over time, for example in transitions such as changing schools (Schunk & Meece, 2006).

There may be intervention effects in the writing process which are not detectable in text quality ratings. Writing processes can be measured with keystroke logging data on for example pauses and revisions. Data from the master's thesis on students with NH in Grade 5 (described in Method development) show that after intervention, some students used less time for pauses (Hammarstedt & Karan, 2016). Further, some of the students made a higher proportion of deep revisions (e.g. finding synonyms) relative to superficial revisions (e.g. correcting typos) after intervention. The authors also showed that deep revision was not always linked to higher text quality ratings. Emergent improving strategies are thus not always apparent in a finished text. Future studies on observational learning intervention should therefore explore effects on writing processes.

Conclusions

This dissertation aimed at developing and evaluating an observational learning intervention and a self-efficacy scale for supporting narrative writing in elementary school students with NH and with HL. The intervention was well received by the students and could be implemented in Grade 5 for students with NH and in Grades 5–8 for students with HL. For students with NH, there were positive effects of intervention on narrative text quality but the effects were not sustained. Students with higher scores on reading comprehension had more benefit from the intervention. For students with HL, text quality ratings did not increase. The students with HL with the highest text quality ratings were younger at amplification than the students with HL with the lowest ratings, indicating that early diagnosis and intervention promotes literacy development. Gender was the most important predictor of text quality for all students. Self-efficacy ratings (students with NH) were stable between different statements and increased after intervention, indicating that the observational learning intervention increased the students' beliefs in their capabilities.

This dissertation has demonstrated that an observational learning intervention for writing and a corresponding self-efficacy scale could be developed and implemented in Swedish elementary school classes for students with NH in Grade 5 and for students with HL in Grades 5–8, as a result of careful planning and development partly in collaboration with schools. It was possible to do so in line with the Swedish curriculum and to adapt it to students' school grade and age.

The observational learning intervention had a positive effect on text quality for students with NH. The students with higher reading comprehension scores had a stronger effect of intervention. Students who are better readers often read more and may thus have more experience of narrative texts. Young students and other students who have less reading experience may benefit from more explicit instruction methods than was offered in the observational learning intervention in the studies described in this dissertation. Observational learning may complement other strategy instruction and other structured writing interventions supporting young writers.

The positive effects of intervention were not sustained over time. Measuring effects over time is important in intervention research (Graham & Harris, 2014). The reinforcement of new skills may be important for sustained effects, especially for

students with poorer language skills. For the students with HL with the lowest text quality ratings i.e. the boys in both Grade 5–6 and Grade 7–8, text quality ratings were lower for texts written at follow-up than the earlier texts, although the effect was not statistically significant. The students with HL wrote the follow-up text after summer vacation. This supports research which has found that students with weak language skills are more at risk for "summer loss" of language skills after a long vacation (Cooper et al., 1996).

There were no indications, when individual cases were inspected, that more severe HL was associated with poorer text quality or text quality changes. This corroborates earlier studies of HL and language skills difficulties (Marschark & Knoors, 2012, 2018a; Tharpe, 2008; Yoshinaga-Itano et al., 1998). On the contrary, some students with profound HL received some of the highest text quality ratings. These students had been diagnosed with HL earlier than the students with the lowest text quality ratings. A conclusion of this could be that the quality and quantity of linguistic input for children with HL influences language skills years later. The results thus underline the importance of early diagnosis of HL for optimal audiological, linguistic and pedagogical support.

Cognitive and linguistic factors, which were measured with tests of WM capacity, language comprehension and reading comprehension, contributed to some extent to text quality in students with NH. For students with HL, age (school grade) influenced text quality. Performing the intervention in classes with slightly older students may result in stronger effects. Students with weaker WM, for example young students, may benefit from more explicit instruction than was offered in the observational learning intervention in the studies described in this dissertation. A greater number of lessons is another possible way to develop the method further. Observational learning may complement other strategy instruction methods and other structured writing interventions supporting young writers.

Gender is an important factor influencing writing, shown by the fact that girls with NH and with HL received higher text quality ratings than boys did. Despite this, boys had similar self-efficacy to girls with NH. The boys in this study may have had self-efficacy beliefs for writing which were slightly too high for optimal development (Bandura, 1997). Knowledge about students' self-efficacy ought to be of high interest for teachers. Writing motivation may increase by observing and identifying with models. Self-efficacy scales should be adapted for the specific context, taking age differences, educational systems, and genre into account.

Whether teaching students with, or at risk for LD or not, the saying 'one size does not fit all' must be emphasized. Thus, adapting, evidence-basing and implementing new teaching practices as well as evidence-basing current methods is important in improving the teaching of writing and strengthening language outcomes. To implement new evidence-based instruction methods in the classroom setting, school management should provide teachers with plenty of support and time for planning

and collaboration with the professionals who are introducing the methods (Ebbels et al., 2019; Markussen-Brown et al., 2017 van Schaik et al., 2019). For sustained effects in groups who have or are at risk for LD, continuous support after initial implementation is needed (Ebbels et al., 2019). The continuous support of students' writing development is a question of facilitating their academic achievement in school and their participation in and contribution to a sustainable society.

Tack

Ett stort tack till er som har gjort detta avhandlingsarbete möjligt att genomföra. En del har varit direkt involverade i arbetet och andra har stöttat mig utan att vara delaktiga i själva forskningsarbetet.

Till att börja med vill jag rikta ett stort tack till **Stiftelsen Marcus och Amalia Wallenbergs minnesfond** som finansierat projektet.

Tack till er **elever** som deltagit i lektionerna och till era **vårdnadshavare** som har hjälpt till med information. Tack till **lärare, rektorer** och **annan personal** som gett oss möjlighet att komma ut till era skolor. Tack till er **elever** som kom till oss och lät sig bli filmade i skrivaktiviteter och tack till era **vårdnadshavare**. Tack till **audionomer** som samlat och skickat information om eleverna med hörselnedsättning.

I would like to express my gratitude to **Dr Martine Braaksma** and **Prof. Huub van den Bergh**, both involved in several of the described observational learning and self-efficacy studies, for your generous advice on design, analysis and evaluation and for access to your self-efficacy scale for argumentative writing.

Tack till **NORPCI (Nordic Researchers in Pediatric Cochlear Implants)**, till **Research in children with nonsyndromal and syndromal hearing impairment/deafness with hearing aids or cochlear implants** och till Linnémiljön **Thinking in Time: Cognition, Communication and Learning**, för givande samtal. Tack till **Humanistlaboratoriet** som vi använt vid inspelning av filmeleverna. **Ulf Södergren** – tack för klippning och textning och tack **Johan Frid** för hjälp med programvaran ScriptLog.

Tack **David Whitling** för din säkra språkgranskning av flera artiklar och kappatext. Du har förbättrat läsbarheten mycket – alla nuvarande otydligheter har jag lagt till själv.

Mina handledare – ett stort tack! **Birgitta, Victoria, Joost, Tina** och **Lena** – ni har bidragit med så mycket av era kunskaper och förmedlat era färdigheter under vad som ibland har varit mycket tålamodsprövande omständigheter.

Viktoria, ett stort tack för allt ditt arbete i början av projektet. Tänk så mycket papper vi burit och så mycket kaffe och bussar och externa hårddiskar som vi brukat och förbrukat.

Mina doktorandrumskompisar förr och ”nu”, **Ida, Suvi, Karolina, Sebastian, Susanna, Ketty** och **Olof**: tack för samvaro och support inom alla möjliga områden – statistikprat, gemensam frustration över den ständiga Reviewer 2, pausgympa, trams och skoj, näsdukar laget runt och en gång akutleverans av inneskor och godis till BUS.

Helena – tack för gemenskap i humor, fuldans med mycket mera.

Andra **doktorandkollegor**, **kollegor** och **arbetskamrater**, nuvarande och tidigare, som varit viktiga på olika och ibland oväntade sätt, bland andra, men inte begränsat till, Jonas B, Kristina H, Sanna H, Anders J, Johan M och Susanne S – tack till er!

Ett stort tack till mina **vänner** - jag ser fram emot husmorssemester och honkräftor.

Tack också till **Anna A, Eivor E, Eva H** och **Ann K** för hjälp och stöd i helt olika kapaciteter.

Daniel, **Cajsa** och annan **familj och släkt** som jag ibland knappt träffat men som finns där ändå – tack! Jag ser fram emot att träffas på släktträffar som verkar bli mer och mer skolutvecklingsrelaterade varje gång. Särskilt tack Cajsa som läst den svenska sammanfattningen från mellanstadielärarperspektiv.

Signe, som hjälper mig fokusera på annat än jobbet, tack! Du är en så fin människa att vara med, helt objektivt sett, och du lär mig så mycket. Nu är min bok färdig, för det måste inte bli ”pafekt”!

Ett stort stort tack till **mamma Marie** för all uppmuntran (verbal, livsmedelsbaserad med mycket mera), alla utflykter med Signe för att ge mig extra skrivdagar och all möjlig markservice och för många insikter om hörselnedsättning.

Slutligen hade jag önskat att jag hade haft möjlighet att säga tack till min **pappa Jan**, som betytt mycket för mig och för det här arbetet men även för andra inom hörselområdet. Så många samtal vi haft om öron och hörsel (som ganska liten ville jag vid nattning helst höra om dagens operationer) och så många gånger jag fått testa nya audiometerfunktioner. Jag tänker också på min **farmor Brita** som gav mig en liten bok om ”Skriftens historia under 5 000 år” för väldigt många år sen och på min **mormor Ingela**, den virtuosa berättaren och även skoljournalisten.

Sammanfattning

Lärande genom observation för att utveckla skriftlig berättarförmåga hos elever med eller utan hörselnedsättning – textkvalitet och självskattad förmåga

Inledning

Skrivandet genomsyrar vårt samhälle, oavsett intressen eller yrke. På arbetet, i vårt sociala liv och inte minst i skolan behöver vi goda skrivfärdigheter. Man har också visat att goda skrivfärdigheter ger starkare skolprestationer. Alla elever behöver således stöd i sin skrivutveckling genom hela skolan. Ännu mer stöd behöver elever som har språkliga svårigheter, t.ex. barn med hörselnedsättning eller språkstörning. Skrivfärdighet är sammansatt av många språkliga förmågor (t.ex. lexikal förmåga, grammatisk förmåga, språkförståelse och läsförståelse). Det krävs att vi håller många saker i huvudet och mentalt kan röra oss mellan dem.

Studier från bl.a. Nederländerna har visat att elever kan stärkas i sin skrivförmåga genom att titta på hur (filmade) jämnåriga kamrater gör olika skrivaktiviteter. På så sätt kan de få insikter om hur de själva gör – och skulle kunna göra – när de skriver. Att se kamrater klara av skrivuppgifter kan göra att elever får ökad tilltro till sin egen förmåga att klara av liknande uppgifter. När ett mål verkar möjligt att nå ökar motivationen att jobba för det.

I denna avhandling undersöker och utvärderar jag en skrivundervisningsmetod, *lärande genom observation* (observational learning) för att se om den stärker elevers skriftliga berättarförmåga. Jag har undersökt elevers textkvalitet och deras *självskattade förmåga* (self-efficacy) före och efter en lektionsserie (intervention). Jag har även undersökt deras arbetsminne, hörförståelse och läsförståelse. För eleverna med hörselnedsättning har jag samlat in data om deras hörsel. De normalhörande eleverna gick i skolår 5 och eleverna med hörselnedsättning gick i skolår 5–8.

Språkutveckling

Barns språkutveckling beror på bland annat socialt sammanhang, allmän kognitiv förmåga, motorisk utveckling, social förmåga och förmåga att uppfatta

sinnesintryck. Hörseln är ett sinne som utvecklas redan i fosterlivet och spädbarn reagerar mer på tal än på andra ljud. Språket är sammansatt av dels grundläggande färdigheter som kunskap om ljudsystemet, böjning av ord, ordföljd och ordförråd, dels komplexa färdigheter som berättande i tal och skrift, samtal och lyssnande. En god språkförståelse (förståelse av ord, grammatiskt innehåll, hela yttranden) är viktigt för den fortsatta språkutvecklingen, berättarförmågan och läsförståelsen. Bristande språkförståelse är en riskfaktor för språkstörning. Förmågan att berätta kräver bland annat att man kan tänka och tala om dåtid, nutid och framtid och förståelse för vad lyssnaren vet och behöver veta. Berättandet är således en krävande aktivitet som utvecklas under uppväxten. Barns berättarförmåga så tidigt som i förskoleåldern är kopplad till deras läsförmåga och skolresultat många år senare.

Skrivande i skolåldern

Skrivande kräver *planering* (vad ska jag skriva om, i vilken ordning, i vilken sorts text), *överföring* (från tanke till ord i huvudet, till nedskrivna ord) och *granskning* (läsning och revidering av det man har skrivit). Skribentens fokus förflyttas hela tiden mellan dessa olika processer. Ju mer tankearbete som krävs för de ”lägre” processerna, såsom att forma eller hitta bokstäver, stava rätt och sätta punkt och frågetecken, desto mindre utrymme finns det kvar för ”högre” processer, som att planera texten, att anpassa sig till läsaren, att skriva lagom långa meningar och att läsa och förbättra sin text. En ung elev som just lärt sig skriva kanske lägger ner lika mycket av sin kapacitet som en äldre elev, men den äldre eleven kan hålla mer i huvudet samtidigt och har också automatiserat några av de ”lägre” processerna.

Skrivundervisning

Några undervisningssätt som man vet fungerar för att förbättra skrivfärdighet är att låta elever se hur någon annan skriver och göra efter, att låta eleverna hjälpas åt att skriva och att lära sig särskilda strategier för skrivande. Att öka elevers motivation genom att exempelvis ge dem mer tilltro till sin egen skrivförmåga fungerar också för att förbättra skrivfärdighet.

Lärare inom den svenska skolan har fått lära sig olika mycket om skrivundervisning under sin utbildning beroende på exempelvis när och vid vilket lärosäte de har utbildat sig. Olika kommuner och skolor använder också olika undervisningsmetoder och fokuserar lärarnas kompetensutveckling på olika saker. Detta medför att det kan finnas stora skillnader i vad elever får lära sig om skrivande och hur. Medan det är ganska fritt att välja metoder så är läroplanens mål tydliga. Lärarna vet alltså vart eleverna ska, men får inte alltid konkret stöd i hur eleverna ska komma dit.

Barn med hörselnedsättning

Runt femtusen elever inom den svenska grundskolan behöver stöd på grund av svårigheter med sin hörsel. De allra flesta, snart 95 procent, går i ”vanliga klasser”, medan några går i hörselklasser och ett fåtal går i teckenspråkig specialskola eller i anpassad grundskola. I hörselklasser är det få elever och det finns specifika anpassningar i klassrummen, såsom mikrofoner och slinga som kan koppla ljudet direkt till elevernas hörhjälpmedel.

Tidig upptäckt av hörselnedsättning är viktigt för språkutvecklingen. Därför görs numera hörseltest på alla nyfödda barn i Sverige. Sen upptäckt av hörselnedsättning förekommer dock fortfarande. Barn med hörselnedsättning får ofta försenad utveckling av både grundläggande och komplexa språkfärdigheter. Runt var femte barn med hörselnedsättning uppfyller kriterierna för språkstörning. Trots det så finns inga bestämda riktlinjer för uppföljning av språkutveckling hos barn med hörselnedsättning. De språkliga svårigheterna se olika ut beroende på exempelvis orsak till hörselnedsättningen, grad av hörselnedsättning, ålder vid upptäckt och vid vilken ålder som hörhjälpmedel satts in. Typen av hörhjälpmedel och hur mycket de används är viktigt. Mer hörapparatanvändning (större andel av den vakna tiden) ger större ordförråd och bättre grammatisk förmåga. Det spelar också stor roll vilket stöd barnet och familjen får från skolan (t.ex. tillgång till logoped och hörselpedagog) och hur de sociala förhållandena ser ut. Det finns alltså både risk- och friskfaktorer för språkutveckling hos barn med hörselnedsättning.

Man vet ganska lite om vad som är bra skrivundervisning för elever med hörselnedsättning, trots att många av dem behöver extra mycket stöd.

Lärande genom observation

En fungerande skrivundervisningsmetod är att låta eleverna observera hur en *modell* (förebild) gör och sedan göra på samma sätt. Modellen är ofta en lärare, men kan också vara en jämnårig kamrat. Att se en jämnårig kan vara särskilt motiverande, eftersom eleven då lättare kan identifiera sig med modellen. Observationen är en aktiv process där eleven bland annat uppmärksammar vad kamraten gör, bedömer om det är ett bra sätt och lägger det hen ser på minnet. Man tror att lärande genom observation minskar den kognitiva belastningen genom att man inte skriver under tiden man lär sig det nya. Om målet är att eleverna exempelvis ska lära sig att tänka på att ta läsarens perspektiv är det gynnsamt att de kan reflektera över och lära sig det utan att behöva skriva samtidigt. Då kommer inte arbetet med de ”lägre” skrivprocesserna (som exempelvis stavning och interpunktion) ta överhanden över högre processer (exempelvis att anpassa sig till läsaren och att skriva lagom långa meningar).

Självskattad förmåga

Den tilltro man har till sin förmåga inom ett område kallas självskattad förmåga (self-efficacy). Begreppet är myntat av socialpsykologen Albert Bandura. Den som skattar sin förmåga högt ser en svår uppgift som en utmaning som kan klaras av, vilket ökar motivationen till att genomföra den. Den som däremot skattar sin förmåga lågt ser uppgiften som nästan omöjlig att genomföra vilket gör att hen lägger mindre kraft på den. Då är det lättare att ge upp, vilket bekräftar ens låga tilltro till förmågan. Enligt Bandura beror självskattad förmåga på fyra saker: 1) Egen erfarenhet av att klara av respektive inte klara av en liknande uppgift, 2) indirekt erfarenhet (att se någon annan klara av eller inte klara av en uppgift), 3) socialt övertygande, (att en trovärdig person uppmuntrar eller avråder från att försöka) och 4) kroppsligt och känslomässigt tillstånd (att känna sig väl till mods kan göra att man ser en uppgift som möjlig att klara, medan man kan misstolka att känna sig illa till mods som att man inte har en färdighet). Självskattad förmåga är alltså en viktig aspekt av motivation. Försäljares säljförmåga, rökstopp och matematikfärdigheter är några av de områden där man sett att självskattad förmåga har betydelse. När det gäller skrivundervisning har man sett att lärande genom observation kan ge högre självskattad förmåga. Den självskattade förmågan hos yngre barn är ofta god och olika typer av förmågor och olika delar av en förmåga kan skattas lika. Under uppväxten börjar barn och ungdomar skatta sig olika inom olika förmågor och även inom olika aspekter av en förmåga såsom att exempelvis kunna skriva en berättelse. Den självskattade förmågan brukar också sjunka med åren, exempelvis vid skolbyten.

Syfte och forskningsfrågor

Mitt övergripande syfte är att utveckla och utvärdera en intervention (i form av en lektionsserie) baserad på lärande genom observation för skriftligt berättande, samt att utveckla och utvärdera ett självskattningsformulär om skriftligt berättande. Interventionen ska prövas för elever i skolår 5 med normal hörsel och elever i skolår 5–8 i hörselklasser. Självskattningsformuläret ska prövas för eleverna med normal hörsel.

Mina forskningsfrågor är följande:

1) Hur kan interventionen och självskattningformuläret utvecklas och implementeras i svenska grundskola?

2) Hur påverkar interventionen elevernas textkvalitet?

3) Hur är kognitiva, språkliga (arbetsminne, språkförståelse och läsförståelse) och demografiska faktorer (kön och för elever med hörselnedsättning även skolår och hörselfaktorer) associerade med textkvalitet före och efter intervention?

4) Hur förändras den självskattade förmågan efter intervention hos elever med normal hörsel och hur är den associerad med textkvalitet och kön?

Metod

Femtiofem elever med normal hörsel (i vanliga femteklasser) och elva elever med hörselnedsättning (i skolår fem till åtta i hörselklasser) deltog. De fick skriva berättelser vid fyra tillfällen, göra test av arbetsminne, språkförståelse och läsförståelse och delta i en intervention i form av fem lektioner som jag höll i tillsammans med en forskningsassistent. Eleverna med normal hörsel gjorde även en självskattning av sin skriftliga berättarförmåga vid två tillfällen.

Under lektionerna fick eleverna se filmer av barn i liknande åldrar på storskärm. De filmade barnen gjorde skrivaktiviteter, exempelvis läste någon annans berättelse och pratade om texten eller förbättrade en färdig text. I andra filmer såg eleverna texten växa fram på storskärmen i en skärminspelning medan ett barn skrev och tänkte högt. Eleverna fick tänka på och diskutera vad de såg på filmerna och skriva ned vad de tyckte var viktigt, i form av råd till någon annan elev som skulle öva på att skriva berättelser (”någon som börjar i femte klass nästa år”).

Lektionerna och självskattningen utvecklades utifrån tidigare studier och anpassades till svenska förhållanden, elevernas ålder och berättandegenren. Tidigare studier har ofta undersökt högstadieelever som skrivit argumenterande text. Kvaliteten på elevernas skriftliga berättelser bedömdes av personer som fått öva på att bedöma motsvarande berättelser på ett likvärdigt sätt.

Resultat och slutsatser

Vi utvecklade en intervention i form av en lektionsserie på fem 40-minuterslektioner, baserad på lärande genom observation, som fungerade i svensk skola utifrån den svenska läroplanen och skolmiljön. Vi utvecklade också ett självskattningsformulär om skriftligt berättande som kunde användas. För elever med normal hörsel blev textkvaliteten högre efter interventionen, men inte för elever med hörselnedsättning (på gruppnivå). För både hörande elever och elever med hörselnedsättning visade det sig att flickor hade högre textkvalitet än pojkar, på gruppnivå. Eleverna som hade mest nytta av interventionen (de som förbättrade sin

textkvalitet mest) var de som hade bättre läsförståelse. En anledning kan vara att de eleverna hade läst mer och därmed hade mer kunskap om berättelser som de kunde använda. När man mäter interventionseffekter är det viktigt att se om de kvarstår över tid. Nästa termin hade textkvaliteten sjunkit till den ursprungliga nivån. Lärande genom observation kan komplettera annan skrivundervisning i svensk skola, men vidare utveckling behövs, inte minst för att det ska vara användbart för elever med hörselnedsättning.

Det är viktigt att upptäcka hörselnedsättning tidigt och sätta in hjälpmedel och pedagogiska åtgärder snabbt. För elever med hörselnedsättning såg man tecken på att tidig upptäckt av hörselnedsättningen var kopplat till hög textkvalitet, medan sen upptäckt var kopplat till låg textkvalitet som dessutom sjönk något till nästa termin. Elever med språkliga svårigheter kan vara särskilt sårbara för uppehåll i undervisningen såsom ett långt sommarlov.

Eleverna med normal hörsel skattade sin förmåga att berätta i skrift högt. Efter lektionsserien skattade de sig ännu högre, de hade alltså fått en stärkt tilltro till sin förmåga. Pojkar och flickor skattade på gruppnivå sin förmåga lika högt, trots att texterna som skrivits av flickor bedömdes som betydligt bättre. Motivationsfaktorer som självskattad förmåga, kan vara viktigt att undersöka och påverka.

Eftersom elevers behov ser olika ut, behövs delvis olika skrivundervisningsmetoder. För att stärka elevers skolresultat är det därför viktigt att dels anpassa och utvärdera nya metoder, dels evidensbasera de metoder som används idag. Tidigare studier visar att det behövs gott om tid och stöd från skolledningen för att lärare ska ha möjlighet till planering och samarbete med de som introducerar nya metoder (som kan vara t.ex. andra lärare och/eller forskare). När det gäller elever med språkliga svårigheter behövs kontinuerligt stöd till lärare för att eleverna ska ha nytta av den nya metoden.

För att elever ska kunna uppnå goda skolresultat och kunna delta i och bidra till ett hållbart samhälle behöver vi stötta alla elevers skrivutveckling genom hela skolgången.

References

Andersson, K., Hansson, K., Rosqvist, I., Åhlander, V. L., Sahlén, B., Sandgren, O., & Lyberg Åhlander, V. (2019). The contribution of bilingualism, parental education, and school characteristics to performance on the clinical evaluation of language fundamentals: Fourth edition, Swedish [Article]. *Frontiers In Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.01586

Applebee, A. N. (1978). *The child's concept of story: ages two to seventeen*. University of Chicago Press.

Arfé, B., Dockrell, J. E., & Berninger, V. (Eds.). (2014). *Writing Development in Children with Hearing Loss, Dyslexia, or Oral Language Problems : Implications for Assessment and Instruction* [Book]. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199827282.001.0001.

Arfé, B., Rossi, C., & Sicoli, S. (2015). The Contribution of Verbal Working Memory to Deaf Children's Oral and Written Production. *Journal of Deaf Studies & Deaf Education*, *20*(3), 203-214. https://doi.org/10.1093/deafed/env005

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items [Article]. *Journal of Memory and Language*, *59*(4), 390-412. https://doi.org/10.1016/j.jml.2007.12.005

Baker, S., Gersten, R., & Graham, S. (2003). Teaching Expressive Writing to Students with Learning Disabilities: Research-Based Applications and Examples [Article]. *Journal of Learning Disabilities*, *36*(2), 109. https://doi.org/10.1177/002221940303600204

Bandura, A. (1977). *Social learning theory*. Prentice-Hall.

Bandura, A. (1986). *Social foundations of thought and action: a social cognitive theory*. Prentice-Hall.

Bandura, A. (1997). *Self-efficacy: the exercise of control*. W. H. Freeman.

Bandura, A. (2006). Guide for constructing self-efficacy scales. In F. Pajares & T. Urdan (Eds.), *Self-efficacy Beliefs of Adolescents*. Information Age Publishing.

Beard, R., Myhill, D., Riley, J., & Nystrand, M. (2009). *The SAGE Handbook of Writing Development*. SAGE Publications.

Berman, R. A., & Slobin, D. I. (1994). *Relating events in narrative: a crosslinguistic developmental study* [Non-fiction]. Lawrence Erlbaum Associates; Psychology Press.

Berninger, V., Cartwright, A. C., Yates, C. M., Swanson, H. L., & Abbott, R. D. (1994). Developmental skills related to writing and reading acquisition in the intermediate grades: Shared and unique functional systems. *Reading and Writing: An Interdisciplinary Journal*, *6*(2), 161-196. https://doi.org/10.1007/BF01026911

Berninger, V., & Swanson, H. L. (1994). Modifying Hayes & Flower's model of skilled writing to explain beginning and developing writing. In E. Butterfield (Ed.), *Children's writing: Toward a process theory of development of skilled writing* (pp. 57–81). JAI Press.

Berninger, V., Whitaker, D., Feng, Y., Swanson, H. L., & Abbott, R. D. (1996). Assessment of planning, translating, and revising in junior high writers. *Journal of School Psychology*, *34*(1), 23-52. https://doi.org/10.1016/0022-4405(95)00024-0

Berninger, V., Yates, C., Cartwright, A., Rutberg, J., Remy, E., & Abbott, R. (1992). Lower-level developmental skills in beginning writing. *Reading and Writing: An Interdisciplinary Journal*, *4*(3), 257-280. https://doi.org/10.1007/BF01027151

Bishop, D. V. M. (2009). *Test for reception of grammar, version 2. Swedish version.* Pearson Education.

Bishop, D. V. M., Snowling, M. J., Thompson, P. A., Greenhalgh, T., Adams, C., Archibald, L., Baird, G., Bauer, A., Bellair, J., Boyle, C., Brownlie, E., Carter, G., Clark, B., Clegg, J., Cohen, N., Conti - Ramsden, G., Dockrell, J., Dunn, J., Ebbels, S., & Gallagher, A. (2017). Phase 2 of CATALISE: a multinational and multidisciplinary Delphi consensus study of problems with language development: Terminology [Article]. *Journal of Child Psychology & Psychiatry*, *58*(10), 1068-1080. https://doi.org/10.1111/jcpp.12721

Bishop, D. V. M., Snowling, M. J., Thompson, P. A., Greenhalgh, T., & CATALISE-consortium. (2016). CATALISE: A Multinational and Multidisciplinary Delphi Consensus Study. Identifying Language Impairments in Children [Article]. *PLOS ONE*, *11*(7), 1-26. https://doi.org/10.1371/journal.pone.0158753

Blom, E., & Boerma, T. (2016). Why do children with language impairment have difficulties with narrative macrostructure? *Research in Developmental Disabilities*, *55*(Supplement C), 301-311. https://doi.org/https://doi.org/10.1016/j.ridd.2016.05.001

Bouwer, R., Koster, M., & van den Bergh, H. (2017). Effects of a Strategy-Focused Instructional Program on the Writing Quality of Upper Elementary Students in the Netherlands. *Journal of Educational Psychology*. https://doi.org/10.1037/edu0000206

Braaksma, M., Rijlaarsdam, G., & van den Bergh, H. (2002). Observational learning and the effects of model-observer similarity [Article]. *Journal of Educational Psychology*, *94*(2), 405.

Braaksma, M., Rijlaarsdam, G., & Van den Bergh, H. (2018). Effects of hypertext writing and observational learning on content knowledge acquisition, self-efficacy, and text quality: Two experimental studies exploring aptitude treatment interactions. *Journal of Writing Research*, *9*(3), 259-300. https://doi.org/10.17239/jowr-2018.09.03.02

Braaksma, M., Rijlaarsdam, G., van den Bergh, H., & van Hout-Wolters, B. H. A. M. (2004). Observational learning and its effects on the orchestration of writing processes [research-article]. *Cognition and Instruction*(1). https://doi.org/doi.org/10.1207/s1532690Xci2201_1

Braaksma, M. A. H., Rijlaarsdam, G., & van den Bergh, H. (2012). Observational learning in argumentative writing. In M. Torrance, D. Alamargot, M. Castelló, F. Ganier, O. Kruse, A. Mangen, L. Tolchinsky, & L. v. Waes (Eds.), *Learning to Write Effectively: Current Trends in European Research* (pp. 123-126). Brill. https://doi.org/https://doi.org/10.1163/9781780529295_027

Brindle, M., Graham, S., Harris, K. R., & Hebert, M. (2016). Third and fourth grade teacher's classroom practices in writing: a national survey. *Reading and Writing*, *29*(5), 929-954. https://doi.org/10.1007/s11145-015-9604-x

Briscoe, J., Bishop, D. V. M., & Norbury, C. F. (2001). Phonological Processing, Language, and Literacy: A Comparison of Children with Mild-to-moderate Sensorineural Hearing Loss and Those with Specific Language Impairment [Article]. *Journal of Child Psychology & Psychiatry & Allied Disciplines*, *42*(3), 329. https://doi.org/10.1111/1469-7610.00726

Brown, D. E. (2013). *Human Universals* [Topic overview]. Sage Publications, Inc.

Bruning, R., Dempsey, M., Kauffman, D. F., McKim, C., & Zumbrunn, S. (2013a). Examining dimensions of self-efficacy for writing. *Journal of Educational Psychology*, *105*(1), 25-38. https://doi.org/10.1037/a0029692

Bruning, R., Dempsey, M., Kauffman, D. F., McKim, C., & Zumbrunn, S. (2013b). *Self-Efficacy for Writing Scale*. https://doi.org/10.1037/t19869-000

Brännström, K. J., von Lochow, H., Åhlander, V. L., & Sahlén, B. (2018). Immediate Passage Comprehension and Encoding of Information Into Long-Term Memory in Children With Normal Hearing: The Effect of Voice Quality and Multitalker Babble Noise. *American Journal Of Audiology*, *27*(2), 231-237. https://doi.org/10.1044/2018_AJA-17-0061

Bulut, P. (2017). The Effect of Primary School Students' Writing Attitudes and Writing Self-Efficacy Beliefs on Their Summary Writing Achievement. *International Electronic Journal of Elementary Education*, *10*(2), 281-285.

Callinan, C. J., van der Zee, E., & Wilson, G. (2018). Developing Essay Writing Skills: An Evaluation of the Modelling Behaviour Method and the Influence of Student Self-Efficacy. *Journal of Further and Higher Education*, *42*(5), 608-622. https://doi.org/10.1080/0309877X.2017.1302564

Carretti, B., Borella, E., Cornoldi, C., & De Beni, R. (2009). Role of working memory in explaining the performance of individuals with specific reading comprehension difficulties: A meta-analysis [Article]. *Learning and Individual Differences*, *19*(2), 246-251. https://doi.org/10.1016/j.lindif.2008.10.002

Casals, M., Girabent-Farrés, M., & Carrasco, J. L. (2014). Methodological Quality and Reporting of Generalized Linear Mixed Models in Clinical Medicine (2000–2012): A Systematic Review. *PLOS ONE*, *9*(11), e112653. https://doi.org/10.1371/journal.pone.0112653

Catts, H. W., Fey, M. E., Tomblin, J. B., & Zhang, X. (2002). A longitudinal investigation of reading outcomes in children with language impairments. *Journal of Speech, Language & Hearing Research*, *45*(6), 1142-1157. https://doi.org/10.1044/1092-4388(2002/093)

Catts, H. W., Hogan, T. P., & Fey, M. E. (2003). Subgrouping Poor Readers on the Basis of Individual Differences in Reading-Related Abilities. *Journal of Learning Disabilities*, *36*(2), 151-164. https://doi.org/10.1177/002221940303600208

Clark, E. (2016). *First Language Acquisition* (3 ed.). Cambridge University Press. https://doi.org/10.1017/CBO9781316534175

Clark, J. G. (1981). Uses and abuses of hearing loss classification. *ASHA*, *23*(7), 493-500.

Cooper, H., Nye, B., Charlton, K., Lindsay, J., & Greathouse, S. (1996). The effects of summer vacation on achievement test scores: a narrative and meta-analytic review [research-article]. *Review of Educational Research*, *66*(3), 227.

Coutinho, S. A., & Neuman, G. (2008). A model of metacognition, achievement goal orientation, learning style and self-efficacy. *Learning Environments Research*, *11*(2), 131-151. https://doi.org/10.1007/s10984-008-9042-7

Couzijn, M. (1999). Learning to write by observation of writing and reading processes: effects on learning and transfer. *Learning and Instruction*, *9*(2), 109-142. https://doi.org/10.1016/S0959-4752(98)00040-1

De Smedt, F., Merchie, E., Barendse, M., Rosseel, Y., Van Keer, H., & De Naeghel, J. (2018). Cognitive and Motivational Challenges in Writing: Studying the Relation With Writing Performance Across Students' Gender and Achievement Level [Article]. *Reading Research Quarterly*, *53*(2), 249-272. https://doi.org/10.1002/rrq.193

Dostal, H. M., & Wolbers, K. A. (2016). Examining student writing proficiencies across genres: results of an intervention study. *Deafness & Education*

International, *18*(3), 159-169. https://doi.org/10.1080/14643154.2016.1230415

Drijbooms, E., Groen, M. A., & Verhoeven, L. (2015). The Contribution of Executive Functions to Narrative Writing in Fourth Grade Children. *Reading and Writing: An Interdisciplinary Journal*, *28*(7), 989-1011. http://dx.doi.org/10.1007/s11145-015-9558-z

Ebbels, S. H. (2017). Intervention research: Appraising study designs, interpreting findings and creating research in clinical practice. *International Journal of Speech-Language Pathology*, *19*(3), 218-231. https://doi.org/10.1080/17549507.2016.1276215

Ebbels, S. H., McCartney, E., Slonims, V., Dockrell, J. E., & Norbury, C. F. (2019). Evidence-based pathways to intervention for children with language disorders. *International Journal of Language & Communication Disorders*, *54*(1), 3-19. https://doi.org/10.1111/1460-6984.12387

Eisenberg, L. S., Fisher, L. M., Johnson, K. C., Ganguly, D. H., Grace, T., & Niparko, J. K. (2016). Sentence Recognition in Quiet and Noise by Pediatric Cochlear Implant Users: Relationships to Spoken Language. *Otology & neurotology : official publication of the American Otological Society, American Neurotology Society [and] European Academy of Otology and Neurotology*, *37*(2), e75-e81. https://doi.org/10.1097/MAO.0000000000000910

Fayol, M., Hupet, M., & Largy, P. (1999). The acquisition of subject-verb agreement in written French: From novices to experts' errors. *Reading and Writing*, *11*(2), 153-174. https://doi.org/10.1023/A:1008038127807

Feagans, L., & Appelbaum, M. I. (1986). Validation of language subtypes in learning disabled children. *Journal of Educational Psychology*, *78*(5), 358-364. https://doi.org/10.1037/0022-0663.78.5.358

Fidalgo, R., Torrance, M., & García, J.-N. (2008). The long-term effects of strategy-focussed writing instruction for grade six students [Article]. *Contemporary Educational Psychology*, *33*, 672-693. https://doi.org/10.1016/j.cedpsych.2007.09.001

Fidalgo, R., Torrance, M., Rijlaarsdam, G., van den Bergh, H., & Lourdes Álvarez, M. (2015). Strategy-focused writing instruction: Just observing and reflecting on a model benefits 6th grade students. *Contemporary Educational Psychology*, *41*, 37-50. https://doi.org/10.1016/j.cedpsych.2014.11.004

Garberoglio, C. L., Palmer, J. L., Cawthon, S., & Sales, A. (2019). *Deaf People and Educational Attainment in the United States*. U.S. Department of Education, Office of Special Education Programs, National Deaf Center on Postsecondary Outcomes.

Geers, A. E., Moog, J. S., Biedenstein, J., Brenner, C., & Hayes, H. (2009). Spoken language scores of children using cochlear implants compared to

hearing age-mates at school entry. *Journal of Deaf Studies & Deaf Education*, *14*(3), 371-385. https://doi.org/deafed/enn046
Geers, A. E., Nicholas, J., Tobey, E., & Davidson, L. (2016). Persistent language delay versus late language emergence in children with early cochlear implantation. *Journal of Speech, Language & Hearing Research*, *59*(1), 155-170. https://doi.org/10.1044/2015_JSLHR-H-14-0173
Geiger, M., A., & Cooper, E., A. (1995). Predicting Academic Performance: The Impact of Expectancy and Needs Theory [research-article]. *The Journal of Experimental Education*, *63*(3), 251-262. https://www.jstor.org/stable/20152454
Gelfand, S. A. (2016). *Essentials of audiology. [Elektronisk resurs]* (Fourth ed.) [Online Non-fiction]. Thieme.
Gilbert, J., & Graham, S. (2010). Teaching Writing to Elementary Students in Grades 4–6: A National Survey. *The Elementary School Journal*, *110*(4), 494-518. https://doi.org/10.1086/651193
Gough, P. B., & Tunmer, W. E. (1986). Decoding, Reading, and Reading Disability. *Remedial & Special Education*, *7*(1), 6.
Graham, S., Daley, S. G., Aitken, A. A., Harris, K. R., & Robinson, K. H. (2018). Do writing motivational beliefs predict middle school students' writing performance? *Journal of Research in Reading*, *41*(4), 642-656.
Graham, S., & Harris, K. (2014). Conducting High Quality Writing Intervention Research: Twelve Recommendations. *Journal of Writing Research*, *6*(2), 89-123. https://doi.org/10.17239/jowr-2014.06.02.1
Graham, S., & Harris, K. R. (2018). Evidence-Based Writing Practices: A Meta-Analysis of Existing Meta-Analyses. In *Design Principles for Teaching Effective Writing: Theoretical and Empirical Grounded Principles* (pp. 13-37). Brill Academic Publishers.
Graham, S., Liu, X., Bartlett, B., Ng, C., Harris, K. R., Aitken, A., Barkel, A., Kavanaugh, C., & Talukdar, J. (2018). Reading for Writing: A Meta-Analysis of the Impact of Reading Interventions on Writing. *Review of Educational Research*, *88*(2), 243-284. http://dx.doi.org/10.3102/0034654317746927
Graham, S., McKeown, D., Kiuhara, S., & Harris, K. R. (2012). A meta-analysis of writing instruction for students in the elementary grades. *Journal of Educational Psychology*, *104*(4), 879-896. https://doi.org/10.1037/a0029185
Grenner, J. (2017). Hörsel [Hearing]. In M. Jägervall & J. Lundgren (Eds.), *Barnneurologi [Child neurology]* (1 ed.). Studentlitteratur.
Griffin, T. M., Hemphill, L., Camp, L., & Wolf, D. P. (2004). Oral Discourse in the Preschool Years and Later Literacy Skills. *First Language*, *24*(2 [71]), 123-147. https://doi.org/10.1177/0142723704042369
Gärdenfors, M., Johansson, V., & Schönström, K. (2019). Spelling in Deaf, Hard of Hearing and Hearing Children With Sign Language Knowledge

[Original Research]. *Frontiers In Psychology*, *10*(2463). https://doi.org/10.3389/fpsyg.2019.02463

Hammarstedt, M., & Karan, V. S. (2016). *Vilken effekt har lärande genom observation på elevers narrativa skrivande i årskurs 5? En studie av skrivprodukter och skrivprocesser (What effect does learning by observation have on narrative writing in students in schoolyear 5?)* [Student paper, Lund University]. Lund.

Hansson, K., Ibertsson, T., Asker-Árnason, L., & Sahlén, B. (2017). Phonological processing, grammar and sentence comprehension in older and younger generations of Swedish children with cochlear implants [article]. *Autism & Developmental Language Impairments*. https://doi.org/10.1177/2396941517692809

Harris, K. R., Graham, S., & Mason, L. H. (2006). Improving the Writing, Knowledge, and Motivation of Struggling Young Writers: Effects of Self-Regulated Strategy Development with and without Peer Support [research-article]. *American Educational Research Journal*(2), 295.

Hayes, J. (1996). A new framework for understanding cognition and affect in writing. In C. M. Levy & S. Ransdell (Eds.), *The Science of Writing: Theories, Methods, Individual Differences, and Applications* (pp. 1-27).

Hayes, J., & Berninger, V. (2014). *Cognitive Processes in Writing: A Framework* [Chapter]. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199827282.003.0001

Hayes, J. R., & Flower, L. S. (1980). Identifying the organization of writing processes. In L. W. Gregg & E. R. Steinberg (Eds.), *Cognitive processes in writing* (pp. 3-30). Erlbaum.

HRF. (2007). *Årsrapport [Yearly report]*. HRF Hörselskadades riksförbund [Swedish association of persons with hearing loss].

Johansson, V. (2009). *Developmental aspects of text production in writing and speech* Lund University]. Lund. https://portal.research.lu.se/portal/files/5221582/1487260.pdf

Kanaris, A. (1999). Gendered Journeys: Children's Writing and the Construction of Gender. *Language and Education*, *13*(4), 254-268. https://doi.org/10.1080/09500789908666772

Karjalainen, S., Sahlén, B., Falck, A., Brännström, J., & Lyberg-Åhlander, V. (2019). Implementation and evaluation of a teacher intervention program on classroom communication. *Logopedics Phoniatrics Vocology*, 1-13. https://doi.org/10.1080/14015439.2019.1595131

Karmiloff-Smith, A. (1981). The grammatical marking of thematic structure in the development of language production. In W. Deutsch (Ed.), *The child's conception of language* (pp. 121-147). Academic Press.

Kellogg, R. T. (1996). A model of working memory in writing. In C. M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences, and application* (pp. 57-71). Lawrence Erlbaum Associates.

Kellogg, R. T., Whiteford, A. P., Turner, C. E., Cahill, M., & Mertens, A. (2013). Working Memory in Written Composition: A Progress Report. *Journal of Writing Research*, *5*(2), 159-190. https://doi.org/10.17239/jowr-2013.05.02.1

Koster, M., Tribushinina, E., Jong, P. F. d., & Bergh, H. v. d. (2015). Teaching children to write: A meta-analysis of writing intervention research. *Journal of Writing Research*, *7*(2), 249-274. https://doi.org/10.17239/jowr-2015.07.02.2

Lagercrantz, H. (2017). Hjärnans utveckling [Brain development]. In M. Jägervall & J. Lundgren (Eds.), *Barnneurologi [Child neurology]* (1 ed.). Studentlitteratur.

Lewis, D., Schmid, K., O'Leary, S., Spalding, J., Heinrichs-Graham, E., & High, R. (2016). Effects of Noise on Speech Recognition and Listening Effort in Children With Normal Hearing and Children With Mild Bilateral or Unilateral Hearing Loss. *Journal of Speech, Language & Hearing Research*, *59*(5), 1218-1232. https://doi.org/10.1044/2016_JSLHR-H-15-0207

Limpo, T., & Alves, R. A. (2017). Relating beliefs in writing skill malleability to writing performance: The mediating role of achievement goals and self-efficacy. *Journal of Writing Research*, *9*(volume 9 issue 2), 97-125. https://doi.org/10.17239/jowr-2017.09.02.01

Limpo, T., Alves, R. A., & Fidalgo, R. (2014). Children's high-level writing skills: Development of planning and revising and their contribution to writing quality [Article]. *British Journal of Educational Psychology*, *84*(2), 177-193. https://doi.org/10.1111/bjep.12020

Lundgren, U. P., Säljö, R., & Liberg, C. (2017). *Lärande, skola, bildning* (4 ed.) [Non-fiction]. Natur & kultur.

Magnusson, E., & Nauclér, K. (2010). *SL40: läsförståelse: meningar [SL40: reading comprehension: sentences]*. Studentlitteratur AB.

Markussen-Brown, J., Juhl, C. B., Piasta, S. B., Bleses, D., Højen, A., & Justice, L. M. (2017). The effects of language- and literacy-focused professional development on early educators and children: A best-evidence meta-analysis [Article]. *Early Childhood Research Quarterly*, *38*, 97-115. https://doi.org/10.1016/j.ecresq.2016.07.002

Marschark, M., & Hauser, P. C. (Eds.). (2008). *Deaf cognition: foundations and outcomes*. Oxford University Press.

Marschark, M., & Knoors, H. (2012). Educating Deaf Children: Language, Cognition, and Learning. *Deafness & Education International*, *14*(3), 136-160. https://doi.org/10.1179/1557069X12Y.0000000010

Marschark, M., & Knoors, H. (2018a). *Evidence-based practices in deaf education* [Non-fiction]. Oxford University Press.

Marschark, M., & Knoors, H. (2018b). Sleuthing the 93% Solution in Deaf Education. In *Evidence-Based Practices in Deaf Education*. Oxford University Press. https://doi.org/10.1093/oso/9780190880545.003.0001

Mattys, S. L., Davis, M. H., Bradlow, A. R., & Scott, S. K. (2012). Speech recognition in adverse conditions: A review [Article]. *Language & Cognitive Processes*, *27*(7/8), 953-978. https://doi.org/10.1080/01690965.2012.705006

McCutchen, D. (2000). Knowledge, Processing, and Working Memory: Implications for a Theory of Writing [Article]. *Educational Psychologist*, *35*(1), 13-23. https://doi.org/10.1207/S15326985EP3501_3

McCutchen, D. (2011). From Novice to Expert: Implications of Language Skills and Writing - Relevant Knowledge for Memory during the Development of Writing Skill. *Journal of Writing Research*, *3*.

McFadden, B., & Pittman, A. (2008). Effect of minimal hearing loss on children's ability to multitask in quiet and in noise. *Language, Speech, and Hearing Services in Schools*, *39*(3), 342-351. https://doi.org/10.1044/0161-1461(2008/032)

Meteyard, L., & Davies, R. A. I. (2020). Best practice guidance for linear mixed-effects models in psychological science [Article]. *Journal of Memory and Language*, *112*. https://doi.org/10.1016/j.jml.2020.104092

Miniscalco, C., Hagberg, B., Kadesjö, B., Westerlund, M., & Gillberg, C. (2007). Narrative skills, cognitive profiles and neuropsychiatric disorders in 7–8-year-old children with late developing language. *International Journal of Language & Communication Disorders*, *42*(6), 665-681. https://doi.org/10.1080/13682820601084428

Muenks, K., Wigfield, A., & Eccles, J. S. (2018). I can do this! The development and calibration of children's expectations for success and competence beliefs. *Developmental Review*, *48*, 24-39. https://doi.org/10.1016/j.dr.2018.04.001

Myhill, D. (2008). Towards a Linguistic Model of Sentence Development in Writing. *Language and Education*, *22*(5), 271-288.

Myhill, D. (2009a). Becoming a designer: trajectories of linguistic development. In R. Beard, D. Myhill, M. Nystrand, & J. Riley (Eds.), *The SAGE Handbook of Writing Development*. SAGE Publications Ltd. https://doi.org/10.4135/9780857021069

Myhill, D. (2009b). Becoming a designer: trajectories of linguistic development. In R. Beard, D. Myhill, M. Nystrand, & J. Riley (Eds.), *SAGE Handbook of Writing Development*. SAGE Publications.

Myhill, D., & Jones, S. (2007). More Than Just Error Correction: Students' Perspectives on Their Revision Processes during Writing. *Written Communication*, *24*(4), 323-343. https://doi.org/10.1177/0741088307305976

Neuman, S. B., Kaefer, T., & Pinkham, A. M. (2018). A double dose of disadvantage: Language experiences for low-income children in home and school. *Journal of Educational Psychology*, *110*(1), 102-118. https://doi.org/10.1037/edu0000201

Nordqvist Palviainen, Å. (2001). *Speech about speech: a developmental study on form and function of direct and indirect speech* [Thesis, edslib. Göteborg.

OECD. (2013). *PISA 2012 Results (Volume 1)*. https://doi.org/10.1787/19963777

OECD. (2016). *PISA 2015 Results (Volume I)*. https://doi.org/10.1787/9789264266490-en

OECD. (2019). *PISA 2018 Results (Volume I)*. https://doi.org/10.1787/5f07c754-en

Olusanya, B. O., Davis, A. C., & Hoffman, H. J. (2019). Hearing loss grades and the International classification of functioning, disability and health. *Bulletin of the World Health Organization*, *97*(10), 725-728. https://doi.org/10.2471/BLT.19.230367

Pajares, F. (1996). Self-Efficacy Beliefs in Academic Settings [research-article]. *Review of Educational Research*, *66*(4), 543.

Pajares, F. (2003). Self-efficacy beliefs, motivation, and achievement in writing: a review of the literature. *Reading & Writing Quarterly*, *19*(2), 139.

Pajares, F. (2007). Empirical Properties of a Scale to Assess Writing Self-Efficacy in School Contexts [Article]. *Measurement & Evaluation in Counseling & Development*, *39*(4), 239-249. https://doi.org/10.1080/07481756.2007.11909801

Pajares, F., Hartley, J., & Valiante, G. (2001). Response Format in Writing Self-Efficacy Assessment: Greater Discrimination Increases Prediction [Article]. *Measurement & Evaluation in Counseling & Development*, *33*(4), 214. https://doi.org/10.1080/07481756.2001.12069012

Pajares, F., Miller, M. D., & Johnson, M. J. (1999). Gender differences in writing self-beliefs of elementary school students. *Journal of Educational Psychology*, *91*(1), 50-61. https://doi.org/10.1037/0022-0663.91.1.50

Pajares, F., & Valiante, G. (1999). Grade Level and Gender Differences in the Writing Self-Beliefs of Middle School Students. *Contemporary Educational Psychology*, *24*(4), 390-405. https://doi.org/https://doi.org/10.1006/ceps.1998.0995

Pajares, F., & Valiante, G. (2001). Gender Differences in Writing Motivation and Achievement of Middle School Students: A Function of Gender Orientation? *Contemporary Educational Psychology*, *26*(3), 366-381. https://doi.org/https://doi.org/10.1006/ceps.2000.1069

Paul, R., Norbury, C., & Gosse, C. (2018). *Language disorders from infancy through adolescence : listening, speaking, reading, writing, and communicating* (Fifth ed.) [Bibliographies Non-fiction]. Elsevier.

Quené, H., & van den Bergh, H. (2004). On multi-level modeling of data from repeated measures designs: A tutorial. *Speech Communication*, *43*(1-2), 103-121. https://doi.org/10.1016/j.specom.2004.02.004

Raedts, M., Rijlaarsdam, G., van Waes, L., & Daems, F. (2007). Observational learning through video-based models: impact on students' accuracy of self-efficacy beliefs, task knowledge and writing performances. In S. Hidi & P. Boscolo (Eds.), *Writing and Motivation* (pp. 219-238). Elsevier. https://doi.org/doi:https://doi.org/10.1163/9781849508216_013

Raedts, M., Steendam, E., Grez, L., Hendrickx, J., & Masui, C. (2017). The effect of different types of video modelling on undergraduate students' motivation and learning in an academic writing course. *Journal of Writing Research*, *8*, 399-435. https://doi.org/10.17239/jowr-2017.08.03.01

Reuterskiöld, C., Hansson, K., & Sahlén, B. (2011). Narrative skills in Swedish children with language impairment. *Journal of Communication Disorders*, *44*(6), 733-744. https://doi.org/https://doi.org/10.1016/j.jcomdis.2011.04.010

Richardson, M., Abraham, C., & Bond, R. (2012). Psychological Correlates of University Students' Academic Performance: A Systematic Review and Meta-Analysis. *Psychological Bulletin*, *138*(2), 353-387. http://dx.doi.org/10.1037/a0026838

Rijlaarsdam, G., Bergh, H. v. d., & Couzijn, M. (2005). *Effective learning and teaching of writing: a handbook of writing in education* (2 ed.) [Non-fiction]. Kluwer.

Rijlaarsdam, G., Braaksma, M., Couzijn, M., Janssen, T., Raedts, M., Van Steendam, E., Toorenaar, A., & Van den Bergh, H. (2008). Observation of peers in learning to write: Practice and research [article]. *Journal of Writing Research , Vol 1, Iss 1, Pp 53-83 (2008)*(1), 53.

Robbins, S. B., Lauver, K., Le, H., Davis, D., Langley, R., & Carlstrom, A. (2004). Do Psychosocial and Study Skill Factors Predict College Outcomes? A Meta-Analysis. *Psychological Bulletin*, *130*(2), 261-288. https://doi.org/10.1037/0033-2909.130.2.261

Roe, A., & Tengberg, M. (2016). *Utvärdering av Läslyftet Delrapport 1: De första 13 modulerna (Evaluation of Läslyftet CPD, Part 1: The first 13 modules)*.

Roos, C., & Allard, K. (2016). *Framgångsfaktorer i läs- och skrivlärande för döva barn och barn med hörselnedsättning: en systematisk litteraturstudie*. Fakulteten för humaniora och samhällsvetenskapt, Pedagogiskt arbete, Karlstads universitet.

Rosén, B., Kalnak, N., Kullebjörk, P., Klintö, K., Karlsson, M., & Jogby, A. (2019). *Genomlysning 2019 Logopedmottagningarna i Region Skåne (Review 2019 SLP clinics in Region Skåne)*. Region Skåne.

Rosqvist, I., Sandgren, O., Andersson, K., Hansson, K., Lyberg-Åhlander, V., & Sahlén, B. (2019). Children's development of semantic verbal fluency

during summer vacation versus during formal schooling. *Logopedics Phoniatrics Vocology*. https://doi.org/10.1080/14015439.2019.1637456
Runnion, E., & Gray, S. (2019). What clinicians need to know about early literacy development in children with hearing loss. *Language, Speech, and Hearing Services in Schools*, *50*(1), 16-33. https://doi.org/10.1044/2018_LSHSS-18-0015
Rydberg, E., Gellerstedt, L., C. , & Danermark, B. (2009). Toward an Equal Level of Educational Attainment Between Deaf and Hearing People in Sweden? [research-article]. *Journal of Deaf Studies and Deaf Education*, *14*(3), 312. https://doi.org/10.1093/deafed/enp001
Rönnberg, J., Lunner, T., Zekveld, A., Sörqvist, P., Danielsson, H., Lyxell, B., Dahlström, Ö., Signoret, C., Stenfelt, S., Pichora-Fuller, M. K., & Rudner, M. (2013). The Ease of Language Understanding (ELU) model: theoretical, empirical, and clinical advances [Review]. *Frontiers in Systems Neuroscience*, *7*(31). https://doi.org/10.3389/fnsys.2013.00031
Sahlén, B., Hansson, K., Lyberg-Åhlander, V., & Brännström, J. (2018). *Spoken Language and Language Impairment in Deaf and Hard-of-Hearing Children : Fostering Classroom Environments for Mainstreamed Children* [Chapter]. Oxford University Press. https://doi.org/10.1093/oso/9780190880545.003.0006
Sandgren, O., Hansson, K., & Sahlén, B. (2015). Working memory and referential communication—multimodal aspects of interaction between children with sensorineural hearing impairment and normal hearing peers [Mini Review]. *Frontiers In Psychology*, *6*(242). https://doi.org/10.3389/fpsyg.2015.00242
SCB. (2015). *Årsrapport (Yearly report)*.
Schunk, D. (1991). Self-Efficacy and Academic Motivation [Article]. *Educational Psychologist*, *26*(3/4), 207.
Schunk, D. (1994). *Goal and Self Evaluative Influences during Children's Mathematical Skill Acquisition*.
Schunk, D., & Meece, J. L. (2006). Self-efficacy development in adolescences. In F. Pajares & T. Urdan (Eds.), *Self-Efficacy Beliefs of Adolescents* (pp. 71-96). Information Age Publishing.
Schunk, D., & Swartz, C. W. (1993). Goals and Progress Feedback: Effects on Self-Efficacy and Writing Achievement. *Contemporary Educational Psychology*, *18*(3), 337-354. https://doi.org/https://doi.org/10.1006/ceps.1993.1024
Schunk, D. H. (2003). Self-efficacy for reading and writing: Influence of modeling, goal setting, and self-evaluation [Article]. *Reading and Writing Quarterly*, *19*(2), 159-172. https://doi.org/10.1080/10573560308219
Schunk, D. H., Hanson, A. R., & Cox, P. D. (1987). Peer-model attributes and children's achievement behaviors. *Journal of Educational Psychology*, *79*(1), 54-61. https://doi.org/10.1037/0022-0663.79.1.54

Skolverket. (2018). *Läroplan för förskolan (Curriculum for the Preschool)*. Skolverket [Swedish National Agency for Education].
Skolverket. (2019a). *Läroplan för grundskolan, förskoleklassen och fritidshemmet (Curriculum for the compulsory school, preschool class and school-age educare)* (2011, revised 2019 ed.). Skolverket [Swedish National Agency for Education].
Skolverket. (2019b). *Slutbetyg i grundskolan våren 2019*. Skolverket [Swedish National Agency for Education]. https://www.skolverket.se/publikationer?id=4771
Skolverket. (2020a). *Läslyftet (Continuing education in reading and writing development)*. Skolverket [Swedish National Agency for Education]. Retrieved 20200625 from https://www.skolverket.se/skolutveckling/kurser-och-utbildningar/laslyftet-i-skolan
Skolverket. (2020b). *Previous curricula for the compulsory school*. Skolverket [Swedish National Agency for Education].
SOU. (2011:30). *Med rätt att välja - flexibel utbildning för elever som tillhör specialskolans målgrupp (The right to choose - flexible education for students in the compulsory special needs schools)*. Elanders.
SOU. (2016:46). *Samordning, ansvar och kommunikation – vägen till ökad kvalitet i utbildningen för elever med vissa funktionsnedsättningar (Coordination and responsibility - increased quality in education for students with certain disabilities)*. Wolters-Kluwer.
SOU. (2020). *En mer likvärdig skola – minskad skolsegregation och förbättrad resurstilldelning [School with more equity - decreased school segregation and increased resource allocation]*. Wolters-Kluwer.
SPSM. (2008). *Måluppfyllelse för döva och hörselskadade i skolan (Academic achievement for students with deafness or hearing loss)*. Specialpedagogiska skolmyndigheten (Swedish National agency for special needs education and schools).
Stipek, D. (2002). Chapter 12 - Good Instruction Is Motivating. In A. Wigfield & J. S. Eccles (Eds.), *Development of Achievement Motivation* (pp. 309-332). Academic Press. https://doi.org/10.1016/B978-012750053-9/50014-0
Strassman, B. K., & Schirmer, B. (2012). Teaching writing to deaf students: does research offer evidence for practice? *Remedial and Special Education*, *34*(3), 166-179.
Strömqvist, S. (1996). *Discourse flow and linguistic information structuring : explorations in speech and writing* [Non-fiction]. Göteborg : Univ., 1996.
Sundström, S., Löfkvist, U., Lyxell, B., & Samuelsson, C. (2018). Phonological and grammatical production in children with developmental language disorder and children with hearing impairment. *Child Language Teaching & Therapy*, *34*(3), 289-302. https://doi.org/10.1177/0265659018805202

Swanson, H. L., & Alloway, T. P. (2012). Working memory, learning, and academic achievement. In *APA educational psychology handbook, Vol 1: Theories, constructs, and critical issues.* (pp. 327-366). American Psychological Association. https://doi.org/10.1037/13273-012

Tharpe, A. M. (2008). Unilateral and Mild Bilateral Hearing Loss in Children: Past and Current Perspectives. *Trends in Amplification*, *12*(1), 7.

Tillema, M., van den Bergh, H., Rijlaarsdam, G., & Sanders, T. (2012). Quantifying the quality difference between L1 and L2 essays: A rating procedure with bilingual raters and L1 and L2 benchmark essays. *Language Testing*, *30*(1), 71-97. https://doi.org/10.1177/0265532212442647

Titz, C., & Karbach, J. (2014). Working memory and executive functions: effects of training on academic achievement. *Psychological Research*, *78*(6), 852-868. https://doi.org/10.1007/s00426-013-0537-1

Tomblin, J. B., Harrison, M., Ambrose, S. E., Walker, E. A., Oleson, J. J., & Moeller, M. P. (2015). Language Outcomes in Young Children with Mild to Severe Hearing Loss. *Ear Hear*, *36*, 76S-91S. https://doi.org/10.1097/aud.0000000000000219

Tomblin, J. B., Oleson, J., Ambrose, S. E., Walker, E. A., & Moeller, M. P. (2020). Early Literacy Predictors and Second-Grade Outcomes in Children Who Are Hard of Hearing. *Child Development*, *91*(1). http://dx.doi.org/10.1111/cdev.13158

Torkildsen, J. v. K., Hitchins, A., Myhrum, M., & Wie, O. B. (2019). Speech-in-Noise Perception in Children With Cochlear Implants, Hearing Aids, Developmental Language Disorder and Typical Development: The Effects of Linguistic and Cognitive Abilities [Original Research]. *Frontiers In Psychology*, *10*(2530). https://doi.org/10.3389/fpsyg.2019.02530

Walker, E. A., Holte, L., McCreery, R. W., Spratford, M., Page, T., & Moeller, M. P. (2015). The Influence of Hearing Aid Use on Outcomes of Children With Mild Hearing Loss. *Journal of Speech, Language & Hearing Research*, *58*(5), 1611-1625. https://doi.org/10.1044/2015_JSLHR-H-15-0043

van de Weijer, J., Åkerlund, V., Johansson, V., & Sahlén, B. (2018). Writing intervention in university students with normal hearing and in those with hearing impairment: can observational learning improve argumentative text writing? *Logopedics Phoniatrics Vocology*, 1-9. https://doi.org/10.1080/14015439.2017.1418427

van Schaik, P., Volman, M., Admiraal, W., & Schenke, W. (2019). Approaches to co-construction of knowledge in teacher learning groups [Article]. *Teaching and Teacher Education*, *84*, 30-43. https://doi.org/10.1016/j.tate.2019.04.019

Westby, C., & Culatta, B. (2016). Telling Tales: Personal Event Narratives and Life Stories. *Language, Speech & Hearing Services in Schools*, *47*(4), 260-282. https://doi.org/10.1044/2016_LSHSS-15-0073

WHO. *Deafness and hearing loss*. https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss

Wolbers, K., Dostal, H. M., Graham, S., Branum-Martin, L., Kilpatrick, J. R., & Saulsburry, R. (2018). Strategic and interactive writing instruction: an efficacy study in grades 3-5. *Journal of Educational and Developmental Psychology*, *8*(1).

Wolbers, K. A., Dostal, H. M., Graham, S., Cihak, D., Kilpatrick, J. R., & Saulsburry, R. (2015). The Writing Performance of Elementary Students Receiving Strategic and Interactive Writing Instruction. *Journal of Deaf Studies & Deaf Education*.

Wolff, U. (2010). *Lilla DUVAN*. Hogrefe Psykologiförlaget AB.

Vygotsky, L. S. (1978). *Mind in Society: The Development of Higher Psychological Processes* [Book].

Yassin Falk, D. (2017). *Skrivundervisning i grundskolans årskurs 3* [Thesis]. Örebro University.

Yoshinaga-Itano, C., Sedey, A. L., Coulter, D. K., & Mehl, A. L. (1998). Language of early- and later-identified children with hearing loss. *Pediatrics*, *102*(5), 1161-1171.

Zimmerman, B. J., & Kitsantas, A. (2002). Acquiring writing revision and self-regulatory skill through observation and emulation. *Journal of Educational Psychology*, *94*(4), 660-668. https://doi.org/10.1037/0022-0663.94.4.660

Paper I

EDUCATIONAL REVIEW
https://doi.org/10.1080/00131911.2018.1536035

Routledge
Taylor & Francis Group

Improving narrative writing skills through observational learning: a study of Swedish 5th-grade students

Emily Grenner[a], Viktoria Åkerlund[a], Lena Asker-Árnason[a], Joost van de Weijer[b], Victoria Johansson[b] and Birgitta Sahlén[a]

[a]Faculty of Medicine, Department of Clinical Sciences Lund, Logopedics, Phoniatrics, Audiology, Lund University, Lund, Sweden; [b]Centre for Languages and Literature, Lund University, Lund, Sweden

ABSTRACT

Observational learning is a successful method for improving writing skills in various genres. We explore effects of a five lesson intervention series based on peer observation. Fifty-five Swedish 5th-grade students aged 10–12 years followed this intervention programme. The students watched short film clips with peers working with texts. Each lesson was organised according to a theme: reader's perception of the text, ordering of events, how to begin a story, how to end a story and how to edit a text. The students wrote four texts during the intervention. The quality of these texts was assessed by a panel of trained raters. Additionally, the language and reading comprehension and working memory capacity were tested. The results show that average text quality had significantly improved at the end of the intervention, and that this improvement was modulated by reading and language comprehension. Three months later, however, text quality was significantly decreased.

ARTICLE HISTORY
Received 5 June 2018
Accepted 3 October 2018

KEYWORDS
Observational learning; personal narratives; writing improvement; intervention; elementary education; intervention for writing

Introduction

Background to the study

Good writing skills are more important than ever before in today's society. To quote Deborah Brandt (2015, p. 3): "For perhaps the first time in the history of mass literacy, writing seems to be eclipsing reading as the literate experience of consequence". Writing proficiency is not only a fundamental tool for educational achievements, it is also a key feature for success in professional and social life. Consequently, the training of writing skills in early years is necessary, and so is the development of good methods for improving writing skills for students at different levels in the school system.

During recent years, the results of international comparative studies of students' achievements (PISA, PIRLS) have suggested that literacy skills (notably reading comprehension) in Swedish children are declining (Skolverket, 2016). In addition, there have been requests for methods of writing instruction that target the needs of specific populations, for instance, children with hearing impairment. In response to these

CONTACT Emily Grenner emily.grenner@med.lu.se Department of Logopedics, Phoniatrics and Audiology, Lund University, Skåne University hospital, Lund, Sweden

developments, this study was carried out as part of a larger project. In this study, the method of observational learning (e.g. Raedts, van Steendam, Grez, Hendrickx, & Masui, 2017; Rijlaarsdam et al., 2008) was implemented as an intervention programme to improve narrative writing in 11-year-olds (fifth graders). Observational learning has repeatedly been shown to be a successful method for the improvement of writing skills. For this reason, we tested it with a group of hearing students, the results of which we report here. There are no well-established guidelines for writing education in Sweden, and therefore we do not evaluate observational learning against another method. In a future study, we will also report on how the method worked with students with hearing impairment of comparable age.

Writing education

Writing development in 11–16 year-olds has been studied less than writing in younger children, or students at higher education (Myhill, 2009). The studies of the linguistic development in this age range are, according to Myhill (2008, p. 403), "patchy", and there is no comprehensive understanding of how language and writing develop during these years. The writing of 11-year-olds "continues to be influenced by their oral and reading experiences and their social experiences and emotional development" (Myhill, 2008, p. 403). The variation in lexical choice and sentence structures continue to develop over the school years (e.g. Myhill, 2008). A recurrent finding is that girls are better writers than boys. Kanaris (1999), for instance, examined the writing of children 8 to 10 years of age and reported that texts written by girls were generally longer, more complex, and more focused on description and elaboration than boys' texts.

In Sweden, the school curriculum defines which knowledge the children should have acquired by the end of some, but not all grades, for instance grade 3 and grade 6. For this study, the knowledge required by the end of grade 3 is especially interesting, since that is the last certain checkpoint for the children in our study. Here we find that writing of narrative texts is one of the central objectives for grades 1–3 (age 7–9; Skolverket, 2016, pp. 260–261), while those for grades 4–6 are that students broaden their repertoire of knowledge about texts and text production, and learn to adapt the structure and linguistic features typical for other genres, notably descriptions, instructions, argumentative texts, expository texts, advertisements, and letters to the editor. Students should further learn not only different ways of processing and revising the form and content of their own texts, but also to give feedback on texts written by others. In addition, the curriculum states that they should learn how to write by hand and on the computer, and how to organise and edit a text. The linguistic structures they should master include subordinate clauses, parts-of-speech, morphology, spelling rules, punctuation, and cohesion (Skolverket, 2016, pp. 260–262). We can thus expect that the teaching for the children in this study is occupied with the topics above, and the knowledge in this area will be assessed against the criteria in the curriculum by the end of grade six. However, the curriculum does not provide teachers with any specifications regarding teaching strategies and methods that can be used to obtain the goals in the curriculum; instead, the teachers are free to choose among methods and didactic choices that are available to them. Generally, through teacher education, and practice, teachers have experienced methods sprung from a socio-cultural perspective, which has emphasised, for example, the importance of exposing children to good literacy

practices, through exposure to a variety of genres to develop deeper knowledge about domain-specific language, and encouraging them to form an identity as readers and writers (cf. Lundgren, Säljö, & Liberg, 2010). At the same time, however, the curriculum does not provide teachers with any specifications regarding strategies and methods that can be used to obtain the goals in the curriculum. The teachers are free to choose among methods and didactic choices they are familiar with. Which of these strategies are most common in Swedish schools has not been studied systematically. However, the need for better practices in writing skills have been highlighted through the Swedish National Agency for Education's initiative, an ongoing (2015–2018) nation-wide project with the aim to promote good teaching practices. This initiative has provided Swedish teachers with many new approaches to teaching in the literacy field, targeting specific areas, such as early writing development, or domain-specific language. Against this background, we recognise a need for developing targeted and informed teaching methods in the area of writing for this age group in Sweden. With this study, we want to contribute with a method and a design that has not previously been applied as a teaching method in Sweden.

Effective models of writing education

Supporting students' writing skills is considered to be of high priority in education. However, the practices of teaching writing differ considerably between countries, and there are substantial differences in the time allotted for teachers' writing instruction and practicing writing (Graham & Rijlaarsdam, 2016). One recurrent theme in the education of writing is that observing models is a successful strategy. This model can be a teacher (Harris, Graham, & Mason, 2006) but is typically a peer, i.e. a student of similar age and level as the observer (Fidalgo, Torrance, & García, 2008). This strategy has been implemented in the observational learning paradigm, and has been applied to writing instruction in many previous studies (e.g. Braaksma, Rijlaarsdam, & van den Bergh, 2018; Raedts, Rijlaarsdam, van Waes, & Daems, 2006; van de Weijer, Åkerlund, Johansson, & Sahlén, 2018). Learning through observation is not just imitating (Bandura, 1997). It requires several higher-order cognitive activities: attention, retention, production and motivation. Observers must pay attention to the crucial details in a model's actions, and store them in memory for future use (retention). Further, observers must be able to implement, coordinate and reproduce the new skills that were learned from the model, and be motivated (e.g. identify the need and have the grit) to reproduce them.

The evidence for the effectiveness of observational learning in writing is growing and different genres have been studied. For example, Couzijn (1999) investigated the difference between "learning by doing" and "learning by observation" in a sample of 120 Dutch students' argumentative writing. The group that was trained using observational learning outperformed the other group, which suggests that observation is more beneficial than practice-only methods for complicated tasks such as writing and reading.

The effectiveness of peers as models was demonstrated by Hillocks (1986) who found larger effect sizes when feedback came from peers, rather than from teachers. The same principle is applied in collaborative learning (Graham & Perin, 2007). Rijlaarsdam et al. (2008) state that instructions must aim to stimulate students' learning-to-write capacities in order to enhance their possibilities to observe and evaluate relevant processes involved in text composition. These processes are part of the seminal, cognitive models of writing by

Hayes and Flower (1980) and Hayes (1996) and include planning, translating and reviewing, reading and focusing on the reader, and communication processes between writers and readers. Many of these processes are automatically incorporated when watching someone else producing or reflecting on a text, which explains part of the strength of the observational learning paradigm in the education of writing (Rijlaarsdam et al., 2008). Additionally, learners who watch someone else successfully perform a task may become motivated to try the same task themselves (Raedts et al., 2006).

Three factors have been identified that explain the advantage of observational learning compared to performative training (Braaksma, Rijlaarsdam, van den Bergh, & van Hout-Wolters, 2004). The first is that watching models think aloud while performing writing tasks in real-time triggers "slumbering powers in the observing students to perform such orchestration as well" (Braaksma et al., 2004, p. 3). The second is that the observing students do not write during the observational learning, so that their cognitive resources can be fully allocated to the observation. In doing this, the students can focus on the learning aspect, instead of the writing aspect (cf. Rijlaarsdam et al., 2008). The third is that observational learning students are stimulated to use observation, reflection and evaluation which helps them gain information that changes their knowledge. This reconceptualising and re-evaluating behaviour can be used in new writing tasks, and is, therefore, one reason why text quality aspects have been shown to increase after an intervention with observational learning.

Writing intervention with observational learning has been adapted for different ages, school subjects and genres. The evaluation of these studies has typically been based on the assessment of text quality before and after intervention. Some examples are argumentative writing in 10th grade students (de Smet, 2014), weak and strong writers of argumentative texts in 8th grade (Braaksma, Rijlaarsdam, & van den Bergh, 2002), argumentative writing in 9th grade (Braaksma et al., 2004; Couzijn, 1999), argumentative text writing in first year university students (Raedts et al., 2006), elementary school students' strategies for writing and organising text structure (Bouwer & Koster, 2016), and argumentative text writing in university students with and without hearing impairment (van de Weijer et al., 2018).

Aims and motivation for this study

This study has two main aims. First, we examine whether the quality of written narrative texts in 10–12 year-old secondary school students can be improved through training according to the observational learning paradigm and, if so, whether such effects are lasting. Second, we ask if intervention effects, if any, are dependent on the students' individual predispositions. The positive effects of observational learning on writing performance have been established in a variety of studies, some of which were mentioned above. It is, however, unclear whether variability in the students' predispositions may explain some of the variation found in the effectiveness of the learning paradigm. The predispositions that we measured were working memory capacity, language comprehension, reading comprehension, and gender.

The focus on narrative writing in this study is motivated by research showing that the establishment of narratives is not only a prerequisite for the development of other genres, such as expository or argumentative structures, but also for results during higher education

studies (Feagans & Appelbaum, 1986). The schema of a narrative is being learned in typically developing preschool children (Karmiloff-Smith, 1981), and is well established in children around nine years of age (Nordqvist Palviainen, 2001; Berman & Slobin, 1994; Westby & Culatta, 2016). Consequently, 11-year-old students (our target group) are expected to be aware of the structure of a narrative text but their ability to write such a text is still under development. This was demonstrated by Johansson (2009), who investigated genre development in narrative and expository texts from 10 years up to adulthood.

Working memory capacity is not often measured in writing studies that investigate observational learning (but see van de Weijer et al., 2018), but it has been identified as a crucial component for the simultaneous coordination of linguistic and cognitive activities in writing, and it is important for the retrieval and structuring of procedural and episodic information from long-term memory (McCutchen, 2011). When lower level processes, such as the transcription, have been automatised, a writer can make cognitive space available for the engagement in higher-level processes (Fayol, Hupet, & Largy, 1999). Research (e.g. Drijbooms, Groen, & Verhoeven, 2015) further suggests that transcription skills of children of the same age as those in this study play a role in the ability to coordinate actions during writing. Likewise, the ability to produce oral narratives correlates with language comprehension (Blom & Boerma, 2016; Reuterskiöld, Hansson, & Sahlén, 2011) and general school performance. Reading comprehension is another factor in writing development and has been shown to be associated with working memory capacity (Sahlén, Hansson, Ibertsson, & Reuterskiöld, 2004).

The effects of writing instruction using observational learning are generally positive, but are limited to students at higher education, primarily writing argumentative or expository texts. Therefore, we ask whether we observe similar beneficial effects in a younger age group, writing a narrative text, and whether these effects persist a period of time after the intervention. Additionally, it is an unexplored issue whether some of the variability in these effects can be explained on the basis of students' individual predispositions related to writing performance. Therefore, we explore whether the effects of our intervention are modulated by the students' gender, working memory capacity, and their reading and language comprehension. To summarise, in this study we hope to provide answers to the following research questions:

RQ (1) Does text quality improve as a consequence of the intervention?

RQ (2) Do any effects last?

RQ (3) Can working memory capacity, language comprehension, reading comprehension, or gender predict text quality or intervention effects?

Method

Participants

Six schools in a school district in the south of Sweden were initially asked to participate in the study. Two schools in areas of similar socioeconomic status accepted. There were a total of 79 students in four classes (two in each school), 35 and 44 in the two schools, respectively. All students who took part in ordinary whole-class activities were considered suitable candidates for the study.[1] Informed consent was received from 59 (34 plus

25) students and their parents. Four of these were subsequently excluded because they missed more than one intervention lesson or had too limited knowledge of Swedish. This reduced the participant group to 55 students, 32 and 23 from each school; 29 of them were girls, 26 were boys. Their average age was 11;2 (years;months) (range 10;9–11;9). For an overview of the participants, see Table 1.

The students' cognitive and linguistic predispositions were measured with standardised or norm-referenced tasks of language comprehension, reading comprehension and verbal working memory capacity (see also Table 1). Working memory capacity was measured with a 36-items subset from a standardised classroom screening test (*"Lilla Duvan"*; Wolff, 2010). The task has a process component consisting of answering general knowledge yes-or-no-questions, and a recall component consisting of remembering letters. Language comprehension was measured with the Test for Reception of Grammar-2 (TROG-2), adapted for Swedish (Bishop, 2009). This task consists of 80 sentences that need to be matched with one of four corresponding images. In an ordinary one-to-one situation, a researcher or a clinician reads the sentences and the child points to the corresponding image in a booklet. In our study, the task was administered in the classroom. The images were projected on a screen, and the students marked the matching picture in a booklet. Reading comprehension, finally, was assessed by the SL40 (Magnusson & Nauclér, 2010). This task is used to assess written sentence comprehension. The maximum possible score is 40 points. Student outcomes on each of these three tasks and expected values for this target group are displayed in Table 1. Most, but not all, students scored within the normal range, but, not unexpectedly, there were some students with poor values (i.e. more than two standard deviations below the norm) on each of the three tasks. More specifically, there were seven students who had low working-memory scores, one of whom also had a low score on the language comprehension task, and two of whom also had low scores on reading comprehension. These seven students were more or less evenly distributed over the four classes. No students performed low on all three tasks. The scores on the working memory task correlated moderately with those on language comprehension ($r = 0.50$) and on reading comprehension ($r = 0.50$). The correlation between the scores on reading comprehension and language comprehension was less strong ($r = 0.22$). The relatively large size of these correlation coefficients, was without doubt, due to the above-mentioned seven students whose scores were well below those of the remaining students.

Intervention

We developed a series of five peer observation lessons, designed to improve narrative writing. In each lesson, several short video clips were presented showing unknown

Table 1. Participant overview.

	Mean	Quartiles	Norm
		0; 0.25; 0.50; 0.75; 1	
Age (years;months)	11;2		
Gender	30F, 25M		
Working memory capacity	31.1	14; 29.5; 34; 36; 36	Mean 31.7; SD 5.8
Language comprehension	16.5	5; 16; 17; 18; 19	Mean 17; SD 3
Reading comprehension	37.5	15; 37.5; 39; 39; 40	37–39 equals percentiles 25–75

Note: Age, gender, and results on tasks of working memory capacity, language comprehension, and reading comprehension for the 55 participants are shown. Means and quartiles are displayed, as well as expected values (norm).

students, aged 10–13 years, who read, wrote and discussed other students' texts. The students in the film clips were of comparable age as the observers and had varying levels of writing skills. This was a deliberate choice, in the hope that the participants in the study would recognise the topics and the writing problems that were discussed in the videos, and would find a peer to identify with amongst the students in the film clips. This recognition has been shown to be a determining feature in learning from peers (Graham & Perin, 2007; Rijlaarsdam et al., 2008). Video peers were recorded alone or in pairs in a lab environment. Interesting sequences on different levels and on a variety of topics were selected as intervention material. During the making of the clips, the students were regularly encouraged by the researchers to "think out loud" while they performed different reading and writing activities. This was done to allow for insights into their thoughts and reasoning when reading someone else's text, suggesting improvements, or while writing their own texts. This design was inspired by Braaksma et al. (2004). Each lesson was centred around one of five writing themes, respectively: (1) the reader's perception: What does the reader find important in a story?; (2) structure: different ways to start a story; In what order should the events unfold?; (3) conclusion: How does one finish a story?; (4) editing someone else's text; (5) editing during writing: What changes do writers make while they write? Each of these themes is listed as target skills in the curriculum for grades 4–6 (Skolverket, 2016).

During the lessons, the students were divided into groups of two to six students. The topic of the lesson was always first introduced by one of the researchers. Then, the students were asked to watch four to six short (no longer than two minutes) film clips related to this topic, and to reflect on the things they agreed or disagreed with ("observation stage"). After each clip, the students discussed the videos in pairs or in groups, guided by the researchers ("reflection stage"). After having watched and discussed all the clips, students within each group summarised and shared their reflections with those in the other groups. At the end of each lesson, the students wrote down what they had learnt ("learning stage"). They formulated these learning outcomes in the form of "advice for next year's grade 5 students". The contents of the lessons were structured and scripted in advance, to ensure replicability in all four classes. All lessons were coordinated by the same two researchers, one who guided the conversations and one who took notes and administered the presentation of the film clips. More details on the intervention are provided in the Appendix. The intervention focused on narrative writing, and, consequently, narrative texts written by the students were used as an instrument of assessing the outcome of the intervention. During the course of the intervention programme, each student wrote four texts about one of the following four topics: (1) *Write about one time when you were saved from a jam you had got into, or when you saved somebody else from a jam*; (2) *Write about one time when you were hurt*; (3) *Write about one time when you were afraid*; (4) *Write about one time when you made somebody happy*. Topics 1 and 3 have been used before, and were found suitable for the age group (Johansson, 2009; Strömqvist, 1996). Topics 2 and 4 were new for this study, and chosen to be similarly easy to write about. Students wrote their texts on a laptop using a keystroke logging program (Frid, Johansson, Johansson, & Wengelin, 2014) resembling a basic word processing interface. The final texts were also converted to CHAT format for automatic analysis of lexical text characteristics (MacWhinney, 2000).

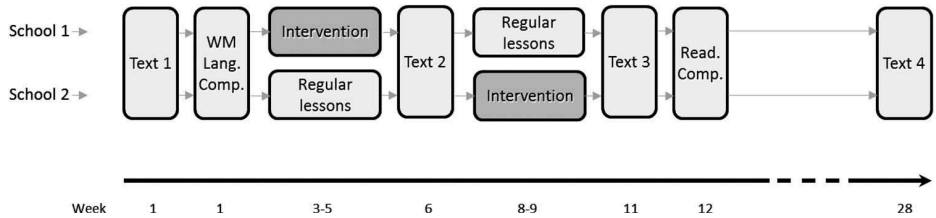

Figure 1. Schematic overview of the intervention programme in the two schools.

The students of each school followed one horizontal path. 'WM' refers to the working memory capacity task, 'Lang. Comp.' to the language comprehension task, and 'Read. Comp.' to the reading comprehension task.

Study design

The design of the study (illustrated in Figure 1) was partially determined by methodological considerations, and partially by practical constraints. Students of one of the two schools had the observational learning intervention first and regular lessons second, while those of the other school had regular lessons first, and the observational learning intervention second (cross-over design). In this design, students are "their own controls" as two classes cross from intervention to regular lessons (and the other two classes cross from regular lessons to intervention). For ethical reasons, we did not include a school to gather data but not implement the education, as observational learning has been shown to be an effective method for developing written language.

The intervention and full data collection (excluding the follow-up) took 13 weeks in total, which is short in comparison to many other intervention programmes. We chose this relatively short period so that the programme could be completed within a single term. We also chose that the intervention would take place in the children's natural environment, i.e. their own classroom, to maximise the study's generalisability (Graham & Harris, 2014). During the first week, all students wrote the first narrative text, and performed the working memory capacity task and the language comprehension task. After week 1, the students from the two schools followed different paths. Students from one school attended the five intervention lessons during weeks 3–5 (i.e. relatively early during the intervention period), while those from the other school attended these lessons during weeks 8–9 (i.e. relatively late during the intervention period). In week 6, all students wrote their second narrative text, which was after the intervention lessons for students from one of the two schools, and before the intervention lessons for students from the other school. In week 11, all students wrote their third narrative text, and took the reading comprehension task. Nineteen weeks after the last intervention week, the students wrote, as a follow-up on the intervention, the fourth narrative text (cf. Stinner & Tennent, 2012).

The tasks of working memory capacity, language comprehension and reading comprehension were administered at two occasions in order not to exhaust the students. All tasks related to the intervention were administered by the researchers. The regular Swedish lessons that took part during the period of "regular lessons" were given by the students' own teachers. During these lessons, the teachers were explicitly asked not to focus on writing instruction.

Text quality

Text quality can be described through different aspects, such as grammatical and lexical features, text organisation, or reader adaptation (Schriver, 1989; McNamara, Crossley, & McCarthy, 2010), and there is no universally used method to judge text quality. The Swedish school system does not provide any set criteria or reference texts that can be used for assessment in this age group. In this study, text quality was assessed on the basis of benchmark texts, i.e. example texts comparable to the ones written by the students, which were given low- to high-quality ratings together with a motivation for the given ratings. This method has also been used in other writing intervention studies using observational learning (see Raedts et al., 2017), and has been tested and validated by Tillema, van den Bergh, Rijlaarsdam, and Sanders (2012).

Six university students were trained to assess text quality on a scale from 0 to 100. During the training, the raters were given four benchmark texts rated 25, 40, 50 and 95, together with a motivation for each of these ratings. The four benchmark texts had been written by children from a comparable age group on similar topics, and had been collected earlier during a pilot study. The motivations for the ratings of the benchmark texts were based on aspects of genre, structure, organisation, punctuation, grammar, spelling, text length and content. They could be positive (e.g. "Punctuation is well used and sentences are well-formed") or negative (e.g. "The text feels unfinished and is difficult to understand for a reader"). During their training, the raters rated six new texts accordingly, that is, with a higher rating if they considered a text better than one of the benchmark texts, or lower when they considered it worse.

The texts collected during the intervention period were assigned at random to the six raters with the restrictions that each text was rated three times, that texts written by one student would be rated by different raters, and that each rater would rate texts from before, as well as after, the intervention lessons. All raters were told that the topics of the texts varied, but they did not know that one and the same person had produced more than one text, or in which order the texts were written, nor that the texts were part of an intervention study. The raters were told to take into consideration that the children who had written the texts were between 10 and 13 years old, and that the texts were written during 30 minutes on a computer without access to an automatic spell checker.

Analyses

The analysis focuses on the intervention effect, and whether or not this effect was modulated by any of the students' cognitive or linguistic predispositions. The intervention appeared in both schools, although at different times. The students came from four different classes from two schools. The two schools were comparable with each other, and so were the classes within the schools. The data were structured hierarchically, and a mixed-effects regression analysis was performed on the data. The outcome variable was text quality, i.e. the ratings given to the texts by the six raters. Since each text was rated by three raters and each student wrote four texts, there were 12 quality ratings for most students. Students, school classes, and raters were added as random effects. Students were nested within the four classes. Raters

were crossed with the students. Intervention was the main predictor of interest, and was coded with three levels: before intervention (i.e. text 1 for one of the two schools and texts 1 and 2 for the other school), after intervention (i.e. texts 2 and 3 for one of the schools, and text 3 for the other school), and follow-up (text 4 for both schools). Additionally, the following predictors were considered: the students' cognitive and linguistic predispositions (i.e. working memory capacity, reading comprehension, and language comprehension), and student gender. These latter four variables were added in order to establish whether any of them interacted with the effect of intervention, and to establish whether intervention would still have an effect after the effects of the other predictors had been partialled out. Since regression outputs normally show predicted estimates at zero-values of the predictors, the values of the non-categorical predictors (i.e. working memory capacity, language comprehension and reading comprehension) were centred at their medians (cf. Table 1) so that interpretation of the regression coefficients for these variables would be the predicted text quality for students with median scores on each of the variables. The intervention predictor was contrast-coded such that the regression coefficients for this predictor corresponded to the difference between pre- and post-intervention, and between post-intervention and follow-up, respectively (repeated contrasts). The analysis was done in R (version 3.4.3, R core team, 2017) using the package lme4 (Bates et al. 2015).

Results

Descriptive data of text quality

A total of 221 texts were written by the students. Three of them wrote only three texts, all the others wrote four. The texts were, for the most part, form focused, and linear in structure, i.e. they merely stated events as they occurred, and did not differentiate background from foreground. This is rather typical for texts produced in this age group. The texts varied in length from 14 to 340 words, 2 to 66 clauses, which, on average, is shorter than written narratives from comparable age groups collected in other studies (Johansson, 2009; Walldén & Åkerlund, 2008). One likely cause of this difference is that participants in these other studies wrote texts in a lab setting (supervised by the experiment leader), while the students in this study wrote texts in the classroom setting.

We did not quite succeed in receiving exactly three ratings per text. The quality of 209 texts was rated by three raters (our target number); that of the remaining 10 texts was rated by one (1), two (8) or four (1) raters. This brought the total number of ratings to 648. The ratings varied along almost the entire scale with a minimum value of 1 and a maximum value of 93 out of the possible 100 points. The ratings were highly correlated with text length (r = 0.81): shorter texts received lower ratings than longer texts. This finding is consistent with that of other studies on the text quality within this age group (Malvern, Richards, Chipere, & Durán, 2004), and presumably contributed to the fact that the inter-rater reliability (calculated on the averages of the per-text ratings) was high: Cronbach's alpha = 0.90.

Figure 2 shows how the ratings were distributed across the four texts within the four classes (labelled A to D in the figure). The figure shows that the values were positively skewed (i.e. relatively more values at the lower end of the scale), and that a few cases were considered outliers, marked as single dots above the whiskers. Furthermore, the

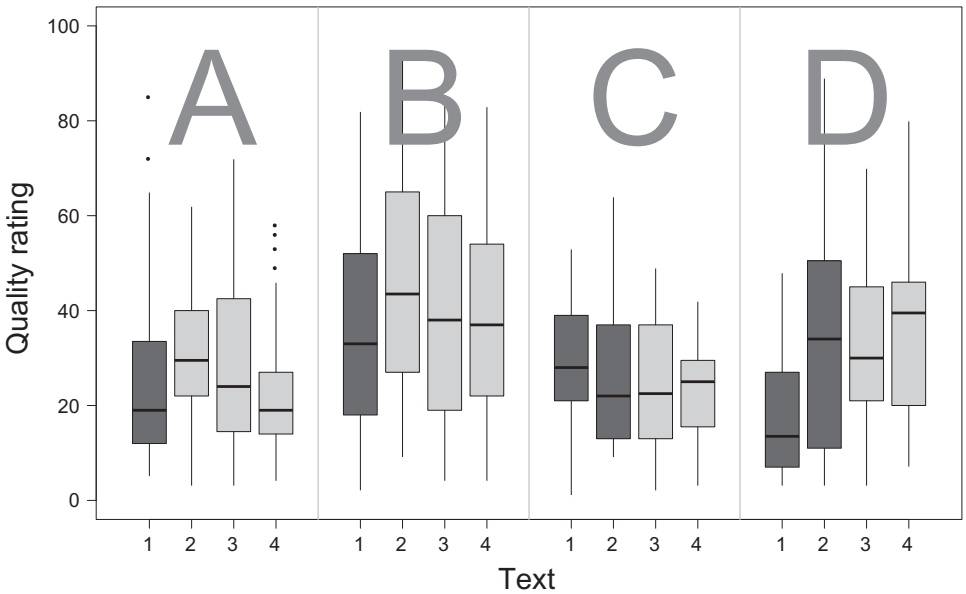

Figure 2. Boxplots of text quality ratings in the four classes.

Classes A and B are from the school with intervention first (between texts 1 and 2), and classes C and D are from the school with intervention second (between texts 2 and 3).

variability between the four classes was larger than that within the four classes. In itself, this is an interesting observation, as it implies that student achievements are not only determined by individual predispositions, but also by external factors. Most extreme ratings, on average, were obtained in classes B (highest) and C (lowest). The effect of intervention on text quality is indicated in the figure by the colour of the boxes (i.e. dark grey is pre-intervention, light grey is post intervention or follow-up). Also here, there appears to be relatively large variability between the four classes. In classes A and B, text 2, written directly after the intervention, received higher ratings than texts written directly before it. This was not the case for text 3 in classes C and D.

Intervention effects and interactions

Two choices were made for the selection of the final regression model. The first choice concerned the components (subject, school class, rater), and the structure (random intercepts or random slopes) of the random effects of the model. We compared models with different random effect structures using likelihood ratio tests. The best structure had random slopes for subjects, and random intercepts for raters and school class. In other words, all three random predictors contributed with non-zero variances to the data. There were differences in intercept due to classroom and to raters. The school class differences can be observed in Figure 2, which shows, for instance, that the ratings for the first texts written by the students in class D received lower ratings than those written by students from the other classes. Since the same protocol was followed in all classes at all times, we have no explanation for this variation.

The second choice concerned the structure of the fixed effects (gender, working memory capacity, language comprehension, reading comprehension). Several regression models (including the random effects structure described above) were considered, ranging from a model without any predictor effects to a complex one with all two-way interactions of the four predictors and intervention, including all logically possible models intermediate between these two. The models were compared on the basis of their predictive accuracy (corrected AIC values, cf. Long, 2012). The one ranked as the best (i.e. having the lowest corrected AIC value) contained main effects of gender and working memory capacity and interaction effects of intervention with language comprehension and with reading comprehension. A summary of this model is given in Table 2, and described below. We note that the second-best fitting model (with an AIC value that was only slightly different from the best) contained the interaction of intervention with working memory capacity instead of language comprehension. We mention this because the discussion below is focused on the interaction between intervention and language comprehension, but given the observed correlation between language comprehension and working memory capacity, we could have focused on the interaction of intervention and working memory capacity instead.

The column labelled "Estimates", under the block "Fixed effects" in Table 2, shows the estimated effects of the predictors. The first estimate, labelled "Intercept", is the predicted text quality before the intervention for an "average" student (i.e. a student with a median score on the tasks of working memory capacity, reading comprehension and language comprehension, and whose gender is unknown). The value is about 33 points. The amount of variability in the intercept due to the random predictors is provided in the block "Random effects". The

Table 2. Regression output.

Random effects				
Groups	Name	Variance	Std. Dev.	Corr.
Subject	Intercept	119.07	10.912	
	Intervention (pre-post)	81.33	9.018	−0.31
	Intervention (post-follow-up)	56.26	7.501	0.02–0.14
Rater	Intercept	15.15	3.892	
School class (A-D)	Intercept	23.04	4.800	
Residual	121.94	11.043		

Fixed effects			
	Estimate	Std. Error	*t*-value
Intercept	32.930	3.367	9.780
Gender	11.958	3.167	3.775
Intervention (pre-post)	6.893	1.781	3.871
Intervention (post-follow-up)	−4.259	1.645	−2.590
Language comprehension	0.278	0.798	0.348
Reading comprehension	0.068	0.460	0.147
Working memory capacity	0.814	0.304	2.679
Intervention (pre-post) × language comprehension	−0.393	0.703	−0.560
Intervention (post-follow-up) × language comprehension	−1.202	0.672	−1.789
Intervention (pre-post) × reading comprehension	1.587	0.400	3.964
intervention (post-follow-up) × reading comprehension	−0.343	0.371	−0.925

Note: Exact *p*-values are not given since there is no consensus on how they are to be determined. Significant predictors are those with an absolute *t*-value larger than 2.

variability due to subject is largest (approximately 11 points), while the variability due to classroom and rater was around four points.

The estimate of the intervention effect is given on the second row of the table. The value is a bit over six points, which, once more, is the estimated intervention effect for the same average student. The *t*-value indicates that this effect is significant. The second intervention effect is negative and indicates that the estimated text quality of the follow-up texts was significantly lower by approximately four points than that of the texts written after the intervention.

The next two rows in the table indicate how text quality was affected by gender and working memory capacity. Student gender had a particularly large effect: girls wrote texts that were rated more than 12 points higher than boys. Working memory capacity had a positive effect. The predicted text quality for two students with minimum and maximum working memory capacity (cf. Table 1) are 18.1 and 34.5, respectively.

The next two effects of language and reading comprehension are simple main effects. They indicate that texts from before the intervention written by students who scored high on these tasks were somewhat better than those written by students with low scores. Neither of these two effects is significant, however.

The last four rows in the table indicate how the effect of intervention was modulated by the language comprehension and working memory capacity. The two estimates of intervention with language comprehension are both negative. They indicate that students with higher scores on the language comprehension task improved less after the intervention and were worse at the follow up than students with lower scores. In other words, students with a relatively poor language comprehension predisposition profited more from the intervention than students with relatively good language predisposition. This effect, however, was not significant. A different pattern is observed for reading comprehension. The first interaction term for reading comprehension is positive, indicating that students with high scores on the reading comprehension task improved more from the intervention than students with low scores. This effect was clearly significant. The second interaction term is negative, suggesting that students with higher reading scores deteriorate more than students with low scores. This last term was not significant, however.

The effects of the four predictors are shown in Figure 3. This figure shows the estimated values based on the output shown in Table 2 for selected values of the predictors. Each panel shows two specific values of the individual predictors of interest: boys or girls (top left panel), low or high values (remaining three panels). The low and high values were chosen to be the lower and the upper boundary of the interquartile range of the variable. The exact values are given in Table 1. The values of the three remaining predictors not shown in a panel were the medians in each case. The intervention is included in every panel. The two upper panels show the main effects of gender and working memory capacity. The two lower panels show the interaction effects of language and reading comprehension. The size of the reading effect appears to be relatively small, which is caused by the fact that the interquartile range was only two points (cf. Table 1).

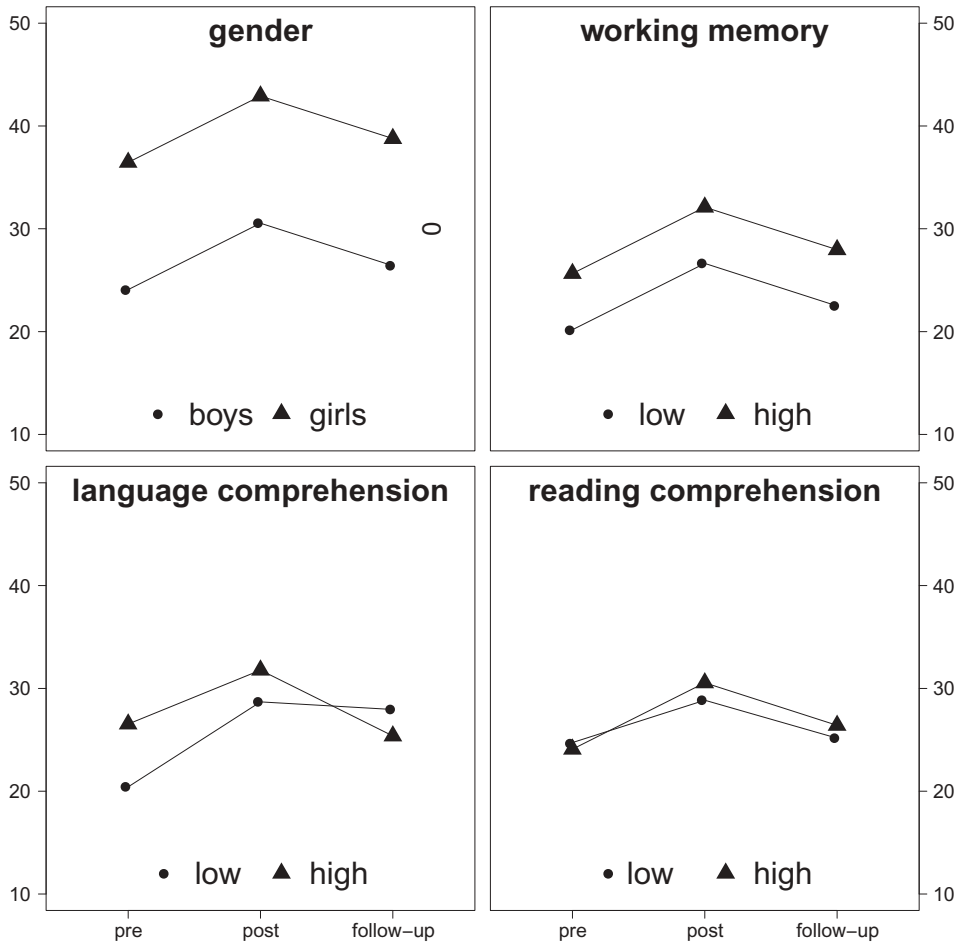

Figure 3. Estimated intervention effects.

Each panel shows estimated values for the predictors at pre- and post-intervention and at follow-up: boys or girls in the top left panel, and low or high values in the remaining three panels.

Discussion

In this study, we tested the effect of a writing intervention programme based on the observational learning paradigm. The paradigm was applied to 10–12-year-old students in grade 5 who wrote personal narrative texts about various topics. We found that the quality of texts written after the intervention was rated significantly higher than that of those written before the intervention (RQ 1). However, the improvement due to the intervention was not maintained when measured at a follow-up, three months after the intervention programme had ended. On the contrary, it had decreased significantly (RQ 2). In addition to the intervention effect, we found that girls wrote better texts than boys, and students with high scores on the working memory task wrote better texts than those with low scores. Additionally, we found that students with high scores on the reading comprehension task profited more from the intervention than those with low

scores. Finally, we found that students with lower scores on the language comprehension task profited more from the intervention than students with higher scores. This final result, though, was not statistically significant (RQ 3).

There was relatively large variability between the four school classes included in the sample. This variability concerned the overall quality of the texts written by students within each class and also the observed effect of the intervention within each class. None of these classes, though, differed from one another in any clearly observable respect. They were selected from two comparable public schools, and we did not see any indication that students within any of the four classes differed from those in the other classes, at least not in the variables that we measured. In other words, the variation that was observed should be considered random variation due to factors that were not controlled in our study.

The average quality of the texts written by the students was relatively low, and the effect of the intervention was modest. Nevertheless, some students appeared to benefit more from it than others. We found that students with high reading comprehension scores benefited significantly more from the intervention than those with low scores. In the selected model, this was the only predictor that interacted significantly with the intervention. We did not observe a direct relationship between reading comprehension and text quality: Texts written before the intervention by students with high reading scores were not rated significantly higher than texts written by students with low reading scores. These two observations suggest that students who are good readers are not necessarily also good writers. It also demonstrates the necessity of good writing instructions in order to improve writing skills in these students.

The second variable included in the selected statistical model, and which interacted with intervention, was language comprehension. The effect of language comprehension on writing intervention is tentative, since the coefficients were not significant. The results suggest that children with high language comprehension scores gained less from the intervention and their scores at follow up went down more than those of students with low scores. We think that this finding deserves further investigation. It seems that, at the age range of our sample, language comprehension is far ahead of written language production. Children at this age can understand sentences more complex than they are likely to use in writing. Less extremely stated, even children with relatively poor language comprehension scores understand sentence constructions that are complex enough to constitute sentences in narrative writing.

The other two predictors (gender and working memory capacity) did not interact significantly with the intervention, but they both had relatively strong effects on text quality. The gender effect is not unexpected. In Sweden, 11-year-old girls, on average, have better linguistic and writing skills than boys of the same age, and, as a consequence, do better at school. The effect of gender has been frequently observed in other studies as well (e.g. Kanaris, 1999), and several explanations of the effect have been proposed, either related to overall maturity or to social expectations.

There is ample evidence that working memory capacity plays a role in writing (e.g. McCutchen, 2011), especially for children who have not yet automatised lower-level writing skills, such as transcription (Drijbooms et al., 2015). Thus, the connection between working memory capacity and text quality is not an unexpected finding in

our results. Children with better working memory capacity are able to write texts with better text quality, due to their increased capacity to coordinate processes during writing. The effect of intervention was not modulated by working memory capacity in the statistical model that we selected, but it was in many other models that we considered. The results suggest that in future studies we must explore the correlation between working memory capacity, transcription skills, text length, and text quality. In addition, future development of interventions of this kind should perhaps more carefully consider aspects of working memory capacity in this age group. Many other studies using observational learning (cf. Raedts et al., 2017), have applied their interventions to older age groups, where basic lower-level writing processes, such as transcription, is, to a greater extent, automatised and established.

The overall relapse that we observed at follow-up may be discouraging for the evaluation of the intervention design, and it calls for some rethinking of the value of a short intervention period, or of the content and focus of the intervention lessons. There are several possible reasons why the effect was not lasting. The instructional method is not well known in Sweden, and unknown to the teachers of the schools that took part in the intervention. It is possible that the intervention effect would have had a longer-lasting effect if it was carried out under the supervision of the regular teachers. We have preliminary evidence that the intervention effects of our lessons are larger when performed by teachers who were trained in the method (and who used the same scripted instructions, film-clips and lesson design) instead of researchers (Hammarstedt & Karan, 2016), but we do not know now whether these effects would also be longer lasting. We believe that including a follow-up constitutes one of the assets in our study, and that it represents a guiding factor to develop better methods for writing instruction. After further studies, observational learning may become a tool that can make up for the lack of standardised writing methods that currently exist in Sweden.

Conclusion

The study shows that it is possible to implement a short observational learning intervention for 10–12 year-olds. The paradigm worked well as an educational tool for the improvement of narrative writing. Text quality improved, and this improvement was statistically significant in spite of large variation between the four classes that participated in the study. Somewhat discouragingly, however, the effects seemed to have disappeared three months after the end of the intervention. This suggests that writing education needs to be implemented into the school curriculum on a regular basis, something which, unfortunately, is not the case at present in the Swedish schools. Our results also suggest that individual cognitive and linguistic predisposition may interact with intervention effects. We think that this is an interesting finding that warrants further investigation. As a first step, this finding should be replicated. Next, the question needs to be asked why some students gain more from observational learning intervention than others, and finally, how the method can be adapted so that the student group who showed a comparably weak improvement can benefit more from an intervention than what they did in this study. We leave these questions for future research.

Note

1. Following the advice from the regional ethics board (EPN dnr. 2013/270) no students were *a priori* excluded.

Acknowledgements

The study was designed in collaboration with Martine Braaksma and Huub van den Bergh, who both gave valuable input on design and statistical analyses. We also thank the Linnaeus environment Cognition, Communication and Learning (CCL) at Lund University for fruitful discussions and financial support, as well as the Lund University Humanities Lab for use of facilities and equipment. For preparation of film material we are grateful to Ulf Södergren.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the Marcus and Amalia Wallenberg foundation under Grant 2012.0038.

ORCID

Emily Grenner http://orcid.org/0000-0002-6818-1093
Lena Asker-Árnason http://orcid.org/0000-0002-5630-3145
Joost van de Weijer http://orcid.org/0000-0001-9843-3143
Victoria Johansson http://orcid.org/0000-0002-0684-6491
Birgitta Sahlén http://orcid.org/0000-0002-8468-0546

References

Bandura, A. (1997). *Self-efficacy: The exercise of control*. Basingstoke: W. H. Freeman.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

Berman, R. A., & Slobin, D. I. (1994). *Relating events in narrative: A crosslinguistic developmental study*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Bishop, D. V. M. (2009). *Test for reception of grammar, version 2. Swedish version*. Stockholm: Pearson Education.

Blom, E., & Boerma, T. (2016). Why do children with language impairment have difficulties with narrative macrostructure? *Research in Developmental Disabilities*, *55*(Supplement C), 301–311.

Bouwer, R., & Koster, M. (2016). *Bringing writing research into the classroom. The effectiveness of tekster, a newly developed writing program for elementary students* (Unpublished dissertation). Utrecht University, Utrecht.

Braaksma, M. A. H., Rijlaarsdam, G., & van den Bergh, H. (2002). Observational learning and the effects of model-observer similarity. *Journal of Educational Psychology*, *94*(2), 405.

Braaksma, M. A. H., Rijlaarsdam, G., van den Bergh, H., & van Hout-Wolters, B. H. A. M. (2004). Observational learning and its effects on the orchestration of writing processes. *Cognition and Instruction*, (1), 1.

Braaksma, M. A. H., Rijlaarsdam, G., & van den Bergh, H. H. (2018). Effects of hypertext writing and observational learning on content knowledge acquisition, self-efficacy, and text quality: Two experimental studies exploring aptitude treatment interactions. *Journal of Writing Research*, *9*(3), 259–300.

Brandt, D. (2015). *The rise of writing: Redefining mass literacy*. Cambridge: Cambridge University Press.

Couzijn, M. (1999). Learning to write by observation of writing and reading processes: Effects on learning and transfer. *Learning and Instruction*, *9*(2), 109–142.

de Smet, M. J. R. (2014). *Composing the unwritten text: Effects of electronic outlining on students' argumentative writing*. Proefschrift Open Universiteit. Heerlen: Open Universiteit.

Drijbooms, E., Groen, M. A., & Verhoeven, L. (2015). The contribution of executive functions to narrative writing in fourth grade children. *Reading and Writing: an Interdisciplinary Journal*, *28*(7), 989–1011.

Fayol, M., Hupet, M., & Largy, P. (1999). The acquisition of subject-verb agreement in written French: From novices to experts' errors. *Reading and Writing*, *11*(2), 153–174.

Feagans, L., & Appelbaum, M. I. (1986). Validation of language subtypes in learning disabled children. *Journal of Educational Psychology*, *78*(5), 358–364.

Fidalgo, R., Torrance, M., & García, J.-N. (2008). The long-term effects of strategy-focussed writing instruction for grade six students. *Contemporary Educational Psychology*, *33*, 672–693.

Frid, J., Johansson, V., Johansson, R., & Wengelin, Å. (2014, February). *Developing a keystroke logging program into a writing experiment environment*. Presentation at Writing Across Borders, WRAB, Paris.

Graham, S., & Harris, K. (2014). Conducting high quality writing intervention research: Twelve recommendations. *Journal of Writing Research*, *6*(2), 89–123.

Graham, S., & Perin, D. (2007). A meta-analysis of writing instruction for adolescent students. *Journal of Educational Psychology*, *99*(3), 445–476.

Graham, S., & Rijlaarsdam, G. (2016). Writing education around the globe: Introduction and call for a new global analysis. *Reading and Writing: an Interdisciplinary Journal*, *29*(5), 781–792.

Hammarstedt, M., & Karan, V. S. (2016). *Vilken effekt har lärande genom observation på elevers narrativa skrivande i årskurs 5? En studie av skrivprodukter och skrivprocesser* [What effect does learning by observation have on narrative writing in students in schoolyear 5?]. (Master thesis). Department of Logopedics, Phoniatrics and Audiology, Lund University, Lund.

Harris, K. R., Graham, S., & Mason, L. H. (2006). Improving the writing, knowledge, and motivation of struggling young writers: Effects of self-regulated strategy development with and without peer support. *American Educational Research Journal*, *2*, 295.

Hayes, J. R. (1996). A new framework for understanding cognition and affect in writing. In C. M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences, and applications* (pp. 1–27). Mahwah, NJ: Lawrence Erlbaum Associates.

Hayes, J. R., & Flower, L. S. (1980). Identifying the organization of writing processes. In L. W. Gregg & E. R. Steinberg (Eds.), *Cognitive processes in writing* (pp. 3–30). Hillsdale, NJ: Erlbaum.

Hillocks, G., Jr. (1986). *Research on written composition: New directions for teaching. National Conference on Research in English*, Urbana, IL: Eric Clearinghouse on Reading Communication Skills.

Johansson, V. (2009). *Developmental aspects of text production in writing and speech*. (Thesis). Department of Linguistics and Phonetics, Centre for Languages and Literature, Lund University, Lund.

Kanaris, A. (1999). Gendered journeys: Children's writing and the construction of gender. *Language and Education*, *13*(4), 254–268.

Karmiloff-Smith, A. (1981). The grammatical marking of thematic structure in the development of language production. In W. Deutsch (Eds.), *The child's conception of language* (pp. 121–147). London: Academic Press.

Long, J. D. (2012). *Longitudinal data analysis for the behavioural sciences using R*. London: Sage Publications, Inc.

Lundgren, U. P., Säljö, R., & Liberg, C. (Eds.). (2010). *Lärande, Skola, Bildning*. Stockholm: Natur & Kultur.

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum.
Magnusson, E., & Nauclér, K. (2010). *SL40: läsförståelse: meningar* [SL40: Reading comprehension: Sentences] (2nd ed.). Lund: Studentlitteratur AB.
Malvern, D. B., Richards, J., Chipere, N., & Durán, P. (2004). *Lexical diversity and language development: Quantification and assessment*. Basingstoke: Palgrave Macmillan.
McCutchen, D. (2011). From novice to expert: Implications of language skills and writing- relevant knowledge for memory during the development of writing skill. *Journal of Writing Research*, *3*, 51–68.
McNamara, D., Crossley, S., & McCarthy, P. (2010). Linguistic features of writing quality. *Written Communication*, *27*, 57–86.
Myhill, D. (2008). Towards a linguistic model of sentence development in writing. *Language and Education*, *22*(5), 271–288.
Myhill, D. (2009). Becoming a designer: Trajectories of linguistic development. In R. Beard, D. Myhill, M. Nystrand, & J. Riley (Eds.), *SAGE handbook of writing development* (pp. 402–414). London: SAGE Publications.
Nordqvist Palviainen, A. (2001). *Speech about speech: A developmental study on form and function of direct and indirect speech*. (Thesis). EBSCOhost.
Raedts, M., Rijlaarsdam, G., van Waes, L., & Daems, F. (2006). Observational learning through video-based models: Impact on students' accuracy of self-efficacy beliefs, task knowledge and writing performances. In G. Rijlaarsdam (Series Ed.), & P. Boscolo & S. Hidi (Vol. Eds.), *Studies in writing, Volume 19: Writing and motivation* (pp. 219–238). Oxford: Elsevier.
Raedts, M., van Steendam, E., Grez, L., Hendrickx, J., & Masui, C. (2017). The effect of different types of video modelling on undergraduate students' motivation and learning in an academic writing course. *Journal of Writing Research*, *8*(3), 399–435.
Reuterskiöld, C., Hansson, K., & Sahlén, B. (2011). Narrative skills in Swedish children with language impairment. *Journal of Communication Disorders*, *44*(6), 733–744.
Rijlaarsdam, G., Braaksma, M., Couzijn, M., Janssen, T., Raedts, M., Van Steendam, E., … van den Bergh, H. (2008). Observation of peers in learning to write: Practice and research. *Journal of Writing Research*, *1*(1), 53–83.
Sahlén, B., Hansson, K., Ibertsson, T., & Reuterskiöld, C. (2004). Reading in children of primary school age - A comparative study of children with hearing impairment and children with specific language impairment. *Acta Neuropsychologica*, *4*, 393.
Schriver, K. A. (1989). Document design from 1980 to 1989: Challenges that remain. *Technical Communication*, *4*, 316.
Skolverket. (2016). *Läroplan för grundskolan, förskoleklassen och fritidshemmet 2011: Reviderad 2016*. Stockholm: Author.
Stinner, D. J., & Tennent, D. J. (2012). Losses to follow-up present a risk to study validity. *AAOS Now*, *6*(2), 38.
Strömqvist, S. (1996). *Discourse flow and linguistic information structuring: Explorations in speech and writing*. Gothenburg papers in theoretical linguistics, 78. Göteborg: University of Göteborg.
Tillema, M., van den Bergh, H., Rijlaarsdam, G., & Sanders, T. (2012). Quantifying the quality difference between L1 and L2 essays: A rating procedure with bilingual raters and L1 and L2 benchmark essays. *Language Testing*, *30*(1), 71–97.
van de Weijer, J., Åkerlund, V., Johansson, V., & Sahlén, B. (2018). Writing intervention in university students with normal hearing and in those with hearing impairment: Can observational learning improve argumentative text writing? *Logopedics Phoniatrics Vocology*, 1–9.
Walldén, R.-M., & Åkerlund, V. (2008). *Muntlig och skriftlig berättarförmåga samt läsförståelse hos barn med typisk språkutveckling i år 5 och 6* [Oral and written narrative skills and reading comprehension in children with typical language development in school years 5-6]. (Master thesis). Department of Logopedics, Phoniatrics and Audiology, Lund University, Lund.
Westby, C., & Culatta, B. (2016). Telling tales: Personal event narratives and life stories. *Language, Speech & Hearing Services in Schools*, *47*(4), 260–282.
Wolff, U. (2010). *Lilla DUVAN*. Stockholm: Hogrefe Psykologiförlaget AB.

Appendix: Procedure overview

Narrative writing assessments (Texts), formal tasks, intervention lessons (Int.) and regular lessons. Students from one of the two schools followed the schedule below, whereas students from the other school had regular lessons weeks 2–4 and intervention lessons during weeks 8–9. The written material was collected by the researchers, edited for clarity, and given back as a printed booklet of the students' own writing advice.

Week	Assessment	Lesson	Content
1	Formal tasks		Working memory capacity task and language comprehension task.
1	Text 1		"Write about one time when you were saved from a jam you had got into, or when you saved somebody else from a jam".
2		Int. 1	**Theme**: Reader perception: What does the reader find important in a story? **Observation**: Film peers read and discuss what made them interested in, moved by, or disinterested in, unmoved by a story. **Reflection in groups of 3–4**: Following several rounds of film clips, students discuss the films. Relating reflections to the whole class. **Learning**: Students write down their best advice for "next year's fifth-graders".
2		Int. 2	**Theme**: Structure: Different ways to start a story; in what order should the events unfold? **Observation**: Film peers read narratives that have been edited to start at the beginning, at a pivotal point, or at the end. They discuss stories with different structures: beginning – middle – end; middle – beginning – end; end – beginning – middle – beginning. **Reflection in pairs**: What happens to the reader when a story begins somewhere that is not the beginning? Would you try it? **Learning**: Students are asked to discuss in pairs how to start an exciting story, and to write down the first sentences of it. Writing advice for "next year's fifth-graders".
3		Int. 3	**Theme**: The end: How to close a story **Observation**: Film peers think aloud while writing the ends of their narratives. **Reflection in groups of 4–5**: Students reflect on what a good ending of a narrative entails. Within their groups, students rank the examples they have seen/heard in the film clips. Groups relate their lists to the class, describing how ranking was carried out. **Learning**: In pairs, listing criteria of what makes a good ending.
3		Int. 4	**Theme**: Editing someone else's text **Observation**: Film peers think aloud while reading a friend's text, proposing revisions on content and form. **Reflection in groups of 3–4**: What was better and poorer advice? **Learning**: Students decide what advice they will pick up for next time they write.
4		Int. 5	**Theme**: Editing during writing: What changes do writers make while they write? **Observation**: Films show narrative texts as they are being written by the film peers. Examples of editing of typos, changing content words to improve the story, and explaining events in more detail after reading through the text. **Reflection in groups of 3–4**: Students reflect on what changes were made, and why. **Learning**: Writing down "advice for next year's fifth-graders", on editing and revising, as well as any advice from intervention lessons 1–4.
6	Text 2		"Write about one time when you were hurt".
8–9			*(Intervention in the other school)*
11	Text 3		"Write about one time when you were afraid".
12	Formal task		Reading comprehension task.
13–27			Regular lessons, not focusing on narrative writing.
~28	Text 4		"Write about one time when you made somebody happy".

Paper II

LOGOPEDICS PHONIATRICS VOCOLOGY
https://doi.org/10.1080/14015439.2021.1881613

Taylor & Francis
Taylor & Francis Group

RESEARCH ARTICLE

OPEN ACCESS Check for updates

Predictors of narrative text quality in students with hearing loss

Emily Grenner[a], Joost van de Weijer[b], Victoria Johansson[c] and Birgitta Sahlén[a]

[a]Department of Clinical Sciences, Logopedics, Phoniatrics and Audiology, Lund University, Lund, Sweden; [b]Lund University Humanities Lab, Lund University, Lund, Sweden; [c]Centre for Languages and Literature, Lund University, Lund, Sweden

ABSTRACT

Aim: Students with hearing loss (HL) often fall behind hearing peers in complex language tasks such as narrative writing. This study explored the effects of school grade, gender, cognitive and linguistic predisposition and audiological factors on narrative text quality in this target group.

Method: Eleven students with HL in Grades 5–6 and 7–8 (age 12–15) who took part in a writing intervention wrote four narrative texts over six months. A trained panel rated text quality. The effects of the students' working memory capacity, language comprehension, reading comprehension, school grade and gender and the intervention were analyzed as a mixed-effects regression model. Audiological factors were considered separately.

Results: The analysis showed that throughout the period, texts written by female students in Grade 7–8 received the highest text quality ratings, while those written by male students in Grade 7–8 received the lowest ratings. There was no effect of the intervention, or of the linguistic and cognitive measures. The students with the lowest text quality ratings received amplification later than those with high ratings, but HL severity was not associated with text quality.

Conclusion: Hearing loss severity was not a decisive factor in narrative text quality. The intervention which the students took part in is potentially effective, with some adaptation to the special needs of students with HL. The strong gender effects are discussed.

ARTICLE HISTORY
Received 27 January 2020
Revised 18 January 2021
Accepted 20 January 2021

KEYWORDS
Secondary school; narrative writing; observational learning; peer observation

Introduction

In spite of substantial pedagogical, technical and medical advances in the field, many students with hearing loss (HL) lag behind peers with normal hearing (NH) in complex language skills, notably reading and writing [1–3]. This may be due to slow development of language skills such as spoken language comprehension, reading comprehension and spoken narration, all essential for the development of writing skills. It is not surprising then that students with HL do not reach the goals stated in school curriculum to the same extent as students with NH [4]. Only 10–15% of Swedish high school graduates with HL proceed to higher education, compared to around 60% of graduates with NH [5,6].

The present study focuses on writing skills in students with HL in Sweden. References to grades and learning goals apply to the Swedish school system and curriculum. One writing skill that is part of the school curriculum is the production of narrative texts. Early establishment of a narrative structure is essential for academic success [7–10]. Moreover, being able to produce a written narrative is a prerequisite for the development of other genres, such as expository or argumentative texts [8]. In spoken language, the narrative structure is well-established in six-year-olds with typical language development [11]. At the age of nine, children generally have access to the schema of a well-formed personal narrative [12–14]. They know that it consists of an introduction, a sequence of events and an ending, and a description of the characters. The curriculum states that knowledge of the narrative structure should be established by the end of Grade 3, when children are nine to ten years of age [15]. It is however not specified to what extent students should be able to demonstrate this knowledge in written narratives [15]. The curriculum furthermore states that students from Grades 4–9 (age 10–16), should expand their writing skills to other genres, process and revise their text, learn to give and receive feedback, write by hand and on the computer, and learn to organize and edit a text [15]. In addition, they should learn to correctly use subordinate clauses, parts-of-speech, morphology, spelling rules, punctuation, and text cohesion [15]. Further, they should learn which features, such as content and lexicon, are typical for different genres [15]. In other words, the processes involved in writing are demanding, even for writers with fundamental transcription skills [16]. Swedish female students reach the curriculum goals to a greater extent than male students do [17]. This is corroborated by for instance Kanaris [18] who noticed that girls often are good writers and boys are under-achievers. Similarly, Myhill [19] reported that 8- to 10-year-old girls' texts were comparatively longer and more

CONTACT Emily Grenner emily.grenner@med.lu.se Department of Clinical Sciences/Logopedics, Phoniatrics and Audiology, Lund University hospital, Lund, S-221 85, Sweden

complex, and focused more on description and elaboration than boys' texts.

At present, it is unclear how the curricular goals in writing may be obtained, despite recent meta-analyses of writing intervention [20]. Neither the curriculum nor the teacher training specifies what type of writing instruction should be used, whether for students with NH or HL. Teachers and schools are free to choose their own methods of instruction. The effects of HL on narrative writing are not clearly identified. One reason for this is the heterogeneity in audiological, cognitive and linguistic predispositions in students with HL, and the complex relationships between these variables. Approximately 20–30% of all school children with HL have language learning problems comparable to those of children with NH diagnosed with a developmental language disorder [21,22]. The degree of HL is one relevant factor in this, but, at the same time, it is seldomly proportional to the severity of language learning difficulties. This suggests that cognitive and linguistic factors need to be taken into consideration to explain why certain students with HL perform better than others. For instance, cognitive resources are particularly taxed by a degraded speech signal due to noise, speaker's voice or poor amplification. Thus, students with HL have less resources left for a school task [23]. There is ample evidence that working memory plays a crucial role in writing [24,25]. The writer of a narrative text must recall what happened, plan and organize the events, make lexical choices, formulate sentences correctly and at the same time think about a range of formal aspects like spelling and punctuation. The comprehension of spoken language has been associated with spoken narrative skills as well as with reading comprehension and school performance [7,26] in students with NH. Reading comprehension and working memory capacity have been shown to be associated with each other in students with HL [27,28].

Research on writing in students with HL is scarce. Some aspects of written narratives were studied in 11- to 19-year-old students with HL and controls with NH [29]. The students with HL wrote fewer complex sentences and used fewer function words, but they were not significantly different from a control group in spelling accuracy. Students with HL had to allocate most of their cognitive resources to language processing and did not have sufficient resources available for the organisational and formal aspects of a text, the authors concluded [29]. Only a few documented and evidence-based writing intervention models are available for this target group. Strassman and Schirmer [30] reviewed teaching practices for students with HL and found that methods for teaching writing fell into four categories: teaching the writing process itself, looking at properties of finished texts, writing to facilitate content learning, and feedback on writing. There was, however, no clear evidence of effects [30]. More recently, it has been suggested that a combination of three elements (i.e. strategy instruction, teacher–student dialogue, and teaching language skills and metalinguistic awareness) improves writing in students with HL [31,32]. One intervention method which includes some of these elements and that others found effective for improving writing [20] is observational learning, which has been found to boost writing in students with NH [33–37].

To summarize, the ability to produce a written narrative is a prerequisite for the production of other text genres and for academic achievement. The results from previous studies of texts written by students with NH suggest that girls are better writers than boys. There is a lack of research on associations between HL and narrative writing. One reason is the large heterogeneity among students with HL. Another is the complicated interaction of HL with other linguistic and cognitive predispositions which may also influence text writing. Finally, neither writing skills in students with HL nor methods for teaching writing for these students have been studied extensively. With this in mind, the current study was carried out.

The present study

The aim of the present study is to identify possible predictors of narrative text writing in students with HL. Eleven students from two school classes, Grades 5–6 and 7–8, with varying HL severity wrote four texts which were graded by a rater panel. The students' cognitive and linguistic predispositions (i.e. working memory capacity, language comprehension, and reading comprehension) were assessed, and data on audiological factors were collected. The present study is part of a comprehensive study in which students with NH and students with HL followed the same writing intervention [36].

The analysis is guided by the following questions:

- What are the effects of working memory capacity, language comprehension, reading comprehension, school grade and gender on narrative text quality in students with HL?
- What associations are there between degree of HL and age at amplification and text quality?
- Is observational learning suitable for the training of narrative text writing in students with HL?

Method

Students

Head teachers and teachers of classes exclusively for students with HL were contacted. In two classes, one a combined Grades 5–6 and the other a combined 7–8, the teachers accepted to participate. The students and teachers in the classes communicated with spoken language sometimes supported with sign language. The classrooms were equipped with hearing loops (FM system) with microphones for students and teachers. The students followed the same curriculum as NH students. The total number of students in the two classes was 19. No student was a priori excluded. Six students chose not to participate in the study, but were nevertheless present during the data collection, as recommended by the Regional Ethical Review Board. Data from these students were discarded. In addition, two students

missed two or more intervention lessons and their data were also excluded from analysis. Thus, the results of 11 students (5 girls and 6 boys) were used for this study. Table 1 provides an overview of the group.

The students in Grade 5–6 were between 12;5 and 13;8 years old, and those in Grade 7–8 between 13;4 and 15;3 years old. The parents of ten students provided information about what (spoken or signed) languages were used at home, and some audiological data of their children. Additionally, audiological records of nine students were consulted, with the consent of the parents. There was thus considerable variation in the detail of available audiological data. For most students, data on better ear hearing level (BEHL) and type of amplification was obtained. The students' degree of HL had been categorized according to the classification by [38]. This means that a mild HL constitutes a BEHL of 20–40 dB, moderate, 40–70 dB, severe, 71–90, and a profound HL constitutes a BEHL over 90 dB. Two students had unilateral HL with severe HL on the afflicted ear, while the remaining nine had bilateral HL varying from mild to profound. The age at diagnosis varied from three months to ten years. One student had no amplification. The other students had one or two hearing aids (HA) or bimodal amplification, i.e. one HA and one cochlear implant (CI). Ten students spoke Swedish as their first language and one had another European spoken first language. Some students were exposed to one or more additional spoken languages or sign language, as well as signing as a form of alternative and augmentative communication.

This study was carried out in accordance with the recommendations of The Swedish Ethical Review Authority and the protocol was approved by the Regional Ethical Review Board in Lund (Dnr. 2013/270). The parents and students gave written informed consent in accordance with the Declaration of Helsinki.

Narrative texts and text quality ratings

The students' writing performance was assessed four times: one week before and one week after the writing intervention (see below), after another six weeks, and once more after the summer vacation five months later. The same time intervals between the first three texts were also used by Grenner et al. [36]. The topics of the four texts were, respectively: 1) Write about a time when you were saved from a jam, or when you saved someone else from a jam; 2) Write about a time when you were hurt; 3) Write about a time when you were afraid; 4) Write about a time when you made somebody happy. These topics have also been used in previous studies and found suitable for the age group [14,36,39]. The students wrote the texts on a laptop using ScriptLog [40], a keystroke logging program with a basic word processing interface. The topic for the narratives was written on a slide in the classroom and were also read aloud to the students. At each time, the students had 30 min for writing.

The quality of the texts was subsequently rated by a panel of six raters with a method validated by [41] and used in comparable studies of writing [36,42]. Before the start of the rating procedure, the raters were shown several benchmark texts on the same topics with quality ratings marked on a 100-mm visual analogue scale, together with a motivation why each text received a specific rating. The rating was holistic, but based on structure and organization, content, grammar and spelling. In the rating procedure, each text was rated in the same way by three of the six raters, yielding three scores for each text that could range from 0 to 100. The raters were unaware of the order in which the texts had been written and they were not informed that the texts had been written by students with HL or that each student had written several texts. Interrater reliability was calculated on a larger dataset including the texts in the present study and other texts on the same topics and was found high (Cronbach's alpha .90). More details on the rating procedure are provided in [36].

Cognitive and linguistic tasks

The students were given norm-referenced or standardized tests of verbal working memory capacity [43], language comprehension [44] and reading comprehension [45]. For practical reasons, the tests of working memory capacity and language comprehension were administered after the first text was written, while reading comprehension was tested after the second text was written. All tests were administered by the researchers in the classroom in the absence of the teachers.

Working memory capacity was assessed with a 36-items subset from a classroom screening test [43]. The items were pre-recorded and were presented to the students *via* the hearing loop. The working memory test has a process component (general knowledge yes/no-questions), and a recall component (remembering letters). As an example, the students heard the letter "B" and were asked "Is France larger than Denmark?". They responded by holding up a YES or a NO sign. Then they would hear the letter "J" and were asked "Is a bird a mammal?". Again they held up a YES or NO sign, and wrote down the two letters, in the right order. The average for students with NH in Grade 5 is 31.7 (SD 5.9) and 32.9 (SD 4.9) for students in Grade 7. Language (listening) comprehension was measured with the Test for Reception of Grammar–2, adapted for Swedish [44]. This test consists of 80 spoken sentences that each needs to be matched with one of four pictures. The sentences are divided into 20 blocks of four sentences which are scored as correct if all responses within a block are correct. This yields a possible maximum score of 20, and the expected score for normal-hearing students is approximately 17 for grade 5 and 18 for grade 7. In the present study, a research assistant read each sentence aloud, using a microphone connected to the hearing loop. The pictures were projected on a screen, and the students marked the matching picture in a booklet. Reading comprehension, finally, was assessed using the SL40 [45]. In this test, the student reads one sentence at a time and chooses a corresponding picture. The maximum score is 40 points, and normal-hearing students in grade 5

Table 1. Student overview.

Case	Gender	Class	Hearing loss	Age at diagnosis	Amplification	Age at amplification	Working memory	Language comprehension	Reading comprehension
1	F	5–6	NH/severe, unilateral	10	1 HA	10	34	16	34
2	M	5–6	Moderate, bilateral	1	2 HA	4	34	12	38
3	M	5–6	Mild, bilateral	4	2 HA	4	34	16	35
4	M	5–6	Unknown, bilateral	6	2 HA	6	34	12	37
5	F	5–6	NH/severe, unilateral	4	none	–	22	10	25
6	F	7–8	Profound, bilateral	0:3	HA + CI	1	26	17	40
7	F	7–8	Severe, bilateral	3	HA + CI	3	15	15	40
8	F	7–8	Moderate, bilateral	4	2 HA	4	36	18	37
9	M	7–8	Moderate, bilateral	5	2 HA	6	36	17	38
10	M	7–8	Unknown, bilateral	N/A	2 HA	3	32	11	36
11	M	7–8	Moderate, bilateral	10	2 HA	10	28	12	20

Students' school class (Grade 5–6 or Grade 7–8), degree of hearing loss, age at diagnosis (in years) and age at amplification (in years), and their results on tests of working memory, language (listening) comprehension and reading comprehension.

normally have a score of at least 38 items correct. Older students are expected to score all items correct.

An overview of the scores on the three tests is displayed in Table 1. Seven students had scores on the working memory test which were higher than or equal to the reference value [43]. The scores of the remaining four, though, were well below this value. The average score from the Grade 7–8 students was not higher than that from the Grade 5–6 students. There was a slight tendency that the boys had higher scores than the girls. All scores from the students in Grade 5–6 were below the age norms for students with NH on the language comprehension test, while three out of the six students in Grade 7–8 had scores which were below the age norms [44]. This suggests that the level of language comprehension was, relatively speaking, somewhat less behind in the higher grade than in the lower grade. Five students had scores below the 10th percentile which is a common cut-off for language disorder. The average score of the Grade 5–6 students was slightly below that of the Grade 7–8 students. Only four students had a score on the reading comprehension test of 38 or more, i.e. the norms of fifth-grade students with NH [45].

Writing intervention

The students took part in a writing intervention based on observational learning [36]. In observational learning, learners watch films of models (usually peers) who perform and comment upon a writing task [35]. In this way, observation and reflection are separated from writing and practice [35]. As a consequence, learners do not have to draw on cognitive resources while simultaneously performing the target skill. The paradigm has been found to have a positive effect on writing skills [33–37]. In a recent study, 55 students with NH from Grade 5 took part in the lessons. A modest but significant increase in text quality was found after the intervention, and this effect was somewhat more pronounced in students with relatively low language comprehension scores [36]. In the present study, five 40 min lessons were given during three consecutive weeks. Each lesson focused mainly on one aspect of narrative writing, targeted in the curriculum [15]. See [36] for a description of the content and the execution of the lessons.

Table 2. Individual text quality ratings.

ID	Text 1	Text 2	Text 3	Text 4	Average
1	43	55	52	24	44
2	32	14	29	19	23
3	31	35	23	21	28
4	20	–	9	11	13
5	19	15	21	39	24
6	75	73	79	80	77
7	74	64	75	89	76
8	49	54	45	66	54
9	26	8	12	6	13
10	21	20	24	22	22
11	13	15	20	7	14

Text quality ratings for each student and text, and average of each student's text quality ratings.

Analyses

Various mixed effects models were used to estimate the effects of six predictors: time of measurement (four texts), Grade (5–6 or 7–8), gender, working memory, language comprehension and reading comprehension. Interaction effects between time of measurement and the remaining measures were also considered. Models were compared on the basis of AIC values. The computations were done in R version 3.5.3 [46], using the package lme4 [47]. The degree of HL was not used as a predictor in the statistical analysis as the information was missing for some students. The effects of HL will be presented and discussed separately.

Results

One student only wrote three texts and one other text was not rated. Consequently, 126 quality ratings of 42 texts were collected and used for data analysis. The individual quality ratings for each text are shown in Table 2. The average ratings for each text varied from 13 to 77 suggesting large differences in writing skills among the students. Across the four texts, the ratings within each student were relatively consistent, which suggest a constant performance not affected by the writing intervention or general development to the follow-up the next semester. The texts written by the Grade 7–8 students received on average considerably higher ratings (10–16 points) than those written by the Grade 5–6 students. The values in Table 2 show that most students were given a lower score on Text 2 than on Text 1.

Several regression models with different groupings of the six predictors (text, grade, gender, working memory, language comprehension and reading comprehension) were evaluated. The predictor Text was never excluded from these models, but any of the other five predictors could be used to see if the exclusion of this predictor made the model fit significantly worse or not. The list included models with only main effects as well as models with two-way interactions between the predictors. An overview of the models that were compared is given in Appendix, with the chosen model indicated in boldface. The random effects in all models were random intercepts for students and for raters. Repeated contrasts were applied to the Text predictor, so that the coefficients represented the successive differences over time, between texts 1 and 2, 2 and 3, 3 and 4. The continuous predictors (i.e. working memory, language comprehension and reading comprehension) were centered at these variables' median values in order to enhance the interpretability of the regression outcome. The effects of these three variables, however, were not significant in any of the models that were tested. The model that was chosen among those that were evaluated contained the predictors Text, gender, and grade, including interactions of Text and gender and of grade and gender. These two-way interactions were both statistically significant, and this model had a lower AIC value than any of the other considered models had. The output of this model is shown in Table 3.

Table 3. Regression output.

	Estimate	Standard error	*df*	*t*	*p* Value
Intercept	33.500	7.327	7.403	4.572	.002
Text 2–1	0.556	3.967	106.251	0.140	.889
Text 3–2	2.905	3.972	106.517	0.731	.466
Text 4–3	4.996	3.972	106.498	1.258	.211
Gender male	−12.424	9.305	6.994	−1.335	.224
Grade 7–8	35.576	9.281	6.924	3.833	.007
Text 2–1 × gender male	−4.753	5.636	108.472	−0.843	.401
Text 3–2 × gender male	−1.644	5.589	106.889	−0.294	.769
Text 4–3 × gender male	−10.905	5.372	106.283	−2.030	.045
Gender male × grade 7–8	−40.479	12.478	6.981	−3.244	.014

Estimated text quality rating for Text 1 for a typical female Grade 5–6 student (intercept), followed by estimated changes in text quality between texts, male students and Grade 7–8 students. The last four rows show interactions between texts, gender, and Grade.

The estimate, labeled "Intercept," is the overall estimated text quality of Text 1 for a female student in Grade 5–6. The value is just above 33 points. The next three lines indicate how text quality changes for Grade 5–6 female students with slight but non-significant increases across the four texts. The following two rows describe the effect of gender and grade on text quality. Boys from Grade 5–6 wrote texts that were rated 12 points lower than girls' texts from the same grade. The girls in Grade 7–8 had an estimated 36 points higher text quality rating than the girls from Grade 5–6. The next three lines indicate how the differences between the four texts written by the boys differ from those written by the girls. The boys differ significantly from the girls on the rating of text four, which received low ratings. The last row in the table shows the interaction between gender and school grade. The boys from Grade 7–8 wrote texts that were actually rated lower than the texts written by the boys from Grade 5–6. Figure 1 shows the predicted text quality of girls and boys in Grades 5–6 and 7–8 for each text. It shows that group differences were larger than any effects over time. Both lines representing girls increase over time, whereas the lines representing boys decrease, most notably to Text 4, which was written after the summer vacation.

In sum, the quality of the texts written by girls increased slightly but non-significantly across the four texts, and was significantly better than the quality of texts written by boys (approximately 12 points for the students in Grade 5–6, and 40 points for the students in Grade 7–8). The quality of boys' texts did not increase over time. On the contrary their text quality decreased (approximately 5 points). In addition, there was a grade effect, but this effect was only significant

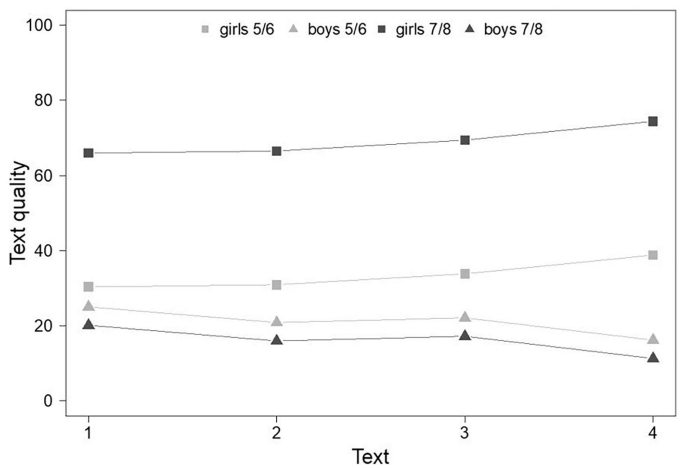

Figure 1. Predicted text quality values on the four texts separated by gender and grade.

for girls (approximately 35 points), and not for boys (approximately 5 points).

Audiological factors

The considerable heterogeneity of the audiological factors motivated the following inspection of individual results for students with the highest and the lowest text quality. The three students with the highest average text quality ratings were all girls in Grade 7–8. The first of these three (case 6) was a student who was diagnosed with profound HL when she was a few months old and did not receive HA until 17 months later. At the time of the study, she was bimodally aided with a CI and a HA. In spite of the profound HL, language comprehension was comparable to norms for students with NH, but working memory and reading comprehension were not. The second student (case 7) had a bilateral severe HL for which she was bimodally aided and she was diagnosed at age three. Her reading comprehension, language comprehension and working memory results were limited compared to norms. The third student (case 8) had a moderate HL and was diagnosed at four years old after which she immediately received bilateral HA. She had adequate results on working memory and language comprehension, but low reading comprehension results. In sum, none of the three students with the best text quality was diagnosed early or received amplification early – the earliest at age one and the latest at age four. All three were girls in Grade 7–8, and their results on the working memory test and language and reading comprehension tests varied.

The three students who received the lowest average text quality ratings were a boy in Grade 5–6 (case 4) and two boys in Grade 7–8 (cases 9 and 11). The boy in Grade 5–6 was diagnosed with a HL (of unknown degree to the authors) at six years of age. He received bilateral HA directly after diagnosis. One boy in Grade 7–8 (case 9) had a moderate HL. He was diagnosed at age five and amplified bilaterally with HA at age six. The other boy in Grade 7–8 (case 11) had a moderate HL and was diagnosed at age 10. In sum, the three students with the lowest text quality ratings were amplified very late. Two of these students had limited language comprehension (below the 10th percentile) and limited reading comprehension, and one of them had low scores on the working memory capacity test. From this, no clear relation between degree of HL and narrative writing skills is apparent. In fact, the three students with the highest ratings had more severe HL than at least two of the students with the lowest ratings. The students who had the lowest text quality ratings were, however, amplified considerably later than the students with the highest text quality ratings.

Discussion

In the present study, possible predictors of narrative text quality in students with HL were investigated over the course of four written narratives and a writing intervention. The results showed effects of gender and grade but not of working memory, reading comprehension, or spoken language comprehension. Nor did text quality ratings change after the intervention. Studies on students with HL all emphasize the great heterogeneity of the population [1,48,49]. The present study is no exception. The individual variability between students in the sample of 11 students was considerable. Results on the formally assessed linguistic and cognitive tests differed considerably. The students had exposure to one or more spoken languages or to spoken and signed language. Hearing sensitivity and time factors (degree of HL, age at diagnosis and age at amplification) also varied greatly. The students were born before neonatal hearing screening had been implemented in Sweden. Age at diagnosis and amplification was thus late for a majority of them. Consequently, listening abilities had been challenged for these students for a long time, by degraded speech signals and limited language skills, which may have affected language comprehension and learning adversely [23].

The regression model did not indicate changes in text quality ratings over time, but there were interaction effects showing that the text quality was significantly lower in the boys' texts than the girls', and that the text quality of the boys' texts was significantly lower at the fourth text. The fourth text was written a month into the semester after the summer vacation. One possible explanation is the "summer loss" in academic results. In a review by Cooper et al. [50], the authors found evidence of setbacks of up to one month after the summer vacation in some studies, and gender differences were found inconclusive between studies. A recent study [51] showed that 6- to 9-year-old students with NH may be set back in semantic verbal vocabulary fluency after the nine-week long summer vacation, but had regained that loss by the end of the fall semester. The decrease in text quality ratings for boys' texts at the follow-up text after summer vacation in the present study suggests that boys may be more affected by a summer loss than girls.

The three students who had the highest text quality ratings were all girls in Grade 7–8. While the effect of working memory was not significant in the statistical analysis, it is striking that among the three students with high text quality ratings were two students with low results on the working memory test. Two of the three students received the highest results of the eleven students on the reading comprehension test. Reading comprehension and narrative skills are associated [52], and these results also suggest that reading comprehension is an important factor to take into consideration in studies of narrative writing. Further, good reading comprehension during writing requires automatized reading processes, leaving more capacity for higher level processes of writing.

The observed gender differences are consistent with previous findings that girls outperform boys in narrative text writing [18,19,36]. The three students with the lowest average text quality ratings were three boys, one in Grade 5–6 and two in Grade 7–8. Two had language comprehension below the 10th percentile compared to reference values on the test, a common cut-off for language disorder. One also had low results on the working memory test, and all three

had low results on the reading comprehension test. Poor spoken language comprehension has been indicated as an early predictor of developmental language disorders in children with NH [53,54]. In addition, the students in the present study may not have received optimal audiological intervention, with a late diagnosis of HL and lack of proper amplification. Although the three students with the highest text quality were not amplified early, at 1, 3, and 4 years old, they were all amplified considerably earlier than the students with the lowest text quality, who received their HA at 6, 6, and 10 years old. This may have played a role for the development of language skills of the students. Early identification and intervention are crucial for language development [1,23,55].

An unexpected finding was that the boys' texts from the higher grade did not receive higher quality ratings (in fact, even somewhat lower) than the boys from the lower grade. A possible interpretation is that students with HL who perform well in classes for students with HL may move to mainstream schools between grades 6 and 7, when many students change schools, or that students struggling in mainstream schools move to classes for students with HL [56]. Another possible interpretation is that the writing teaching strategies in the higher grade were geared more towards girls than towards boys. However, in the absence of more precise information on what these strategies were, this conclusion is very tentative and may be addressed in future studies.

The students in the present study responded well to the writing intervention, even though this did not result in noticeable improvements in text quality ratings. Some aspects of the writing intervention may be suitable for students with HL. When listening is challenged, as it is in students with HL, a clear and recurrent structure (observation, reflection and learning) could support listening and thus the comprehension of instructions. Further, the "film peers" could be simultaneously seen, heard and read (by subtitles), and even reiterated, which may relieve the students' listening effort. On the other hand, the relatively implicit nature of observational learning may prove to be too abstract for students with HL, as their linguistic and cognitive skills are often not on par with those of age peers with NH. It may be supplemented with, for instance, individual feedback on students' written texts with reference to the themes addressed during the lessons or other explicit writing instruction. It may pose a challenge to design an intervention long enough to be effective, but short enough to fit into one semester and to spare time for other curricular goals.

Conclusions

In the present study, possible predictors of narrative writing skills in students with HL were explored. Girls wrote better texts than boys, and school grade had a positive effect on texts written by girls but not on texts written by boys. Instead, a "summer loss" was observed in texts written by boys but not by girls. Age at amplification seemed more important for text quality than severity of HL. There were no statistically significant effects of the students' working memory, language comprehension, or reading comprehension on text quality. The absence of effects of these predictors may have been due to limitations in sample size which inevitably reduced statistical power in the study. For that reason, these conclusions are tentative at present. Finally, the writing intervention in the context of which the data were collected is a promising paradigm but should be further adapted to the special needs of students with HL.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the Marcus and Amalia Wallenberg foundation (Grant no. 2012.0038).

Notes on contributors

Emily Grenner is a speech-language therapist and PhD student at the Department of Clinical Sciences, Logopedics, Phoniatrics and Audiology, Lund University, Sweden. Her PhD studies concern processes and products of narrative writing in school children with normal hearing and school children with hearing loss.

Joost van de Weijer is Associate Professor in Linguistics at the Centre for Languages and Literature at Lund University. His research focuses on the perception of foreign-accented speech. He is also affiliated as methodologist with the Lund University Humanities Lab, where he mainly works with the analysis of experimental data.

Victoria Johansson is Associate Professor in Linguistics at the Centre for Languages and Literature and Deputy Director at the Lund University Humanities Lab, Lund University, Sweden. Her research focus is on language development through the lifespan, primarily cognitive aspects of writing.

Birgitta Sahlén is Professor in Speech Pathology at the Department of Clinical Sciences, Logopedics, Phoniatrics and Audiology, Lund University, Sweden. Her research focus is on cognition and communication in children with language disorder and/or hearing loss. comprehension, listening effort, motivation and learning in relation to children's cognitive capacity. She is currently heading comprehensive intervention projects aiming at improving children's narrative writing by observational learning, and teachers' communicative techniques for language learning interaction in the classroom.

ORCID

Emily Grenner http://orcid.org/0000-0002-6818-1093
Joost van de Weijer http://orcid.org/0000-0001-9843-3143
Victoria Johansson http://orcid.org/0000-0002-0684-6491
Birgitta Sahlén http://orcid.org/0000-0002-8468-0546

References

[1] Geers AE, Nicholas J, Tobey E, et al. Persistent language delay versus late language emergence in children with early cochlear implantation. J Speech Lang Hear Res. 2016;59:155–170.
[2] Sandgren O, Hansson K, Sahlén B. Working memory and referential communication-multimodal aspects of interaction

between children with sensorineural hearing impairment and normal hearing peers. Front Psychol. 2015;6:242.
[3] Sahlén B, Hansson K, Lyberg-Åhlander V, et al. Spoken language and language impairment in deaf and hard-of-hearing children: fostering classroom environments for mainstreamed children. New York (NY): Oxford University Press; 2018.
[4] SOU. Samordning, ansvar och kommunikation – vägen till ökad kvalitet i utbildningen för elever med vissa funktionsnedsättningar (Coordination and responsibility – increased quality in education for students with certain disabilities). Stockholm: Wolters-Kluwer; 2016. 46.
[5] HRF. Årsrapport [Yearly report]. 2007.
[6] SCB. Årsrapport [Yearly report]. 2015.
[7] Reuterskiöld C, Hansson K, Sahlén B. Narrative skills in Swedish children with language impairment. J Commun Disord. 2011;44:733–744.
[8] Feagans L, Appelbaum MI. Validation of language subtypes in learning disabled children. J Educ Psychol. 1986;78:358–364.
[9] Catts HW, Hogan TP, Fey ME. Subgrouping poor readers on the basis of individual differences in reading-related abilities. J Learn Disabil. 2003;36:151–164.
[10] Griffin TM, Hemphill L, Camp L, et al. Oral discourse in the preschool years and later literacy skills. First Language. 2004;24:123–147.
[11] Karmiloff-Smith A. The grammatical marking of thematic structure in the development of language production. In: Deutsch W, editor. The child's conception of language. London (UK): Academic Press; 1981. p. 121–147.
[12] Berman RA, Slobin DI. Relating events in narrative: a crosslinguistic developmental study. Hillsdale (NJ): Lawrence Erlbaum Associates, New York (NY): Psychology Press; 1994.
[13] Nordqvist Palvianen Å. Speech about speech: a developmental study on form and function of direct and indirect speech [Theses]. Göteborg; 2001.
[14] Strömqvist S. Discourse flow and linguistic information structuring: explorations in speech and writing. Göteborg: Univ.; 1996. (Gothenburg papers in theoretical linguistics: 78).
[15] Skolverket. Läroplan för grundskolan, förskoleklassen och fritidshemmet (Curriculum for the compulsory school, preschool class and school-age educare). 2011. revised 2019 ed. 2019.
[16] Bourdin B, Fayol M. Even in adults, written production is still more costly than oral production. Int J Psychol. 2002;37:219–227.
[17] Skolverket. Slutbetyg i grundskolan våren 2019:1342. [Board of education, final grades of the compulsory school, spring. 2019].
[18] Kanaris A. Gendered journeys: children's writing and the construction of gender. Lang Educ. 1999;13:254–268.
[19] Myhill D. Towards a linguistic model of sentence development in writing. Lang Educ. 2008;22:271–288.
[20] Graham S, Harris KR. Evidence-based writing practices: a meta-analysis of existing meta-analyses. In: Redondo RF, Harris K, Braaksma M, editors. Design principles for teaching effective writing: theoretical and empirical grounded principles. Vol. 34, Studies in writing. Leiden (the Netherlands): Brill Academic Publishers; 2018. p. 33–37.
[21] Briscoe J, Bishop DVM, Norbury CF. Phonological processing, language, and literacy: a comparison of children with mild-to-moderate sensorineural hearing loss and those with specific language impairment. J Child Psychol Psychiatry. 2001;42:329–340.
[22] Geers AE, Moog JS, Biedenstein J, et al. Spoken language scores of children using cochlear implants compared to hearing age-mates at school entry. J Deaf Stud Deaf Educ. 2009;14:371–385.
[23] Mattys SL, Davis MH, Bradlow AR, et al. Speech recognition in adverse conditions: a review. Lang Cogn Process. 2012;27:953–978.
[24] McCutchen D. Knowledge, processing, and working memory: implications for a theory of writing. Educ Psychol. 2000;35:13–23.
[25] McCutchen D. From novice to expert: implications of language skills and writing-relevant knowledge for memory during the development of writing. J Writ Res. 2011;3:51–68.
[26] Blom E, Boerma T. Why do children with language impairment have difficulties with narrative macrostructure? Res Dev Disabil. 2016;55:301–311.
[27] Torkildsen JK, Arciuli J, Haukedal CL, et al. Does a lack of auditory experience affect sequential learning? Cognition. 2018;170:123–129.
[28] Sahlén B, Hansson K, Ibertsson T, et al. Reading in children of primary school age – a comparative study of children with hearing impairment and children with specific language impairment. Acta Neuropsychol. 2004;2:393.
[29] Asker-Árnason L, Åkerlund V, Skoglund C, et al. Spoken and written narratives in Swedish children and adolescents with hearing impairment. Commun Disord Q. 2012;33:131–145.
[30] Strassman BK, Schirmer B. Teaching writing to deaf students: does research offer evidence for practice? Remedial Spec Educ. 2013;34:166–179.
[31] Dostal HM, Wolbers KA. Examining student writing proficiencies across genres: results of an intervention study. Deafness Educ Int. 2016;18:159–169.
[32] Wolbers K, Dostal HM, Graham S, et al. Strategic and interactive writing instruction: an efficacy study in grades 3-5. JEDP. 2018;8:99.
[33] Braaksma M, Rijlaarsdam G, van den Bergh H. Observational learning and the effects of model-observer similarity. J Educ Psychol. 2002;94:405–415.
[34] Braaksma M, Rijlaarsdam G, van den Bergh H, et al. Observational learning and its effects on the orchestration of writing processes. Cogn Instr. 2004; 22:1–36.
[35] Rijlaarsdam G, Braaksma M, Couzijn M, et al. Observation of peers in learning to write: practice and research. J Writ Res. 2008;1:53–83.
[36] Grenner E, Åkerlund V, Asker-Árnason L, et al. Improving narrative writing skills through observational learning: a study of Swedish 5th-grade students. Educ Rev. 2020;72:691–710.
[37] van de Weijer J, Åkerlund V, Johansson V, et al. Writing intervention in university students with normal hearing and in those with hearing impairment: can observational learning improve argumentative text writing? Logoped Phoniatr Vocol. 2019;44:115–119.
[38] Clark JG. Uses and abuses of hearing loss classification. ASHA. 1981;23:493–500.
[39] Johansson V. Developmental aspects of text production in writing and speech. Lund (Sweden): Lund University; 2009.
[40] Frid J, Johansson V, Johansson R, et al. Developing a keystroke logging program into a writing experiment environment. Writing across borders; 2014.
[41] Tillema M, van den Bergh H, Rijlaarsdam G, et al. Quantifying the quality difference between L1 and L2 essays: a rating procedure with bilingual raters and L1 and L2 benchmark essays. Lang Test. 2013;30:71–97.
[42] Raedts M, Steendam E, Grez L, et al. The effect of different types of video modelling on undergraduate students' motivation and learning in an academic writing course. J Writ Res. 2017;8:399–435.
[43] Wolff U. Lilla DUVAN. Stockholm (Sweden): Hogrefe Psykologiförlaget AB; 2010.
[44] Bishop DVM. Test for reception of grammar, version 2. Swedish version. Stockholm (Sweden): Pearson Education; 2009.
[45] Nielsen JC, Jensen SE, Magnusson E, et al. Ordläsningsprov OS64 och OS120, meningsläsningsprov SL60 och SL40 [Word reading test OS64 and OS120, sentence reading test SL60 and SL40]. Malmö: Pedagogisk design, 1997.
[46] Team RC. A language and environment for statistical computing. Vienna (Austria): R Foundation for Statistical Computing; 2014.

[47] Bates D, Mächler M, Bolker B, et al. Fitting linear mixed-effects models using lme4 [sparse matrix methods; linear mixed models; penalized least squares; Cholesky decomposition]. J Stat Softw. 2015;67:48.
[48] Sundström S, Löfkvist U, Lyxell B, et al. Phonological and grammatical production in children with developmental language disorder and children with hearing impairment. Child Lang Teach Ther. 2018;34:289–302.
[49] Hansson K, Ibertsson T, Asker-Árnason L, et al. Phonological processing, grammar and sentence comprehension in older and younger generations of Swedish children with cochlear implants. Autism Dev Lang Impair. 2017;2: 239694151769280.
[50] Cooper H, Nye B, Charlton K, et al. The effects of summer vacation on achievement test scores: a narrative and meta-analytic review. Rev Educ Res. 1996;66:227–268.
[51] Rosqvist I, Sandgren O, Andersson K, et al. Children's development of semantic verbal fluency during summer vacation versus during formal schooling. Logoped Phoniatr Vocol. 2020;45: 134–142.
[52] Suggate S, Schaughency E, McAnally H, et al. From infancy to adolescence: the longitudinal links between vocabulary, early literacy skills, oral narrative, and reading comprehension. Cognitive Development. 2018;47: 82–95.
[53] Bishop D, Edmundson A. Specific language impairment as a maturational lag: evidence from longitudinal data on language and motor development. Dev Med Child Neurol. 1987;29: 442–459.
[54] Bruce B, Kornfält R, Radeborg K, et al. Identifying children at risk for language impairment: screening of communication at 18 months. Acta Paediatr. 2003;92:1090–1095.
[55] Yoshinaga-Itano C, Sedey AL, Coulter DK, et al. Language of early- and later-identified children with hearing loss. Pediatrics. 1998;102:1161–1171.
[56] SPSM. Måluppfyllelse för döva och hörselskadade i skolan [Academic achievement for students with deafness or hearing loss]. Specialpedagogiska skolmyndigheten (Swedish National agency for special needs education and schools). 2008.

Appendix.

List of regression models, with the chosen model in bold text.

Predictors	AIC
Text	1024.3459
Text + language comprehension	1020.7463
Text + working memory	1021.4931
Text + reading comprehension	1021.7042
Text + grade	1017.8478
Text + gender	1010.5815
Text + language comprehension × gender	1005.5341
Text + language comprehension × reading comprehension	1019.7526
Text + language comprehension × working memory	1016.0177
Text + language comprehension × grade	1012.2727
Text + working memory × gender	1008.4402
Text + working memory × reading comprehension	1019.1912
Text + working memory × grade	1012.8378
Text + grade × gender	992.1998
Text + grade × reading comprehension	1012.7642
Text × gender + grade × gender	**972.4090**
Text + gender × reading comprehension	996.7998

Paper III

LOGOPEDICS PHONIATRICS VOCOLOGY
https://doi.org/10.1080/14015439.2019.1709539

Taylor & Francis
Taylor & Francis Group

RESEARCH ARTICLE

OPEN ACCESS Check for updates

Effects of intervention on self-efficacy and text quality in elementary school students' narrative writing

Emily Grenner[a], Victoria Johansson[b], Joost van de Weijer[b] and Birgitta Sahlén[a]

[a]Department of Clinical Sciences, Logopedics, Phoniatrics and Audiology, Lund University, Lund, Sweden; [b]Centre for Languages and Literature, Lund University, Lund, Sweden

ABSTRACT

Aim: Self-efficacy for writing is an important motivational factor and considered to predict writing performance. Self-efficacy for narrative writing has been sparsely studied, and few studies focus on the effects of writing intervention on self-efficacy. Additionally, there is a lack of validated measures of self-efficacy for elementary school students. In a previous study, we found that a trained panel rated personal narrative text quality higher for girls than for boys, which led to our aim: to investigate boys' and girls' self-efficacy for narrative writing before and after an intervention, and to explore associations between self-efficacy and text quality.

Methods: An 18-item self-efficacy scale was developed. Fifty-five fifth-grade students (*M* 11:2 years, *SD* 3.7 months) filled out the scale before and after a five-lesson observational learning intervention. Self-efficacy was then related to writing performance as measured by holistic text quality ratings.

Results: The students demonstrated strong self-efficacy, which increased significantly post-intervention. Girls and boys demonstrated similar self-efficacy, despite girls' higher text quality. There were moderate correlations between self-efficacy and writing performance pre- and post-intervention.

Conclusions: The results support previous findings of strong self-efficacy at this age. The interaction between writing self-efficacy and performance is complex. Young students may not be able to differentiate between self-efficacy, general writing skills, task performance, and self-regulation. Self-efficacy scales should thus be carefully constructed with respect to age, genre, instruction, and to students' general educational context.

ARTICLE HISTORY
Received 20 February 2019
Revised 11 December 2019
Accepted 17 December 2019

KEYWORDS
Self-efficacy; observational learning; writing improvement; intervention; elementary education

Introduction

Self-efficacy is the belief in one's own capability to perform successfully within a field. According to social cognitive theory [1], self-efficacy varies widely between domains, in different activities within a domain, and under different task demands. Bandura states that stronger self-efficacy leads to a better performance regardless of skill: a person with stronger beliefs in her or his capabilities, i.e. self-efficacy, will approach a difficult task as an achievable challenge, while a person with weaker self-efficacy will view it as a threat [2,3]. To see an achievable challenge will in turn motivate more effort, while a threatening task leads to less effort and a higher risk of giving up [3]. Self-efficacy has been shown to predict outcome in diverse domains, including sales performance, health, and academic performance [4]. For instance, students with strong self-efficacy participating in math groups of different levels had better performance than group-members with weak self-efficacy [5]. Bandura further states four sources of self-efficacy: (1) enactive experience, i.e. an earlier successful performance within the field or a similar field; (2) vicarious experience, i.e. watching "similar others" perform a comparable task; (3) social persuasion, i.e. others expressing beliefs in or providing adequate feedback on one's performance; (4) the emotional and physiological state of the person performing the task [3,6].

Research on self-efficacy for writing has mainly followed one of two paths [7]. One path is represented by correlational studies, in which factors that potentially influence self-efficacy have been explored. The other path, which is taken here, includes experimental studies in which the effects of writing intervention on self-efficacy are tested. These studies have explored the influence of various instructional approaches on self-efficacy outcomes. One such approach is the observational learning paradigm [8,9]. The writing instruction in the present study is based on this paradigm, with a design in which students observed peer models working with texts, and thereafter discussed the models' behavior with their classmates. This method is considered to increase student motivation [10], since it offers students structured opportunities to discuss the observations in groups which induces a comparison of their own writing performance to that of the peer models (i.e. the above-mentioned "vicarious experience"). There is some evidence of

CONTACT Emily Grenner emily.grenner@med.lu.se Department of Clinical Sciences, Logopedics, Phoniatrics and Audiology, Universitetssjukhuset, Lund, S-221 85, Sweden

effects on writing performance. For instance, in one study on academic writing, an observational learning intervention increased task knowledge in university students who observed films of dyads of peers writing a literature review. Controls, who did the same exercise with traditional writing instruction did not show a similar positive effect. In parallel, the increased task knowledge led to an increase in self-efficacy in the experimental group, but not in the control group [11].

Self-efficacy development

Young children's self-efficacy is often strong, in that they rate their capability as high in relation to their performance, and holistic, meaning that they do not analyze what they think they can do in different aspects or components [12]. From the age of 13 or even earlier, students' beliefs about their performance weaken [12,13], and simultaneously become more diverse across domains. The earlier, more holistic self-efficacy may include a combination of appraisal of their general skills, their performance at a specific task and their perceived effort, but the links between these factors may not be clear to them [14]. Self-efficacy, from the social cognitive point of view, depends on an interaction of cognitive, physical, and social development [2–4,15]. During adolescence, students will begin to experience others' evaluations of their ability through different forms of feedback, to a greater extent than in younger ages. They will therefore become more analytic and begin to differentiate between different aspects of a certain skill and result in a better correspondence between self-efficacy and actual performance [14]. These associations, sometimes referred to as calibration of self-efficacy or self-efficacy accuracy, are complex. When self-efficacy is much too strong or if it is too weak, learning is impeded. In the first case, it may lead to the student believing she or he masters the task and thereby does not fully engage in it, and in the second case, that she or he focuses too much on basic concepts to be able to move on to complex matters [12,16,17]. For example in writing, a student may focus too much on the spelling of single words, instead of engaging in writing a text with cohesion. In addition, a perfect correspondence between self-efficacy and performance may also be detrimental to development. Judging one's capability too accurately may block creativity and, as a consequence, limit learning by "trial and error". When self-efficacy is slightly lower than performance, the performance can still be adequate. However, it may lead to an anxious and perfection-seeking student, worrying that she or he is not sufficiently capable [3,12,18]. The most productive self-efficacy is found in individuals with a slightly stronger belief in their capabilities than their actual performance [3,6].

Measuring self-efficacy

Self-efficacy varies between domains, and must thus be measured in a task-specific way. In his recommendations for construction of self-efficacy scales, Bandura [18] describes how the statements in a self-efficacy scale should tap into the many different skills which are important for the domain of functioning, but not into other skills or general abilities. They should concern self-perceived beliefs in one's own capability to perform a task, as opposed to constructions targeting intent or comparison to others' performance. Further, the scale should reflect the genre, but not a specific topic or subgenre, and performance assessment (e.g. text quality) should correspond to the content of the scale. It should not have too few intermediate steps [18]. When the present study was initiated in 2013, validated self-efficacy measures were lacking for the chosen age group, for the narrative genre, and for students in a Swedish school context (for more recent scales measuring self-efficacy in younger students, see Bruning et al. [19,20]). A scale for students in a Dutch context developed by Braaksma et al. [9] was used as a point of departure, since the writing intervention in Braaksma's study was based on an observational learning paradigm as was the present study. The scale was designed for teenagers writing argumentative texts and we adapted it to fit 10–12-year-olds writing narrative texts within the Swedish curriculum.

Gender differences in writing performance are often reported, e.g. in the OECD assessment PISA (The Program for International Student Assessment) which shows higher performance in girls [21]. Research on gender differences in self-efficacy for writing has shown contradictory results. In two studies by Pajares et al., similar self-efficacy for writing was found between girls and boys, while girls' performance was assessed as better [22,23]. The students were also asked to compare their own writing ability to the writing ability of other girls and boys. In the first study, on 8–11-year-old students (grades 3–5), both girls and boys on average considered themselves as better writers than classmates of the other gender, but girls did so to a higher degree [22]. In the second study on students aged 11–14 (grades 6–8), girls again considered themselves better writers than boys, while the boys considered themselves poorer writers than girls [23]. Pajares and Valiante concluded that even if self-efficacy is similar for girls and boys, statements in a scale may be judged differently between genders, girls answering more cautiously [23]. Pajares and Valiante further found that gender differences were non-significant when controlling for what they called "gender-stereotypic beliefs" [24]. In a review however, Pajares found that several studies show that girls report stronger self-efficacy than boys during earlier school years, but that the differences even out or reverse later on [25].

Text quality

The Swedish curriculum does not provide set criteria for assessment for the age group 10–12 years old. Instead, a method of holistic text quality ratings based on benchmark texts was chosen for the present study. This method has been used in previous observational learning studies [11] as well as other writing intervention studies, and was tested and validated by Tillema et al. [26]. A previous study has

explored writing performance measured by text quality for the same students as in the present study [27]. This was done through measuring text quality in personal narrative texts on repeated occasions in a waiting control design, before and after the observational learning intervention. The narrative genre was chosen as it is a prerequisite for developing other genres and predicts results in higher education [28]. The students in the age-group of the present study are thus expected to be aware of the structure of a narrative text but still to be developing their abilities for creating such texts [29]. The results showed that the estimated text quality for boys increased from approximately 33 (on a scale from 0 to 100) before the intervention to approximately 40, as a consequence of intervention. For girls, the estimated text quality increased from 45 to 52. Thus, girls had considerably higher text quality than boys, pre- as well as post-intervention, while intervention effects were similar at about seven points on the 0–100 scale.

Cognitive and linguistic abilities are important prerequisites for writing performance, at micro- as well as at macro-level [30]. Measures of working memory capacity, language comprehension, and reading comprehension were collected for our previous study [27], and are presented as demographic data in the present study.

Aim

In our previous study, the trained raters found that girls' texts had higher text quality than boys' texts [27]. Considering this difference, the aim of the present study is to explore boys' and girls' self-efficacy for narrative writing, before and after intervention. Further, considering that intervention effects on text quality were similar for boys and girls, a secondary aim is to explore associations between self-efficacy and text quality.

The study addresses two research questions:

1a. Does self-efficacy for narrative writing change after intervention, and if it does, in what way? 1b. Are there any gender differences?
2. Is self-efficacy related to text quality, and if it is, in what way?

Methods

Data regarding the participating students, their results on cognitive and linguistic tasks, the texts and text quality, and intervention were also described in our previous study [27].

Design and procedure

Before and after an observational learning intervention, students wrote personal narratives and filled out self-efficacy scales. An overview of the design is found in Figure 1. The students wrote their personal narratives on laptops. A few days or up to a week after the first writing assignment, the students filled out a self-efficacy scale, which is found in Table 1, and were given a working memory test and a language comprehension test. The following week, the intervention (led by researchers) started for one school, replacing the regular Swedish lessons. The other school had regular Swedish lessons with their teachers during this time. The five intervention lessons were given over three weeks. In the second school, intervention started after nine weeks (week 10 in Figure 1). After the intervention period, all students wrote another personal narrative. A few days or up to a week later, they filled out the self-efficacy scale for the

Table 1. Self-efficacy statements.

Self-efficacy statements
1. I can quickly write a text on the computer.
2. I can find all the letters on the keyboard.
3. I can write a text without spelling errors.
4. I can use periods, question marks and exclamation marks in the right places.
5. I can use commas in the right places.
6. I can partition my text into paragraphs.
7. I can write a heading that fits with the content.
8. I can write a story in such a way that the reader understands where the story took place and who was in it.
9. I can divide my text into beginning, middle, and end.
10. I can write a story with an unexpected ending.
11. I can write a story in such a way that the reader understands what happened.
12. I can begin the story in an exciting way, so that the reader wants to continue reading.
13. I can change one or several words in my text if I am not satisfied with what I had written initially.
14. I can write both long and short sentences.
15. I can decide which parts of my story need to be explained to the reader.
16. I can read through my text and correct spelling mistakes.
17. I can read through my text and make changes to improve it.
18. I can write a text which is one page long.

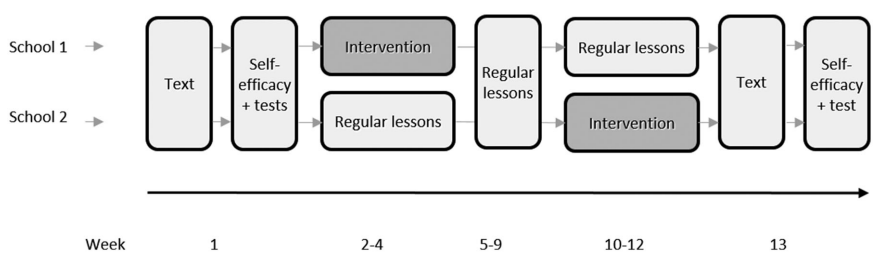

Figure 1. Overview of data collection and intervention. Narrative writing assignments pre- and post-intervention (Text), self-efficacy scales pre- and post-intervention (Self-efficacy), tests of working memory and language comprehension pre-intervention (tests), intervention, regular lessons, and test of reading comprehension (test).

second time and participated in a reading comprehension task. For practical reasons, working memory and language comprehension tasks were administered pre-intervention, and the reading comprehension task post-intervention. The whole data collection, including intervention, took 13 weeks. Teachers were explicitly asked not to work on written narratives with their students during this time.

The full data collection and intervention took place in the classroom. All students who typically attended the classes were included, as this study was meant to test the intervention in a natural school environment. All data collection and intervention was carried out by the first author and a research assistant.

Participants

All students came from two schools in areas of similar socio-economic status. With 35 students in fifth grade in one school and 44 in the other, there was a total of 79 students. Informed consent was collected from 59 students and their parents. Following the advice from the regional ethics board (EPN dnr. 2013/270), no students were a priori excluded from participation in the intervention or data collection. The inclusion criteria were, aside from written consent: taking part in at least four of the five intervention lessons, adequate Swedish listening comprehension and speech production skills, and being regularly present for whole-class activities. Four students, two from each school, did not meet these criteria and were excluded from the dataset. Thus, 55 students (30 girls and 25 boys) remained, 32 from one school and 23 from the other. Their age was 10:9 to 11:9 years (*M* 11:2 years, *SD* 3.7 months) and five students had other first languages in addition to Swedish. One student was not present for writing the pre-intervention personal narrative, two students' post-intervention narratives were lost due to software issues, and one student was not present for filling out the post-intervention self-efficacy scale. Additionally, two students did not complete the self-efficacy scales. Thus, the total number of students in the analyses varies between 49 and 55.

Cognitive and linguistic tasks

The students performed according to age norms and results were similar across the two schools and between boys and girls on norm-referenced or standardized tests of working memory [31], language comprehension [32], and reading comprehension [33]. As reported in our previous study [27], the students' mean result was 31.2 (*SD* 6.5) on the working memory task Lilla Duvan, ranging from 14 to the maximum 36. The norms for the fifth grade are 31.7 (*SD* 5.8) [31]. The mean was 16.5 (*SD* 2.3) on the language comprehension task TROG-2, ranging from 5 to 19 (maximum score is 20). Scores 15–19 represent percentiles 23–73 according to age norms [32]. In the reading comprehension task SL40, students read sentences and chose the corresponding pictures. The mean result was 37.5 (*SD* 4.1), ranging from 15 to the maximum 40. Scores 37–39 equal percentiles 25–75 [33].

Self-efficacy scale

The self-efficacy scale was adapted from the scale by Braaksma et al. [9] to fit the narrative genre as well as the goals for writing stated in the curriculum for the subject of Swedish [34] for the age group. There were 18 statements which are listed in Table 1. Six of them were identical or very similar (statements 1, 3, 6, 9, 12, and 18) to the original scale. The 18 statements reflected aspects of form as well as content, of writing processes as well as the finished text. For example, statement 2, "I can find all the letters on the keyboard" merely concerned low-level writing processes, while statement 9, "I can divide my text into beginning, middle, and end" concerned *form* (the structure of a narrative) and *content* (relevant content in the beginning, middle, and end of a story), and to some degree also *writing processes* (how to go about writing a story with these elements). The students filled out the self-efficacy scale once before and once after the intervention. The instructions were read aloud to the class and written in the booklet containing the self-efficacy scale:

> Imagine the following scenario: In school, you get a writing assignment where you are to write a narrative story about something you have experienced. For example, it could be writing a story about the most exciting time you had during the summer vacation. It should be written so that somebody in your class could understand it, and the text should be about one page long. You are not going to write this text yourself, but please answer some questions about what writing such an assignment would be like. Answer each question by marking the horizontal line beneath each question with a vertical line. The further to the right your mark is, the more you agree with the statement.

Beneath each statement was a 100-mm, visual-analogue scale [35]. The VAS scale was marked with the phrases "not at all" and "yes, completely" below the left and right endpoints. One of the researchers explained how students should mark the scales according to their beliefs, with illustrations on the whiteboard. This included demonstrating that the students should put a mark in the middle of the scale if they believed their ability to be average. The marks on the scales were measured resulting in possible values ranging from 0 to 100 for each of the 18 statements.

Personal narrative texts

Personal narratives were written before and after the intervention (Figure 1). The students wrote on laptops with a basic word-processing interface, in the classroom. Students were asked to write a personal narrative about *"one time you were saved from a jam you had got into, or when you saved somebody else from a jam"* (the first text) and about *"one time when you were afraid"* (the second text). The topics of the narratives have been used and found suitable for the age group [29,36]. Students were not given feedback after the writing tasks.

Text quality

Text quality was assessed by raters who were trained by assessing comparable texts. Four benchmark texts written by age peers were given to the raters, six university students. These texts were given a holistic score, and rated 25, 40, 50, and 95 on a 0–100 scale. The scores of the benchmark texts were motivated by short, written summaries, describing aspects of content, structure, genre, organization, grammar, spelling, punctuation, and text length. Following this, the raters were practicing in the method of holistic scoring through the rating of another six texts in the same way. In the final data set, each text was rated by three or in some cases two raters.Their inter-rater reliability, calculated on the averages of ratings of the whole rated dataset (including about 150 texts written by students who were not included in the present study) was high (Cronbach's alpha = 0.90). The raters and texts were distributed at random with the restrictions that each text was rated three times, texts written by one student would be rated by different raters, and each rater would rate texts from pre- and post-intervention. Raters were not aware of intervention or that one student had produced more than one text.

Intervention

The five-lesson intervention was developed within the observational learning paradigm, a method relying on vicarious experience and structured reflection. The intervention is thoroughly described in Grenner et al. [27]. Each lesson had a different theme, which in design and content was based on the Swedish curriculum [34] and writing development for the age group: (1) the reader's perception – what does the reader find important in a story? (2) structure – different ways to start a story; in what order should the events unfold? (3) conclusion – how to finish a story; (4) editing someone else's text; (5) editing during writing – what changes do writers make while they write? The lessons were structured around short video clips of 12 different peers. The students were not given any information about the proficiency of the "film peers" (low, average, or advanced) they were watching. Thus, instruction was implicit.

Analyses

Pearson's bivariate correlations were calculated between each statement and the self-efficacy mean pre- and post-intervention. Means and standard deviations for the self-efficacy statements and the self-efficacy mean pre- and post-intervention were calculated. Repeated measures ANOVA were calculated to test intervention effects of each self-efficacy statement, including interactions with gender. Each self-efficacy statement and self-efficacy mean pre-intervention was correlated to text quality pre-intervention by using Pearson's bivariate correlations. The same was done for these measures post-intervention. The alpha level was set at .05 for all analyses.

Results

Internal consistency within the self-efficacy scale

To explore the internal consistency of the self-efficacy scale, each statement (Table 1) pre-intervention was correlated to the self-efficacy mean pre-intervention. The same procedure was followed for the statements post-intervention. The results are presented in Table 2 (pre-intervention) and Table 3 (post-intervention). The correlations between the individual statements and the self-efficacy mean varied between $r = .463$ and $r = .858$. All correlations were statistically significant indicating strong internal consistency, and suggesting that the students had consistent beliefs in their ability, across the

Table 2. Pearson bivariate correlations between self-efficacy statements before intervention and mean self-efficacy before intervention, text quality before intervention, mean self-efficacy after intervention, and text quality after intervention.

Statement	Mean SE pre	Text quality pre	Mean SE post	Text quality post
1	.624**	.323*	.531**	.085
2	.601**	.353**	.555**	.245
3	.491**	.256	.510**	.258
4	.595**	.188	.530**	.353*
5	.614**	.062	.490**	.343*
6	.706**	.222	.519**	.320*
7	.521**	.155	.437**	.266
8	.806**	.328*	.653**	.161
9	.791**	.314*	.501**	.252
10	.677**	.251	.458**	.108
11	.858**	.379**	.630**	.264
12	.616**	.396**	.525**	.220
13	.727**	.362**	.548**	.237
14	.676**	.192	.478**	.109
15	.820**	.386**	.683**	.227
16	.604**	.145	.573**	.163
17	.795**	.247	.672**	.266
18	.716**	.328*	.589**	.170
Mean		.394**	.816**	.327*

*Significant correlations $p < .05$.
**Significant correlations $p < .01$.

Table 3. Pearson bivariate correlations between self-efficacy statements after intervention and mean self-efficacy before intervention, text quality before intervention, mean self-efficacy after intervention, and text quality after intervention.

Statement	SE mean pre	Text quality pre	SE mean post	Text quality post
1	.640**	.319*	.565**	.178
2	.507**	.166	.603**	.086
3	.557**	.151	.633**	.260
4	.480**	.270	.658**	.386**
5	.682**	.291*	.653**	.251
6	.597**	.337*	.651**	.296*
7	.490**	.314*	.612**	.345*
8	.693**	.340*	.790**	.170
9	.569**	.130	.592**	–.046
10	.387**	.322*	.572**	.130
11	.640**	.256	.733**	.098
12	.497**	.313*	.637**	.134
13	.327*	.050	.580**	.161
14	.676**	.168	.742**	–.026
15	.612**	.012	.755**	.111
16	.482**	.144	.703**	.186
17	.325*	.080	.589**	.259
18	.433**	.261	.581**	.246
Mean	.816*	.394**		.322*

*Significant correlations $p < .05$.
**Significant correlations $p < .01$.

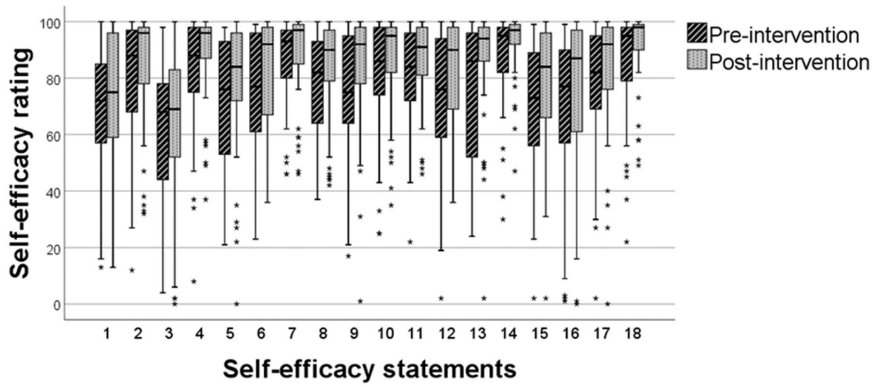

Figure 2. Boxplots of self-efficacy for each of the 18 statements, pre-intervention (striped) and post-intervention (dotted). Asterisks are outliers (values > 1.5 times the height of the box). The rating from 0 to 100 on the *y*-axis represents the 100 mm VAS scale.

Table 4. Means and standard deviations of self-efficacy statements and self-efficacy mean; effects of intervention, gender effects, and interaction effects between intervention and gender.

		Boys				Girls						Interaction
		Pre-int.		Post-int.		Pre-int.		Post-int.				
	N	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	Intervention effect	Gender effect	Intervention × gender
1	53	65.0	23.9	74.6	18.1	69.0	23.9	73.8	24.3	**$F(1, 51) = 7.240$, $p = .010$, $\eta_p^2 = .124$**	$F(1, 51) = 0.076$, $p = .784$	$F(1, 51) = 0.823$, $p = .369$
2	52	81.7	22.0	90.2	14.1	77.6	25.3	82.1	22.3	**$F(1, 50) = 7.725$, $p = .008$, $\eta_p^2 = .134$**	$F(1, 50) = 1.208$, $p = .277$	$F(1, 50) = 0.729$, $p = .397$
3	53	67.0	18.9	70.6	24.3	55.6	27.3	57.1	30.9	$F(1, 51) = 0.898$, $p = .348$	$F(1, 51) = 3.470$, $p = .068$	$F(1, 51) = 0.147$, $p = .703$
4	53	83.5	19.3	86.7	23.8	81.6	20.8	87.3	15.1	**$F(1, 51) = 4.010$, $p = .051$, $\eta_p^2 = .073$**	$F(1, 51) = 0.017$, $p = .896$	$F(1, 51) = 0.311$, $p = .580$
5	53	70.2	23.1	75.8	21.0	68.7	25.8	79.7	25.0	**$F(1, 51) = 8.582$, $p = .005$, $\eta_p^2 = .144$**	$F(1, 51) = 0.040$, $p = .843$	$F(1, 51) = 0.931$, $p = .339$
6	53	75.4	24.6	75.6	27.3	75.6	20.7	82.1	19.6	$F(1, 51) = 1.119$, $p = .295$	$F(1, 51) = 0.374$, $p = .544$	$F(1, 51) = 0.985$, $p = .326$
7	53	84.2	15.1	88.6	16.9	88.5	14.3	87.8	16.0	$F(1, 51) = 0.872$, $p = .355$	$F(1, 51) = 0.209$, $p = .650$	$F(1, 51) = 1.647$, $p = .205$
8	53	78.2	20.0	82.1	18.2	76.9	18.4	82.1	19.1	**$F(1, 51) = 7.231$, $p = .010$, $\eta_p^2 = .124$**	$F(1, 51) = 0.018$, $p = .894$	$F(1, 51) = 0.162$, $p = .689$
9	53	73.0	20.5	85.3	15.9	75.7	22.8	80.8	23.9	**$F(1, 51) = 6.886$, $p = .011$, $\eta_p^2 = .119$**	$F(1, 51) = 0.036$, $p = .850$	$F(1, 51) = 1.168$, $p = .285$
10	51	82.4	20.9	89.2	12.1	80.9	20.4	84.6	19.6	**$F(1, 49) = 3.944$, $p = .053$, $\eta_p^2 = .074$**	$F(1, 49) = 0.449$, $p = .506$	$F(1, 49) = 0.360$, $p = .551$
11	51	78.5	17.7	85.2	14.6	84.4	17.8	85.7	15.6	**$F(1, 49) = 4.245$, $p = .045$, $\eta_p^2 = .080$**	$F(1, 49) = 0.615$, $p = .437$	$F(1, 49) = 1.688$, $p = .200$
12	51	67.1	23.5	81.2	17.9	77.9	24.9	83.2	18.7	**$F(1, 49) = 11.541$, $p = .001$, $\eta_p^2 = .191$**	$F(1, 49) = 1.422$, $p = .239$	$F(1, 49) = 2.324$, $p = .134$
13	51	73.0	24.0	91.9	8.7	77.2	22.0	82.3	24.1	**$F(1, 49) = 9.858$, $p = .003$, $\eta_p^2 = .167$**	$F(1, 49) = 0.342$, $p = .561$	$F(1, 49) = 3.284$, $p = .076$
14	51	89.4	15.2	94.6	7.0	87.0	16.7	91.6	12.9	**$F(1, 49) = 6.405$, $p = .015$, $\eta_p^2 = .116$**	$F(1, 49) = 0.625$, $p = .433$	$F(1, 49) = 0.030$, $p = .863$
15	49	70.9	19.5	80.2	16.2	68.9	24.8	72.0	27.8	**$F(1, 47) = 5.652$, $p = .022$, $\eta_p^2 = .107$**	$F(1, 47) = 0.721$, $p = .400$	$F(1, 47) = 1.385$, $p = .245$
16	49	74.5	21.2	82.0	17.8	63.8	32.3	69.7	33.3	**$F(1, 47) = 6.076$, $p = .017$, $\eta_p^2 = .114$**	$F(1, 47) = 2.359$, $p = .131$	$F(1, 47) = 0.075$, $p = .785$
17	49	77.7	20.6	84.3	16.7	76.4	25.0	83.4	24.2	**$F(1, 47) = 4.760$, $p = .034$, $\eta_p^2 = .092$**	$F(1, 47) = 0.040$, $p = .842$	$F(1, 47) = 0.004$, $p = .948$
18	49	86.3	19.9	90.8	14.8	86.9	17.0	92.0	12.7	**$F(1, 47) = 4.650$, $p = .036$, $\eta_p^2 = .090$**	$F(1, 47) = 0.042$, $p = .838$	$F(1, 47) = 0.019$, $p = .892$
Mean	53	76.0	13.9	82.5	12.5	76.4	15.2	81.0	14.5	**$F(1, 51) = 22.423$, $p < .000$, $\eta_p^2 = .305$**	$F(1, 51) = 0.018$, $p = .893$	$F(1, 51) = 0.671$, $p = .416$

Values that were missing either pre- or post-intervention were excluded from the analysis. Statistically significant effects are indicated by bold text. Numbers 1–18 in the left column indicate the self-efficacy statements, and mean is the self-efficacy mean.

different statements. The variability between the various self-efficacy statements is illustrated in Figure 2.

Self-efficacy increased post-intervention

Mean values for the 18 self-efficacy statements varied between 60.8 and 88.0 pre-intervention and between 63.2 and 92.8 post-intervention on the scale with possible values from 0 to 100. All mean values post-intervention were higher than those pre-intervention. Pre-intervention values were $M = 76.2$, $SD = 14.5$ for the whole group of students (boys and girls), and post-intervention, the self-efficacy mean had increased to $M = 81.7$, $SD = 13.5$ for the whole group of students. Means and standard deviations for each statement and for the self-efficacy mean for boys and girls are found in Table 4. Some students had considerably lower self-efficacy mean values than the rest of the group, which is illustrated in Figure 3.

For each self-efficacy statement as well as for self-efficacy mean, a repeated measures ANOVA was run with intervention and gender as the independent variables, and change in self-efficacy (from pre- to post-intervention) as the dependent variable. The results are presented in Table 4. The self-efficacy mean increased significantly from pre- to post-intervention, $F(1, 51) = 22.423$, $p < .000$, $\eta_p^2 = .305$. There was no significant effect of gender and no significant interaction between gender and intervention. The increase in self-efficacy was significant in 13 of the 18 statements (numbers 1, 2, 6, 9, 10, 12, 13, 14, 15, 16, 17, 18, and 19) but not in the remaining five statements. For the individual statements, there were no main effects of gender, nor interactions with gender. This means that boys' and girls' self-efficacy increased similarly.

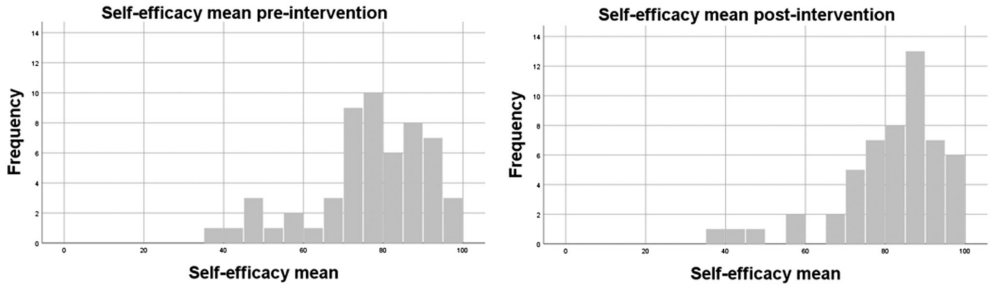

Figure 3. Histograms of self-efficacy mean values of the 55 students pre- and post-intervention.

Self-efficacy and text quality

Correlations were calculated between the self-efficacy mean pre- and post-intervention and the text quality pre- and post-intervention. Results are found in Table 2 (pre-intervention) and Table 3 (post-intervention). Self-efficacy mean pre-intervention had a strong, significant correlation to self-efficacy mean post-intervention ($r = .816$, $p = .000$). Self-efficacy pre-intervention and text quality pre-intervention had a significant, moderate correlation ($r = .394$, $p = .003$), and the same was found for self-efficacy and text quality post-intervention ($r = .322$, $p = .021$). Thus, self-efficacy and text quality had an association, but not as pronounced as the association of self-efficacy pre- and post-intervention. Text quality as measured by trained raters was considerably lower than students' perception of their skills as shown by their self-efficacy [2]. This means that the students showed a general "overestimation" of their capability compared to their actual performance as measured by text quality.

Correlations between self-efficacy statements (Table 1) and text quality were calculated. Pre-intervention, nine of the 18 self-efficacy statements were moderately, statistically significantly correlated with text quality (Table 2). These nine statements were all included in the 13 statements in which self-efficacy increased significantly (Table 4). Post-intervention, statements 4 (I can use periods, question marks and exclamation marks in the right places), 6 (I can partition my text into paragraphs), and 7 (I can write a heading that fits with the content) had a statistically significant correlation with text quality (Table 4). None of the three was significantly correlated with text quality pre-intervention. When there is little variation in the values, correlations will not be significant.

Summary of results

The students demonstrated strong self-efficacy pre- as well as post-intervention, however with considerable individual variability. The self-efficacy mean increased significantly after intervention, even though the effects of the intervention on performance, i.e. the increased text quality, were mild, as our previous study showed [27]. Scores on 13 out of the 18 self-efficacy statements increased significantly. There were no interactions between gender and intervention effects for individual statements or for the self-efficacy mean. Although girls' texts were assessed as having higher text quality [27], boys and girls demonstrated similar self-efficacy. At group level, there were moderate, statistically significant correlations between the self-efficacy mean and text quality, pre- as well as post-intervention. There were statistically significant correlations between half of the self-efficacy statements and text quality pre-intervention, but only three statistically significant correlations post-intervention.

Discussion

In the present study, our aim was to explore self-efficacy for narrative writing in girls and boys, before and after an observational learning intervention. We also aimed to explore associations between self-efficacy and text quality in boys and girls, as our previous study on the same students showed higher text quality in girls' texts than boys' texts [27]. The students wrote personal narratives and filled out a self-efficacy scale before and after a short observational learning intervention. Our first research question was whether and in what way self-efficacy for narrative writing changes after intervention. Most students displayed strong self-efficacy already before intervention, and the self-efficacy ratings increased after intervention. There were, however, large individual variations. To sum up, at group level self-efficacy was strong, especially in relation to the text quality ratings made by trained raters. Strong self-efficacy in this age group is consistent with earlier research [19,25]. Social cognitive theory postulates several sources of self-efficacy [1,2,18]. Our intervention design embraces several of these sources. Observing video clips of 12 different peers offered the students many examples of how peers work with narrative writing ("vicarious experience"). Thus, the students had access to a richer and wider range of skills than using film clips of only an expert writer, which has been tried in e.g. a study by van de Weijer et al. [37] or showing only two contrasting film peers, e.g. as in the study by Braaksma et al. [9], in which the video clips showed only two students at a time working with texts. During the intervention in the present study, one important component consisted of time for

structured reflection, where the participating students were encouraged to talk about the filmed peer models' oral or written contributions. This was done in small groups and shared with the whole class and can be considered as a form of "social persuasion". Although students were not receiving direct feedback on their written texts, they were given "enactive experience" when they wrote them. The second part of the first research question was whether there were gender differences in self-efficacy ratings. No such differences were found.

Our second research question was whether, and in what way, self-efficacy was associated with text quality. The self-efficacy results were compared to intervention effects on text quality from our previous study [27] on the same students. There was a statistically significant, moderate correlation between self-efficacy mean and text quality before intervention. Post-intervention, the correlation was less pronounced, but still moderate. The importance of self-efficacy for writing performance has been discussed by several researchers, and evidence remains inconclusive. For example, in a recent study by Graham et al. [38] of older secondary school students, it was shown that writing attitudes and self-efficacy accounted for statistically significant and unique variance in essay-writing after a range of other variables were controlled for, i.e. gender, eligibility for free lunches, reading self-efficacy, and first language. These findings indicate that motivational factors such as self-efficacy and attitudes towards writing must be taken into account in research on students' writing and writing development. Nine self-efficacy statements had statistically significant correlations with text quality pre-intervention. Post-intervention, three statements had statistically significant correlations with text quality. These concerned punctuation (4), partitioning one's text into paragraphs (6), and writing headings (7). The three statements did not have statistically significant correlations with text quality pre-intervention. Interestingly, few students actually wrote headings for their narrative texts and very few divided their texts into paragraphs. If there is a lack of distribution along a scale, correlations will not be significant, which may explain that there are only three statistically significant correlations post-intervention, as at least statements 4 and 7 had very high values. Another possible explanation may be that the students got insights into their own (and peers') competence in these aspects of writing from the intervention, and were able to rate their self-efficacy more accurately post-intervention, though they did not, for some reason, demonstrate these particular aspects of writing in their texts. Our results show that students had strong beliefs about their writing capabilities which is in accordance with previous studies [9,11,19]. Previous research also indicates that an "overestimation" of skills, i.e. stronger self-efficacy than (raters' assessment of) writing performance, can be expected in this age group [7,14]. In the present study, no students' self-efficacy mean decreased after intervention. Instead, the self-efficacy mean and the score on most self-efficacy statements increased significantly. We found no significant differences in self-efficacy between girls and boys, neither before nor after intervention. However, as our earlier study showed that girls' text quality was higher [27], the girls did not overestimate their capability in relation to performance (text quality) to the same degree that the boys did. The argument that girls may be more cautious when reporting self-efficacy [23] may thus apply to the results in this study.

There are some drawbacks in the design that may explain why no student had decreased their self-efficacy score. For logistic and practical reasons, the same two researchers distributed self-efficacy scales, administered the narrative writing tasks and performed the intervention lessons. This may have influenced students trying to be compliant towards the researchers by stating increased self-efficacy. The increase in self-efficacy due to study participation per se, i.e. so-called Hawthorne effects [39], could also explain the increase. Just performing the intervention could make students feel more confident in their abilities. However, the text quality as measured by trained raters not aware of intervention or that each student wrote more than one text, also increased significantly (around seven points on the 0–100 scale) [27].

The validity of the scale merits some further considerations. One issue is internal consistency between statements in the self-efficacy scale. In our study, the correlations between statements in the self-efficacy scale and the self-efficacy mean were strong pre- as well as post-intervention, indicating that the students' self-efficacy for narrative writing was quite holistic. The mean values of the 18 self-efficacy statements were higher at post-intervention than pre-intervention (with similar standard deviations). Some of the self-efficacy statements were related to specific aspects of narrative writing (e.g. statement 2, "I can find all the letters on the keyboard"). In other statements, there was an overlap between form, content, and organization or between process and product aspects of narrative writing. As an example, statement 9, "I can divide my story into beginning, middle, and end", may concern writing processes more in this age group than in older adolescents, as they still have not mastered lower-level writing processes fully [40,41]. Thus, we cannot conclude with certainty that self-efficacy for specific aspects of writing increased as an effect of the intervention. This may, however, be further explored by relating each self-efficacy statement to the other statements.

The difficulty in deciding the level of task-specificity has been addressed repeatedly [9,11,18]. The level of specificity which was chosen for the scale in this study aligned with Bandura's recommendations [18]. Further, the scale should concern belief in one's capability ("I can … "), rather than e.g. one's intent or judgment of performance [18]. Our measure of self-efficacy for narrative writing for 10–12-year-olds was adapted from Braaksma et al. [9]. Their scale was targeting self-efficacy for argumentative writing in teenagers, but had the same number of statements, which were similar in their specificity for the genre. Five statements which were independent of genre were similar or identical, e.g. "I can write a text which is one page long" (statement 18) and "I can partition my text into paragraphs" (statement 6). The genre- and age-specific statements were adapted to suit the narrative genre and the age group of the present study,

basing the content on the Swedish curriculum. For example, the statement "I can present the arguments and subordinate arguments in a structured way in my text" from the original scale was adapted to "I can divide my text into beginning, middle and end" (statement 9). Thus, it was not our intention to make the statements more specific for the task at hand, as such scales may end up in a pure self-assessment scale. The students in the present study had not been made explicitly aware of all the skills mentioned in the statements during intervention, and they were not told which peers showed stronger or poorer writing abilities.

Despite using a different scale than previous researchers, our results corroborate earlier findings reporting a somewhat poor calibration, i.e. accuracy of self-efficacy, in relation to writing performance [9,11,19]. A self-efficacy for writing scale ("SEWS") for middle and high school students by Bruning et al. comprises statements representing three levels of self-efficacy: *ideation* (capability to generate ideas), *convention* (capability to express ideas with language-related tools), and *self-regulation* (capability to manage one's behavior and writing decisions during writing) [19,20]. This scale was also used by De Smedt et al. [16]. SEWS was not available when the present study was planned and piloted in 2013, but still, it might not have been a better choice of scale for the current study. A third of the statements in SEWS regard self-regulation, which may require metacognitive skills not yet developed in students of our age group. In our scale, we wanted to avoid tapping meta-cognitive skills such as self-assessment or self-regulation. However, the question remains whether we succeeded. It is a challenging task to construct valid and reliable self-efficacy scales in younger students. This can offer an explanation to why children in this age span perhaps have difficulties in differentiating their skills [12,19]. If they consider their abilities in a holistic way, this may be a reason why it is difficult to isolate clear findings in this study. It is not straightforward to disentangle whether students report on their general capability to write good narratives (i.e. self-efficacy) or if they report their self-regulation capacity or their recent performance on a particular narrative, when they are responding to statements in self-efficacy scales.

In conclusion, the students in the present study had strong self-efficacy for narrative writing, which increased after intervention. This supports previous findings. The students may still have a holistic view of their capabilities, not separating their self-efficacy from related factors such as assessment of performance, self-regulation, or general writing skills. Self-efficacy had significant, moderate correlations to performance as measured by text quality, pre- as well as post-intervention. No interaction with gender was found. Constructing self-efficacy scales demands careful consideration. Different age groups, educational systems, and genres demand different scales, which have to be properly adapted to the specific educational context. Several research topics should be addressed in future studies including self-efficacy scales for writing pre- and post-intervention; the level of task specificity, the age-group, cognitive and linguistic prerequisites and explicit or implicit instruction. Exploring the relationship between different self-efficacy statements, and between specific aspects of writing and self-efficacy for those aspects may add to our understanding of young students' change from holistic to differentiated self-efficacy.

Acknowledgments

We thank Martine Braaksma for fruitful discussions on observational learning intervention and self-efficacy scales. We also thank the Linnaeus environment Cognition, Communication and Learning (CCL) at Lund University and Lund University Humanities Lab where the 'film peers" for the observational learning intervention were filmed. Lastly, we thank the anonymous reviewers, whose comments have greatly improved this manuscript.

Disclosure statement

The authors report no conflict of interest.

Funding

The present work was financially supported by The Marcus and Amalia Wallenberg foundation [2012.0038].

Notes on contributors

Emily Grenner is a speech-language therapist and PhD student at the Department of Clinical Sciences, Logopedics, Phoniatrics and Audiology, Lund University, Sweden. Her PhD studies concern processes and products of narrative writing in school children with normal hearing and school children with hearing loss.

Victoria Johansson is an Associate Professor in Linguistics at the Centre for Languages and Literature and Deputy Director at the Lund University Humanities Lab, Lund University, Sweden. Her research focus is on language development through the lifespan, primarily cognitive aspects of writing.

Joost van de Weijer is an Associate Professor in Linguistics at the Centre for Languages and Literature at Lund University. His research focuses on the perception of foreign-accented speech. He is also affiliated as methodologist with the Lund University Humanities Lab, where he mainly works with the analysis of experimental data.

Birgitta Sahlén is a Professor in Speech Pathology at the Department of Clinical Sciences, Logopedics, Phoniatrics and Audiology, Lund University, Sweden. Her research focus is on cognition and communication in children with language disorder and/or hearing loss. Comprehension, listening effort, motivation, and learning in relation to children's cognitive capacity. She is currently heading comprehensive intervention projects aiming at improving children's narrative writing by observational learning, and teachers' communicative techniques for language learning interaction in the classroom.

ORCID

Emily Grenner http://orcid.org/0000-0002-6818-1093
Victoria Johansson http://orcid.org/0000-0002-0684-6491
Joost van de Weijer http://orcid.org/0000-0001-9843-3143
Birgitta Sahlén http://orcid.org/0000-0002-8468-0546

References

[1] Bandura A. Social foundations of thought and action: a social cognitive theory. Englewood Cliffs (NJ): Prentice-Hall; 1986.

[2] Bandura A. Social learning theory. Englewood Cliffs (NJ): Prentice-Hall; 1977.

[3] Bandura A. Self-efficacy: the exercise of control. Basingstoke: W. H. Freeman; 1997.

[4] Schunk DH. Self-efficacy and academic motivation. Educ Psychol. 1991;26:207–231.

[5] Coutinho SA, Neuman G. A model of metacognition, achievement goal orientation, learning style and self-efficacy. Learn Environ Res. 2008;11:131–151.

[6] Bandura A. Perceived self-efficacy in cognitive development and functioning. Educ Psychol. 1993;28:117–148.

[7] Bruning R, Kauffman DF. Self-efficacy beliefs and motivation in writing development. In: MacArthur C, Graham S, Fitzgerald J. Handbook of writing research. New York: Guilford Publications; 2016.

[8] Rijlaarsdam G, Braaksma M, Couzijn M, et al. Observation of peers in learning to write: practice and research. J Writ Res. 2008;1:53–83.

[9] Braaksma MAH, Rijlaarsdam G, van den Bergh H. Observational learning and the effects of model-observer similarity. J Educ Psychol. 2002;94:405–415.

[10] Raedts M, Rijlaarsdam G, van Waes L, et al. Observational learning through video-based models: impact on students' accuracy of self-efficacy beliefs, task knowledge and writing performances. In: Hidi S, Boscolo P, editors. Writing and motivation. Oxford: Elsevier; 2007. p. 219–238.

[11] Raedts M, Steendam E, Grez L, et al. The effect of different types of video modelling on undergraduate students' motivation and learning in an academic writing course. J Writ Res. 2017;8:399–435.

[12] Muenks K, Wigfield A, Eccles JS. I can do this! The development and calibration of children's expectations for success and competence beliefs. Dev Rev. 2018;48:24–39.

[13] Pajares F. Empirical properties of a scale to assess writing self-efficacy in school contexts. Meas Eval Counsel Dev. 2007;39:239–249.

[14] Stipek D. Good instruction is motivating. In: Wigfield A, Eccles JS, editors. Development of achievement motivation. San Diego (CA): Academic Press; 2002. p. 309–332.

[15] Schunk DH, Meece JL. Self-efficacy development in adolescences. In: Pajares F, Urdan T, editors. Self-efficacy beliefs of adolescents. Greenwich (CT). Information Age Publishing; 2006. p. 71–96.

[16] De Smedt F, Merchie E, Barendse M, et al. Cognitive and motivational challenges in writing: studying the relation with writing performance across students' gender and achievement level. Read Res Q. 2018;53:249–272.

[17] Klassen R. A question of calibration: a review of the self-efficacy beliefs of students with learning disabilities. Learn Disabil Quart. 2002;25:88–102.

[18] Bandura A. Guide for constructing self-efficacy scales. In: Pajares F, Urdan T, editors. Self-efficacy beliefs of adolescents. Greenwich (CT): Information Age Publishing; 2006. p. 307–337.

[19] Bruning R, Dempsey M, Kauffman DF, et al. Examining dimensions of self-efficacy for writing. J Educ Psychol. 2013;105:25–38.

[20] Bruning R, Dempsey M, Kauffman DF, et al. Self-efficacy for writing scale. Database record retrieved from PsycTESTS; 2013.

[21] OECD. PISA 2015 results. (Volume I), Excellence and Equity in Education, PISA. Paris: OECD Publishing; 2016. doi:10.1787/9789264266490-en

[22] Pajares F, Miller MD, Johnson MJ. Gender differences in writing self-beliefs of elementary school students. J Educ Psychol. 1999;91:50–61.

[23] Pajares F, Valiante G. Grade level and gender differences in the writing self-beliefs of middle school students. Contemp Educ Psychol. 1999;24:390–405.

[24] Pajares F, Valiante G. Gender differences in writing motivation and achievement of middle school students: a function of gender orientation? Contemp Educ Psychol. 2001;26:366–381.

[25] Pajares F. Self-efficacy beliefs, motivation, and achievement in writing: a review of the literature. Read Writ Quart. 2003;19:139–158.

[26] Tillema M, van den Bergh H, Rijlaarsdam G, et al. Quantifying the quality difference between L1 and L2 essays: a rating procedure with bilingual raters and L1 and L2 benchmark essays. Lang Test. 2013;30:71–97.

[27] Grenner E, Åkerlund V, Asker-Árnason L, et al. Improving narrative writing skills through observational learning: a study of Swedish 5th-grade students. Educ Rev. 2018. doi:10.1080/00131911.2018.1536035

[28] Feagans L, Appelbaum MI. Validation of language subtypes in learning disabled children. J Educ Psychol. 1986;78:358–364.

[29] Johansson V. Developmental aspects of text production in writing and speech [dissertation]. Lund: Department of Linguistics and Phonetics, Centre for Languages and Literature, Lund University; 2009.

[30] McCutchen D. From novice to expert: implications of language skills and writing-relevant knowledge for memory during the development of writing. J Writ Res. 2011;3:51–68.

[31] Wolff U. Lilla DUVAN. Stockholm: Hogrefe Psykologiförlaget AB; 2010.

[32] Bishop D. Test for reception of grammar. version 2. Swedish version. Stockholm: Pearson Education; 2009.

[33] Magnusson E, Nauclér K. SL40: läsförståelse: meningar [SL40: reading comprehension: sentences]. Lund: Studentlitteratur AB; 2010.

[34] Skolverket. Curriculum for the compulsory school, preschool class and school-age educare 2011. Stockholm: Norsteds juridik; 2018.

[35] Wewers ME, Lowe NK. A critical review of visual analogue scales in the measurement of clinical phenomena. Res Nurs Health. 1990;13:227–236.

[36] Strömqvist S. Discourse flow and linguistic information structuring: explorations in speech and writing (Gothenburg papers in theoretical linguistics: 78). Gothenburg: Göteborg University; 1996.

[37] van de Weijer J, Åkerlund V, Johansson V, et al. Writing intervention in university students with normal hearing and in those with hearing impairment: can observational learning improve argumentative text writing? Logoped Phoniatr Vocol. 2018;44:1–9.

[38] Graham S, Daley SG, Aitken AA, et al. Do writing motivational beliefs predict middle school students' writing performance? J Res Read. 2018;41:642–656.

[39] Schwartz D, Fischhoff B, Krishnamurti T, et al. The Hawthorne effect and energy awareness. Proc Natl Acad Sci USA. 2013;38:15242–15246.

[40] Bourdin B, Fayol M. Even in adults, written production is still more costly than oral production. Int J Psychol. 2002;37:219–227.

[41] Chanquoy L. Revision processes. In: Beard R, Myhill D, Nystrand M, editors. The SAGE handbook of writing development. London: Sage Publications; 2009. p. 80–97.