

Bente Rigmor Walgermo

Universitetet i Stavanger

Per Henning Uppstad

Universitetet i Stavanger

Kjersti Lundetræ

Universitetet i Stavanger

Finn Egil Tønnessen

Universitetet i Stavanger

Oddny Judith Solheim

Universitetet i Stavanger

DOI: <http://dx.doi.org/10.5617/adno.8136>

Screening Tests of Reading: Time for a Rethink?¹

Abstract

At present, the national Norwegian quality-assessment system for basic education (NKVS) includes a range of reading tests with different purposes and designs, to be administered at various points during a 13-year period of schooling. In this article, we trace the historical background to the introduction of screening tests of reading in early education and identify a number of circumstances that call for a rethink of the overall test concept. Using data from a longitudinal study, we show that scores on a brief task administered towards the end of grade 1 predict reading-comprehension difficulties in grade 3. Taking that finding as our starting point, we discuss an idea for a possible new test concept that might have the potential to (a) improve the match between reading theory and reading-skill measurements, (b) enable longitudinal prediction, and (c) take less time to administer and be more useful in an educational context. The central elements of this new test concept are a short initial test meeting the primary objective of identifying students at risk of developing reading difficulties, and a follow-up explorative part to be carried out one-on-one, which will provide the teacher with information about the nature of each student's difficulties.

Keywords: reading assessments, screening for reading difficulties, longitudinal prediction, early efforts, dyslexia, reading and writing difficulties, special-needs education

Kartleggingsprøver i lesing: Tid for nytenking?

Sammendrag

Som en del av det nasjonale kvalitetsvurderingssystemet for grunnsopplæringen (NKVS) i norsk utdanning har vi i dag forskjellige leseprøver med ulike formål og utforminger

¹ A Norwegian version of this article has formerly been published in this journal: Walgermo, B. R., Uppstad, P. H., Lundetræ, K., Tønnessen, F. E., & Solheim, O. J. (2018). Kartleggingsprøver i lesing – tid for nytenking? *Acta Didactica Norge*, 12(4), Art. 7. <https://doi.org/10.5617/adno.6499>

for bruk i det 13-årige skoleløpet. I denne artikkelen tegner vi opp en historisk bakgrunn for innføringen av kartleggingsprøvene i lesing i begynneropplæringen, og peker på flere forhold som krever at man tenker nytt om dette prøvekonseptet. I artikkelen viser vi ved hjelp av data fra en longitudinell studie hvordan en kort oppgave gjennomført i slutten av 1. klasse forutsier vansker med leseforståelse i 3. klasse. Med utgangspunkt i disse resultatene drøfter vi kimen til et mulig nytt prøvekonsept med potensial for a) bedre samsvar mellom teori om lesing og måling av lesing, b) longitudinell prediksjon, og c) en prøve med forbedret pedagogisk potensial som kan gjennomføres på kortere tid. Sentralt i dette forslaget står tanken om en kort inngangsprøve som oppfyller hovedformålet om å identifisere de elevene som står i fare for å utvikle vansker med lesing. Denne følges av en utforskende del som gjennomføres en-til-en og gir læreren informasjon om hvordan vanskene arter seg.

Nøkkelord: kartleggingsprøver i lesing, longitudinell prediksjon, tidlig innsats, dysleksi, lese- og skrivevansker, intensiv opplæring

Introduction

At present, the national Norwegian quality-assessment system for basic education (NKVS) includes a number of reading tests to be administered at various points during a 13-year period of schooling: there are screening tests for reading (grades 1–3), national reading tests (grades 5, 8 and 9) and reading tests intended as support for learning (grades 5–7). The compulsory tests – the screening and national ones – have been a cornerstone of Norwegian education policy and education management since the early years of the 21st century.

In this article, we first trace the historical background to the objective and design of the grade 1 screening test of reading. After pointing out some of the challenges associated with the present version of the test, we go on to investigate the extent to which the various sub-tests of the present grade 1 screening test are necessary to meet the primary objective of that test: identifying those students who will need extra help with reading during their first years of school. Based on that analysis, we then discuss the characteristics of a possible alternative test concept retaining that primary objective.

Background to the objective and design of the tests

Ever since work to design screening tests began in the early 1990s, the objective has been to find those students who need extra support in their reading development. This work was prompted by teachers' need for a screening test that they could administer in the classroom and that could investigate as many aspects as possible of students' reading skills in order to reveal any difficulties, but without probing as deeply as diagnostic tests do (Tønnessen & Solheim, 1998).

The national tests aim not only to give teachers information enabling them to assess their own instruction, but also to provide principals, local education authorities, private owners and regional authorities with data to inform their

efforts to improve and develop schools. For this reason, the national tests are designed to yield information about the reading skills of students *at all skill levels*. By contrast, the screening tests are first and foremost a tool to identify students needing additional follow-up of their reading development:

The objective of the screening tests is to help schools and teachers find those students who need extra help and support in their reading development during the first years of school. (Directorate for Education and Training, 2018; English translation by the present authors)

Hence, the screening tests are designed to provide as much information as possible about the reading skills of the *lowest-performing* students.

The first three screening tests of reading, for grades 1, 2 and 6, were developed between 1992 and 1995. At that time they were not compulsory, but they proved so popular that work immediately began to develop corresponding tests for grades 5 and 9 as well. In 1997, the school-starting age in Norway was lowered from seven to six years and the duration of compulsory school was correspondingly increased from nine to ten years. This entailed major changes in early reading instruction. The idea was for the new first year (grade 1) to prepare students for school proper, meaning that formal instruction would not begin until grade 2. This made it necessary to reassess the existing screening materials to determine at what points they should be administered in the new schooling structure. The 1997 reform stressed that systematic work to learn the letters should not begin before (the new) grade 2, and this strongly influenced the decision taken in 2000 to administer compulsory screening only after grade 2 (and also in grade 7). The grade 1 test was to remain voluntary until 2009, and the grade 3 screening test did not become compulsory until 2010.

The ‘PISA shock’ of 2001 also had an impact on early reading instruction. This shock was the aftermath of the clash between the reality of Norwegian 15-year-olds scoring only in line with the OECD average for reading on the PISA tests and the widespread perception among those involved in Norwegian school policy that their country’s schools were excellent. The calls for better learning outcomes and a reorientation of education policy that the PISA shock prompted have entailed that formal reading and writing instruction now begins in grade 1 at most Norwegian schools. Gabrielsen and Lundetræ (2017) point out that the introduction, in 2009, of compulsory screening tests in grade 1 must be seen as “a clear signal that teachers should introduce the alphabet and work systematically with reading and writing even in grade 1” (Gabrielsen & Lundetræ, 2017, p. 209; English translation by the present authors).

As can be seen from the above, the ideas about what grade 1 should encompass have changed since the introduction of the screening tests. Even so, there has been a strong awareness throughout that period that grade 1 students exhibit wide variation in terms of their reading skills. For example, in relation to the first generation of screening tests, it was claimed that it was pointless to design a test

for grade 1 students that would yield normally distributed scores, because those students have such highly different starting points:

Reading skills are developing in grades 1 and 2, meaning that there is little point in normalising tests based on the statistical distribution of scores along a Gaussian curve. What this would yield is in fact a seemingly exact measure of a highly unstable skill. It was therefore decided that the purpose of the tests in the first two grades would not be to make a fine-grained assessment of the most accomplished students' skills but first and foremost to help identify the 15–20% weakest readers. (Tønnessen & Solheim, 1998; English translation by the present authors)

Hence, it was concluded that the tests in grades 1 and 2 would not aim to provide fine-grained information about all students but only identify the weakest readers. A 'concern cut-off' encompassing the 20% lowest-performing students was set, based on teaching experience as well as research findings to the effect that about 20% of the student population are at particular risk of developing different types of reading difficulties (Shaywitz, Fletcher, Holahan, & Shaywitz, 1992; Tønnessen & Solheim, 1998). However, one inevitable corollary of having the test measure various aspects of reading with the best possible accuracy for the weakest students, was that there would be a ceiling effect. Specifically, many students would obtain the top score on the sub-tests, and hence the best readers would not have the opportunity to show how good they really were. Another reason for choosing the test design in question was to ensure that the test would not take too long to administer. A test providing a comprehensive measure of reading skill would necessarily have to be more extensive. Further, it was also explicitly stated that the tests should help give the weakest readers a sense of mastery. This could obviously not be accomplished if those students were confronted with a large number of tasks that were far beyond their capacity. The idea of focusing on identifying the weakest students in a test with a ceiling effect made it an obvious choice to have a fairly large number of sub-tests, but to let each sub-test include a fairly small number of items, so that the test could be administered during a 45-minute classroom session. The existing test is administered during a pre-determined period in March–April and the time set aside for it is 60 minutes.

The screening tests thus serve as a tool to identify students in need of extra support for their reading development. A further intended purpose of these tests has been to guide the development of teaching practices. In a doctoral thesis based on the work to design the first screening tests of reading for grades 1 and 2, Engen (1999) described how those tests were also supposed to provide guidance about good reading instruction in classrooms:

We wanted [the tests] to help draw teachers' attention to key aspects of the reading process, to give them knowledge about the skills that were important to observe and advice about how this could be done. [...] For this reason, we assumed that the aspects focused upon in the tests would eventually also have an impact on reading instruction in general and that the tests would thus be of indirect importance for the planning of

teaching as well as in preventive efforts. (Engen, 1999, p. 31; English translation by the present authors)

Hence, the tests were supposed to provide information that would be of assistance in the design and performance of instruction for beginning readers and students with special needs. They were also supposed to be experienced as useful in the assessments carried out by schools to determine whether a given student needed to be referred to a school psychologist (Tønnessen & Solheim, 1998). However, with the exception of a handful of Master's dissertations (see, e.g., Barkved, 2012) establishing that the tests do fill a need, no empirical studies have thoroughly investigated how the introduction of screening tests for reading have changed practices and knowledge at schools. In the present article, we claim that it is now time for a rethink. While part of our reason for doing so is that we are aware that the tests have had an unfortunate effect on classroom practices, a further rationale is that the existing screening tests strongly reflect and implement theoretical perspectives that were dominant at a certain time and within a certain tradition but have since been forcefully challenged.

Theoretical background

The work to develop the screening tests of reading for grade 1 built mainly on research into initial reading instruction and reading difficulties (Engen, 1999). Given that one explicit aim was to provide information that could support teachers' instruction, it was decided to focus on certain aspects of the reading process which, according to a range of studies, were crucial for learning how to read – and for which it had also been found that difficulties with the skills in question predicted reading difficulties (Schneider, Roth, & Ennemoser, 2000). Based on Høien and Lundberg (1991), it was attempted to apply an analytical view of the reading process in that the various sub-tests were chosen to represent different components of the reading process (Engen, 1999). Alongside sub-tests targeting students' actual reading skills, other sub-tests therefore focused on their letter knowledge and phonemic awareness. The purpose of the latter sub-tests was to draw attention to areas that it was considered important for teachers to focus on during initial reading instruction, based on the idea that it was possible to target interventions directly to those areas. Hence, the screening tests of reading for grade 1 can be said to reflect a view of reading as consisting of sub-processes where reading is assumed to require specific knowledge and skills (Høien & Lundberg, 1991; Høien, 2012; Engen, 1999). This view can still be seen in the present design of the screening test for grade 1, which consists of six sub-tests for the following aspects: letter knowledge; positional (phonological) analysis; phonological synthesis; spelling of simple, regularly spelled monosyllabic words; word reading; and reading and comprehension of simple sentences.

Challenges of today's screening tests

An account for the 20-year history of the design of the screening test should also describe the experiences gained from using the test as well as the specific challenges in relation to the administration of the test in the classroom, the content of the test and the teachers' interpretation of test scores. Those challenges can be said to concern the following: (1) test length; (2) unintended consequences for classroom practices; (3) incorrect use of test scores; (4) varying rates of progression in initial reading instruction; and (5) a high threshold for using test scores.

One obvious challenge associated with the screening test relates to its *length* and hence to the time required to administer it. The test may be particularly taxing on those students who struggle the most – 60 minutes is a long time to spend working on something that you do not really feel that you have mastered. It is also unfortunate that all students perform an extensive test which is only designed for some of them in that it aims to yield information in areas that are relevant mainly for those students who struggle with reading. Further, the fact that the test is administered in a whole-class setting, makes it difficult to obtain a nuanced picture of exactly why the weakest readers are struggling. There is a wide range of possible reasons why a given student does not answer an item or gives the wrong answer to it. With group administration, teachers have little opportunity to observe how the weakest readers solve the items. In other words, it is possible to claim that the screening test is too extensive both for the strongest readers and for the weakest ones – but for different reasons. The new test design should be such that the test is shorter for those that it is not intended for while providing more information about those that it is meant for and also giving them a stronger sense of mastery.

A second challenge is that the design of the screening test has had *unintended consequences for classroom practices*. In particular, this is because the test is based on a theory which places a strong focus on the isolated training of phonological skills. More recent research has shown that the expectations placed on such training were excessive (Gustafson, Samuelsson, & Rönnerberg, 2000). The best effect is in fact attained by combining elements focusing on wholes with elements focusing on parts, that is, by focusing on both decoding and comprehension (National Reading Panel, 2000; Suggate, 2010; Ehri et al., 2001; Scammacca et al., 2007). What is more, while one of the original aims of the screening test was to promote good classroom practices (Engen, 1999), experience shows that teachers often have a limited understanding of what the individual sub-tests measure, of how a certain sub-skill contributes to overall reading skill and of how they might adjust their instruction in the light of their students' test scores. For example, the sub-test of phonological synthesis could be a good predictor of difficulties, but there would be little point in actually practising what that sub-test measures. In fact, many students who are struggling to gain phonemic awareness will reach that milestone only when they are introduced to the letters and discover

the links between them and language sounds (Cunningham, 1990). It is therefore highly unfortunate that there are reports of students being made to practise phonological synthesis for extended periods of time without working on letters at the same time, on the ground that this is targeted by one of the sub-tests of the screening test in grade 1. Such training of phonological skills in isolation from work on letters, letter patterns and words is an example of an unfortunate practice (as is ‘teaching to the test’ in general). To counteract such practices, there is a need for a more functional test design which is capable of harnessing the feedback effect that tests often have on classroom practices (Johnson & Kress, 2003) to actually help teachers become better at *teaching students how to read*. In other words, a new test concept should be so designed that teachers’ classroom practices are guided in a more intuitive way towards the core of good reading instruction, that is, towards actual reading.

A third challenge relates to *incorrect use of test scores*. Ever since the screening test came into being in the 1990s, it has consistently proved very difficult to convey information about its specific objective and limitations to teachers and principals alike. There has been a clear tendency for them to believe that the objective of the screening test is the same as that of the national tests. Hence, although it is clear from its guidance documents that the screening test is intended as a tool to identify students in need of extra help and support in their reading development, test scores are very often used to say something about the reading performance of all students in a classroom, for example by providing all parents with information about their children’s test scores at a parents’ meeting. Here, however, part of the blame for the misunderstanding of the purpose of the test may also be laid on the Norwegian name for it, literally ‘mapping test’, which (unlike perhaps the English word ‘screening’ with its connotations of picking some out among many) has been wrongly taken to mean that the test yields some sort of ‘map’ of each student taking it – when it actually provides next to no information about 60–70% or so of them. It should be noted that a misunderstanding of the objectives of the test is also reflected in the above-mentioned practice of extensive exercises geared specifically to the test. It is unclear whether that misunderstanding is to be found mainly with the teachers themselves or with their principals.

A fourth challenge relates to *variation in the rate of progression during grade 1*. This is because all students in Norway are screened at a pre-determined time, regardless of the progression in terms of letter introduction and reading instruction that characterises the teaching they have received. The national curriculum does not set any competence objectives for reading and writing until for the end of grade 2 (Ministry of Education and Research, 2017), and teachers are free to choose the rate of progression for their initial reading instruction that they believe will be the most effective in attaining those competence targets. Based on a questionnaire survey to which a large number of Norwegian teachers responded, Rasmussen (2013) reported that the rate of progression in initial reading

instruction tended to be relatively rapid and relatively homogeneous. The vast majority of teachers had introduced all letters of the alphabet before Easter in grade 1 – but a small proportion of them did not complete the introduction of the letters until grade 2. This challenge obviously also has an effect on the validity of the screening test. Moreover, there is much to indicate that the rate of progression for the introduction of the letters is at present subject to rapid change (Educational Research, 2017). There is an ongoing large-scale shift towards greater use of tablet computers and new methodologies involving more rapid introduction of the letters, and this is likely to further increase the diversity of practices found across grade 1 classrooms.

Finally, the fifth challenge pertains to an excessively *high threshold for using test scores*. This does not imply that the tests are bad, only that their users need a certain competence to be able to translate test scores into good practices. The guidance documents of the test do describe to a large extent how to go from score to intervention, but many teachers report making little use of those documents (Barkved, 2012). This means that the road to educational usefulness is long. It is clear from discussions with teachers about the screening test that they would prefer the road from test to action to be as short as possible. Teachers' lack of competence, as well as their sense of being under surveillance, may also cause test scores to obtain an unfortunate authoritative status, isolated from the real function of the test. Hence it is important to use all means available to gear a new test concept to the original purpose of the test.

The present study

The challenges associated with today's screening test give us concrete indications of the needs that a new test must meet. In the following, we investigate whether it is possible to radically reduce the extent of the test by studying how first-graders' scores on a test of sentence reading predicts their reading-comprehension difficulties in grade 3. This has not been possible to study before, owing to a lack of longitudinal data where the current sub-tests of the screening test are used as measures. This, in turn, is due to the fact that screening-test scores are not much used in research because, on their own, they do not provide good information about the reading skills of all students, only about those of the lowest-performing ones. In the present project, the original test was used along with an additional component which makes it possible to measure all students' skills while using the original test format (see Solheim et al., 2018; Lundetræ, Solheim, Schwippert, & Uppstad, 2017). The reason for investigating the predictive value of the grade 1 test is an aim to identify, early on and with good precision, those students about whom the teacher needs more information in order to be able to carry out early interventions against reading difficulties.

By taking the sentence-reading sub-test as our starting point, we have chosen, among the various aspects addressed by the grade 1 test, the one that most directly concerns reading comprehension. Then we have compared scores on that sub-test

with a measure of reading comprehension included in the screening test administered in grade 3. Hence, this also represents an attempt to emphasise those aspects of reading that are in line with more recent views on reading. Those views involve a shift away from a focus on isolated training of individual components of the reading process towards a perception of reading as an interpretive skill (Tønnessen & Uppstad, 2015). This is not a shift *away* from the parts, in favour of the whole. On the contrary, this shift actually underscores the importance of the parts. The point is that they should not be worked on in isolation; rather, we should work on parts *and* wholes in an integrated manner. The principles underpinning this way of working and thinking come from the field of hermeneutics (Gadamer, 1960; Tønnessen & Uppstad, 2015).

Methodological perspectives

The data used in the present study come from a randomised controlled study called 'On Track' ['På sporet'] which was carried out under the aegis of the Stavanger Reading Centre (Lundetræ et al., 2017). The main objective of the On Track study is to help reduce the percentage of students who develop reading and writing difficulties by designing (i) screening tools that teachers can use to identify at-risk students at an early stage of their schooling, and (ii) research-based interventions for that group of students. The On Track intervention consisted of four weekly 45-minute sessions over a 25-week period. During those sessions, the remaining students in the classrooms concerned took part in station-based instruction relating to reading and writing activities. The intervention groups were led by a teacher who had received training in administering the intervention programme. A teacher's manual with detailed instructions for each session was developed as part of the project. Each of the 100 intervention sessions consisted of four ten-minute sub-sessions targeting ABC, guided reading, spelling and reading aloud, respectively. For more information about the intervention design, see Solheim et al. (2018).

The On Track project follows a large sample of students from 2014, when they first started school, to 2018, when they sat the grade 5 national reading tests. The data reported in the present study come from the screenings of the students in April of their first academic year and in April of the spring when they were in grade 3. All schools participating in the project agreed to have introduced their first-graders to all letters of the alphabet before Easter in grade 1.

Sample

The sample of the On Track study is a convenience sample consisting of 19 schools in the southern part of the Region of Western Norway, all located relatively close to the Reading Centre. A total of 1,199 students first started school in the autumn of 2014 at the 19 participating schools, and the parents of 97.7% of

them consented to having their child participate in the study. This yielded a sample of 1,171 first-graders (50.8% girls), of whom 18.6% had parents who spoke a language other than Norwegian at home and 14.1% had at least one parent who reported having (had) reading difficulties (Lundetræ et al., 2017). By the end of grade 3, the sample had shrunk slightly to 1,105 students because of relocation. Finally, one of the schools taught Nynorsk as its main variety of Norwegian (while all the others taught the other variety, Bokmål). That school was excluded from the sample, yielding a remaining sample of 996 students.

Procedures and measures

The screening tests of reading for grades 1 and 3 were used to obtain information about the students' reading skill at the end of the respective grade. The tests were administered by researchers from the Reading Centre in the students' own classrooms. Below is a presentation of the various sub-tests included in the grade 1 screening test as well as the reading-comprehension sub-test of the grade 3 screening test. In both cases, the time limit for the administration of the screening test of reading is 60 minutes.

Grade 1

The screening test of reading for grade 1 consists of six sub-tests targeting the following aspects: letter knowledge, positional analysis, phonological synthesis, spelling, word reading, and the ability to read and understand simple sentences. To ensure that students understand the task design, practice items are administered before the real items for each sub-test. Table 1 shows statistics for the individual sub-tests with regard to the sample of the present study.

Letter knowledge

This sub-test targets students' letter knowledge, or more precisely their letter recognition. The students are first shown a picture and then asked to write the first letter of the word for the object represented by the picture in a box next to the picture. Example prompt: "In the picture you see a car². What is the first letter of the word *car*? Write the first letter of the word *car* in the box next to the picture of the car."

Phonological awareness

Two aspects of students' phonological awareness are measured by two different sub-tests. In the first of them, relating to positional analysis, the test administrator reads out a simple word with regular spelling while the students are shown a picture of the object that the word refers to. For each item, the students are then asked to identify the first, middle *or* last phoneme of the word and to write the corresponding letter in a box next to the drawing. Example prompt: "What is the

² The Norwegian words used in the sub-tests have here simply been translated into English. No effort has been made to find English words matching the orthographic, phonological, etc., qualities of those Norwegian words.

last sound of the word *town*? Write the letter for the last sound of the word *town*.” In the second sub-test, targeting phonological synthesis, the students are shown four different drawings. The test administrator first says the word for each of the objects represented by the four pictures and then reads out, in correct order, the letter sounds of one of those words, one letter sound per second. The students’ task is to combine those letter sounds into a word and identify the picture corresponding to that word. Example prompt: “Here you see pictures of a horse, a hedge, a witch and biscuits. Listen carefully and put a cross on the picture that matches w-i-t-c-h.”

Spelling

The sub-test measuring students’ spelling skill consists of 14 words with regular spelling. The test administrator first reads out a sentence and then repeats one word from that sentence, asking the students to write it down. Example prompt: “Lise likes ice-cream. *Ice-cream* – write *ice-cream*.”

Word reading

The ‘From pictures to words’ sub-test, which is used to test students’ word-reading skills, contains 14 items. Each item consists of a picture followed by four similar words, of which only one corresponds to the picture (for example: *øks* ‘axe’, *økt* ‘work shift’, *øke* ‘increase’, *øse* ‘bail out water’). The students are asked to look at the picture, read the words and mark the word that best matches the picture. The idea for this sub-test originally comes from the working cards and reading booklets of Asheim (1981). The students are given five minutes for independent work on this sub-test. The distractor words (i.e., the non-correct options) have been chosen with a view to providing teachers with information about what is causing the difficulties experienced by those students who are struggling with word reading. Words that are semantically related to the target word have been included to reveal any naming difficulties. Words beginning with the same syllable or letter sequence as the target word have been included to help teachers identify students who are struggling with the perception of letter sequences, which may reflect a weakness in their phonological strategy. In addition, some distractors may resemble the target word visually, so as to uncover students who are guessing or who read logographically.

Sentence reading

Each item of the sentence-reading sub-test presents the students with a sentence consisting of simple words with regular spelling, accompanied by four pictures with inter-related content. The students’ task is to put a cross on the picture that best matches the meaning of the sentence. The idea for this sub-test comes from two Danish reading tests, SL 60 and SL 40 (Nielsen, Kreiner, Poulsen, & Sjøgaard, 1986). The students are given ten minutes to read the sentences and perform the tasks.

Table 1. Descriptive statistics relating to the sub-tests for letter recognition, positional analysis, phonological synthesis, spelling, word reading and sentence reading in the screening test for grade 1 ($N = 996$)

Sub-test	Letter recognition	Positional analysis	Synthesis	Spelling	Word reading	Sentence reading
Number of items	13	14	12	14	14	10
Average score	12.54	12.66	10.28	11.09	12.10	8.71
Standard deviation	1.38	2.30	2.06	3.45	2.73	2.35
Minimum/maximum score	0–13	0–14	0–12	0–14	0–14	0–10
Internal consistency (Cronbach's alpha)	.75	.79	.68	.84	.74	.82
'Concern cut-off'	< 12	< 11	< 9	< 8	< 9	< 5

Grade 3

The screening test of reading for grade 3 consists of sub-tests for the following four aspects: word reading, spelling, vocabulary, and reading comprehension. For the purposes of the present study, only the reading-comprehension measure is of interest.

Reading comprehension

The sub-test measuring reading comprehension in grade 3 consists of two texts. For descriptive statistics, see Table 2. The students read and answer multiple-choice items for one non-fiction text (*Dyr i ørkenen* 'Animals of the Desert') and one fiction text (*Hundevalpen* 'The Puppy'). Before the first part of the sub-test, the test administrator goes through a practice item showing how the students need to use the text to find the right answer to the items. According to the test guidelines, the 'concern cut-off' is fewer than eleven correct answers (of the students in the present sample, 260 (22.9%) obtained scores below that cut-off).

Table 2. Descriptive statistics relating to the reading-comprehension sub-test in the screening test for grade 3 ($N = 996$)

Sub-test	Reading comprehension
Number of items	15
Average score	13.04
Standard deviation	2.8
Minimum/maximum score	0–15
Internal consistency (Cronbach's alpha)	.74
'Concern cut-off'	< 11

Analysis

To investigate the extent to which students' scores on the individual sub-tests of the screening test of reading for grade 1 help to predict whether those students will score below the 'concern cut-off' for reading comprehension in grade 3, we carried out a logistic regression analysis using version 25 of the SPSS statistical

software. In a (binary) logistic regression, the dependent variable has two values – in this case 0 for students scoring above the cut-off and 1 for students scoring below it. Furthermore, the measure of the overall effect is called the odds ratio. If it is above 1, the likelihood of scoring below the cut-off is increased, and if it is below 1, the likelihood of scoring below the cut-off is reduced. For example, we would expect a high score for word reading in grade 1 to reduce the probability of scoring below the cut-off for reading comprehension in grade 3 (meaning that the odds ratio should be below 1). As the sample studied was part of an intervention study, the logistic regressions controlled for whether students had received an intervention or not. However, no special attention was paid in the analysis to linguistic background or family risk of dyslexia.

Results

Table 3 shows the results of a logistic regression analysis with ‘scoring at or below the 20th percentile’ (i.e., being among the 20% lowest scorers) for reading comprehension in grade 3 as the dependent variable. Step 1 of the model shows that the sentence-reading sub-test in grade 1 predicts, on its own, 24.1% of the students who will manifest reading-comprehension difficulties at the end of grade 3. Step 2 then shows that the remaining sub-tests of the grade 1 test, taken together, predict a further 0.7% of the students who will have reading-comprehension difficulties two years later. Hence the total explanatory power of the model is 24.8%, and the sentence-reading sub-test clearly provides the lion’s share of that power.

Table 3. Logistic regression of the relationships between scores on the screening test in grade 1 and scores below the ‘concern cut-off’ for reading comprehension in grade 3 ($N = 996$)

		B	SE	OR	95% confidence interval
Step 1^a	Constant	0.48	0.32	1.62	
	Sentence reading	-0.31	0.04	0.73	[0.69–0.79]
Step 2^b	Constant	1.92	0.90	6.79	
	Sentence reading	-0.23	0.05	0.79	[0.72–0.88]
	Letter knowledge	-0.08	0.08	0.92	[0.79–1.08]
	Analysis	-0.02	0.06	0.98	[0.87–1.10]
	Synthesis	-0.02	0.05	0.98	[0.89–1.09]
	Word reading	-0.05	0.05	0.95	[0.86–1.04]
	Spelling	0.00	0.04	1.00	[0.92–1.09]

B = beta; SE = standard error; OR = odds ratio.

^aNagelkerke R square = 0.241. ^b Nagelkerke R square = 0.248.

Table 4 shows correlations between certain sub-tests of the screening tests for grades 1 and 3 (including two additional grade 3 sub-tests). It is clear that the sentence-reading sub-test for year 1 shows substantial co-variation with both reading comprehension, word reading and spelling in grade 3.

Table 4. Correlations between scores on sub-tests of the screening tests for grades 1 and 3 ($N = 996$)

		Letter 1	Analysis 1	Synthesis 1	Word r. 1	Spelling 1	Sent. r. 1	RC 3	W. ch. 3
Letter	1	–							
Analysis	1	.68**	–						
Synthesis	1	.33**	.49**	–					
Word r.	1	.56**	.63**	.45**	–				
Spelling	1	.61**	.71**	.52**	.68**	–			
Sent. r.	1	.59**	.63**	.42**	.72**	.68**	–		
RC	3	.44**	.46**	.34**	.48**	.46**	.49**	–	
W. ch.	3	.43**	.47**	.31**	.54**	.48**	.51**	.51**	–
Spelling	3	.45**	.49**	.39**	.50**	.57**	.51**	.50**	.62**

1 = grade 1; 3 = grade 3; r. = reading; Sent. = sentence; RC = reading comprehension; W. ch. = word reading. ** = statistically significant at the .05 level.

In brief, these results tell us that the scores on the sentence-reading sub-test administered in grade 1 are sufficient to predict reading-comprehension difficulties at the end of grade 3. In other words, the other five grade 1 sub-tests add only marginally to the precision of the prediction of such difficulties.

Discussion

Meeting the challenges with a new test

We have shown above how a short and simple sentence-reading test administered at the end of grade 1 has a good predictive value in relation to reading-comprehension difficulties at the end of grade 3. This finding places us in a good position to design a new test concept addressing key aspects of each of the challenges associated with the present test, as described above, namely (1) test length; (2) unintended consequences for classroom practices; (3) incorrect use of test scores; (4) varying rates of progression in initial reading instruction; and (5) a high threshold for using test scores.

To begin with, if only that sentence-reading test is used, the *length* of the test will be reduced to about one-fifth of the time. In addition, actual reading – as an interpretive skill – becomes central to the test. Further, if there is no longer any reporting on individual components of reading skill, this will reduce the likelihood of *unintended consequences for classroom practices* such as having students practise those skills (for example, phonological synthesis) in isolation from letter training in meaningful contexts. However, the new test concept must still provide clear guidance on how those components should be addressed in the reading of real words.

When it comes to *incorrect use of test scores*, it will be easier to understand the purpose of a simple test which measures reading comprehension alone and which has a well-documented predictive value, and this might address the present challenge consisting in teachers' frequent use of screening-test scores as a basis

for assessing the skill level of all their students. The challenge associated with *varying rates of progression* in the learning of letters can be addressed by administering the test when all letters have been introduced to the students. Making the test voluntary would give teachers greater flexibility to choose such a meaningful time for it. Here it is worth pointing out that the predictive value reported in this study is based on a sample of students who had had all letters introduced to them by Easter in grade 1. Finally, the challenge of a *high threshold for using test scores* must be addressed on several levels. Even so, introducing a simple test with a clear predictive value in relation to later reading difficulties will take us a considerable part of the way. The new test will also have greater utility value in that it offers clear longitudinal prediction which can form the basis for early intensive interventions. One possible objection is that making the test shorter will entail that the teachers obtain less information and knowledge about their students – but here it should be stressed that the new test concept must also include a *second part*, to be administered after the initial test to a smaller group of students. That second part must be designed to provide teachers with information about each student's reading behaviour which can be easily linked to the teachers' knowledge about good reading instruction.

Bringing the test into step with more recent theories of reading

According to Alexander and Fox (2004), research into initial reading instruction has clung to a view of reading as the mastery of sub-skills. They consider this idea to be closely associated with an understanding of reading which is heavily influenced by behaviourist ways of thinking, focusing on the external and observable. In that tradition, reading is studied as the acquisition of behaviours, and learning as the result of repeated, controlled stimuli from the surrounding world and the subsequent expected responses. When it comes to special-needs education, the view of reading as a set of sub-skills manifests itself in the use of materials intended to drill students' sub-skills, with a focus on automatization. Reading is understood as the competent performance of a series of related yet separate sub-skills which can be both trained and tested in isolation (Pearson & Stevens, 1994). Such ideas are widespread in Norway, and there is one theory in particular which has set the premises for the content of initial and special-needs reading instruction: the phonological theory of cognitive psychology. According to that theory, reading consists of a set of components which can be trained separately, and some of those components must be trained or mastered before any actual reading can take place. One unintended consequence of the design of the present test is that many students are made to work on isolated components of reading – for example by reading syllables – without receiving any explicit instructions about how they should use those skills when reading texts. It would be hard to refute the claim that present practices in initial and special-needs reading instruction are dominated by such ideas. By contrast, more recent theories of reading emphasise the movement into – and through – the text (see Kintsch,

1994; Alexander, 2005; Langer, 2015). In other words, reading is seen as a process and as an interpretive skill.

The gap between these two views on reading gives rise to a kind of disharmony which is particularly noticeable in connection with the screening tests but is not necessarily *caused* by those tests. That disharmony is linked by Pressley (2006) to what is referred to in the reading-research literature as the dichotomy between the *bottom-up* and *top-down* approaches (which involves some simplification) or between *phonics* and *whole language* (which more resembles caricature). Pressley advocates a ‘balanced view’ of reading based on the idea that both the parts and the wholes are important, but beyond stressing the importance of bringing the two schools of thought together, he contributes little to a theory for such a balanced view and provides only limited explanations for how that view could be implemented in practice. However, the above-mentioned disharmony between the two approaches – as reflected in the clash between the view on reading underpinning the screening tests and a newer view of reading otherwise practised in present-day teaching – shows that it is difficult to bridge the gap between those views or create a dynamic interaction between them. In other words, teachers lack both the tools and the theoretical foundation that they would need to reflect on how the combination of those views manifests itself in classroom practices. What is more, the disharmony is further reinforced by the fact that the older view remains largely prevalent in the tradition of special-needs education while the newer one now dominates the tradition of general education.

One possible solution to this problem could be to define reading as an *interpretive skill* (Tønnessen & Uppstad, 2015). Such a definition could in fact work for a variety of more recent views on reading. Here, ‘skill’ refers to a flexible combination of automaticity and awareness while ‘interpretation’ involves constantly seeing the whole and the parts in the light of each other, as in hermeneutics (Gadamer, 1960). The idea of a flexible combination reflects the fact that the amount of awareness required to find the meaning of a word or paragraph varies. When reading is defined as an interpretive skill in this way, it no longer makes sense to draw a clear line between decoding and comprehension. This is because everything in a text, from the individual letters to its deeper meaning, must be interpreted, although the relative proportion of automaticity and awareness will vary. Further, given that the ‘phonics’ tradition has tended primarily to emphasise decoding while the ‘whole language’ approach has focused mainly on comprehension, it is also clear that the conflict between them will no longer be relevant if reading is defined as an interpretive skill. Hence, the idea is not to reject the components – or parts – of reading. On the contrary, the parts are crucial for reading, but not in the same way as in the traditional view. Rather, the parts must constantly be exposed to the whole, because the interaction between the parts and the whole is what makes learning and interpretation happen (Tønnessen & Uppstad, 2015). ‘Reading’ and reading skill, when defined as an interpretive skill in this approach, differ from ‘reading comprehension’ in that the latter concept

implies a *final, fixed product* of interpretation, whereas, on the view of reading as an interpretive skill, any interpretation can itself be part of a subsequent, broader interpretation.

The need for further development

We have shown above how a simple sentence-reading test can predict reading difficulties in grade 3. However, this should be seen only as a promising idea for a new test design, not as a fully-blown design. The main focus of our further work must be on how to gather more information about those students who are identified as being at risk of developing reading difficulties. That information must be obtained in a way which ensures that the end user of the scores – the teacher – remains in touch with reading as such, with the interpretive skill. In today's tests, the actual scores are assigned excessive importance and teachers seem to find it difficult to go from interpreting those scores on group-administered tests to implementing effective interventions. In addition, there may also be an opposite causal link: a further reason why great importance is attached to the score is that teachers do not really understand it.

Our proposal is that the second part of the test should be administered by the teachers to their students one-on-one, and that maximum effort should be devoted to shifting the focus away from the score and towards the teacher's observation of the difficulties encountered by a student as he or she answers items or reads. In this way, we make it possible to link the parts and the whole (top-down/bottom-up) in a different manner from before – by asking, “What exactly is this student struggling with when he or she is moving back and forth between the parts and the whole?”

As Tveitnes (2018) points out with regard to school psychology services, the authority attributed to test scores is often excessive and those who are to make use of the scores often do not really understand them. Even so, the second part of the new test design could include some of the same sub-tests as the existing test design, considering that they are designed to provide detailed information about the weakest readers. Since the initial part of the test will now fulfil the identification function by itself, the second part can focus on specific aspects of reading to provide teachers with information which they can translate into effective interventions. Hence, the two parts of the new test will each perform a different function. The idea is that the new methodology and philosophy will give the sub-tests to be included in the second part of the test a clear purpose. It seems like a good idea to design the second part as an ‘adaptive’ test to be performed on tablet computers. Given that the test will be administered one-on-one, the additional equipment cost to schools will be limited, or none at all (a single tablet will be needed). An adaptive test adjusts to the individual student's skill level, ensuring that he or she is given tasks that are neither too easy nor too hard. This yields much more information about the students' skill level than a standardised test,

which is the same for all students, but on the other hand, considerable resources are required to develop an adaptive test.

To summarise, in the present article we stress the need to use a different methodology – and philosophy – for a test intended to identify children likely to fall behind in reading. In brief, the *methodology* we propose is to perform a screening test administered to all students which is so short and inconspicuous that teachers will not consider it meaningful to report the students' scores to their parents or others. Further, this methodology will yield additional information only about those students where more information is really needed. The teacher will obtain that information from sitting one-on-one with each student, watching him or her solve the tasks. This offers an opportunity to study reading as an interpretive skill. The new *philosophy* to be introduced involves a shift away from the naked numbers or scores yielded by the present group-administered test towards the teacher's qualitative understanding of the difficulties encountered by each student as he or she is solving the tasks. In line with the ideas of phenomenology, this opens up for a different, more comprehensive understanding. The underlying assumption is that such an understanding will make teachers better able to make interventions and to adjust their instruction. In brief, the idea is to remove as much as possible of the unfortunate aspects of the authoritative test while assigning a higher value to the teacher's judgement.

Conclusion

The present screening tests of reading were designed at a time when research was documenting that struggling readers performed much worse than good readers with regard to a variety of 'components' associated with the reading process. The idea was that any interventions should focus on training within the precise area of difficulty. From such a perspective, students' performance in relation to the individual components of reading clearly serves a purpose, namely as the basis for interventions. Today, while research has further confirmed the existence of such characteristic differences between struggling and good readers, it is no longer considered that interventions should focus exclusively on the components where difficulties manifest themselves (Scammacca et al., 2007; National Reading Panel, 2000). There has been a shift from a view of reading as the sum of a set of components to a view of reading as *more* than the sum of its components – to what might be referred to as holistic views of reading (see, e.g., Tønnessen & Uppstad, 2015; Alexander, 2005; Langer, 2015). From such a perspective, the basis for interventions becomes different in that the interpretive process is instead used as the starting point for any efforts undertaken to address the difficulties suggested by scores for various components. However, while those components no longer take pride of place in newer definitions of reading, they do remain important domains for the characterisation of students with reading difficulties.

The aim of the present study was to investigate the extent to which the sub-tests (relating to various such components) of the present screening test are necessary to meet the primary objective of that test: identifying students at risk of developing reading and writing difficulties. We did this by examining which of the sub-tests of the present test predict reading difficulties at the end of grade 3. The study shows that such prediction is possible using only a single one of the sub-tests of the present test. In line with the changes to the views of reading described above, a reduced focus on components will be a good fit with more recent definitions of reading. In addition, our finding has the potential to bring about progress at schools in that the new test design that we propose will direct teachers' attention towards reading skill as such and towards conditions for reading.

About the authors

Bente Rigmor Walgermo is Associate Professor in Special Needs Education at the Norwegian Reading Centre, University of Stavanger. She specialises in relationships among interest, motivational beliefs and reading skill, and early reading and writing assessment.

Institutional affiliation: Norwegian Reading Centre, University of Stavanger, Professor Olav Hanssens vei 10, 4021 Stavanger, Norway.

E-mail: bente.r.walgermo@uis.no

Per Henning Uppstad is Professor in Special Needs Education at the Norwegian Reading Centre, University of Stavanger, and Professor II at Volda University College. His research interests include technology enhanced learning, reading and writing, educational assessment and early literacy intervention.

Institutional affiliation: Norwegian Reading Centre, University of Stavanger, Professor Olav Hanssens vei 10, 4021 Stavanger, Norway.

E-mail: per.h.uppstad@uis.no

Kjersti Lundetræ is Professor in Special Needs Education and Director of the Norwegian Reading Centre. She is an early-intervention researcher with special research interests related to reading and writing difficulties, early reading instruction and assessment of reading.

Institutional affiliation: Norwegian Reading Centre, University of Stavanger, Professor Olav Hanssens vei 10, 4021 Stavanger, Norway.

E-mail: kjersti.lundetre@uis.no

Finn Egil Tønnessen, PhD, is an emeritus professor at the Norwegian Reading Centre, University of Stavanger. His research interests are language, reading and dyslexia; psycholinguistics; behavioural problems; and motor coordination difficulties.

Institutional affiliation: Norwegian Reading Centre, University of Stavanger, Professor Olav Hanssens vei 10, 4021 Stavanger, Norway.

E-mail: finn-egil.tonnessen@uis.no

Oddny Judith Solheim is Professor in Special Needs Education at the Norwegian Reading Centre, University of Stavanger. She specialises in the fields of engaging literacy instruction, reading and writing difficulties and assessment of reading.

Institutional affiliation: Norwegian Reading Centre, University of Stavanger, Professor Olav Hanssens vei 10, 4021 Stavanger, Norway.

E-mail: odny.j.solheim@uis.no

References

- Alexander, P. A. (2005). The path to competence: A lifespan developmental perspective on reading. *Journal of Literacy Research*, 37(4), 413–436.
https://doi.org/10.1207/s15548430jlr3704_1
- Alexander, P. A., & Fox, E. (2004). A historical perspective on reading research and practice. *Theoretical Models and Processes of Reading*, 5, 33–68.
- Asheim, S. (1981). *Ny giv for lesesvake skolebegynnere* [A new deal for beginning students who are poor readers]. Oslo: Universitetsforlaget.
- Barkved, E. (2012). *Kartlegging av leseferdighet og vurdering for læring* [Surveying of reading skill and evaluation for learning]. Master's thesis, University of Stavanger.
<https://uis.brage.unit.no/uis-xmlui/bitstream/handle/11250/185727/Barkved%2CEdle.pdf?sequence=1&isAllowed=y>
- Cunningham, A. E. (1990). Explicit versus implicit instruction in phonemic awareness. *Journal of Experimental Child Psychology*, 50(3), 429–444.
[https://doi.org/10.1016/0022-0965\(90\)90079-N](https://doi.org/10.1016/0022-0965(90)90079-N)
- Directorate for Education and Training [Utdanningsdirektoratet] (2018). *Rammeverk for kartleggingsprøver på 1.–4. trinn* [Framework for screening tests at grades 1–4].
<https://www.udir.no/eksamen-og-prover/prover/rammeverk-for-kartleggingsprover-pa-1.-4.-trinn/>
- Ehri, L. C., Nunes, S. R., Willows, D. M., Schuster, B. V., Yaghoub-Zadeh, Z., & Shanahan, T. (2001). Phonemic awareness instruction helps children learn to read: Evidence from the National Reading Panel's meta-analysis. *Reading Research Quarterly*, 36(3), 250–287. <https://doi.org/10.1598/RRQ.36.3.2>
- Engen, L. (1999). *Kartlegging av leseferdighet på småskoletrinnet og vurdering av faktorer som kan være av betydning for optimal leseutvikling: En beskrivelse av den faglige prosessen med å utvikle nasjonale kartleggingsprøver for småskoletrinnet, og en vurdering av forholdet mellom fonologiske delferdigheter, ordlesings- og tekstlesingsferdigheter blant elever i 1. og 2. klasse* [Surveying of reading skill in lower-primary school and evaluation of factors that might be of importance for an optimal reading development: A description of the professional process of developing national screening tests for lower-primary school and an evaluation of the relationship between phonological sub-skills, word-reading skills and text-reading skills among students in grades 1 and 2]. Unpublished doctoral thesis, University of Bergen.

- Gabrielsen, E., & Lundetræ, K. (2017). Indikerer de norske PIRLS-resultatene et behov for å justere retningslinjene for skolestartsalder? [Do the Norwegian PIRLS scores indicate a need to adjust the guidelines for school-starting age?]. I: E. Gabrielsen (Ed.), *Klar framgang. Leseferdighet på 4. og 5. trinn i et femtenårsperspektiv* [Clear success. Reading skill in grades 4 and 5 from a fifteen-year perspective] (pp. 204–221). Oslo: Universitetsforlaget. <https://doi.org/10.18261/9788215030258-2017-12>
- Gadamer, H. G. (1960). *Wahrheit und Methode: Grundzüge einer philosophischen Hermeneutik* [Truth and method. Outlines of a philosophical hermeneutics]. Tübingen: Gesammelte Werke J, 5.
- Gustafson, S., Samuelsson, S., & Rönnerberg, J. (2000). Why do some resist phonological intervention? A Swedish longitudinal study of poor readers in grade 4. *Scandinavian Journal of Educational Research*, 44(2), 145–162. <https://doi.org/10.1080/713696666>
- Høien, T. (2012). *Håndbok til LOGOS: teoribasert diagnostisering av lesevansker* [LOGOS handbook. Theory-based diagnosis of reading difficulties]. Bryne: Logometrica.
- Høien, T., & Lundberg, I. (1991). *Dysleksi. Fra teori til praksis* [Dyslexia. From theory to practice]. Oslo: Gyldendal Norsk Forlag A/S.
- Johnson, D., & Kress, G. (2003). Globalisation, literacy and society: Redesigning pedagogy and assessment. *Assessment in Education: Principles, Policy & Practice*, 10(1), 5–14. <https://doi.org/10.1080/09695940301697>
- Kintsch, W. (1994). Text comprehension, memory, and learning. *American Psychologist*, 49(4), 294–303. <https://doi.org/10.1037/0003-066X.49.4.294>
- Langer, J. A. (2015). *Envisioning knowledge: Building literacy in the academic disciplines*. Teachers College Press.
- Lundetræ, K., Solheim, O. J., Schwippert, K., & Uppstad, P. H. (2017). Protocol: ‘On Track’, a group-randomized controlled trial of an early reading intervention. *International Journal of Educational Research*, 86, 87–95. <https://doi.org/10.1016/j.ijer.2017.08.011>
- Ministry of Education and Research [Kunnskapsdepartementet] (2017). *Læreplan i norsk* [Curriculum for Norwegian] (NOR1-05). <https://www.udir.no/kl06/NOR1-05>
- National Reading Panel (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. Washington, DC: National Institute for Child Health and Human Development.
- Nielsen, J. C., Kreiner, S., Poulsen, A., & Sjøgaard, A. (1986). *Sætningslæseprøverne SL60 og SL40. SL-håndbog* [The SL60 and SL40 sentence-reading tests. SL handbook]. Copenhagen: Dansk Psykologisk Forlag.
- Pearson, P. D., & Stephens, D. (1994). Learning about Literacy: A 30-Year Journey. In R. B. Ruddel, M. R. Ruddel, & H. Singer (Eds.), *Theoretical Models and Processes of Reading* (4th ed., pp. 22–42). Newark, DE: International Reading Association.
- Pressley, M. (2006). *Reading instruction that works: The case for balanced teaching* (3rd ed.). New York, NY: Guilford Press.
- Rasmussen, A. (2013). *Begynneropplæring i lesing i norske skoler – en undersøkelse av bokstavinnlæring og metodevalg* [Initial reading instruction at Norwegian schools – an investigation into letter learning and method choices]. Master’s thesis, University of Stavanger. <https://uis.brage.unit.no/uis-xmlui/handle/11250/185828>
- Educational research [Utdanningsforskning] (2017). *Førsteklassingene lærer bokstavene raskere* [First-graders are learning the letters faster]. <https://utdanningsforskning.no/artikler/forsteklassingene-larer-bokstavene-raskere/>
- Scammacca, N., Vaughn, S., Roberts, G., Wanzek, J., & Torgesen, J. K. (2007). *Extensive Reading Interventions in Grades K–3: From Research to Practice*. Portsmouth, NH:

- RMC Research Corporation, Center on Instruction.
[http://buildingrti.utexas.org/sites/default/files/booklets/Extensive Reading Interventions K-3.pdf](http://buildingrti.utexas.org/sites/default/files/booklets/Extensive_Reading_Interventions_K-3.pdf)
- Schneider, W., Roth, E., & Ennemoser, M. (2000). Training phonological skills and letter knowledge in children at risk for dyslexia: A comparison of three kindergarten intervention programs. *Journal of Educational Psychology*, 92(2), 284–295.
<https://doi.org/10.1037/0022-0663.92.2.284>
- Shaywitz, B. A., Fletcher, J. M., Holahan, J. M., & Shaywitz, S. E. (1992). Discrepancy compared to low achievement definitions of reading disability: Results from the Connecticut Longitudinal Study. *Journal of Learning Disabilities*, 25(10), 639–648.
- Solheim, O. J., Frijters, J. C., Lundetræ, K., & Uppstad, P. H. (2018). Effectiveness of an early reading intervention in a semi-transparent orthography. A group randomised controlled trial. *Learning and Instruction*, 58, 65–79.
<https://doi.org/10.1016/j.learninstruc.2018.05.004>
- Suggate, S. P. (2010). Why what we teach depends on when: Grade and reading intervention modality moderate effect size. *Developmental Psychology*, 46(6), 1556–1579.
<https://doi.org/10.1037/a0020612>
- Tønnessen, F. E., & Solheim, R. G. (1998). *Kartlegging av leseferdighet og lesevaner på 2. klassetrinn* [Survey of reading skill and reading habits in grade 2]. Oslo: Kirke-, utdannings- og forskningsdepartementet.
<https://www.regjeringen.no/no/dokumenter/kartlegging-av-leseferdighet-og-lesevane/id105518/>
- Tønnessen, F. E., & Uppstad, P. H. (2015). *Can We Read Letters? Reflections on Fundamental Issues in Reading and Dyslexia Research*. Rotterdam: Sense Publishers.
- Tveitnes, M. S. (2018). *Sakkyndighet med mål og mening: En analyse av sakkyndighetens institusjonaliserte kjennetegn; et grunnlag for refleksjon og endring* [Expertise with a purpose and a meaning: An analysis of the institutionalised characteristics of expertise; a basis for reflection and change]. Doctoral thesis, University of Stavanger.
<http://hdl.handle.net/11250/2499868>