

MASTEROPPGAVE

Assessing Strategy Use on the National Reading Comprehension Test in English

Erika Strandberg

May 2023

Master fremmedspråk i skolen

Fakultet for lærerutdanninger og språk
Institutt for språk, litteratur og kultur



Abstract

In Sweden, standardized tests are conducted in year nine to support teachers' assessment of student abilities and to ensure national validity and reliability (Skolverket, 2016). The national test in English is a proficiency test consisting of three parts; speaking, writing and reception. The test aims to assess students' communicative competence and the ability to use strategies is integrated in all parts of the test (Erickson, 2018). Reception is tested in two parts; reading and listening. However, since the aim is to test strategy use as an integrated part of the communicative competence there are no specifications on how strategy use is to be assessed in the separate parts of the test thus leaving teachers to interpret how to analyze strategy use on the national reading comprehension test.

The use of reading comprehension strategies is difficult to research as there are several ways of defining strategies. However, they could be described as actions taken to increase comprehension (CEFR, 2001). Reading comprehension strategies can either be metacognitive which means that the learner can plan and evaluate the reading consciously or they can be cognitive, meaning that they can be used both consciously or unconsciously (Börjesson, 2012). Research on standardized testing and strategy use have shown that although students use various strategies when taking a test it might not be possible to identify or assess whether a student uses efficient strategies in a standardized reading comprehension test (Muijselaar et al, 2017). Furthermore, researchers have different theories on what standardized tests can measure and how the question affects what the test can actually test (Badger & Thomas, 1991).

Since the tests are meant to contribute to equitability in assessment it is crucial to understand how teachers perceive and use the national reading comprehension tests to assess strategy use. In order to examine teachers' perceptions of the national tests in English a survey was conducted in combination with two interviews. The survey was answered by 82 teachers of English, two of whom expressed an interest in participating in a digital interview.

The results indicate that teachers might lack sufficient knowledge of reading comprehension strategies. The respondents have different perceptions of to what extent the national reading comprehension test can test the use of reading comprehension strategies and as to what strategies can be tested in the reading comprehension test. The majority of the respondents

think that the test can test the use of strategies to some or great extent, and the majority identified the use of context to understand content and specific words as strategies that can be tested. However, previous research has shown that using context is one of the strategies used the least in standardized reading comprehension tests and that the use of previous knowledge is one of the most commonly used strategies (Cordon & Day, 1996). Interestingly, very few teachers in this study identified using previous knowledge as a strategy that the national reading comprehension test can test. The national test result is used to varied extent amongst respondents in this study when assessing student abilities and there are significant differences in teachers' perceptions of the assessment manual.

Even though strategy use is not part of the grading criteria as of autumn 2022, the results of the study seem to suggest that teachers' understanding of reading comprehension strategies need to increase. Furthermore, test constructors might benefit from revising assessment manuals and ways of evaluating the national tests in order to improve test validity and reliability.

Acknowledgements

As a teacher of English, I have assessed several national tests with wonderful colleagues which is what has inspired the topic of this thesis and I would like to thank all of you for great discussions and my professional development. Furthermore, I would like to thank all teachers who generously took the time to participate in the study conducted for the thesis. The survey responses and interview responses made this project possible, and I am extremely grateful.

Second, I would also like to thank my supervisor, Henrik Bøhn, for all your guidance. Your comments, advice and interest in my thesis has been very valuable.

Lastly, I would like to thank friends and family whose support has made the writing of the thesis possible. To my parents, for your support and always believing in me. To Jim, for all the useful input and perspectives, and to Oscar, for the unconditional love and support that made it possible to finish writing.

I am forever grateful.

Erika Strandberg

Trollhättan, May 2022.

Table of content

1. Introduction.....	5
1.1 Purpose and research questions.....	6
1.2 Standardised testing in Sweden.....	7
1.3 The national test in English.....	7
1.4 The national test in English 2022.....	9
2. Theoretical Framework.....	11
2.1 Language strategies.....	11
2.2 Reading comprehension strategies.....	13
2.3 Reading comprehension strategies in the Swedish syllabus.....	16
2.4 Validity in testing.....	17
2.5 Chapter summary.....	19
3. Literature review.....	20
3.1 Assessing strategy use.....	20
3.2 Strategy use on standardized reading comprehension tests.....	21
3.3 Teachers' perceptions and use of standardized tests.....	25
3.4 Chapter summary.....	27
4. Method.....	28
4.1 Research method.....	28
4.2 Samples.....	29
4.3 Data collection.....	30
4.4 Data analysis.....	33
4.5 Validity & Reliability.....	34
4.6 Ethical considerations.....	35
5. Results.....	37
5.1 Survey results.....	37
5.1.1 General findings.....	37
5.1.2 Impact of work experience.....	42
5.2 Interview results.....	46
5.2.1 Perceptions of strategy assessment on the national test.....	46
5.2.2 Use of test result and strategy assessment.....	47

6. Discussion.....	49
6.1 Assessing strategy use.....	49
6.2 Assessing strategy use on the national reading comprehension test.....	50
6.3 The use of test results.....	53
6.4 Chapter summary.....	54
7. Conclusion.....	55
7.1 Limitations.....	56
References.....	58
Appendices.....	64
Appendix 1 – Survey questions.....	65
Appendix 2 – Interview questions.....	68

1. Introduction

English is one of the core subjects in Swedish education and a passing grade in year nine is required to be eligible for upper secondary school. In order to obtain a passing grade students must develop communicative competence in English which is the ability to understand written and spoken English as well as being able to communicate in speech and writing. Furthermore, students must have the ability to use strategies when encountering obstacles in communication (Skolverket, 2011). Teachers conduct formative and summative assessment throughout the school year but at the end of year three, six and nine there are national standardized tests in several subjects, including English. The results of the tests in year nine are to be taken into special consideration when assessing students' abilities in a final grade.

On behalf of Skolverket (the Swedish national agency for education), test constructors at Gothenburg University construct the national test in English which is a proficiency test consisting of three parts; reception, speaking and writing. Reception is divided into two tests one on listening comprehension and one on reading comprehension. In all parts of the test the use of strategies is included in the assessment and in the reading comprehension test the students' ability to adapt the reading to the requirements of the test at hand is what is tested in terms of strategy use (Göteborgs Universitet n.d). The national reading comprehension test in English is a standardized reading comprehension test where students are meant to read different types of texts and answer various questions but the exact construction varies every year. Skolverket (n.d) provides a leaflet of information for educators for every test which states that the test does not aim to assess a students use of reading comprehension strategies explicitly but that the construction of the test includes strategy use thus providing teachers an opportunity to use the result of the test as part of the assessment of reading comprehension strategies. The leaflet does not however provide information on what reading comprehension strategies students could or would have to use when taking the test. As a result, teachers are given the opportunity to use the test result as part of the assessment of strategy use but are not required to or instructed on how that assessment should be done. Consequently, there is a risk of subjective assessment and differences in how national results are viewed and used by individual teachers.

Seeing as teachers in Sweden grade their own students the national tests are important to ensure a national standard and valid grading across the country. Even though teachers grade the national tests of their students, a standardized test with high reliability and validity can be an efficient assessment tool if teachers know what to assess and receive clear guidelines on how to do it. In 2017 a law was passed that established that national test results have to be taken into specific consideration when grading students (SFS 2017:1104). The significance put on the national test results requires that teachers use the test to assess the same abilities and to the same extent in order to ensure validity and reliability.

1.1 Purpose and research questions

Since there is a lack of research on teachers' perceptions and use of the national reading comprehension test in English, this study aims to contribute to the knowledge of what perceptions teachers have of strategy use on the reading comprehension test and to what extent they use the result to assess strategy use. The study aims to answer the following research questions:

1. In terms of strategy use, what do teachers think can be tested in the national reading comprehension test?
2. To what extent do teachers use the national reading comprehension test to assess students' use of reading comprehension strategies?
3. If teachers use the national reading comprehension test to assess strategy use, how do they use them?

1.2 Standardised testing in Sweden

Standardised testing in Sweden is regulated by the Swedish school law which states that all students in year three, six and nine as well as students in upper secondary school are required to take tests in particular subjects. In year nine the students take standardized tests in Maths, English, Swedish, Science studies as well as one social studies subject (SFS 2011:185).

National standardized tests serve two main purposes; to support teachers' assessment of students and to provide information on students' subject knowledge in Sweden on a national level. The aggregated test results from all Swedish students can be used by the government to evaluate whether students meet knowledge requirements and to revise or create new legislation. In Sweden, formative and summative assessment of student abilities are made by individual teachers responsible for the various subjects, which is why the national tests are supposed to contribute to a fair assessment on a national level, making sure that teachers' interpretation of the curriculum corresponds to the national standard. The validity and reliability of grading in Sweden has been a topic of discussion for a long time due to the fact that there has been a significant decrease of equivalence between national test results and student final grades as well as an increased difference between different schools which could be an indication of teachers taking other aspects into consideration when grading student performance (Skolverket, 2016).

1.3 The national test in English

The national test in English in year nine is a proficiency test and consists of three main parts; part A which is a speaking test, part B which tests reception and part C, the writing test. The three tests are taken at separate times where part A is taken at the end of the fall term whereas the B and C tests are taken at two specific dates in the spring set by Skolverket. Even though part B is considered as one test it is actually divided into two different parts where the first part, B1, tests reading comprehension and the second part, B2, tests listening comprehension (Skolverket, 2020). Seeing as this study only focuses on reading comprehension strategies B1 is the only test examined or referred to in this paper.

The national reading comprehension test usually consists of three or four different types of texts with various types of assignments and question items and every text is followed by 6-25 question items. The syllabus states that the student should encounter texts from different

genres and with different topics so the type of texts in the national tests are varied. The test includes at least one longer text where the student's ability to read and comprehend longer texts is tested. This means that the student has to be able to make inferences based on implicit information in the text. To show comprehension the student answers both multiple choice questions and open-ended questions. Furthermore, the test includes shorter texts where one text consists of gap items where the students are meant to fill in a gap with a correct word. Some tests have multiple cloze tests which means that there are alternatives to choose from while other tests include open cloze question items. Finally, the test includes an info seek where the students are meant to find specific information in the text by either answering multiple choice questions, combining shorter texts to statements or by answering open-ended questions (Olsson, Nilsson & Lindqvist, 2018).

After the test has been conducted the teachers at the individual schools are responsible for marking and grading the tests. Skolverket provides answer keys and scoring rubrics for each test. For the constructed responses there are always a few examples of acceptable answers as well as wrong answers, sometimes including a comment to what is necessary for the response to be considered correct. In addition to the answer key and scoring rubrics teachers are provided with information on the test and the construction as well as general guidelines on what to consider when assessing the national tests. One part of the information is what is called the test specification which contains a list of the knowledge requirements and the letters A,B and C have been written next to the knowledge requirements to indicate what part of the national test that tests that specific requirement (GU, n.d.). Next to the knowledge requirement on strategy use, all letters have been written and it states that while none of the tests focus specifically on testing strategy use all three parts offer an opportunity to assess strategy use but does not elaborate on how that assessment should be done.

Erickson (2018) states that the national test tests the usage of strategies, the ability to adapt language to purpose, recipient and situation as well as intercultural competence integrated in the receptive, productive and interactive parts of the tests. The aim is to assess general communicative competence which for the purpose of the test has been divided into parts which is why there is no explicit focus on strategy use on the test in reading comprehension.

The fact that the test consists of various types of texts and question items means that the different texts will require different strategies to solve the problem. Some texts are longer narrative texts and require in-depth reading and the ability to make connections and come to conclusions whereas other texts are shorter and more factual which means that the students will have to adapt the reading and thus use various strategies (Lindkvist, Arvidsson & Nilsson, 2018).

1.4 The national tests in English 2022

Seeing as the study conducted for this paper was done in the spring of 2022 around the time of the national tests taking place, respondents might have been particularly thinking of that test when participating in the study which is why a description of the test for the specific year will be provided. The national reading comprehension test of 2022 is covered by a law of confidentiality which means that the specific test and test instructions can not be discussed in detail. However, the test constructors write a report after each year's test which contains a description of the various parts of the test. The report describes how the reading comprehension test of 2022 consisted of three different texts, one of them being a text retrieved from a British website that provided information and description of places. The text was followed by both open-ended questions where students had to construct responses as well as an item where students were meant to combine statements and texts. The second text of the test was an American story with omitted words thus making it an open-cloze test item. The report states that the test item provides the opportunity for students to demonstrate the ability to understand based on information in the specific clause as well as the overall context. The final text was a longer reading comprehension text extracted from a British novel and was followed by both multiple-choice questions as well as open ended questions. In addition to requiring students to be able to read and understand details, the report also states that the questions on the final text require a deeper understanding and the ability to make conclusions based on what has been read which is referred to as reading between the lines and reading beyond the lines (Göteborgs Universitet, 2022). The report also states that strategic competence is integrated in the various parts but does not elaborate on what strategic competence, what specific parts or questions or how a lack of strategic competence could be identified.

As part of the development of the national tests teachers are encouraged to complete an evaluation survey after the national tests have been conducted each year which aims to investigate whether teachers think that the test worked well and served its purpose. After the national test in 2022 the survey was completed by 290 teachers of English and 94% of the respondents concur completely or to a great extent that the national test in English as a whole was good. In regards to the reading comprehension test, 84% of respondents thought that the test provided a reliable measurement of students' abilities to read and interpret the content of various texts in English and 15% considered the test to some extent provide a reliable measurement. The respondents were also asked to evaluate the three test items separately in relation to the course syllabus. Interestingly 99% of respondents considered the first and third test item to be appropriate or very appropriate in relation to the course syllabus whereas the second test item with the open-cloze was considered appropriate or very appropriate by 95% of respondents (Göteborgs universitet, 2022). The test constructors conclude that the national test in English as a whole worked well and was well-received in general seeing as the majority of respondents reported that the results confirmed their own assessment and the results from the year 2022 is equivalent to results of previous years which contributes to the evaluation and conclusion that the test has worked well and in the way intended (Göteborgs Universitet, 2022).

2. Theoretical framework

The study conducted for this paper is based on definitions of various terms such as strategies, reading comprehension strategies and validity which will be elaborated on in this chapter.

2.1 Language strategies

When learning and using languages speakers will encounter situations where the use of various strategies are required to either increase learning or to use the language in a certain context (Oxford & Amerstorfer, 2018). In second language acquisition research it has also been claimed that learners who are more successful in learning languages might do things differently to those who struggle with learning and using a new language which has been referred to as employing different strategies (Gass, 2020). However, the term strategy is complex and there are multiple ways of defining and categorizing strategies.

In research on second language acquisition the term learning strategies is often used to describe the process in which a learner is using various methods to increase learning or use of a foreign language. Learning strategies are often described as actions a learner takes to learn a specific aspect of a language, such as vocabulary or pronunciation. Seeing as learning is an internal process it is difficult to establish what is a learning strategy, however the aim of learning strategies are often to improve learning and become a more proficient user of the language (Cook, 2016). Learning strategies can be divided into metacognitive, cognitive or socio-affective strategies where metacognitive strategies refer to methods consciously used to plan and monitor the learning of a language whereas cognitive strategies refer to methods used intentionally or unintentionally to improve learning, for instance note-taking or using resources such as dictionaries. Socio-affective strategies however, refer to strategies that involve interaction with others such as asking a teacher for help (Cook, 2016).

When using a language in a specific context one might also use communication strategies which in contrast to learning strategies are not necessarily used to learn something but rather to handle a communicative situation. Communication strategies can either be used to make someone understand what is being said by paraphrasing for instance or strategies can be used to get around an issue such as speaking in the native language. Finally, communication

strategies can also be used to avoid a situation completely by for instance changing the subject or avoiding communication (Cook, 2016).

In addition to strategies for learning and strategies for using a language, research sometimes divides strategies based on skills, for instance strategies for reading, strategies for speaking and strategies for writing (Gass, 2020). Regardless of what classification is used to discuss strategies in language learning, the theory behind language learning strategies is that learners who employ various strategies frequently learn languages better than others, and, if it is possible to determine what strategies make language learners successful these strategies could be taught to learners who are not currently as successful. Research has indicated that strategy instruction can be effective and increase learning for students. However, teaching strategies is not universally beneficial and will not necessarily lead to improvement for all learners. In studies the teaching of strategies has been more successful when targeting reading and speaking than when it has targeted writing or listening and was more efficient when fewer strategies were taught (Gass, 2020). The use of strategies varies between tasks and context but can also vary between individuals which is why one strategy may seem efficient in one context or for one individual but not in a different context or for a different individual (Cook, 2016). To complicate the matter even more, strategy use is extremely difficult to research as it is nearly impossible to research mental behaviors or processes objectively. Research mainly relies on observations or self-reports from learners which means that only intentional and consciously used strategies will be reported on and observations will only be interpretations of a behavior. Furthermore, a strategy can only be proven to be objectively efficient if it can also be proven that poorer language users do not use the strategy and seeing as not all strategies are consciously used, proving that is extremely challenging (Gass, 2020).

The question of how strategies can improve language learning has also been discussed in the Common European Framework of Reference (CEFR). The CEFR is meant to serve as guidelines for language teaching and assessment in Europe, and therefore the Swedish syllabus for English as well as the knowledge requirements are to some extent based on the standard and its assessment criteria (Lundahl, 2019). According to the CEFR (2001), a strategy is an action used for a particular purpose when solving a task using the resources from previous learning experiences. Strategy use should therefore not just be considered a

way of compensating for language gaps but rather a language ability that will vary based on task and context and must therefore be contextually assessed.

2.2 Reading comprehension strategies

To define reading comprehension one must determine what it means to be able to read a text. According to Lundahl (2019), having the ability to read means that one is capable of dealing with different types of texts on several levels. Lundahl (2019) mentions five aspects of reading and what is required: understanding, using, evaluating, reflecting and engaging. The first requirement, understanding means the ability to construct meaning from the words in a text. The second aspect, using, involves the ability to adapt the reading to the purpose and type of text. Thirdly, evaluating is the ability to determine relevance of the content whereas the fourth aspect, reflecting, refers to the ability to relate the content to previous knowledge. Finally, the ability to read includes being able to feel about reading and the content as well as being able to participate in the reading - engaging.

Oakley (2011) uses the first aspect, the ability to construct meaning from a text as the definition of reading comprehension and states that reading comprehension strategies are required in order to understand texts. According to Block (1986) the essence of reading comprehension is in fact what the learner thinks and struggles with when reading and that the ability to use reading comprehension strategies is what makes a good reader. Learners who can monitor their comprehension, are aware of strategies and can use them efficiently, adjust reading to the type of text and its purpose, distinguish between what is important information and details and adjust when encountering inconsistencies are the most efficient readers. Evidently, there are several different ways of defining reading comprehension, however the use of reading comprehension strategies is a crucial part of what constitutes reading comprehension and is thus required in order to read a text and construct meaning from the content. CEFR (2001) describes reading comprehension strategies as methods used to increase understanding of a written text and provides examples such as the ability to identify context, anticipate what is to come, use general knowledge of topic and text type, the ability to identify significant information as well as testing an hypothesis and revising it when necessary.

Similarly to Gass (2020) categorization of learning strategies, Börjesson (2012) also uses the terms metacognitive, cognitive and socio-affective strategies to categorize reading comprehension strategies. Metacognitive reading comprehension strategies are methods where the learner needs to be aware of what is required in order to understand what is being read. For instance, the learner needs to be able to plan the reading, execute the task, evaluate the outcome and make adjustments based on the results. Cognitive strategies however require different abilities but could be used without the learner being aware of it. Using previous knowledge of the topic or type of text can support understanding and being able to make predictions and come to conclusions based on the context are also common cognitive reading comprehension strategies. In a social setting the learner could ask for assistance or receive feedback on language reception and production thus being examples of socio-affective strategies. However, socio-affective strategies are not very common to use in reading comprehension since reception and production require different abilities (Börjesson, 2012).

Reading comprehension strategies can be divided into other types of categories, for instance categories based on the competence required rather than what method is used. Block (1986) presents two alternative categories: *general comprehension strategies* and *local linguistic strategies*. *General comprehension strategies* refer to the methods focusing on the content of what is being read, for instance, the ability to recognize a structure and make predictions based on the information. Furthermore, the ability to interpret information provided in a text as well as information provided in a question is considered general competence and therefore general comprehension strategies. In contrast, *local linguistic strategies* include the use of methods related to the language itself such as solving vocabulary problems, understanding the meaning of specific words, being able to paraphrase and to identify relevant parts to re-read (Block, 1986).

As the notion of linguistic strategies alludes to, an important aspect of reading and interpreting texts is the ability to understand words and as a result strategies for understanding vocabulary can be used in reading comprehension assignments if one specific word is causing difficulties. Cook (2016) describes four common strategies for understanding vocabulary, one of them being guessing from information in the situation or the specific context. When reading certain types of texts such as novels, specific words might not be of

great significance and being able to make an approximate guess of the meaning of a word might be sufficient for the understanding of the context. A second strategy for understanding vocabulary is using resources such as dictionaries to increase understanding. The use of dictionaries are often considered rather controversial and especially in terms of what type of dictionary to use, depending on the purpose of the dictionary (Cook, 2016). However, dictionaries are often not allowed during standardised reading comprehension tests. If encountering an unknown word students could also try to make meaning of the word by making inferences from the form of the word if students have enough linguistic competence. By analyzing for instance prefixes or suffixes in a word students could be able to identify the type of word and based on the context come to a conclusion of the meaning of a word. A useful strategy for students who know other languages could be to link the word to cognates and make connections between different languages they know to make meaning of a word which could be especially efficient if students have knowledge of relationships between languages (Cook, 2016).

2.3 Reading comprehension strategies in the Swedish syllabus

As briefly mentioned above, the overall purpose expressed in the syllabus for English in year nine is to develop communicative competence, which involves the general ability to understand and use English in various contexts. However, a major part of communicative competence is being able to use strategies in order to make meaning or be able to make oneself understood also when comprehension problems and communication breakdowns occur. In addition to the intentions expressed in the syllabus, strategy use is mentioned explicitly in the core content as well as the knowledge requirements. The subject is meant to teach students strategies in order to comprehend details and context in written language by for instance adapting reading based on the form, content and purpose of the text. When assessing students the knowledge requirements states that they should be able to choose and use a strategy to improve understanding for a passing grade. For a higher grade students are expected to use strategies to some extent to improve comprehension (Skolverket, 2011).

In Sweden, all teachers have access to extra material published by Skolverket regarding the syllabus, including comments and elaborations of various terms that occur in the syllabus. The material contains a section on reading comprehension strategies with a description

stating that all means used to understand text could be considered reading comprehension strategies. One strategy mentioned is using a dictionary but there is emphasis on what is referred to as "guessing competence" meaning that students should learn how to guess what a text is about or make predictions on what will happen based on information available at the time or previous knowledge on a topic. Students should be taught to use contextual clues such as pictures, titles or familiar words in order to increase comprehension. Even though there are a few explicit examples, Skolverket (2017) emphasizes in the extra material that strategy use must be assessed in the specific context as one strategy might be very efficient in one context but not another and that there are other strategies than the ones explicitly mentioned in the material.

2.4 Validity in testing

Validity is a frequently used term in relation to both research and education and is a central concept of standardized testing (Bachman & Palmer, 2010). Previously, validity was often defined as a quality aspect of measurement and assessment in education and a test could be considered valid if it measured what was intended and if the conclusions drawn were proven accurate (Johansson, 2015).

Bachman (2004) however, claims that there are various aspects of validity and that the validity of measuring a specific quality is only one aspect of validity - construct validity. As a result, validity should not be used to describe the quality of a specific test or test score. Instead, validity refers to the quality of the interpretations made from and the assessment results (Kane, 2013). Consequently, there is a need to ensure that the process of interpreting scores is clear and that test administrators make similar inferences of what ability the test measures. Furthermore, there is widespread consensus that the use of the score and the consequences of the score use must be established in order to ensure validity (Bachman 2004; Kane, 2012; Messick, 1989). Messick (1989) also states that the process used by a test-taker when taking the test will sometimes have to be taken into consideration when interpreting a score and when they are, they need to be included in the construct definition. If processes are included in the construct definition it must be proved that the processes are used whereas if they are excluded the contrary must be proven, i.e. that the processes do not impact test performance.

Skolverket (2017) is naturally updated on these theoretical perspectives of test quality, stating that a test in itself cannot be discussed in terms of validity. Rather, what is important in the case of the Swedish national tests, according to Skolverket, is the interpretations and uses of the national tests. As for, aspects that are described as threats to the validity of the national test, these are listed in the framework for the national tests. For instance, if a test only partly tests a construct that is difficult to observe or assess there is a risk of underrepresentation of the construct which again could decrease validity. Moreover, if a test could be impacted by other factors than the ones included in the construct definition there is a risk of construct irrelevant variance which means a student could be assessed on aspects that are not relevant to measure the intended construct. Construct irrelevant variance could consequently impact the validity of a test (Skolverket, 2017).

In order to measure an ability a theoretical definition of the ability to be tested is required. This theoretical definition will then need to be operationalized (Fulcher, 2014) which simply means turning the purpose into concrete test items. One way of defining validity could thus be the relationship between theoretical definition and the operationalisation (Skolverket, 2009). For the national tests, the aims expressed in the syllabus for each subject serves as the theoretical definition which test constructors then interpret and use to create tests and assessment manuals. The tests and assessment manuals are then interpreted and used by teachers. When tests are open to interpretation there is a risk of lower validity, but the possible errors that could occur are considered to be the result of assessment processes used by teachers (Skolverket, 2009). Bachman (2004) however, claims that validity will depend partly on how clear the connections between score and construct are as the connection will impact how the tests are interpreted and used by teachers. Consequently, validity could depend on whether test constructors provide clear evidence of what qualities a test can measure and how.

For this thesis, the term validity will refer to the broader definition of validity which includes both construction and the various uses and interpretations made from the test results. Seeing as Skolverket (2017) states that one specific test can not be discussed in terms of validity without including the interpretations and uses of the scores, I find it highly problematic not to

include the interpretations and uses of the national tests in discussions of validity in this thesis.

2.5 Chapter summary

The chapter has introduced the theoretical framework this thesis is based on. The complex term strategy has been described from various perspectives and possible functions and categorizations have been explored. Reading comprehension has been discussed and reading comprehension strategies have been explained and categorized as metacognitive, cognitive and socio-affective, but a contrasting view of reading comprehension strategies in terms of what competence is required has also been presented. In addition, the CEFR views of reading comprehension strategies which the Swedish syllabus for English is based on was presented as well as what is expected of teachers according to syllabus and commentary material available. Finally, the term validity in terms of testing was elaborated on and various aspects that could impact validity were described.

In the following chapter, previous research and theories on assessment of strategy use will be explored as well as different perspectives of what standardized tests can be used to assess. Finally, research on how teachers perceive national tests specifically and what impact standardized tests may have on strategy instruction and assessment will be discussed.

3. Literature review

3.1 Assessing Strategy Use

Assessment can be defined as a process in which several kinds of measurements are collected to understand an individual's learning and can be either formative or summative (Farrell, 2008). Whereas formative assessment refers to an on-going process that aims to provide information that can be used to increase learning, summative assessment is usually conducted at the end of a course to determine what has been learned (Farrell, 2008). Strategy use is often assessed in order to support students' learning and to increase self-regulation which requires the use of methods where the result provides information on what strategies students are using and how they are using them (Gunning & Turner, 2018).

Almasi & Fullerton (2012) argue that when assessing the use of reading comprehension strategies the goal is to assess students' processing skills, i.e how proficiently students use strategies when processing and reading a text. Seeing as strategy use is an internal process that cannot automatically be observed, Amasi & Fullerton (2012) suggest using a method where the five characteristics of being a proficient strategy user are revealed. One of the characteristics of being a proficient strategy user is having knowledge of strategy use, both declarative, procedural as well as conditional knowledge. Second, a student must have the motivation to use strategies and thirdly, metacognitive competence to plan and monitor understanding. The fourth characteristic of a proficient strategy user is the ability to analyze the task at hand. Lastly, a proficient strategy user also has access to a variety of strategies to use if necessary. However, since most assessment tools available rely on measuring a student's comprehension of a text after reading, the processes used during reading are difficult to observe in the most commonly used formats (Magliano & Millis, 2011). Quantitative methods are often limited in terms of what can be tested and strategy use should be assessed in several ways where students are required to reflect on strategy use after reading for instance by writing strategy logs or answering questionnaires (Gunning & Turner, 2018).

Due to the fact that strategy use is of very personal nature (Gunning & Turner, 2018), teachers should systematically have students describe and report on strategy use and provide multiple opportunities for students to demonstrate strategy use in various ways. If teachers

rely too heavily on a few assessment methods, it might not be sufficient to obtain a comprehensive understanding of students' strategy use. However, teachers often ignore strategy assessment if grades are not assigned and/or if teachers lack access to adequate assessment resources (Gunning & Turner, 2018). In addition, teachers tend to become confused regarding assessment of strategy use and the purpose and goals of assessment on standardized tests which may lead to a formative use of summative tests designed for external accountability (Almasi & Fullerton, 2012). Oakley (2011) states that teachers often express uncertainty regarding how assessment of strategy use should be conducted and instead tend to rely on the results of comprehension tests. Lundahl (2019) expresses similar concern in the Swedish context and states that many teachers of English in Sweden mainly use the score of national test results to assess the use of reading comprehension strategies. Oxford & Amerstorfer (2018) concludes that since the assessment of strategy use is highly complex, more research is needed and especially in terms of tools and methods for assessment. Furthermore, the teaching and assessment of language learning strategies are not currently featured in the majority of language programs (Chamot, 2018) and consequently a lot of teachers do not receive sufficient training which is why Oxford & Amerstorfer (2018) state that language learning strategies should be included in teacher education programs.

When assessing strategy use it is important to be aware of the fact that the way reading comprehension or the use of reading comprehension strategies is assessed can affect an individual's performance (Spencer, Gilmour, Miller, Emerson, Saha & Cutting, 2019). Consequently, some knowledge on what standardized tests can assess and how various formats affect strategy use is needed.

3.2 Strategy Use on Standardized Reading Comprehension Tests

The relationship between reading comprehension and reading comprehension strategies has been researched in several ways and at several times. Studies show that knowledge of reading comprehension strategies have a direct impact on learners' reading comprehension in various tasks (Almasi & Fullerton, 2012). However, the relationship between reading comprehension and reading comprehension strategies is not evident in standardized reading comprehension tests (Muijselaar, Swart, Steenbeek-Planting, Droop, Verhoeven & Jong, 2017). Badger & Thomas (1991) state that a standardized test should not be used to assess the use of reading

comprehension strategies if the test construction only provides a product such as multiple choice answers, due to the fact that strategy use is a process and a product cannot demonstrate a process. Spencer et al. (2019) claim that although there are several formats that can be used to assess reading comprehension, the format can affect an individual student's performance and that scores should be interpreted in the context where text complexity, format and type of reading skill assessed must be considered. Even though there are various views on the role of the standardized test as a tool for assessing strategy use, the question format is generally considered as particularly significant and there are multiple perceptions of the various question formats available (Spencer et. al, 2019). As described previously, the national test constructors use a variety of question formats for the national reading comprehension test in English, for instance multiple-choice questions, open-ended questions and cloze test items and an overview of research regarding these various formats is necessary.

One frequently used item is the open-ended question. An open-ended question requires the student to produce a written answer without having any alternatives to choose from (Spencer et. al, 2019). Commonly, the correct answer could be produced in different ways and there might also be more than one acceptable answer. Open-ended questions make it possible to demonstrate reasoning and complex thinking thus providing an insight into the test-taker's interpretation of the text. The item could thus be appropriate to assess a process such as usage of reading comprehension strategies. It is however, important to consider that assessing answers to open-ended questions is much more difficult than assessing multiple choice questions in objective standardized tests as there are several possible ways of answering and conclusions can vary from the intentions of the teacher or test constructor (Badger & Thomas, 1991).

Multiple choice questions are in contrast to open-ended questions very simple to assess as there is one correct answer out of a limited number of options and are because of this very common and popular in standardized reading comprehension tests (Ajidhe & Mozzafarzadeh, 2012). Multiple choice questions as a question item has been discussed intensely lately due to the fact that it is very difficult to determine what a multiple choice question can in fact test. Lim (2019) states that although MC questions are valid in assignments where the purpose is to find explicit information and basic reading comprehension, it is not possible to assess the

use of reading comprehension strategies with the use of such questions. In fact, students can answer multiple choice questions without reading the text properly and instead use test-wiseness strategies such as skimming a text and matching words (Cohen, 2018; Sanders & Garwood, 2022). Furthermore, MC items do not involve any higher cognitive processes that can indicate deep comprehension (Lim 2019; Magliano & Millis, 2011). In contrast, Hoepfl (1994) maintains that MC items do demand cognitive processes. When given several alternatives to choose from, the students are required to process all options and come to a conclusion based on the information the students have as well as their previous knowledge on the topic. However, it should be considered that Hoepfl (1994) does not refer to MC in reading comprehension tests or in terms of reading comprehension strategies but as a question item in general.

Cloze tests however, consists of items that are specifically used for reading comprehension. In a cloze test item, the student is required to fill in a gap where a word or phrase has been omitted from a text, employing either alternatives to choose from (multiple cloze) or no alternatives (open cloze) (Klapwijk, 2013). In order to fill the gap the student will have to use the information in the text, either the context in general or the specific words before or after the gap, which Klapwijk (2013) states is an example of measuring reading comprehension strategies. Ajidhe et. al (2012) states that even though a cloze test might not test reading there is a strong correlation between student performances on cloze test and other types of reading comprehension tests. Moreover, language knowledge is necessary in order to understand the context well enough to produce the correct missing word, especially in an open cloze where the student cannot guess the correct alternative. Farrell (2008) however, claims that the cloze item might not be valid to test reading because a student might be a proficient reader but lack test-taking strategies that are required to perform well on a cloze test.

As seen, several researchers consider the standardized test as more or less appropriate to assess strategy use depending on what question type is used. Wang (2006) however, claims that it is not the question type that determines strategy use but rather the relationship between the question and the answer and the cognitive demands the question requires, i.e strategy use is not determined by the format used to answer but rather what is asked to find. Three categories of questions can be distinguished where the first category is questions where the

answer can be found explicitly in the text. Second, there are questions where the answers can be found in the text but in a much more implicit manner and finally there are questions where the answer is implied scriptically but not expressed explicitly in the text. Questions from the second category require an ability to make inferences and a deeper understanding of the language whereas questions from the third category demand analytical thinking as well (Wang, 2006). As the cognitive demands increase the students' performances decrease indicating that the more complex questions where strategies are necessary might be more difficult which is why Spencer et. al (2019) state that a score needs to be interpreted in the context. In order to assess the use of reading comprehension strategies the teacher would have to examine the relationship between the questions and answers on the specific test for each student rather than looking at an overall score (Wang, 2006). Products on a standardized test are not necessarily related to processes used that lead to the specific answers which is why the assessment of comprehension strategies cannot provide comprehensive information on reading comprehension strategies. Students could use completely different processes but still achieve the same score and assessment, even though one student might be focusing on test-taking strategies rather than reading comprehension (Ardoin, Binder, Zawoyiski, Nimocks & Foster, 2019).

Even though it might be difficult to prove what strategies students are using for particular questions on a standardized test, it has been proven that knowledge of strategy use can improve performance on various comprehension tests (Almasi & Fullerton, 2012). In a study where test-takers were asked either to think aloud while taking a standardized test or to describe their process afterwards a majority reported using multiple strategies such as using pre-existing knowledge and rereading when necessary (Cordón & Day, 1996). In an attempt to investigate strategy use further a study was conducted where one group of students were asked to read a text and identify the main idea and a second group of students were asked to read the text and answer standardized questions on the content while thinking out loud. The results showed that both groups used strategies and the test-taking group to an even greater extent than the other group. Five different types of strategies were identified by the researchers; rereading, paraphrasing, using previous knowledge, summarising and using context however, the test-takers used rereading, paraphrasing and previous knowledge to much greater extent than summarising and using context. Even though strategy use was

evident there was no clear relationship between strategy use and test performance indicating that although strategies can be used when taking standardized reading comprehension test it is not possible to determine how effectively a test-taker is using reading comprehension strategies to complete a task (Cordón & Day, 1996). Almasi & Fullerton (2012) conclude that even though knowledge of strategy use and explicit instruction lead to improved results on standardized as well as non-standardized reading comprehension tests the format of formal standardized assessments cannot measure processes used when reading.

3.3 Teachers' perceptions and use of standardized tests

In the development of standardized tests there are several parties involved with various interests which could impact how teachers perceive the purposes and functions of the standardized test (Dewitz & Graves, 2021). As the importance of standardized tests has been emphasized in many countries, including Sweden, the interest in investigating what effects standardized testing has on teaching has increased. In the U.S the emphasis of testing has led to changes in the curriculum where strategies such as prior knowledge and metacognition is rarely assessed and instead teachers focus on what strategies students would need for specific test items (Dewitz & Graves, 2021). In Canada, strategy assessment was previously done through teaching and then testing however, that was considered contradictory to the nature of strategy use and now the assessment of strategies should rather be used to increase learning. Unfortunately, when strategy use was no longer part of grading, teachers often ignored assessing strategy use completely (Gunning & Turner, 2018).

In an interview study conducted with teachers and students that aimed to research what effect the national tests have on Swedish education in terms of instruction and assessment, teachers stated that the national tests have several benefits (Arensmeier, Bonnevier, Borgsten, Lennqvist-Lidén, Lundahl, Nilsson, Sundberg, Sundhäll, Yassin & Wetterstrand, 2014). For instance, the tests provide confirmation on teachers' own assessments of student abilities. Furthermore, teachers see the tests as an explicit example of the content in the syllabus which is valuable input in terms of planning the content of the subject the rest of the year. The fact that some teachers also assess the tests together with other teachers is also considered beneficial for their professional development and some teachers express that the tests lead to a more equal assessment throughout the country (Arensmeier et. al, 2014).

Even though the national tests in English are generally well-accepted and appreciated amongst teachers, there are aspects of the test that teachers express concern about. One of the major issues expressed is that the tests take a lot of time to plan and prepare for as well as execute and assess afterwards which has a negative impact on the ordinary lessons around the time of the national tests. Teachers of English express that the results of the national tests are used as one part of teachers' general assessment of students and it is up to the individual teacher to decide to what extent the tests should be considered, although the majority of respondents state that since the tests take place in the late spring there is not enough time for any formative use of the results. The receptive part of the national tests is the one where most teachers have objections and the objections are mainly related to the reading comprehension test. Several teachers find the multiple choice-questions on the reading test problematic, however, for different reasons. While some teachers think that the question type makes it easy for students to guess their way to a passing grade, others instead claim that the multiple choice questions are phrased in a way that "fools" students or "sets them up" to misinterpret the alternatives (Arensmeier et. Al, 2014). Even though teachers of English generally claim that the national reading tests are rather straightforward to mark, there are quite significant differences in how assessment is done between schools. While some teachers assess their own students by themselves other schools let teachers assess tests together and some do not let teachers assess their own students. Furthermore, there seem to be different interpretations of the guidelines regarding adjustments for students with special needs which could lead to different opportunities for students at different schools. The fact that the tests can be conducted and assessed in such different ways can have a severe impact on the national tests' validity according to the authors of the report (Arensmeier et al., 2014).

To evaluate what impact the national tests have on Swedish education and how teachers use the national tests a survey was conducted in 2014. Teachers of all subjects where national tests are conducted participated in the survey and 78% of teachers of year nine agree completely or to great extent with the statement that the national tests increase equability in assessment. However, only 39% of teachers in year nine think that the national tests ensure equability in assessment. Teachers of Maths, Swedish and English consider the national tests valid in relation to the knowledge requirements to a greater extent than teachers of other subjects, even though the tests do not cover all knowledge requirements in the syllabus

(Hirsch, 2015). Even though the survey was conducted in 2014 before the law regulating the importance of the national tests when grading students it is evident that the tests were generally considered rather significant when grading student abilities since 19% of teachers in year nine state that it would be difficult to assess student abilities without the national tests. Furthermore, 39% of year nine teachers agree completely or to great extent to the statement saying that the national tests play a heavy part when grading students (Hirsch, 2015). However, it is important to emphasize that the survey was not limited to teachers of English and the results have not been presented for teachers of each subject.

3.4 Chapter Summary

This chapter has explored previous research on the assessment of strategy use in reading comprehension and established that there are different opinions on whether strategies should be assessed, for what purposes and how to conduct assessment on strategy use. Strategy use on standardized tests has been discussed and while some claim that standardized tests should not be used to assess strategy use, others argue that they could depending in question format or question type. However, the fact that strategy use is difficult to assess as it is part of an internal process is evident. Interestingly, the research investigated in this chapter has concluded that test-takers use strategies to a greater extent than non test-takers when reading and that strategic readers often perform better at reading comprehension tests. Even though a relationship could be determined to exist, it might not be possible to determine how effectively a test-taker is using strategies to complete a task. Finally, teachers' perceptions of strategy assessment were mentioned and the fact that teachers often find it difficult to assess strategy use which often lead to teachers either relying on standardized test scores or ignore assessing strategy use.

4. Method

4.1 Research method

In order to answer the research questions for this study, focusing on the perceptions of the extent to which the national test in English assess reading comprehension strategies and to what extent and how teachers use the result to assess strategy use, a combination of methods was used. While surveys are beneficial when wanting information from a larger number of respondents that is simple to analyze, they provide limited information on the contextual issues of the phenomenon being investigated. Furthermore, in-depth reflections on the topic can be quite difficult to collect and analyse through survey responses (McKay, 2006) which is why the study also included interviews with respondents from the survey.

Mixed methods have become more common in several fields and there are several ways in which a mixed method approach study can be conducted (Creswell, 2014). The sequential mixed method is often used when there is a theory or concept to test on a larger population using a quantitative method which is then followed by more detailed exploration through a qualitative method. When investigating something fairly new or where there is very little research present, the sequential explanatory method is often used as the weight is on the quantitative data but the result is used to elaborate the qualitative part of the study (Creswell, 2014).

Seeing as the only research of teachers' perceptions of assessment of strategy use on the national reading comprehension test in English has been conducted in the form of evaluations made by Skolverket a mixed method was used to obtain information from a larger population as well as retrieving in-depth reflections that can help understanding the survey results.

Whereas the survey aimed to answer questions on teachers' perceptions on what the national reading comprehension test can test and to what extent teachers use the national test to assess students' use of strategies the interviews were conducted to obtain more elaborate answers on teachers' opinions and usage of the national reading comprehension test. Furthermore, the answers from the interview could be compared to results from the surveys in order to make meta-inferences, conclusions or explanations that can be made from combining methods (Tashakkori & Teddlie, 2008), and perhaps provide possible explanations as to why or why not teachers use the results of the national test in the assessment of students' use of reading

comprehension strategies. The questions asked in the interviews could have been asked even if the survey had not been a part of the study, however, it would not be possible to make any assumptions of teachers in general based on a few teachers' perceptions. The benefit of using a mixed method is the fact that more can be understood and said about the question being researched than what could be understood by only using a qualitative or quantitative method (Tashakkori & Teddlie, 2008).

Even though surveys can include open-ended questions where respondents can provide information in comments that could result in some in-depth reflections, the data analysis of such surveys can be rather complex and it could be difficult to make comparisons between respondents. Furthermore, extensive surveys can be time consuming for respondents and can instead result in respondents not finishing if the questions are too demanding (Cohen, 2008), in order to obtain a larger number of respondents I decided to use a mixed method to keep the survey brief. The survey used for this study only consisted of five questions where no free responses were needed and respondents who were interested could participate in an interview on the topic.

4.2 Samples

For this study, a sample of convenience was used which means that the respondents are people from the target population that the researcher has access to (McKay, 2006). A description of the study as well as the link to the digital survey questionnaire were posted in two different Facebook groups for teachers of English. In the description of the study there was also a request for teachers as respondents in an interview and how to express interest in participating. The post was published in the week before the national tests in English were to take place in the spring of 2022 and the survey was available to answer for approximately three weeks. The questionnaire was answered by 82 respondents and interviews with two teachers were conducted. Due to geographical distance the interviews were conducted digitally which Creswell (2014) states is beneficial when respondents cannot be observed directly. The audio of the interviews was recorded to enable transcription after the interviews.

4.3 Data collection

In order to obtain questionnaire data which could answer my research questions in a valid and reliable manner, several issues were considered. Firstly, the items in the questionnaire had to operationalize my research questions. Cohen (2008) states that operationalizing is done by turning the purpose of the study into concrete research questions that can be transformed to be used in a survey. As previously mentioned, it was also of importance that the survey was simple to complete and that it would not require too much of the respondents time so I decided to focus on the first two research questions in the survey and investigate the third research question through interviews. Seeing as the first research question relates to how teachers perceive assessment of reading comprehension strategies on the national test, the first question was going to be to what extent they believe that the national reading comprehension test can test students' use of strategies. Cohen (2008) states that when operationalizing the research questions it is important to think about what measures or empirical indicators there are that could be relevant to what is studied. It occurred to me when creating the survey that teachers might think that it can test strategy use to a similar extent but that does not necessarily mean teachers concur on what strategies that the national reading comprehension test can test which is why a question on what strategies the respondents believe to be tested was included. In order to obtain a sufficient number of responses the question was made into a close-ended question with alternatives to select where the alternatives consisted of reading comprehension strategies that were frequently mentioned in previous research, particularly research or definitions related to Swedish education. A list of 13 strategies were included as alternatives to choose from in addition to the alternatives "none of the above" and "other" where respondents could add alternatives or comments. Two respondents used the "other" alternative and constructed their own responses. Finally, a question on the second research question related to the use of the national test result was created and since I wanted to be able to see what possible differences could be caused by, I also included a question on teachers' perceptions of the assessment manual provided by Skolverket.

As a result the survey consisted of five questions in total, one of them being a factual information question about the respondents number of years as a teacher of English. Three of the questions regarded teachers' attitudes and opinions on reading comprehension strategies

on the national test and one question was related to behavioral information in order to understand to what extent teachers use the national test when assessing students' abilities in terms of using reading comprehension strategies. As stated, surveys should generally be simple to complete and not require too much time especially if the aim is to obtain a large number of responses and as a result the survey consisted only of close-ended questions. The majority of questions were Likert-scale questions where respondents rate an item on a scale. An odd number of alternatives on Likert-scale questions could result in respondents choosing the middle alternative in order to avoid taking a stand (McKay, 2006). In order to avoid that pattern an even number of four options was used for all questions with the addition of a fifth alternative "I don't know" to provide respondents with an opportunity to not take a stand. Even though all respondents are teachers of English thus being proficient users of English the survey was written in Swedish. Seeing as all information provided by Skolverket on the national tests is written in Swedish, it is the language that the majority of teachers of English use when receiving information or communicating about the national tests in English which is why the questions for this study were provided in Swedish. However, it is worth noting that questions and answers have been translated from Swedish to English for the writing of this paper but the questions as asked can be found in the appendix.

As the interviews aimed to provide elaborations on the survey questions which were asked in Swedish, the interviews were also conducted in Swedish. Seeing as the interviews were the second part of a sequential explanatory strategy the questions used for the interviews had to be based on the questions asked in the survey. As a result, the majority of questions were phrased as open-ended versions of the questions from the survey. Even though interviews in a sequential explanatory mixed method are used to describe and explain findings from a quantitative method (Creswell, 2014), I wanted to use a standardized open-ended interview structure to ensure that all respondents were asked the same questions which McKay (2006) states will make the result easier to compile and analyze. Furthermore, the interview respondents expressed interest in participating in the interviews and were not chosen by me. Consequently, I had no previous knowledge of what the respondents had answered in the survey and since the survey result had indicated there were very differing opinions it increased the need to avoid bias in the interview questions. Sequential explanatory strategies are often very useful when receiving unexpected results (Creswell, 2014), however, it was of

great importance to not make assumptions of what the respondents had answered in the survey and as a result the initial questions were open-ended versions of the survey questions.

When some knowledge and understanding of the respondents survey responses had been obtained, the interviews could focus more on the third research question as it was not covered to great extent in the survey. To ensure that the respondents were asked the same questions initially, standardised open-ended questions were created regarding teachers' use of the national test results and follow-up questions were planned for in case the questions would not result in interesting information. The follow-up questions were placed in parentheses in the interview guide (see appendix 2) and regarded teachers' use of the national test results more explicitly as the research question on how the respondents use the national test result would be more difficult to obtain answers on through a survey. Finally, seeing as there were a few respondents who stated in the survey that they do not use the national test to assess students' use of reading comprehension strategies, a question on their assessment of reading comprehension strategies in general was included. Even though teachers' assessment of reading comprehension strategies in general is not part of the study for this thesis, the question could however provide knowledge on as to why or why not teachers use the national test results to assess students' use of reading comprehension strategies. As previously mentioned, the interviews were recorded and transcribed shortly after which McKay (2006) states is important in order to analyze the results. According to McKay (2006) note-taking during interviews can also be useful as it can assist the analysis of the transcribed material as well as provide some support in case of technical issues. Notes were taken during the interviews and key words were written for each question. Seeing as transcribing interviews is quite a lot of work and it might not be possible to do right after the interview, McKay (2006) suggests writing a summary after the interview which was done for both interviews conducted for this study. The summary consisted of interesting perspectives or examples of elaborations of the survey questions that seemed interesting.

4.4 Data analysis

Seeing as the study was based on a sequential explanatory mixed method strategy, the quantitative part of the study was conducted before the qualitative interviews since interviews in this type of research aims to explain and describe results from the quantitative part (Creswell, 2014). According to Creswell (2014) it is important to consider the weight of the study when analyzing results and seeing as this study has researched a phenomenon that has not been researched to great extent previously, the weight was put on the quantitative part of the study to obtain knowledge from a larger population.

After the survey was closed, the process of analyzing the data was initiated. The survey was conducted using Google Forms which is a programme in which responses are statistically presented based on frequency of responses automatically. This function provided a first understanding of teachers' perceptions in general and revealed patterns. When analyzing the frequency and percentage of each question it was clear that the responses were extremely varied which led to an analysis based on the number of years as a teacher of English to determine whether years of experience had a general impact on responses. Lastly, answers to different questions were compared to each other to investigate whether there were any patterns in responses between the various questions.

The analysis of the interviews began with reading the summaries from the interviews as well as the transcriptions. The data was then organized through a cross-case analysis where the responses were organized based on the topics in the interviews which McKay (2006) states is particularly useful when wanting to highlight certain aspects of the research topic. For this study it was particularly relevant to highlight what were elaborations and explanations of the survey questions and in particular regarding the third research question focusing on how the national tests are used if the respondents stated that they were indeed used. Mixed methods are particularly useful to follow-up on unexpected results (Creswell, 2014) which is why the intention was to do a content analysis where key topics are identified (McKay, 2006) and in particular content that can provide various explanations of the survey result to get a better understanding of the phenomenon. As a result both similarities and differences between the responses in the two interviews were highlighted.

4.5 Validity & Reliability

When conducting research, the discussion of validity and reliability is of great significance although the terms are used differently in different research methods. In quantitative research validity often refers to which extent the instrument used measures what is intended while qualitative methods often discuss validity in terms of credibility and to what extent bias could have affected the results (McKay, 2006). In mixed methods research, validity refers to the quality of the inferences made from using a combination of methods (Tashakkori & Teddlie, 2008). Seeing as this study aims to provide a more thorough understanding of teachers' perceptions of the assessment of strategy use on the national reading comprehension test, the use of mixed methods must provide a more thorough understanding than one method could do on its own (Tashakkori & Teddlie, 2008). According to Creswell (2014) the weight of each part of the mixed method will also impact the validity of the conclusions or inferences made.

The weight of this study has been put more on the quantitative part of this study than the qualitative part and the interviews should be considered as possible perspectives of respondents. Considering the fact that there were only two participants in the interview, the study does not strive to cover all possible perspectives of survey respondents or to identify the perspectives of the majority of teachers. However, the interviewed respondents point to interesting aspects that can provide some insight to the significant variations in responses from the survey thus providing more information than either method could do on its own. Furthermore, the fact that only two respondents expressed interest to participate in an interview made it difficult to obtain in-depth reflections that represented the significant differences in the survey responses as the two interview respondents had a lot of similarities. If the study had included more interview respondents who had provided different answers to the survey questions, the validity of this study could have increased as it likely would have resulted in even more information. However, it is important to note that qualitative studies generally do not aim to generalize findings but rather provide descriptions based on a certain context (Creswell, 2014).

One important aspect of validity in quantitative research is the one of internal validity which refers to the control of the research design and whether other factors could influence the result (McKay, 2006). Seeing as a sample of convenience was used to obtain respondents for

the study there is always some uncertainty regarding the backgrounds of the respondents as they have not been controlled. Although rather unlikely, the survey could have been answered by people who are not part of the population for the study, i.e. are not teachers of English. However, seeing as the request was posted in two Facebook-groups specifically for teachers of English and that are not open to the public, the majority of group members are active teachers of English and there should be little interest to participate in the study if not a teacher of English. Creswell (2014) states that one of the issues of mixed methods research is the fact that it is time consuming and data collection is a long process which is why a sample of convenience was used. However, since the importance of the national test result has been emphasized and is now regulated by law a mixed method seemed most relevant in order to obtain knowledge on a phenomenon not previously researched to a great extent.

4.6 Ethical considerations

According to Vetenskapsrådet (2017), who provides research guidelines in Sweden, one of the most important aspects of conducting research is the treatment of the respondents participating in the research. In Sweden, research only requires specific permission from a board of ethics in certain instances, for example when sensitive personal details are being treated (Vetenskapsrådet, 2017). For this study no personal details have been saved and instead the survey respondents have been anonymous as the material has been collected but no records of their identities have been saved. In order for respondents to know what the purpose of the study has been and how the information they provide has been treated a letter of information was included in the survey. McKay (2006) states that respondents should be provided with information on what research is being conducted, the purpose of the study, information regarding how data is collected and kept as well as a statement regarding the fact that participation is voluntary and can be withdrawn at any time. Information was provided in the post on Facebook where the link to the survey was published as well as in the introduction of the survey. However, seeing as no records are kept of the identity of the respondents their participation could only be withdrawn before sending the survey form. The responses could after sending not be traced to a certain individual thus making it impossible to withdraw from the study.

The interviewed respondents however have received separate information of the agreement to participate in an interview. In order to keep confidentiality, which means that the information will not be spread and identities will not be revealed (Vetenskapsrådet, 2017), the respondents received a code name at the beginning of the interview "Teacher A" and "Teacher B" and no personal information such as names or contact information has been kept. The respondents have been informed that the recordings of the interviews will be kept for the remainder of the time this thesis is being written and will be deleted after final submission.

5. Results

5.1 Survey result

The survey was answered by 82 respondents all of whom are currently working or have previously worked as teachers of English in year nine and aimed to provide an overview of English teachers' perceptions of reading comprehension strategies on the national reading comprehension test as well as teachers usage of the test in assessment of student abilities. The respondents' work experience was rather varied as can be seen in the table below.

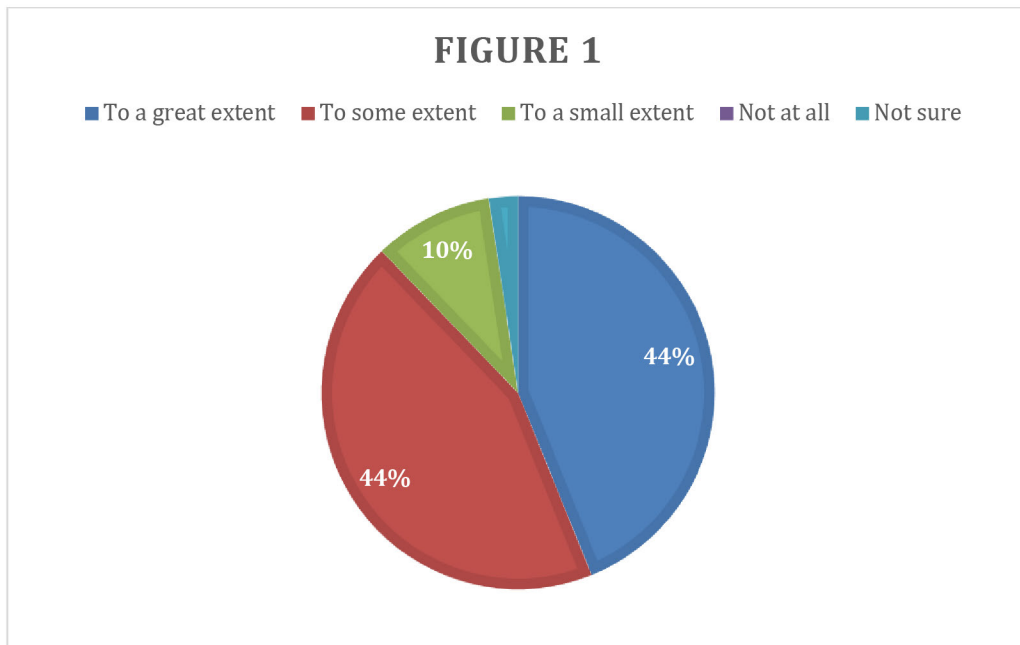
Table 1.

Responses to the question: Number of years in the profession? (n=82)

Years as an English teacher	Number of respondents	Percentage
0-5	12	14,6%
5-10	16	19,5%
10-15	13	15,9%
15-20	16	19,5%
20-25	16	19,5%
25-30	5	6,1%
30+	4	4,9%
Total.	82.	100%

5.1.1 General findings

When asked the question to what extent do you think that the national reading comprehension test can test students' use of reading comprehension strategies, there is clear consensus that it does in fact test use of reading comprehension strategies which can be seen in figure one.



Responses to the question: "To what extent do you think that the national reading comprehension test can test students use of reading comprehension strategies?" (n=82)

The respondents were then provided with a list of strategies and asked to choose the examples of strategies they believe can be tested through the national reading comprehension test. The responses showed that there were significant differences in what strategies the respondents considered to be tested in the national reading comprehension test as can be seen in table two.

Table 2.

Responses to the question: "What strategies do you think that the national reading comprehension test can test? Several alternatives possible" (n=82)

Strategy	Respondents (%)
Identifying keywords and/or most important content	85,4%
Re-reading of relevant parts	57,3%
Using context to understand specific words	70,7%
Using context to understand the general content	80,5%
Using previous knowledge of subject content	37,8%
Using previous knowledge of text type/genre	24,4%

Monitoring understanding	28%
Summarising content	32,9%
Making inferences (connections/conclusions based on information at hand)	61%
Making predictions of content	28%
Planning reading	30,5%
Choosing strategies to improve understanding	41,5%
Evaluating reading/ revising strategy	13,4%

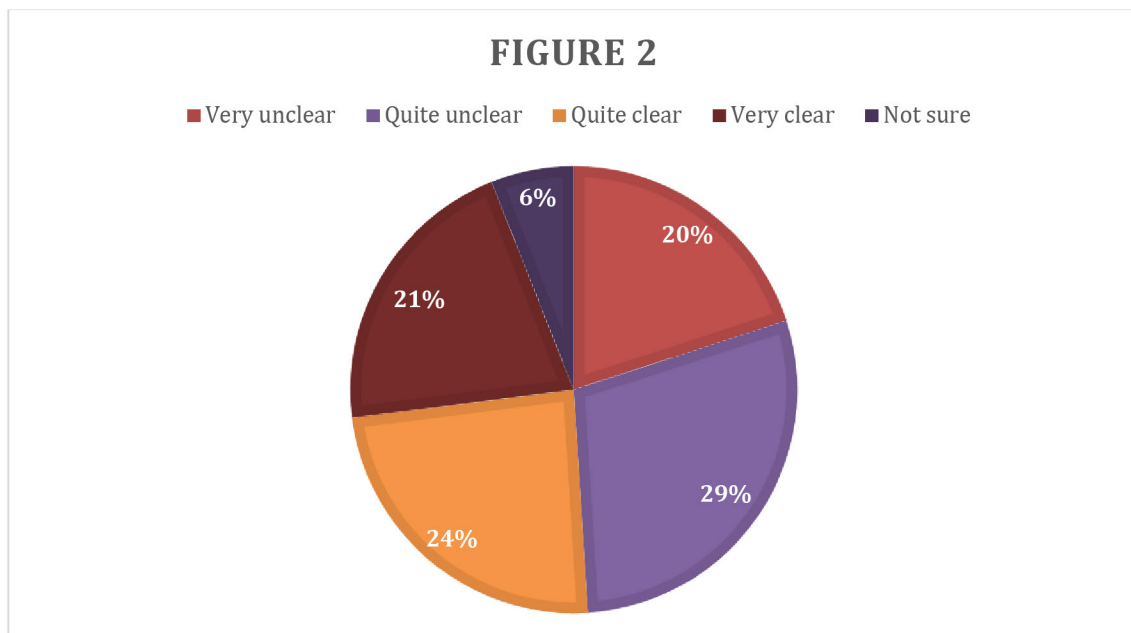
As illustrated in table two, there are significant differences in what strategies the respondents considered to be tested although a majority of respondents considered it possible to test the use of the following strategies:

- Identifying keywords and/or most important content (85,4%)
- Using context to understand the general content (80,5%)
- Using context to understand specific words (70,7%)

In addition to the list provided, respondents could also add strategies if there were other strategies they considered to be tested that were not included in the list, however no respondents added additional strategies. One respondent chose to not choose any of the strategies provided but wrote a comment instead: "The national test can test all types of strategies, the question is whether the students have practiced them and how one can assess the use". A second respondent added a comment in addition to the alternatives marked for the question: *If the students have previous knowledge or not completely depends on the subject in the text AND the individual student's previous knowledge so it is difficult to answer the questions on previous knowledge in general.* The comments from respondents raise some interesting questions, for instance what the first respondent means by *the question is whether the students have practiced them and how one can assess the use*, indicating that practicing would make a difference in terms of assessment. Furthermore, the second comment seem to suggest that previous knowledge is context based and even if that might be true, an interesting question is to what extent subjects in the texts impact individual students' strategy use and how teachers consider previous knowledge when assessing the results of the national

test. When responding to the question, the respondents could choose as many alternatives as wanted and 59 of the respondents chose five alternatives or more but there were significant differences in number of alternatives chosen as well, ranging from nine to one between the respondents. In addition to the alternatives with strategies the respondents could also choose an alternative “none of the above” which was not chosen by any of the respondents, indicating that all respondents to more or less extent consider strategies as part of the assessment of the national reading comprehension test.

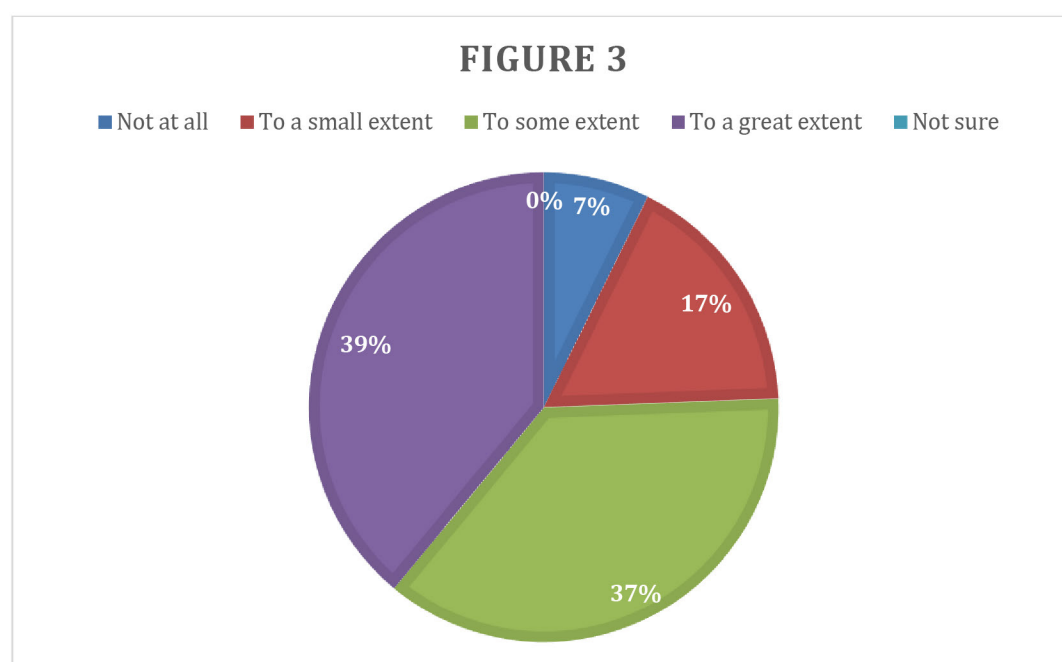
As seen, there seems to be consensus about the fact that the test does test use of strategies but varying opinions on what strategies can be tested and to what extent. Consequently, it becomes interesting to analyze how respondents consider the assessment manual and whether they consider it clear in terms of the assessment of strategy use. The differences seen could indicate that the assessment manual provided by Skolverket might be somewhat unclear on what strategies to assess and how. However, the respondents seem to have different opinions on the clarity of the assessment manual, as can be seen in figure two.



Responses to the question "How clear do you consider the assessment manual provided by Skolverket in terms of assessing strategies on the national reading comprehension test?" (n=82).

As seen, the number of respondents who consider the assessment manual quite or very unclear is very similar to the number of respondents who consider it quite or very clear which raises the question as to why there are such significantly different perceptions of the assessment manuals. Moreover, it is not necessarily the case that those who consider the assessment manual as quite or very clear regarding the assessment of strategies are “correct” since there is no specific definition or explanation of what strategies to assess in the test to compare their responses to.

However, seeing as 48,8% consider the assessment manual quite or very unclear it would be possible to assume that the respondents could be reluctant to use the test results when assessing students' strategy use in general. However, when asked to what extent the respondents use the national test as part of the overall assessment of students' use of reading strategies the results indicate the contrary as seen in figure three.



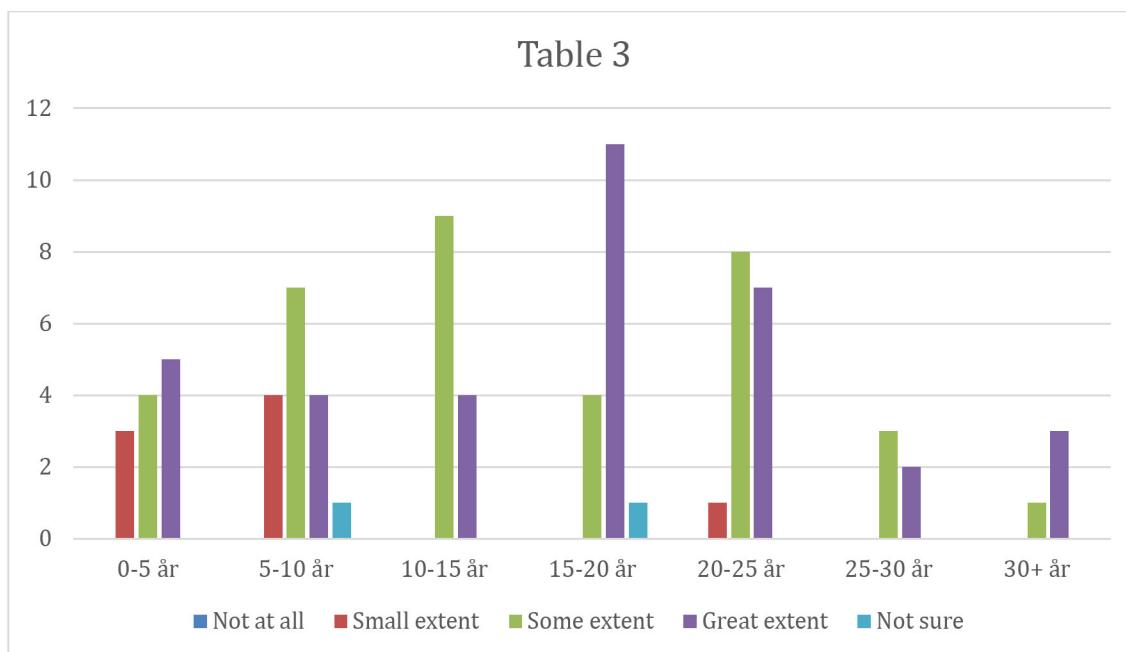
Responses to the question "To what extent do you use the national reading comprehension test to assess students' use of strategies for reading comprehension?" (n=82)

Even though many respondents consider the assessment manual unclear and the opinions on strategy use on the national reading comprehension test vary significantly, 75,6% of respondents use the test to some or great extent when assessing the students reading strategy use in general. Since teachers are required to consider the test results when grading it is

logical that a large number of respondents report using the test to a great extent, although it should be noted that the responses could be impacted by the fact that teachers know they are supposed to. The number might not necessarily be an accurate measure of what teachers are in fact doing but rather what the respondents know they should be doing. It is also possible that since many respondents consider the assessment manual unclear, they might instead rely on the score as an indicator on whether a student is a good strategy user or not. It is interesting to note though that despite it being a requirement, 7,3% of respondents claim to not use the test when assessing strategy use. A reasonable explanation as to why some teachers do not use it could simply be that they do not know how to use it to assess strategy use. When analyzing the other responses of the respondents who state they do not use the national test result to assess strategy use, all have answered that they consider the assessment manual quite or very unclear. Furthermore, two of the respondents have also answered "not sure" to the question on whether the test can be used to assess use of strategies thus indicating that a lack of knowledge of how to assess strategy use on the national reading comprehension test might be the explanation as to why it is not used to assess strategy use.

5.1.2 Impact of work experience

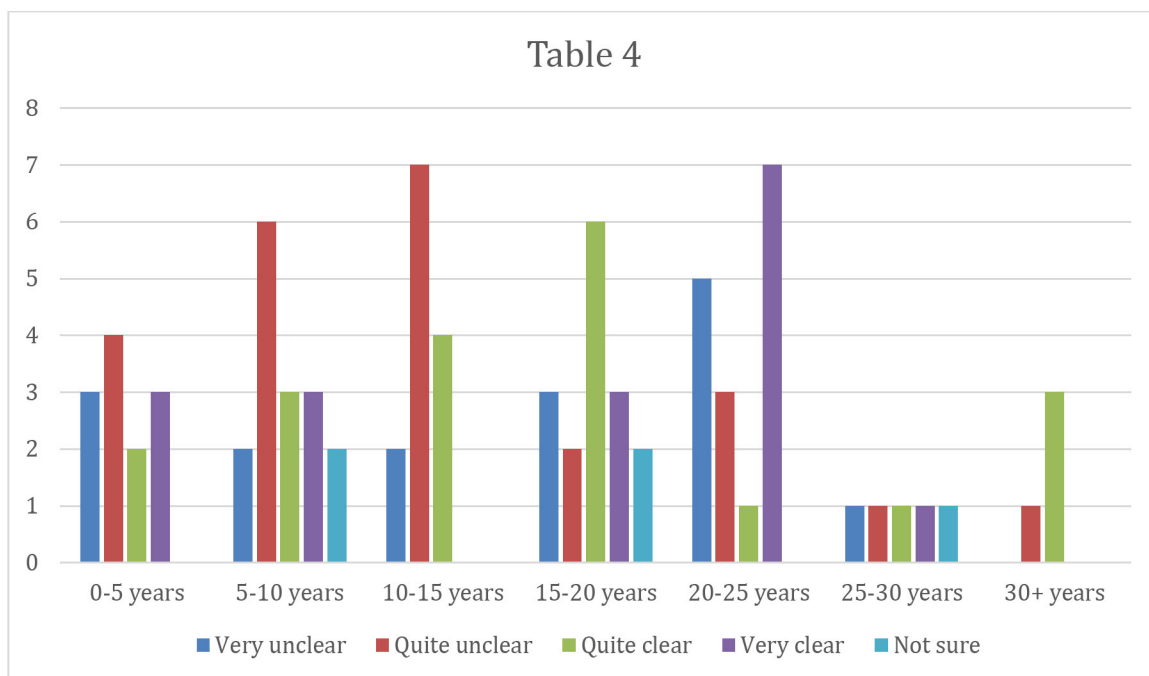
When analyzing respondents' perceptions of strategy assessment on the national test in reading comprehension, it is interesting to investigate how work experience might impact perceptions of the national test and strategy assessment. In table three responses to the question regarding to what extent the national tests can test strategy use have been divided based on number of years as a teacher of English.



Number of responses to the question: "To what extent do you think that the national reading comprehension test can test students use of reading comprehension strategies?" (n=82.)

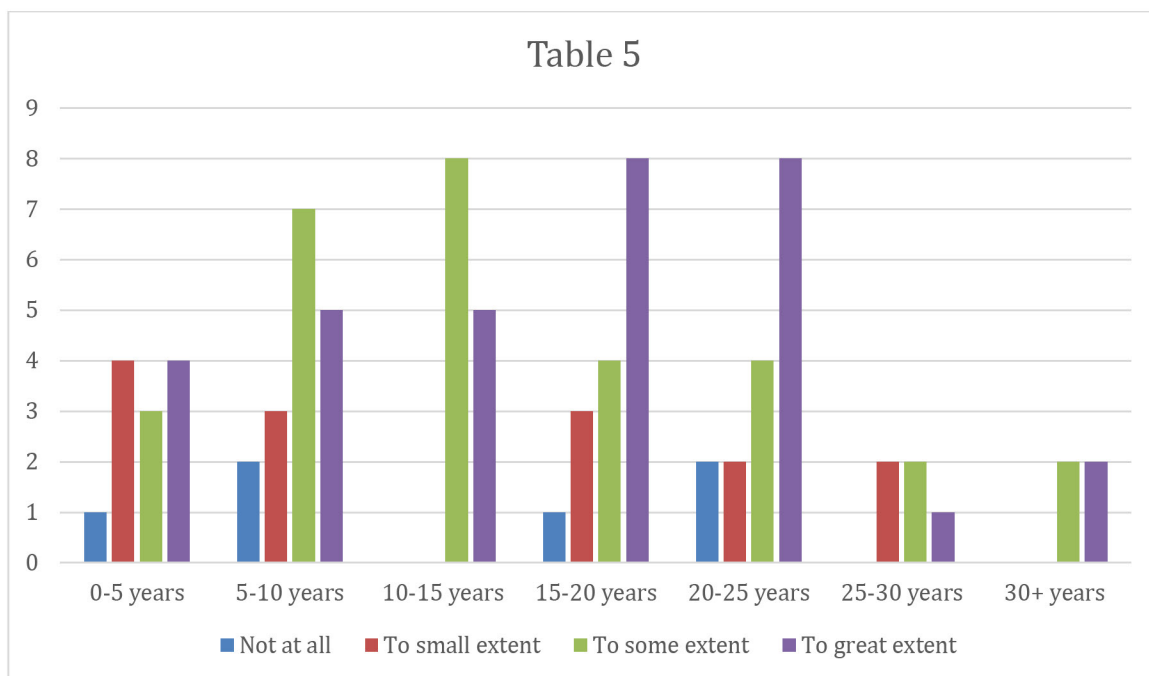
As seen in table three, all responses are demonstrated although it is important to note that there are differences in number of responses represented in each category which impacts what comparisons can be made between categories. However, an interesting pattern is revealed where the majority of respondents who believed that the national test can test strategy use to a small extent have worked as teachers for 0-10 years. Amongst respondents who have worked as teacher for over 10 years, which is 65,9% of respondents, only one respondent considered strategy assessment possible to a small extent, indicating that experience increases the understanding of strategy assessment through the national test.

Seeing as respondents with more experience seem to consider strategy assessment a part of the national test in reading comprehension to greater extent than less experienced respondents, one might assume that more experienced respondents consider the assessment manual clearer. In table four, responses to the question regarding the assessment manual is presented with a division of respondents based on number of years in the profession.



Number of responses to the question "How clear do you consider the assessment manual provided by Skolverket in terms of assessing strategies on the national reading comprehension test?" (n=82).

Interestingly, no significant patterns or differences emerge between respondents' perceptions of the assessment manual. There are respondents from each category who consider the assessment manual unclear to some extent and clear to some extent and even in the category of respondents with 20-25 years of experience where a significant number consider the manual very clear in comparison to other categories, there are also several respondents who consider the manual very unclear. The polarization between the responses is especially interesting when considering the fact that in the category of respondents with 20-25 years of experience, all but one respondent believed that the national test can test strategy use to some or great extent. Consequently, it seems relevant to analyze whether there are differences regarding to what extent respondents use the national reading comprehension test result when assessing students use of strategies, which is demonstrated in table five.



Number of responses to the question "To what extent do you use the national reading comprehension test to assess students' use of strategies for reading comprehension?" (n=82)

As demonstrated in table five, responses vary within each category of numbers of years of experience. The decision to not use the result to assess strategy use do not seem to relate to the respondents' work experience as the response "not at all" can be found in four different categories. Interestingly, if comparing the responses in table five to the responses in table three, the respondents who are not sure on whether strategy use can be assessed in the national test and those who consider it part to a small extent can be found in the same categories as the respondents who claim to not use the test result when assessing strategy use at all.

Even though the survey does not provide insight into what causes the differences that can be seen in the tables or why respondents perceive the national reading comprehension tests the way they do, the results seem to indicate that work experience might not be the determining factor to how the results are used. Instead, patterns emerge when analyzing the responses to the various questions and the general perceptions of strategy assessment through the national test instead seem to affect to what extent the results are used. However, the fact that experience does not seem to affect the perceptions of the assessment manuals indicate that work experience might affect respondents general understanding of strategy use and

assessment but not in terms of what information is provided by Skolverket or how to assess strategy use on the reading comprehension test.

5.2 Interview results

Interviews with two teachers were conducted for the study in order to obtain more elaborate descriptions and reasonings regarding the teachers' perceptions and use of the national test results. The interviewed teachers will be referred to as teacher A and teacher B. Teacher A has four years of work experience and has graded national tests in English reading comprehension four times whereas Teacher B has only worked as a teacher for one year and just graded the national tests for the first time right before the interview was conducted.

5.2.1 Perceptions of strategy assessment on the national test

In the first question both teachers were asked what they think about the national tests in English in general and the reading comprehension test specifically. Both Teacher A and teacher B considered the national tests in English useful and were very positive towards the tests in general, as well as the reading comprehension test in particular. Both respondents were relatively new to the profession and they both mentioned that the national test is a great tool to ensure that their formative assessment of the students' abilities is equivalent to the national standard. Furthermore, both teachers stated that the national tests have been a beneficial learning experience and have provided opportunities for professional development and learning seeing as the marking of the tests has been done with colleagues.

Teacher B, who did the assessment for the first time this year, stated that the grading of the national reading comprehension test was done in a team where several English teachers participated and marked a few pages of each test. The teacher thought that marking the reading comprehension test with an answer key would be straightforward and quite simple but realized that there was still a lot of room for interpretation when discussing the assessment with colleagues, in particular regarding the open-ended-questions. Teacher A mentioned this aspect of the assessment as well and claimed that students sometimes provide answers that the teacher consider to be correct but the answer might not be included in or not accurate according to the answer key. The uncertainty of interpretations made the marking of the national test complicated at times and there were instances where the respondent and

colleagues had different opinions on how to interpret the students' answers and the answer key, while some colleagues wanted to accept a student's answer others thought that since it was not included in the answer key it should automatically be considered inaccurate.

When asked about what strategies the national reading comprehension test aims to test, none of the teachers interviewed could provide an account of specific strategies that are tested according to Skolverket but both teachers stated that they are aware that the test is supposed to test reading strategies to some extent. Teacher B compared the reading comprehension test in English to the national reading comprehension test in Swedish and described how the answer key for the test in Swedish clarifies what each question item is meant to test which made it easier to know what ability is being tested and when to accept an answer. Every question item in the Swedish reading test is given one or several specific purposes, for instance finding information or making connections and coming to conclusions which means that it is possible to identify patterns in what it is that the student is struggling with when marking the test. If the purpose of each question item or for each text was specified in the teacher information, teacher B stated that it would be easier to know when to accept an answer that is not mentioned at all in the answer key. Furthermore, a more specific answer key could make the result easier to interpret when assessing a student's abilities in terms of reading comprehension strategies.

5.2.2 Use of test results and strategy assessment

Both teacher A and B expressed that there is no explicit information on how strategies should be assessed in the test and considered it a very small part of the assessment of students' abilities for the final grade. Teacher A stated that the test does not really test a student's use of strategies since there are not any questions on what the students are thinking. Teacher A claimed that: *They would get a different result if I sat next to them and asked them to explain what they are doing* and elaborated the statement by saying that even though a student might use strategies when taking the test there is no way of knowing when and what strategies the student is using and therefore the test cannot really be used to assess a student's use of reading comprehension strategies. Instead, the general score could indicate if a student has been successful thus making it possible to assume that a student can use strategies when reading in English or the opposite, that a student might lack strategies, but the test cannot provide any

more information than that. Teacher B however, claimed that the test could possibly test a student's use of strategies to some extent, for instance in the part of the test where a student fills in a gap. However, it is not stated explicitly in the answer key thus making it difficult for a new teacher to know how to interpret and use the result, according to teacher B. The teacher must make a general assessment of the test and strategy use is automatically assessed at the same level as the overall score indicates. A student could potentially use efficient strategies at a high level but still get the answer wrong if the language is too complex for the student, but that is not possible for the teacher to identify and there are no instructions on how to use the test result other than looking at the overall score. Teacher B stated that assessing strategy use in general is difficult and that it is likely partly due to inexperience and consequently, teacher B has not done much of it the first year. As a result, the teacher had to rely on the overall national test results in terms of the score since there was not any explicit information on how to assess strategy use.

Teacher A however does not use the national test result to a great extent to assess strategy use and emphasized that strategy use must be assessed through process. It is not the national reading comprehension test in itself that teacher A considered the problem, but it is the fact that there is a need of knowing what a student is thinking or doing while reading in order to be able to use it for assessment. The main assessment of strategy use is done during the lessons where teacher A can ask students to explain and elaborate on what they are thinking while solving the problems. If the national test includes assessment of reading comprehension strategies, there should be information on what parts exactly is testing what strategy in order for the assessment to be valid and reliable. Teacher A expressed concern that teachers are currently making general assumptions about strategy use based on an overall score thinking that the national test provides information on strategy use. Even if strategies are used on the test it is not necessarily possible to assess the use. Teacher A posed the question: *How many teachers have actually reflected on what they can see?*

6. Discussion

6.1 Assessing strategy use

Reading comprehension strategies are evidently highly complex from several perspectives. It has been established that the use of strategies is difficult for researchers to study and consequently research on the topic is limited and often considered controversial (Gass, 2020). Nevertheless, teachers are expected to assess strategy use when teaching English in Sweden. However, the purpose of strategy assessment is often to support students' learning where the result is meant to provide information that can be used to assist students and increase their self-regulation (Gunning & Turner, 2018) which is not the explicit purpose of the national test. As stated previously, the national exams aim to support teachers' assessment of student abilities and to provide information on students' performances on a national level (Skolverket, 2016). Summative assessment has been described as measuring what a student has learned and is usually done at the end of a course (Farrell, 2008) which could be argued to accurately describe what the national tests do. The purpose and relevance of assessing strategy use through standardized tests could therefore be questioned and previous research has indicated that teachers frequently get confused regarding strategy assessment. Consequently, teachers use summative assessments designed for external accountability in a formative manner (Almasi & Fullerton, 2012).

The differences in responses in the survey conducted for this study could indicate that teachers might be confused regarding strategy assessment on the national tests in reading comprehension. In addition, the varying responses in terms of the clarity of the assessment manual seem to suggest that teachers are not provided with relevant information on how to assess strategy use in a standardized reading comprehension test. However, a majority of respondents claimed to use the test result to assess students' ability to use strategies while reading which supports Lundahl's (2019) concern that teachers rely on an overall score on a national test to assess strategy use and that other few methods of assessing strategy use are used. Similarly to Lundahl (2019), Oakley (2011) established that teachers struggle with assessment of reading comprehension strategies and as a result rely on comprehension products to assess strategy use. Seeing as teachers seem to rely on comprehension products or standardized test results to assess students' ability to use reading comprehension strategies it is relevant to discuss what information standardized tests can provide according to teachers.

6.2 Assessing strategy use on the national reading comprehension test

As stated, there seems to be confusion regarding the assessment of strategy use on the national reading comprehension test, however, 87.8% of respondents in the survey conducted for this study think that the national test in reading comprehension can test strategy use to some or great extent. In contrast, Thomas & Badger (1991) state that standardized reading tests are inappropriate to use for the assessment of strategy use if the questions are only related to product and not process. Similarly, Wang (2006) claims that the relationship between the question and the answer it asks for is what determines whether strategies can be assessed or not. Interestingly, teacher A expressed in the interview that there are no questions on the test regarding what the student is thinking thus making it impossible to assess strategy use. However, it could be argued that teacher A referred to strategies in the sense that research refers to metacognitive strategies as they are consciously used strategies and relates to what a student is thinking (Börjesson, 2012). It would seem as though teacher A excluded cognitive strategies which could be used without awareness which is why a test with questions on what students are thinking might not capture cognitive processes used that could have been crucial for the student's understanding. In contrast, teacher B provided examples of question items that could test strategy use such as the cloze test item. Teacher B stated that students have to understand how to use the text to fill in the gap which is similar to what researchers claim in terms of strategy assessment on standardized reading comprehension tests (Klapwijk, 2013). However, when teachers were asked to evaluate the parts of the national test of 2022, the cloze test item was considered appropriate to assess reading abilities by fewer teachers than the two parts consisting of texts and questions (Göteborgs Universitet, 2022), even though the difference was small. Ajidhe et al. (2012) state that cloze test might not test reading comprehension but rather strategy use whereas Farrell (2008) however, claims that the cloze test rather tests students test-taking strategies and that some proficient readers might not benefit from the cloze test. Previous research as well as the results presented in this study suggest that the term strategy is very complex and in order to know whether it should be assessed or not and how to assess strategy use on a standardized reading comprehension test, the purpose of the test and the assessment manual must be clear. As mentioned, there are various perspectives on what standardized tests can measure and how. Multiple choice-questions are often considered controversial and while Wang (2006)

states that the cognitive demand determines what is tested, Lim (2019) claims that multiple choice-questions can only be used for basic comprehension and does not require strategy use. Arensmeier et. al (2014) concluded that multiple choice-questions are controversial amongst teachers as well. While some teachers consider them too simple and think that students can guess their way to a passing grade there are other teachers who instead think that the questions are phrased in a misleading way resulting in students misinterpreting the alternatives. These different perspectives might indicate that teachers do not necessarily receive enough information on test construction and various purposes of questions. Even though the questions referred to by the teachers in that study are unknown and cannot be analyzed in this paper the questions might have different purposes and should thus be interpreted differently by students and teachers. MC questions where students can guess the correct answer might simply be basic comprehension questions. The questions perceived to "fool" students however, might test strategy use and the students getting the answers wrong could be skimming the text and matching words instead of reading for comprehension and using relevant strategies (Sanders & Garwood, 2022). The teachers' perceptions of the MC questions expressed in that study could be interpreted as a lack of understanding of the question item and the test construction in general.

In the interviews both teachers expressed knowledge of the fact that strategies are part of the assessment of the reading comprehension test but none of them can provide examples of what strategies Skolverket aim to test. Interestingly, one of the teachers compared the test in English to the national reading comprehension test in Swedish where each question item has an explicit purpose which the English test does not. Moreover, the teacher suggested that if the purpose of each question item was specified, it would be easier to know when to accept an answer not included in the answer key. The lack of instructions on the assessment of reading comprehension strategies in English is identified in the interviews as an issue as it could potentially impact the validity of the test. When it is up to individual teachers to interpret how to analyze and assess strategy use there is a risk of variations in the interpretations that are not considered or evaluated to a sufficient extent. In order to ensure validity the process of interpreting score and making inferences from the result must be made clear by test constructors (Bachman, 2004; Kane, 2012). When evaluating the validity and reliability of the test in 2022 the test constructors considered it valid as the evaluation showed

that 99% of the respondents stated that the test tested what is aimed to be tested to some or great extent (Göteborgs Universitet, 2022). However, the fact that strategies can and perhaps should be tested in the national reading comprehension test does not necessarily mean that the assessment has been done in a similar way which is crucial to ensure validity (Skolverket, 2017). While Bachman (2004) states that validity will depend on the clarity of connections between score and construct, Skolverket (2017) claims that even though tests that include questions open to interpretation could lower validity the possible errors that could occur are the result of the assessment process used by teachers. The results from this study seem to suggest however that even though perceptions of what the reading comprehension test can measure may vary depending on experience of the profession, the opinions on the clarity of the assessment manual differs regardless of experience.

When respondents of the survey conducted for this study were asked to select strategies that they believe are tested in the national reading comprehension test in English, three strategies were selected to the greatest extent. Identifying keywords and/or most important content, using context to understand the general content and using context to understand specific words were identified by a majority of respondents which is particularly interesting in light of what Cordon & Day (1996) concluded in their study on strategy use on standardized reading comprehension tests. In contrast to the teachers in this study, Cordon & Day (1996) found that even though the use of context occurred during reading comprehension test it was to a significantly less extent than other strategies such as rereading, paraphrasing and using previous knowledge. In this study, using context was identified by most respondents whereas using previous knowledge of subject content was identified by 37,8% and using previous knowledge of text type/genre was only identified by 24,4% of respondents. The fact that few teachers believe that the test can test students' use of previous knowledge could be considered rather surprising since Skolverket (2017) explicitly mentions the use of previous knowledge in the comments to the syllabus, stating that students should develop the ability to use contextual clues and previous knowledge of a topic to increase comprehension. However, considering the comment made by one of the respondents in the survey that it depends on topic, using previous knowledge might be considered a strategy that students could use when taking the test but that is difficult to identify and assess in a standardized reading comprehension test. The use of context to understand the general content and specific words

is probably identified by the majority of respondents due to the construction of the national reading comprehension test in English. In contrast to the comprehension questions used by Cordon & Day (1996), the national reading comprehension test consists of a variety of test items and question types, one of which being the cloze test where context is necessary to complete the task (Klapwijk, 2013) as previously mentioned.

6.3 The use of test results

When analyzing test validity and reliability, the use of the test and test results must be taken into consideration (Skolverket, 2017). Even though the consideration of the results has been regulated by law for the past five years, it has always been considered an important part of grading in Sweden and 19% of teachers claimed in Hirsch study (2015) that it would be difficult to assess student abilities without the national tests. Seeing as the results of the national tests have to be taken into specific consideration when grading students, it is interesting that 7.3% of survey respondents in this study claim that they do not use the national test at all to assess students' use of reading comprehension strategies. However, the fact that all 7.3% have answered that the assessment manual is quite or very unclear or 'not sure' would seem to argue that there is uncertainty on how to assess strategy use which might explain why it is not done. The interviews also indicate that there is uncertainty on how to assess the use of strategies in the national reading comprehension test. Neither teacher A or B can provide examples of what strategies they are expected to assess or how and instead the overall score is used to analyze whether a student can use efficient strategies or not. Similarly to concerns expressed by both Lundahl (2019) and Oakley (2011). Teacher A stated that the result is used to a small extent but that assessment of strategy use is mainly done in other situations and during lessons. Seeing as several methods of strategy assessment is encouraged (Almasi & Fullerton, 2012) assessing strategy use during lessons as well is beneficial. In contrast, teacher B relies on the results of the national test to quite a large extent as teachers often do (Oakley, 2011) due to uncertainty on how to assess strategy use in general. The two different perspectives of the use of the national test result that emerged in the interviews become increasingly relevant when looking at the survey result in general. The responses indicate significant differences in how the results are used since 39% of respondents use the test to a great extent to assess strategy use, 36.6% to some extent and 17.1% to a small extent. If 17.1% of teachers work with assessment of strategy use through other methods, it could

mean that their students get more practice and have several results available that will affect the final grade. In contrast, students who have teachers that rely heavily on the national test results and who do not assess strategy use formatively might not get to develop reading comprehension strategies to the same extent and the national test result will have a greater impact on the overall assessment of student abilities. Wang (2006) discusses the risk of using an overall score to evaluate students' use of reading comprehension strategies and states that in order to assess strategy use on tests, teachers are required to analyze the relationship between questions and answers.

6.4 Chapter summary

This chapter has discussed the relevance of assessing strategy use in a summative standardized test as well as different perspectives on what standardized reading comprehension tests can measure. The results of this study concur with previous research on strategy assessment in terms of establishing that there are several ways of interpreting the term strategy, how to assess strategy use and what standardized tests can be used to assess.

Even though studies show that knowledge of reading comprehension strategies have an impact on learners' reading comprehension in various tasks the relationship between reading comprehension and reading comprehension strategies is not evident in standardized reading comprehension tests (Muijselaar et. Al, 2017). Cordon & Day (1996) conclude that even though strategies are often used when taking standardized tests, it is not possible to determine how effectively a student is using strategies to complete the tasks. Teachers in this study express similar concerns in comments and interviews and teacher A summarizes the concern with the question; *How many teachers have actually reflected on what they can see?*. A valid question based on the results in this study and might be one worth researching further.

7. Conclusion

Reading comprehension strategies are difficult to define and research as it entails understanding an internal process that is not always conscious, yet teachers have been expected to assess students' use of reading comprehension strategies in different ways. One of which being the national reading comprehension test in English which aims to support teachers' assessments and to contribute to reliability and validity in formal grading (Skolverket, 2016).

This study has indicated that using national reading comprehension tests to assess strategy use without providing specifications on how to do so could possibly impact the validity and reliability of the national tests and student grades. Furthermore, it has been established that teachers perceive reading comprehension strategies and the possibilities of the reading comprehension test in different ways even though teachers generally claim that the national test in English works well (Skolverket, 2016) which has previously been the method used to evaluate the validity of the results.

As of autumn 2022 the assessment of students' use of reading comprehension strategies has been omitted from the grading criteria. Although the potential issues of assessing reading comprehension strategies in standardized tests presented in this paper may have been solved while this paper was being written, there are still two aspects that are highly relevant. First and foremost, the issues regarding teachers' perceptions of reading comprehension strategies in general. Even though the summative assessment has been removed the syllabus still states that strategies should be taught. For teachers to know what strategies to teach and how, it might be relevant to further research how teachers understand and teach strategies and how they evaluate students' learning. In the interviews one of the teachers stated that strategies are difficult, and the survey responses do indicate that there might be teachers who are uncertain of how to work with strategy use. Considering the fact that previous research on the usefulness of strategy teaching is rather inconclusive it would be unreasonable to assume that all teachers know how to teach reading comprehension strategies. More research on strategy use in general is needed and in order for students to obtain equal education, all teachers must know how to support students' reading abilities and comprehension.

Second, test construction and the purposes of a test should be clearly communicated to the ones conducting the tests with the students, especially if they are also the ones assessing the tests and analyzing the relevance of the results. Descriptions in general terms may not be sufficient if the aim is for teachers to interpret and understand the tests and their relevance in the same way. In addition, the way in which tests are evaluated might be worth considering since teachers state that the tests work well in evaluations but when asked what the tests can assess the responses vary to great extent.

The national tests in English as well as teachers' use and perceptions of the test might need to be researched further and in terms of specific strengths and weaknesses rather than teachers' general opinions and practical issues. The national tests are important and can be very useful as the majority of teachers attest in various evaluations, it is however important to ensure that teachers perceive and use the tests in the same way in order to ensure validity and reliability.

7.1 Limitations

The purpose of this thesis has been to investigate teachers' perceptions of strategy assessment on the national reading comprehension test in English in Sweden. Due to the very limited amount of research on teachers' perceptions of national tests in general, previous research on standardized testing has been presented in order to demonstrate and discuss the complexity of strategy assessment and standardized testing. However, the intention has not been to evaluate whether teachers' perceptions or use of national tests are accurate in any sense but rather to show how perceptual differences could impact validity. Moreover, a mixed method was used which is beneficial to obtain explanations to results from a survey, however the differences in weight between the two methods results in certain limitations. Although the survey was answered by a relatively large number of respondents, the questions were brief and provided no explanations or in-depth responses and only two respondents participated in an interview. As a result, the explanations provided might not be possible to generalize or base conclusions on. It is rather likely that several additional perspectives exist that have not been accounted for in this thesis.

Even though the results of this study might not provide a comprehensive overview of teachers' perceptions of strategy assessment on the national reading comprehension test in

English, I believe the results contribute to increase the understanding of how national testing in Sweden is conducted and perceived. Hopefully, further research will be conducted on national tests in English in Sweden where teachers' perceptions and use of results can be investigated and considered by test constructors to further increase validity of national testing.

References

- Ajideh, P. & Mozaffarzadeh, S. (2012). C-Test vs. Multiple-Choice Cloze Test as Tests of Reading Comprehension in Iranian EFL Context: Learners' Perspective. *English Language Teaching*, 5(11), 143-150
- Almasi, J.F. & Fullerton, S.K. (2012). *Teaching Strategic Processes in Reading*. Guilford Publications.
- Ardoin, S.P., Binder, K.S., Zawoyiski, A.M., Nimocks, E & Foster, T.E. (2019). Measuring the Behavior of Reading Comprehension Test Takers: What Do They Do and Should They Do It? *Reading Research Quarterly*, 54 (4). 507-529
- Arensmeier, C., Bonnevier, J., Borgström, E., Lennqvist-Lindén, A-S., Lundahl, C., Nilsson, P., Sundberg, B., Sundhäll, M., Yassin, D. & Wetterstrand, F. (2014). *De Nationella proven och deras effekter i årskurs 6 och 9*. Örebro Universitet.
- Bachman, L. (2004). *Statistical Analyses for Language Assessment*. Cambridge University Press.
- Bachman, L. F., & Palmer, A. (2010). *Language Assessment in Practice*. Oxford: Oxford University Press.
- Badger, B.E. & Thomas, B. (1991). Open-Ended Questions in Reading. *Practical Assessment, Research and Evaluation*. Vol. 3(4) DOI: 10.7275/fryf-z04
- Block, E. (1986). The Comprehension Strategies of Second Language Readers. *TESOL Quarterly*, 20(3), 463-494. <https://doi.org/10.2307/3586295>
- Börjesson, L. (2012). *Om strategier i engelska och moderna språk*. Stockholm: Skolverket.

Cohen, D.A. (2018). A Closer Look at Language Learning Strategies. In: Oxford, R. & Amerstorfer, C. (Ed.). *Language Learning Strategies and Individual Learner Characteristics: Situating Strategy Use in Diverse Contexts*. Guilford Publications.

Cohen, L., Manion, L. & Morrison, K. (2018). *Research Methods in Education*. Cambridge: Routledge

Cook, V. (2016). *Second Language Learning and Language Teaching*. Cambridge: Routledge

Cordón A. Luis & Day D. Jeanne (1996). Strategy Use on Standardised Reading Comprehension Tests. *Journal of Educational Psychology*. Vol. 88 No. 2 (288-295)

Council of Europe (2001). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge UK: Cambridge University Press.

<https://www.coe.int/en/web/common-european-framework-reference-languages>

Creswell, J. W. (2014). *Research Design. Qualitative, Quantitative, and Mixed Method Approaches*. SAGE Publications.

Dewitz, P. & Graves, F.M. (2021). The Science of Reading: Four Forces that Modified, Distorted or Ignored the Research Finding on Reading Comprehension. *Reading Research Quarterly*, 56 (1), 131-144.

Erickson, G. (2018). Bedömning av språklig kompetens. I G. Erickson (Red.) *Att Bedöma Språklig Kompetens*. RIPS: Rapporter från institutionen för pedagogik och specialpedagogik nr. 16. Göteborgs Universitet.

Farrell, C.S.T. (Ed.) (2008). *Teaching Reading to English Language Learners: A Reflective Guide*. Corwin Press.

Fulcher, G. (2014). Philosophy and language testing. In A.J Kunnan (Ed.). *The Companion to Language Assessment*. Vol. 3

Gass, S. M., Behney, J., and Plonsky, L. (2020). *Second Language Acquisition: An Introductory Course* (5th ed.). New York: Routledge.

Gunning, P. & Turner, E. C. (2018). The Development of Strategy Assessment Methods and Tools. In: Oxford, R. & Amerstorfer, C. (Ed.). *Language Learning Strategies and Individual Learner Characteristics: Situating Strategy Use in Diverse Contexts*. Guilford Publications.

Göteborgs universitet. (n.d). *Översikt: Bedömning och innehåll delprov B*. Retrieved: 2022-04-14 : <https://www.gu.se/nationella-prov-frammande-sprak/prov-och-bedomningsstod-i-engelska/engelska-arskurs-7-9/nationellt-prov-i-engelska-for-arskurs-9#Uppgiftstyper>

Göteborgs universitet. (n.d). *Provspecifikation*. Retrieved 2022-04-14:
<https://www.gu.se/nationella-prov-frammande-sprak/prov-och-bedomningsstod-i-engelska/engelska-arskurs-7-9/nationellt-prov-i-engelska-for-arskurs-9#Uppgiftstyper>
Stockholm: Skolverket

Göteborgs Universitet (2022). *Nationellt prov i engelska, årskurs 9 2021/2022*. Göteborgs universitet: Projektet Nationella prov i främmande språk.
https://www.gu.se/sites/default/files/2022-12/Rapport_Nationellt_prov_Eng_åk9_2021-22.pdf (Retrieved 2023-02-05)

Göteborgs Universitet (2022). *Lärarenkät för det nationella provet i engelska årskurs 9, 2021/2022*. Göteborgs Universitet: Projektet Nationella prov i främmande språk.
https://www.gu.se/sites/default/files/2022-12/Lärarenkäter_Eng_åk9_2022_resultat.pdf
(Retrieved 2023-02-05)

Hirsh, Å. (2015). *Nationella prov i grundskolan - en studie av hur lärare och rektorer uppfattar & hanterar prov och provresultat*. Högskolan i Jönköping.

Hoepfl, M. (1994). Developing and Evaluating Multiple Choice Tests. *The Technology Teacher*, 53(7), 25-26

Johansson, S. (2015). Validitet och lärares bedömningar. *Pedagogisk forskning i Sverige*. 20 (1-2) 2015.

Kane, M. T. (2012). All validity is construct validity. Or is it? *Measurement: Interdisciplinary Research and Perspectives*. 10(1-2), 66-70.

Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement*.

Klapwijk, N. (2013). Cloze Tests and Word Reading Tests: Enabling Teachers to Measure Learners' Reading-related Abilities. *Per Linguam*, 29(1), 49-62 DOI: <http://dx.doi.org/10.5785/29-1-541>

Lindkvist, A-K., Arvidsson, U & Nilsson, S. (2018). Språkprov på dator - vad händer då? I G. Erickson (Red.) *Att Bedöma Språklig Kompetens*. RIPS: Rapporter från institutionen för pedagogik och specialpedagogik nr. 16. Göteborgs Universitet.

Lim, H. (2019) Test Format Effects: A Componential Approach to Second Language Reading. *Language Testing in Asia*, 9(1), 1-22

Lundahl, B. (2019). *Engelsk språkdidaktik*. Lund: Studentlitteratur

Magliano, J.P & Millis, K.K. (2011). Assessing Comprehension during Reading with the Reading Strategy Assessment Tool (RSAT). *Metacognition Learning*. 6, 131-154.

Marloes M. L. Muijselaar, Nicole M. Swart, Esther G. Steenbeek-Planting, Mienke Droop, Ludo Verhoeven & Peter F. de Jong (2017) Developmental Relations Between Reading Comprehension and Reading Strategies, *Scientific Studies of Reading*, 21:3, 194-209, DOI: 10.1080/10888438.2017.1278763

McKay, L., Sandra (2006). *Researching Second Language Classrooms*. Routledge: New York

Messick, S.A. (1989). Validity. In: Robert. L. Linn (Ed.). *Educational Measurement*. New York: American Council on Education/MacMillan.

Oakley, G. (2011) The assessment of reading comprehension cognitive strategies: practices and perceptions of Western Australian teachers. *Australian Journal of Language and Literacy*, vol.34, No.3 2011, pp 279-293

Olsson, E., Nilsson, S. & Lindqvist A-K. (2018) Test-taker feedback i utvecklingsprocessen av nationella prov i engelska. *Acta Didactica Norge*, 12(4), Art. 14.
<https://doi.org/10.5617/adno.6288>

Oxford, L. R. & Amerstorfer, M. C. (2018). *Language Learning Strategies and Individual Learner Characteristics: Situating Strategy Use in Diverse Contexts*. Bloomsbury Publishing Plc.

Sanders, S. & Garwood, D.J. (2022). Assessment of Effective Strategy Instruction and Reading Comprehension. *Preventing School Failure: Alternative Education for Children and Youth*. Vol. 66 (4) 320-326

SFS 2017:1104. *Ändring i skollagen 2010:800*.

https://www.riksdagen.se/sv/dokument-lagar/dokument/svensk-forfattningssamling/skollag-2010800_sfs-2010-800

Skolverket (2009). *Bedömaröverensstämmelse vid bedömning av nationella prov*. Dnr: 2008:286. Stockholm: Skolverket.

Skolverket. *Engelska*. [Course syllabus] (2022).

<https://www.skolverket.se/undervisning/grundskolan/laroplan-och-kursplaner-for-grundskolan/laroplan-lgr11-for-grundskolan-samt-for-forskoleklassen-och-fritidshemmet;jsessionid=ECFBEE8BE939B39D30C3AFC162E37680?url=1530314731%2Fcompulsorycw%2Fjsp%2Fsubject.htm%3FsubjectCode%3DGRGRENG01%26tos%3Dgr%26p%3Dp&sv.url=12.5dfee44715d35a5cdfa219f>

Skolverket (2016). *Nationella proven i grundskolans årskurs 6 och 9* (rapport 447). Stockholm: Skolverket

Skolverket (2017). *Skolverkets systemramverk för nationella prov*. Stockholm: Skolverket.

Skolverket (2020). *Nationella prov i årskurs nio*. Retrieved 22-03-22
<https://www.skolverket.se/undervisning/grundskolan/nationella-prov-i-grundskolan/provdatum-i-grundskolan#Engelskaiarskurs9>

Spencer, M., Gilmour, F.A., Miller, C.A., Emerson, M.A., Saha, M.M. & Cutting, E.L. (2019). Understanding the Influence of Text Complexity and Question Type on Reading Outcomes. *Reading and Writing*. 32, 603-637.

Vetenskapsrådet (2017). *God forskningssed*. Stockholm: Vetenskapsrådet.

Wang, D (2006). What Can Standardised Reading Tests Tell Us? Question-answer Relationship and Students' Performance. *Journal of College Reading and Learning*, 36(2) 21-37. DOI:10.1080/10790195.2006.1085018

Appendices

Appendix 1 - Survey questions

Appendix 2 - Interview questions

Appendix 1 - Survey

Hej!

Erika Strandberg heter jag och är student vid det nordiska masterprogrammet för språklärare med inriktning Engelska som är ett samarbete mellan Göteborgs universitet, Linnéuniversitetet och Høgskolen i Østfold. I mitt arbete vill jag undersöka lärares uppfattningar om läsförståelsestrategier på det nationella läsförståelseprovet i engelska i årskurs nio.

Denna enkät riktar sig därför till lärare som undervisar i engelska och som genomför samt bedömer nationella prov för årskurs nio.

Syftet med enkäten är att undersöka om och hur lärare använder det nationella läsförståelseprovet i engelska i sin bedömning av elevers användande av strategier för läsförståelse. Enkäten består av 5 korta frågor om din upplevelse av provets konstruktion och hur du använder dig utav provet. Du kommer att vara anonym och kan när som helst under besvarandet välja att avbryta ditt deltagande.

För att få en djupare förståelse kommer också några intervjuer genomföras och om du är intresserad av att delta i en fördjupande intervju är du välkommen att ta kontakt.

Erikastrandberg@outlook.com

1. Antal år i yrket?

- ☐ 0-5 år
- ☐ 5-10 år
- ☐ 10-15 år
- ☐ 15-20 år
- ☐ 20-25 år
- ☐ 25-30 år
- ☐ 30+ år

2. I hur stor utsträckning anser du att det nationella provet i läsförståelse testar elevers användande av läsförståelsestrategier?

- ☐ Inte alls
- ☐ I liten utsträckning
- ☐ I viss utsträckning
- ☐ I stor utsträckning
- ☐ Vet ej

3. Vilka strategier anser du att det nationella provet i läsförståelse kan testa? Flera val möjliga.

Identifiera nyckelord och/eller det viktigaste innehållet

Omläsning av väsentliga delar

Använda kontext för att förstå specifika ord

Använda kontext för att förstå det övergripande innehållet

Använda förkunskaper om ämnesinnehåll

Använda förkunskaper om texttyp/genre

Kontrollera sin egen förståelse

Sammanfatta innehåll

Göra inferenser (kopplingar eller slutledningar baserat på information som finns till hands)

Göra förutsägelser av innehåll

Planera läsning

Välja strategier för att öka förståelse

Utvärdera läsning/revidera strategi

Inget av ovanstående

Övrigt:

4. I hur stor utsträckning använder du det nationella provet i läsförståelse som underlag för bedömning av elevers användande av strategier för läsförståelse?

- ☐ Inte alls
- ☐ I liten utsträckning
- ☐ I viss utsträckning
- ☐ I stor utsträckning
- ☐ Vet ej

5. Hur tydliga anser du att bedömningsanvisningarna är när det gäller bedömning av strategier på det nationella provet i läsförståelse?

- ☐ Mycket otydliga
- ☐ Ganska otydliga
- ☐ Ganska tydliga
- ☐ Mycket tydliga
- ☐ Vet ej

Appendix 2 - Interview questions

1. Hur många år har du arbetat som lärare?
2. Hur många gånger har du rättat det nationella läsförståelseprovet i engelska (uppskattningsvis)?
3. Hur upplever du det nationella provet i läsförståelse?
4. Hur upplever du rättningen/bedömningen av de nationella proven?
5. I hur stor utsträckning anser du att NP läs testar elevers användande av lässtrategier?
6. Vilka strategier skulle du säga testas?
7. I hur stor utsträckning använder du NP läs som underlag för bedömning av strategier?
8. På vilket sätt använder du det som underlag? (Poäng/enskilda uppgifter)
9. Hur tydliga anser du att bedömningsanvisningarna är kring bedömning av lässtrategier?
10. Hur arbetar du med bedömning av lässtrategier utöver NP?

