

9

PhD revisited: The Acid Test

*Does upper secondary EFL instruction effectively prepare Norwegian students for the reading of English textbooks at colleges and universities?*¹

GLENN OLE HELLEKJÆR

University of Oslo

ABSTRACT This quantitative study used an International English Language Testing System (IELTS) Reading Module to examine the academic English reading proficiency of Norwegian upper secondary and university students. About 66% of the senior upper secondary school students scored below IELTS Band 6 compared to 32% of the university students. Their problems were due to slow reading for detail and poor vocabulary. Implications for teaching and further research are discussed.

KEYWORDS Reading English for Academic Purposes | language testing | Content and Language Integrated Learning | International English Language Testing System

-
1. This chapter presents the overall results from my doctoral study (Hellekjær, 2005). The entire thesis can be downloaded from: <http://www.uv.uio.no/ils/forskning/publikasjoner/rapporter-og-avhandlingen/HellekjaerAvhandling%5B1%5D.pdf>

This work is licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0).

To view a copy of this license, visit <https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode>.

INTRODUCTION

The role of English in Norwegian higher education in the 1990s and early 2000s was characterized by three overall factors. First, since Norway is a small language community with a limited book market it has long been necessary to put English texts and textbooks on university reading lists (Dahl, 1998; Hatlevik & Norgård, 2001). Second, an increased focus on internationalization and student mobility (Norwegian Ministry of Education and Research, 2002) led to a growing number of English-Medium Instruction (EMI) courses and programs for Norwegian and international university students. Third, Norwegian higher education institutions relied on upper secondary school English as a Foreign Language (EFL) instruction to provide the skills needed for higher education. The importance of English for Academic Purposes (EAP) reading proficiency and the lack of research in this area led to the doctoral study presented here. This doctoral study aimed to examine: *Whether, and to what extent, Norwegian students in higher education have problems reading the English texts and textbooks on their reading lists.*

After a brief look at what characterized Norwegian reading research of the time, selected data from three quantitative sub-studies of EAP reading proficiency are presented. Two are from higher education, and one from the General Studies (GS) branch of upper secondary school that qualifies for higher education. The focus is put on presenting reading scores, and their covariance with selected variables such as unfamiliar vocabulary, EFL courses, extracurricular reading, and university level study experience.

THEORY: DEFINING READING IN A FOREIGN LANGUAGE

In the present thesis, reading or reading to learn was understood as the active creation of meaning in an interactive process between information in the text on the one hand, and the knowledge of the reader on the other (Bråten, 1997). It drew upon a model of reading as an interactive process involving lower-level (bottom-up) processing and higher-level (top-down) processing (e.g. Alderson 2000; Bråten, 1997; Grabe, 1999; Koda, 2005). The core, bottom-up process involves recognizing the written words in the text along with relevant grammatical information. This hinges upon rapid and automatic word recognition, which provides the basis for higher-level processing, i.e. the creation of meaning in an interactive process between the information in the text being read, and the reader's knowledge of language, content, processing skills and strategy use. For academic reading it also requires the ability "to integrate text and background information

appropriately and efficiently” (Grabe & Stoller, 2002, p. 28). This involves using background knowledge, that is content knowledge on the one hand, and knowledge of the language and text types on the other. It may also involve other cognitive processes, meta-cognitive monitoring in particular, and strategy use.

Reading in a second language (L2) was considered to be similar to in first language (L1), but subject to “a number of additional constraints on reading and its development” (Grabe, 1999, p. 11). The most important of these was vocabulary, since fluent reading depends on rapid and automatic word recognition that leaves as much as possible of the limited working memory free for higher level processing (Rayner & Pollatsek, 1989). This means that having to struggle with unfamiliar words can slow down or even disrupt the reading process, as will excessive dictionary use. There was at the time an on-going debate about what an adequate vocabulary for academic reading meant in practice, and the ability to understand 95% of the words in academic texts was considered a minimum for the fluent reading of academic English as the L2. Such a level would require extensive and systematic vocabulary instruction, along with years of reading practice (Coady, 1997; Day & Bamford, 1998; Grabe, 1999; Paribakht & Wesche, 1997), and was by no means reflected in either syllabus goals or requirements of the then current EFL syllabi.

Another of the key differences between L1 and L2 reading discussed at the time was the claim that readers approached the latter with a dual-language system (Koda, 2005). For many such dual processing could even be an advantage – in the Interdependence Hypothesis, Cummins (2000) argued that “academic proficiency transfers across languages such that students who have developed literacy in their first language will tend to make stronger progress in acquiring literacy in their second language” (p. 173). Bernhardt (2005) proposed a compensatory model of L2 reading that attempted to quantify the importance of L1 literacy, L2 language knowledge and what she terms “unexplained variance”. The latter category comprises content, comprehension strategies, interest and motivation, etc., and Bernhardt (2005) argued that a weakness in one area might be compensated for by knowledge from another (see also Stanovich, 1980). However, poor L2 proficiency could hinder the transfer of skills and strategies to the L2, which was known as the threshold hypothesis (Alderson, 2000; Bernhardt & Kamil, 1995; Carrell, 1991; Laufer, 1997).

Against this background, the reading construct in this thesis draws heavily on Grabe’s (1999, p. 34) description of what is needed in order to be a good reader in either the L1 or L2, researched-based criteria upon which tests such as the Inter-

national English Language Testing System (IELTS) admission tests were operationalized. Grabe's (1999) criteria are listed below:

1. Fluent and automatic word recognition skills, ability to recognize word parts (affixes, word stems, common letter combinations);
2. a large recognition vocabulary;
3. ability to recognize common word combinations (collocations);
4. a reasonably rapid reading rate;
5. knowledge of how the world works (and the L2 culture);
6. ability to recognize anaphoric linkages and lexical linkages;
7. ability to recognize syntactic structures and parts of speech information automatically;
8. ability to recognize text organization and text-structure-signaling;
9. ability to use reading strategies in combination as strategic readers [. . .];
10. ability to concentrate on reading extended texts;
11. ability to use reading to learn new information;
12. ability to determine main ideas of a text;
13. ability to extract and use information, synthesize information, and infer information; and
14. ability to read critically and evaluate text information.

Again, few of these criteria were reflected in the then-current upper secondary school curriculum.

READING RESEARCH IN NORWAY – A BRIEF REVIEW

During the late 90s there was mounting criticism of Norwegian L1 reading instruction for failing to go beyond the decoding stage to the teaching of “reading to learn”, with negative effects on students’ reading practices with regard to strategy use and the ability to adjust reading to reading purpose (Bråten, 1997; Bråten & Olaussen, 1997, 1999). In a study of teacher education students, Fjeldbraaten (1999) found that these used the counterproductive, reading-for-detail approach to their L1 textbooks they had learnt in secondary education. She also found that even a one-year reading strategy course did not change this – when under stress, the students promptly reverted to the reading-for-detail approach they were used to. Despite the growing criticism, there was little debate about reading instruction before the 2001 OECD “Pisa-shock” study showed that Norwegian 15-year-olds were not proficient L1 readers (Lie, Kjærnsli, Roe, & Turmoe, 2001). It was first

and foremost the high, in-class variation in reading scores that led to serious debate and the Knowledge Promotion curriculum reform, which continued when the next tests showed no improvement (Kjærnsli, Lie, Olsen, Roe, & Turmoe, 2004).

Given the transfer from L1 to L2, there was good reason to expect that the weaknesses found in L1 reading proficiency would be present in the L2 – English – as well. There had been only a few relevant but small-scale, studies. The first found that that current EFL teaching was causing upper secondary school students to use a counterproductive, slow and detail-focused approach to the reading of English (Hellekjær, 1992). Later this became a serious problem for the early implementation of Content and Language Integrated (CLIL) instruction in Norwegian upper secondary school, although the students' reading skills improved rapidly when given relevant reading instruction (Hellekjær, 1996). A later survey of 145 first-year university college respondents confirmed that the same slow, careful, word-by-word reading of English texts that was typical of upper secondary school students persisted in higher education (Hellekjær, 1998). Finally, there was the 2004 large-scale European comparative study of English proficiency comprising representative samples of Norwegian, Danish, French, Finnish, Dutch, German and Spanish 16-year-olds (Bonnet, 2004). While Norwegian students did well, closer analysis showed that the Norwegian scores had the highest standard deviations, in particular for reading comprehension. Ibsen (2004, p. 35) concluded that this high, in-class variation for English, mirrored the findings for L1 reading in the recent OECD/PISA survey (see Lie, Kjærnsli, Roe, & Turmoe, 2001).

This brings us to the present study, where, in the light of the importance of EAP reading proficiency for higher education, the available research on L1 and L2 reading clearly indicated the need to further investigate this issue. This led to this doctoral study, where the main research aim was: *Whether, and to what extent, Norwegian students in higher education have problems reading the English texts and textbooks on their reading lists*. As mentioned, it compared university with senior upper secondary school student reading scores.

METHOD

The present quantitative study, comprising three sub-studies with surveys and tests of three different samples, used a quasi-experimental, one-group, post-test research design (see Shadish, Cook, & Campbell, 2002, pp. 106–107). See table 9.1 for a more detailed overview.

MATERIAL

Three questionnaires were used to collect data. All included items tapping into the dependent variables English and Norwegian reading proficiency, into the independent variables expected to co-vary with reading comprehension such as English reading habits, media use, and the handling of unfamiliar words when reading. They also included items providing information about student backgrounds, such as university students' study experience. Sub-study 1 used self-assessment items that could be merged into additive indices to measure reading proficiency; sub-study 2 combined these and an IELTS Academic English Reading Module, part of an English proficiency test used by UK and Australian universities for admission purposes, to validate the self-assessment items; and the IELTS Module was used for sub-study 3.

The use of self-assessment items to measure the reading proficiency of university students in sub-study 1 was to ensure that the survey could be answered quickly during the last ten minutes of a lecture. Research shows that self-assessment surveys can provide a reliable and valid picture of skills and/or levels of proficiency in low-stakes situations, such as in a survey (Bachman, 1990; Oscarson, 1997). Drawing on the construct definition of reading, I therefore developed a number of self-assessment items testing different aspects of the reading process. Items 40, 41 and 42 tapped into bottom-up processing, and 43, 44 and 45 into top-down processing when reading. To allow for comparisons there were identical items for Norwegian and English. Those for English are presented below:

40. How **quickly** do you **read English** texts on your reading lists? (Give only one answer)
41. Indicate on the scale from 1 to 7 **how many words you do not understand** in the **English** texts on your reading lists.
42. Indicate on the scale from 1 to 7 to what extent you find the **sentences in the English texts** difficult to understand.
43. Indicate on the scale from 1 to 7 to what extent you **find the English texts coherent** when reading.
44. Indicate on the scale from 1 to 7 to what extent the **information in the English texts is so densely presented** that it hinders your understanding of the contents.
45. Indicate on the scale from 1 to 7 to what extent you find **the contents of the English texts understandable**.

All the items used seven-point scales from 1 (highly difficult) to 7 (no difficulties). Following reliability analysis (Cronbach's alpha), these six items were merged

into additive indices and used as dependent variables – as measures of academic reading proficiency.

I was granted permission to use an Academic English Reading Module specimen test that IELTS said was identical in difficulty to comparable tests (UCLES, 2001a, 2001b). It comprised a Geology, a Business, and a Technical text, each about 900 words long (UCLES, 2001a). The tasks – in this case, 38 multiple choice items – require respondents to read strategically, to make use of previous knowledge, both of content and text type, to interpret and understand the texts, to draw upon the ability to make correct inferences needed for understanding, and to monitor and realign comprehension while reading. The one-hour time limit served to give an indirect measure of reading speed – the more slowly the respondent reads, the more unanswered items. IELTS uses conversion tables to score the items from Bands 1 to 9, Band 6 being the most common minimum requirement for admission to most UK and Australian universities, although some also require Band 7. The IELTS conversion table with the weightings of the different items was not available; instead I used the IELTS guideline in the accompanying booklet (UCLES, 2001a) to determine band levels. For this specimen text, IELTS set Band 6 to 23 of 38 points.

PROCEDURE AND SAMPLES

For the first university-level study (sub-study 1), I first had to identify undergraduate courses with English texts on their reading lists, which precluded random sampling. Either the lecturer or I handed out the questionnaires to students during the last ten minutes at the end of the lecture. This resulted in 578 undergraduate level respondents from three faculties at the University of Oslo.

Next, for sub-study 2, I contacted Computing, Business and Administration and Political Science students at Østfold University College, Geography and Social Anthropology students at the University of Science and Technology in Trondheim (NTNU), and with the help of an assistant, Biology students at the University of Bergen. The students were given 80 minutes to answer the questionnaire and complete the IELTS Academic Reading Module. Even though the students were paid for their time, it proved difficult to get volunteers, so I stopped after 53 students.

Finally, for sub-study 3, I contacted ten upper secondary schools, three because they had single-subject, sheltered Content and Language Integrated Learning (CLIL) classes. Seven schools agreed to participate. The teachers handed out and collected the questionnaires with the IELTS Academic Reading Module, and the senior GS branch students were given roughly 80 consecutive minutes to answer.

This resulted in a convenience/purposive sample of 217 students. An overview is provided in Table 9.1.

TABLE 9.1. Overview of the survey samples, time of surveying, type of survey, respondent affiliation, respondent numbers and instruments used to assess English reading proficiency.

Sub-study	Time of survey	Type of survey	Respondent affiliation	Number of respondents	Instruments used to assess reading proficiency
1	Spring 2001 and fall 2001	Survey of university level students' reading proficiency	University of Oslo, the Faculties of Education, Social Sciences, and Natural Sciences	578	Self-assessment items
2	Fall 2001	Validation test of university level students' self-assessment scores	Østfold University College and the Universities of Bergen and NTNU, Trondheim	53	Self-assessment items and an IELTS Academic Reading Module
3	Spring 2002	Survey of senior (third year) upper secondary level students' reading proficiency	Seven upper secondary schools from different parts of Norway. 39 students had CLIL instruction, 178 EFL only	217	An IELTS Academic Reading Module

Of the 578 respondents in sub-study 1, 159 (28%) were from the Faculty of Education, 266 (46%) from the Faculty of Mathematics and Natural Sciences, and 53 (26%) from the Faculty of Social Sciences. Of these, 45 (8%) did not have Norwegian as their L1. The majority of students were female, 427 (74%). The courses in questions had 1,125 registered students, giving a 51% reply rate, rising to 65% if the 894 completed examinations were used as a baseline.

Sub-study 2 comprised 53 respondents, 44 from Østfold University College, seven from the Norwegian University of Science and Technology (NTNU), and two from the University of Bergen. Thirty-one (58%) of the respondents were male and 22 (42%) female. Only one respondent (2%) did not have Norwegian as his or her L1. The reply rate could not be determined.

Sub-study 3 comprised senior upper secondary students, 217 in all. Thirty-nine of these had a single, sheltered CLIL subject from three different schools, either History or Physics taught in English. The remaining 178 had different EFL courses, 45 (34%) with the first-year Foundation course, 30 (17%) with English I, and 100 (56%) with English II courses. It was not possible to determine the reply rate.

DATA ANALYSES

The statistical analyses focused on mean scores, score and respondent distributions, and analysing co-variations between dependent and independent variables using correlation analysis (Pearson's r) and multiple regression (linear). I also carried out reliability analyses (Cronbach's alpha) of the self-assessment and IELTS items prior to merging these into additive indices for use as dependent variables.

RESULTS

This doctoral study aimed to examine: Whether, and to what extent, Norwegian students in higher education have problems reading the English texts and textbooks on their reading lists. The three studies showed that about 32% of the university level respondents had difficulties reading English texts and textbooks and that 66% of the upper secondary school GS students also had difficulties, scoring below Band 6 on the IELTS Academic Reading Module. Interestingly, 74% of the students with a CLIL course managed Band 6 or better. A more detailed analysis is presented in the following.

SUB-STUDY 1 – THE MAIN UNIVERSITY LEVEL STUDY

This university-level study used a set of self-assessment items to investigate reading proficiency in English and Norwegian. After a reliability analysis, these were merged into the additive indices *Enindex* and *Noindex* to provide measures of reading difficulty in the two languages. The Cronbach's alpha coefficients were high, $a = .84$ for *Noindex*, and $a = .94$ for *Enindex*. This made it possible to compare the scores for reading difficulty between the two languages, and next to examine reading difficulty in English. The mean score for *Enindex*, $\bar{x} = 4.6$ ($SD=1.1$), was clearly below the $\bar{x} = 5.7$ ($SD=0.73$) of *Noindex*, and with a markedly higher standard deviation than for *Enindex*. The score distribution of is presented in Figure 1. The scores to the left indicate higher levels of reading difficulty, on a scale from 1 (difficult to understand) to 7 (no difficulties).

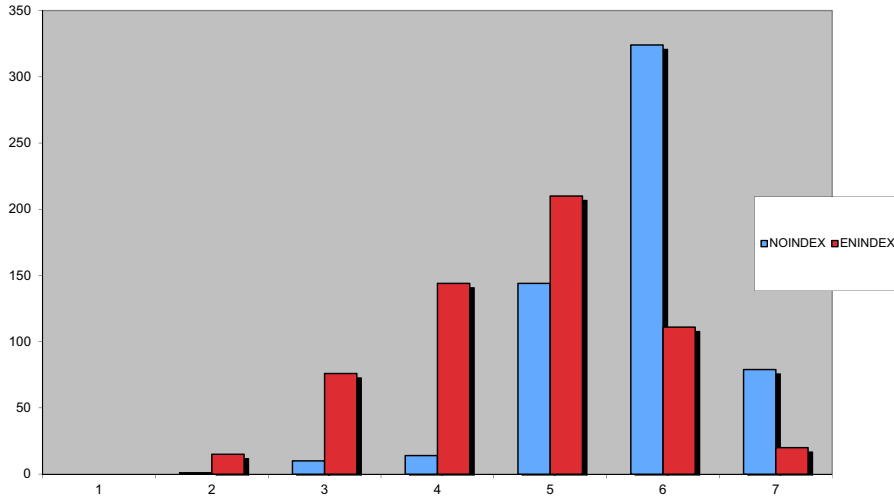


FIGURE 9.1. Distribution of self-reported Noindex and Enindex additive indices scores displaying reading proficiency scores. The scale is from 1 (difficult to understand) to 7 (no difficulties). For display purposes, values from 0 to 1.49 being counted as 1, from 1.5 to 2.49 as 2, etc.

As displayed, while the scores for *Noindex* are skewed to the right around a median value of 5.8, while those for *Enindex* are more evenly distributed with many well below the median value of 4.7 and with a markedly higher SD, clear indications of greater perceived reading difficulties for Norwegian. As I will return to in sub-study 2, a comparison of the IELTS and self-assessment scores in this study indicated that self-assessment scores of 4 or below also fell below IELTS Band 6. In other words, the 185 (32%) of 578 respondents that scored below 4, and who would most probably not achieve Band 6 on the IELTS Reading Module, are those experiencing moderate to serious difficulties reading English texts and textbooks. As for main sources of reading difficulty, I investigated this question through a more detailed comparison of the unmerged items used for the *Noindex* and *Enindex* indices. The comparison is presented in Table 9.2.

TABLE 9.2. Comparison of reading difficulties between English and Norwegian, mean scores and standard deviations (SD) for items 34 to 39 for Norwegian, and the equivalent items for English, 40 to 45. Scoring is on seven-point scales from 1 (lowest) score to 7 (highest).

Items	Norwegian mean scores and standard deviations	English mean scores and standard deviations
40. How quickly do you read the texts on your reading lists?	5.43 (SD= 1.2)	4.31 (SD= 1.4)
41. Indicate on the scale from 1 to 7 how many words you do not understand in the texts on your reading lists.	5.91 (SD= 0.8)	4.47 (SD= 1.1)
42. Indicate on the scale from 1 to 7 to which extent you find the sentences in the texts difficult to understand.	5.81 (SD= .1.0)	4.63 (SD= 1.2)
43. Indicate on the scale from 1 to 7 to which extent you find the texts coherent when reading.	5.83 (SD= 0.9)	4.73 (SD= 1.3)
44. Indicate on the scale from 1 to 7 to which degree information in the texts is so densely presented that it hinders your understanding of the contents.	5.42 (SD= 1.0)	4.58 (SD= 1.3)
45. Indicate on the scale from 1 to 7 to which extent you find the contents of the texts understandable.	5.79 (SD= 0.8)	4.88 (SD= 1.2)

As can be seen in Table 9.3, the fairly low scores for reading speed (Item 40) goes to show that this is a challenge in both languages. The next source of difficulty in English is unfamiliar vocabulary (Item 41), with complex sentences (Item 42) and with texts with densely presented information (Item 44) following close behind. Finally, the higher standard deviations for reading in English give evidence of greater variation in reading proficiency. To further investigate the problem of unfamiliar vocabulary in particular, a number of items asked how the respondents handled unfamiliar items of vocabulary when reading. How these correlate (Pearson's r) with reading proficiency, as measured by *Enindex*, is displayed in Table 9.3.

TABLE 9.3. Bivariate correlations between *Enindex* and ways of handling unfamiliar words when reading in the Study 1 survey. $n=527$ (51 missing answers).

Independent variables	r
Dictionary use	-.17*
Guess meaning of word using subject knowledge	.17*
Guess meaning of word using context	.27*
Ask lecturer	-.01
Ask other students	-.11*
I just keep on reading	.04
I give up reading altogether	-.50*

* $p < .01$

This overview shows that only the ways of handling unfamiliar words that avoided disrupting the reading process, such as guessing word meaning, had positive correlations. Those that disrupted the reading process, such as using a dictionary or asking fellow students, had negative correlations – meaning that the more frequent their use – the lower the *Enindex* score. Interestingly, the highest negative correlation, $r = -.50$, was for the item “I give up reading altogether” is a clear indication of serious reading difficulties that are most probably due to poor English proficiency.

FURTHER ANALYSIS

Given the two indices tapping into reading in the L1 and L2, this gave the opportunity to examine the Interdependence Hypothesis – the expectation that if a respondent read well in Norwegian as the L1, they would also read well in English as the L2. This was supported by a correlation (Pearson’s r) between *Enindex* and *Noindex* of $r=0.43$ ($p>.01$, $N=528$), with the caveat that there is no clear indication of direction. Next, cross tabulation revealed a considerable number of respondents who fell below the Linguistic Threshold Level due to poor English skills, with high scores for *Noindex* but very low for *Enindex*.

There were also a number of background questions: about study experience, advanced English courses taken in upper secondary school, and about extracurricular reading and exposure to English. When correlated with *Enindex*, the item for completing the third advanced upper secondary school English course gave only

a low correlation of $r = 0.13$ ($p = .01$, $N = 572$). With regard to study experience, where reading difficulty might be expected to decrease with time, no significant correlation between study experience and reading proficiency was found. Instead, extracurricular reading gave a positive result. Analysis showed that many students read extensively, about half having read 16 novels or more, and of these 108 (18%) having read 51 or even more. This gave a fairly high and positive correlation with *Enindex* of $r = .47$ ($p = 0.01$, $N = 573$), while for magazine reading the correlation was $r = 0.43$ ($p = 0.01$, $N = 574$), and for internet reading $r = 0.4$ ($p = 0.01$, $N = 576$). Furthermore, multiple regression analyses with these three variables (for reading novels, magazines and periodicals, and on the Internet), with *Enindex* as the dependent variable, gave an explained variance of $R^2 = .29$, accounting for 29% of the variation in the reading scores.

SUB-STUDY 2 – THE VALIDATION STUDY

The main aim of this survey was to validate the self-assessment indices used to measure English reading proficiency in sub-study 1, and if possible, benchmark these against the IELTS Academic Reading Module scores. The questionnaire included the same self-assessment items for reading in English and Norwegian as in sub-study 1, and the same items on reading habits, the handling of unfamiliar vocabulary, upper secondary English courses, and study experience. As in sub-study 1, the self-assessment items for Norwegian and English could be merged into additive indices. For this sample the Cronbach's alpha coefficients for *Enindex* was a high $\alpha = .9$, and for *Noindex* a moderate to high $\alpha = .7$. With a high Cronbach's alpha coefficient of $\alpha = .9$, the 38 IELTS scores could also be merged into the additive index *Alltext*.

The IELTS test results

The distribution of the IELTS Academic Reading Module scores (*Alltext*) for the 53 respondents are displayed in Figure 9.2.

As can be seen, 44 of 53 (83%) respondents scored Band 6 (24 points) or better, with 40% achieving 34 of the 38 points or more, a clear indication of a ceiling effect, and this should be considered an acceptable outcome. However, these were volunteer respondents, as likely as not fairly confident about their English skills – one respondent mentioned that many who did not feel proficient did not volunteer. In addition, the 25 (48%) Computer Science respondents from Østfold University College were at the outset above-average students because of heavy competition

for admittance to their study program. It is therefore probable this particular sample had a higher level of proficiency than would be the case with a representative sample of Norwegian students.

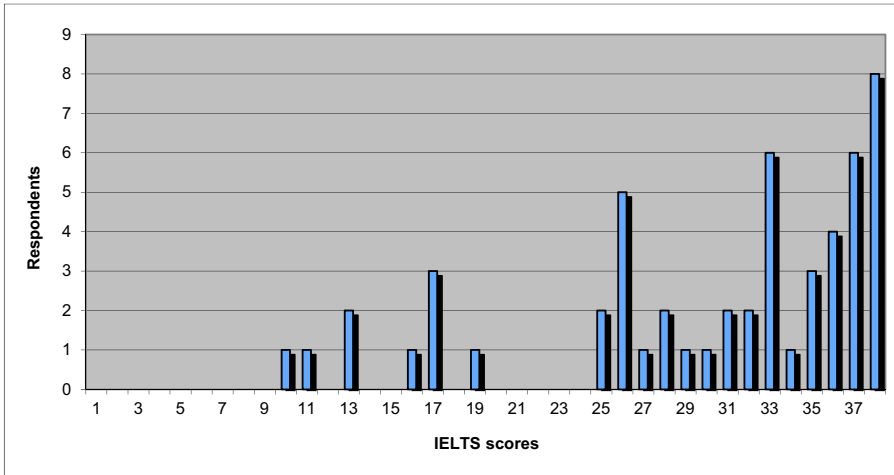


FIGURE 9.2. IELTS Academic Reading Module scores. The maximum score is 38, the mean, $\bar{x} = 30$, and standard deviation, $SD= 8.0$, $N= 53$.

However, the main goal of this study was to validate the self-assessment indices used to measure reading proficiency in sub-study 1. I therefore correlated the index for the English self-assessment items, *Enindex* with the IELTS index *Alltext*, and (Pearson’s r) correlation between these was $r= 0.72$ ($p<.01$, $N= 53$). This fairly high correlation gives reason to claim that the self-assessment index *Enindex* gives a useful and reasonably valid picture of student reading difficulties in English.

Next, in an attempt to benchmark the self-assessment index for English against the IELTS scores, further analysis showed that the (relatively few) respondents who scored below Band 6 on the IELTS test also scored below 4 on *Enindex*. This gives reason to argue that the 32% with *Enindex* scores below 4 in sub-study 1 also fall below the IELTS Band 6 level.

Finally, further analyses gave largely the same results as in sub-study 1. For study experience, correlating the number of completed ECTS credits (29) with *Enindex*, *Noindex*, and IELTS scores gave no meaningful or significant correlations. Once again, only extracurricular reading gave positive correlations; when correlated against IELTS scores (*Alltext*), book reading gave $r=0.58$; periodical reading $r=0.38$; and Internet reading $r=0.47$, all statistically significant ($p=0.01$,

$N=53$). Multiple regression analysis with IELTS scores as the dependent variable and the three items for book, periodical and internet reading as independent variables gave an explained variance of $R^2 = 0.40$, clearly higher than for sub-study 1. With regard to the handling of unfamiliar vocabulary, the pattern largely reflected that of sub-study 1, but with non-significant correlations, most probably due to the small sample.

Sub-study 3–The upper secondary school study

Having found that about a third of the university students had difficulties reading English, the next step was to examine the proficiency of senior GS upper secondary school students. As mentioned, this sample comprised 217 students, of which 39 had a single CLIL (the CLIL sub-sample) while the remainder, 178 respondents had EFL instruction only (the EFL sub-sample). The questionnaire had roughly the same items as in the university surveys and the same IELTS Academic Reading Module used for sub-study 2. The IELTS items were again combined into an additive index (*IELTS*) to serve as a dependent variable. For the EFL sub-sample, reliability according to Cronbach's alpha was high, $\mu = .95$; as for the CLIL sub-sample, $\mu = .92$. An overview of the IELTS scores for the different sub-samples, and according to English course, is presented in Table 9.4.

TABLE 9.4. An overview of the IELTS mean scores with standard deviations, and of the percentages who achieved 24 points/Band 6 or better for the EFL and CLIL sub-samples, and for the EFL sub-sample according to completed English courses. $N=217$

Sample and subsamples	Mean IELTS scores (\bar{x})	SD	N	Percentage who achieve 24 points/ Band 6 or better
EFL sub-sample	21	9.03	178	33 %
Foundation Course	22	9.36	45	47 %
Advanced English I	19	9.17	30	33 %
Advanced English II	21	8.92	99	29 %
CLIL sub-sample	28	7.9	39	74%
Total			217	

As can be seen from this overview, only one third (33%) of the EFL sub-sample achieved the Band 6 level or better. Closer examination of their IELTS scores for

the EFL sub-sample revealed a consistent pattern where respondents apparently read and worked quite slowly, for the most part answering correctly, but leaving many items unanswered.

Furthermore, in the EFL sub-sample completing the Advanced English I and II Courses, each representing one or two extra years of English teaching with five lessons per week, did not give higher mean reading scores ($\bar{x} = 19$ and $\bar{x} = 21$ respectively). In fact, the scores for this group are somewhat lower than those with the Foundation Course only ($\bar{x} = 22$), perhaps due to many of the better students opting for the Natural Science subjects instead of English. In comparison, the CLIL subsample, which had had a single CLIL subject taught in English, scored far higher, with 74% achieving 24 points/Band 6 or better ($\bar{x} = 28$). In Figure 3 the EFL sample scores are compared with those of the CLIL sub-sample.

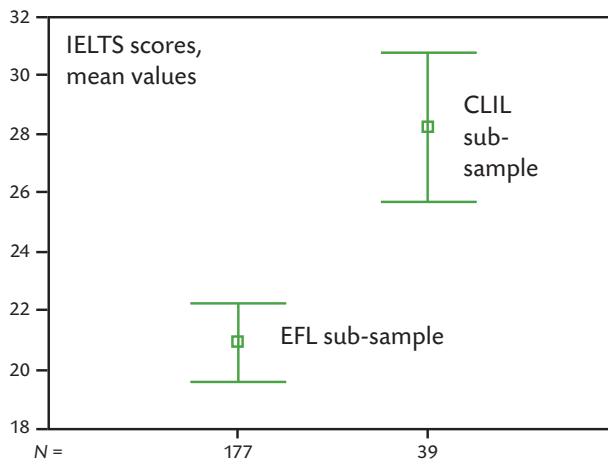


FIGURE 9.3. Confidence intervals for the IELTS scores of the EFL sub-sample with 178 respondents, and the CLIL sub-sample with 39. The interval between the group means is statistically significant at the 95% level of certainty. Maximum IELTS score is 38.

As can be seen, the confidence intervals in Figure 9.3 also show that the interval between the IELTS mean scores for EFL and CLIL sub-samples is about seven points, and this is statistically significant at the 95% level.

Extracurricular reading and unfamiliar vocabulary

With regard to reading habits, further analysis of the EFL sub-sample revealed that the great majority had read little beyond the minimum requirements in the current R94 curriculum, which can probably explain the low correlations for reading

books, $r = 0.21$ ($p = .01$, $n = 177$), periodicals $r = 0.15$ (not significant), and on the Internet $r = 0.21$ ($p = .01$, $n = 177$). With regard to how the EFL students coped with unfamiliar words when reading, the correlations between these items and the IELTS scores are displayed in Table 9.5.

TABLE 9.5. Ways of coping with unfamiliar words correlated with the IELTS test scores in the EFL sample. $N = 177$.

Independent variables	r
Dictionary use	.00
Guess meaning of word using subject knowledge	-.03
Guess meaning of word using context	.10**
Ask teacher	-.20*
Ask parents	-.24*
Ask other students	-.31*
I just keep on reading	-.11**
I give up reading	-.28*

* $p < .01$ ** $p < .05$

As can be seen, the correlations for guessing the meaning of unknown words and otherwise not disrupting the reading process are quite low, lower than for the university students in sub-study 1. There were also higher negative correlations for interrupting the reading process to consult their teachers, peers or parents, but interestingly, not for dictionary use.

DISCUSSION: CONTRIBUTIONS TO THE ENGLISH DIDACTICS FIELD

I would contend that the doctoral study presented in this chapter presented a serious, empirically based challenge to the Norwegian assumptions about the quality of Norwegian EFL instruction in general, and for reading proficiency in particular.

EMPIRICAL CONTRIBUTIONS

One of the main empirical contributions of this study was that upper secondary school English instruction failed to prepare all-too-many students for the reading

of English texts in higher education – as indicated by as many as two thirds of the senior upper secondary school students not achieving the minimum Band 6 IELTS score required for admission to most UK and Australian universities.

The next contribution was to show that that these difficulties persisted in higher education, with both sub-studies 1 and 2 indicating that about one third of the surveyed university level students experienced difficulties. Interestingly, further analysis found no correlation between study experience and reading scores, which means that one cannot take it for granted that student reading proficiency improves with experience, and cannot explain the performance gap between education levels. Instead, either student attrition and/or selection factors stand forth as probable explanations. It may be that many upper secondary school students who are poor readers either do not go on to higher education, choose studies where English textbooks are not used, or drop out of their first university-level courses. Third, the many respondents who struggled with unfamiliar words to the point where they gave up reading altogether clearly show the need to pay greater attention to vocabulary development.

Finally, to return to upper secondary school EFL instruction, the IELTS scores of the EFL sample – one-third achieving Band 6, and that the advanced elective English courses did not contribute to higher scores while a single, sheltered CLIL courses clearly did, clearly showed the need to improve reading instruction. Closer examination showed that the low scoring students at both levels tended to read and work very slowly, often answering IELTS items correctly but leaving many or most unanswered due to the lack of time. Correlation analysis also indicated that many gave up reading due to unfamiliar words, probably either due to inadequate vocabulary knowledge, and/or the inability to handle unfamiliar words without unduly interrupting the reading process. However, while English proficiency, and vocabulary in particular, can in part explain these challenges, another explanation might be weaknesses in the respondents' L1 reading proficiency, given the arguments for transfer from L1 to the L2. One indication was that the 2001 OECD PISA study revealed that many Norwegian 15-year-olds were poor readers in their L1. Another was the contemporary criticism of lack of emphasis put on teaching Nordic students *reading to learn* (Bråten, 2007; Bråten & Olausen, 1997, 1999). Furthermore, the IELTS test data reflected those of Fjeldbraaten's (1999) study, which showed that university level students under pressure used counterproductive reading strategies in the L1, first and foremost the tendency to read carefully for detail. In other words, and paraphrasing Alderson (1984), the reading problems found in this thesis are most probably due to reading as well as language problems.

THEORETICAL CONTRIBUTIONS

One of the theoretical contributions of this study was its support of the Interdependence Hypothesis, given the positive correlation between L1 and L2 reading scores found in sub-study 1. Cross tabulating these same scores also gave support for the Threshold Hypothesis, since it revealed a considerable number of respondents who were good readers in the L1 but among the poorest in the L2. Among other issues there were the clear and consistent positive correlations between reading of books and periodicals and reading scores, a clear argument for the importance of extensive reading. Finally, the lack of any positive correlations between study experience and reading scores at the university level indicates that little incidental acquisition is taking place, unless, as Fjeldbraaten (1999) shows, that this is a reading problem, not a language problem. This in turn clearly indicates the need for consistent, long-term reading strategy instruction in the L1 and L2.

METHODOLOGICAL CONTRIBUTIONS

With regard to the methodological contributions, the main contribution of this doctoral study was the development of, and subsequent validation of, using the IELTS Academic Reading Module, a set of self-assessment items that can quickly and easily be used to measure EAP reading difficulties, in part by benchmarking this against reading in the L1. The other methodological contribution is that it clearly demonstrated the utility of and need to use validated language tests to assess student language proficiency – in this case, reading.

IMPLICATIONS FOR TEACHING ENGLISH AS L2

A key implication of this study for EFL instruction was the need to focus on the teaching of reading strategies, and to teach students to accept uncertainty when they encounter unfamiliar words and guess from context or using their previous knowledge in order to avoid disrupting the reading process. Another was the need for a far more systematic approach to improving students' vocabularies to prepare them for higher education. Improving reading proficiency as well as developing vocabulary entails the need to read far more extensively as part of EFL instruction, and requires going beyond the textbooks to the reading of a variety of longer texts (novels, biographies, documentaries), including technical and academic reading. A crucial part of this approach would be to allow students to choose challenging texts on topics that they are interested in. Last, and by no means least, the study

showed the need for an overhaul of the content and teaching of the Advanced English elective courses.

RECENT DEVELOPMENTS AND SUGGESTIONS FOR FUTURE RESEARCH

This thesis was defended at the same time that a new curriculum (*Knowledge Promotion*) was under development, intended to address the problems found by this doctoral study and the OECD PISA “shock” study (Lie et al., 2001). A crucial change made in the new curriculum was to make reading one of five basic skills that were to be taught across the curriculum in the L1 as well as L2. The new curriculum also designated clear competency aims for reading to be achieved after grades 4, 7, 10, and after the Vg1, Vg2, and Vg3 courses. The need to teach reading strategy use was also made clear, and the curriculum was quite explicit about the students’ learning to adjust their “reading according to reading purpose”. National reading tests for English after grades 5, 8, and the Vg1 level were also developed to support changes in teaching². Finally, the new English curriculum acknowledged far more clearly than did its predecessor R94 the role of English instruction in preparing for higher education.

A number of Norwegian studies have evaluated English reading instruction since 2005, including Master’s theses. At the upper secondary school level, one of the first follow-up studies was a Master’s thesis where Faye-Schøll (2009) found that the teachers interviewed seemed quite unaware of the Knowledge Promotion English curriculum requirements regarding the teaching of reading and strategy use. However, a later doctoral study, Brevik (2015), found clear improvements in reading strategy instruction. She also found that teachers needed help to articulate their tacit knowledge of reading strategy use.

Another follow-up study by Hellekjær and Hopfenbeck (2012) investigated whether upper secondary school GS student reading proficiency had improved following the 2006 curriculum. It used the same IELTS Academic Reading Module and many of the same questionnaire items from sub-study 3 that is reported in this chapter. They found a clear improvement among upper secondary GS students with only EFL instruction, with 57% of the students achieved 24 points/Band 6 compared to the previous 33%. However, despite the new curriculum, they also found that completing the Vg2 and Vg3 elective English courses still did not give improved reading scores, and since these students had the same Vg1 English grades, this effectively countered the common assumption that the poor results

2. The Vg1 tests were so-called mapping tests from 2010–2016, and have since 2017 been replaced by tests designed to support learning.

were due to negative selection. While single-subject, sheltered CLIL classes still gave better results, this varied according to the extent of English use for instruction. Hellekjær and Hopfenbeck (2012) found it difficult to explain the generally improved IELTS scores, first and foremost because they found little increase in student reading. They therefore speculated about the improvement observed being due to increased internet reading, to increased extracurricular exposure to English (e.g., Brevik & Hellekjær, 2018), due to the mandated focus on the teaching of reading strategies at the lower secondary level, or due to greater student familiarity with multiple choice tests due to the introduction of national examinations.

A number of newer studies have highlighted the growing importance of extracurricular English as an explanation for many students' often impressive English skills (Brevik, 2016; Brevik & Hellekjær 2018; Garvoll, 2017). Still, this is an issue in urgent need of further research, in particular to explain the poor outcomes from the Vg2 and Vg3 elective English courses.

In higher education, a number of recent studies have found that many students still struggle with English. With regard to EMI lecture comprehension, Hellekjær (2010) found that as many as 40% of the students had comprehension difficulties. With regard to Academic English reading, a Master's study, Arnsby (2013), used the self-assessment items developed in sub-study 1 in combination with student interviews to investigate university students' EAP reading. She found that the students' reading proficiency had improved only marginally in comparison to sub-study 1. She concluded that further improvements in upper secondary school reading instruction are still needed. Another Master's thesis is by Busby (2015), in which she compared the reading speed, comprehension and vocabulary knowledge of Norwegian students with native English-speaking students, finding that Norwegian students were more likely to have a native speaker-like proficiency in general-language English proficiency than in academic language English, particularly with regard to vocabulary comprehension. She concluded that students needed further help in developing their Academic English proficiency. In her doctoral study, Busby (on-going) continues work on this topic, using a combination of reading and vocabulary tests to investigate Norwegian university students' Academic English proficiency. In a recent article comparing student metacognitive strategy use in the L1 and L2, she found little difference between languages, although she concluded that for the L2 "students may benefit from additional training in the use of higher-level reading strategies to improve their comprehension of L2 academic texts" (Busby, 2018, p. 1).

While some issues have been addressed, further research is still called for. The perhaps most urgent project would be to investigate possible beginner student

attrition in higher education – to find out whether many of the poorer readers drop out of their beginner courses either due to L1 or L2 reading difficulties. Another would be a follow-up of Hellekjær and Hopfenbeck’s (2012) study. By surveying a new sample of upper secondary students with the same IELTS Academic Reading Module that was used in this thesis and in the 2012 study, it would be possible to see whether there has been any improvement. A comparable study of university students would of course also be useful, as would investigating why the advanced elective English courses do not improve reading proficiency. Yet another, and quite ambitious project would be an intervention study where lower or upper secondary students were to read a larger number of longer texts than usual (novels or documentaries), and their development monitored with pre- and post-tests. This would also develop our knowledge about how best to implement additional reading in the English classroom.

REFERENCES

- Alderson, C. J. (2000). *Assessing Reading*. Cambridge: Cambridge University Press.
- Arnsby, E. S. (2013). *How do Norwegian beginner students experience the reading of English course material at university? A mixed-methods study*. (MA thesis), University of Oslo, Oslo.
- Bachman, L. (1990). *Fundamental Considerations in Language Testing* (2nd ed.). Oxford: Oxford University Press.
- Bernhardt, E. (2005). Progress and procrastination in second language reading. *Annual Review of Applied Linguistics*, 25, 133–155.
- Bonnet, G. (Ed.) (2004). *The assessment of pupil's skills in English in eight European countries 2002: The European network of policy makers for the evaluation of educational systems*. <http://cisad.adc.education.fr/revu/>
- Bråten, I. (1997). Leseforståelse. *Nordisk Pedagogikk*, 17(2), 95–110.
- Bråten, I., & Olaussen, B. S. (1997). *Strategisk læring hos norske høyskolestudenter: en foreløpig rapport*. Oslo: Høgskolen i Oslo.
- Bråten, I., & Olaussen, B. S. (1999). *Strategisk læring: teori og pedagogisk anvendelse*. Oslo: Cappelen akademisk forlag.
- Brevik, L. M., & Hellekjær, G. O. (2018). Outliers: Upper secondary school students who read better in the L2 than in L1. *International Journal of Educational Research*, 89, 80–91. doi: [10.1016/j.ijer.2017.10.001](https://doi.org/10.1016/j.ijer.2017.10.001)
- Brevik, L. M., Olsen, R. V., & Hellekjær, G. O. (2016). The Complexity of Second Language Reading: Investigating the L1-L2 Relationship. *Reading in a Foreign Language*, 28(2), 161–182.
- Busby, N. L. (2018). Comparing first and second language reading: the use of metacognitive strategies among Norwegian students. *Acta Didactica Norge*, 12(2), 1–26. doi: <http://dx.doi.org/10.5617/adno.5579>

- Dahl, A., & Busby, N. (2015). *Too cool for school? Sources of English language acquisition, attitudes, and academic reading ability among Norwegian university students*. (MA thesis), NTNU, Trondheim.
- Dahl, J. (1998). *Pensumlitteratur i høyere utdanning – en undersøkelse av forholdet mellom norsk og engelskspråklig pensumlitteratur i fire grunnfag* (14). Oslo: Norsk institutt for studier av forskning og utdanning.
- Garvoll, K. K. (2017). *The Gamer, the Surfer and the Social Media Consumer Vocational students' English use in and out of school*. (MA thesis), Oslo: University of Oslo.
- Grabe, W. (1988). Reassessing the word interactive. In P. L. Carrell, J. Devine, & D. E. Eskey (Eds.), *Interactive Approaches to Second Language Reading* (pp. 56–70). Cambridge: Cambridge University press.
- Grabe, W. (1999). Developments in reading research and their implications for computer adaptive-reading assessment. In M. Chalhoub-Deville (Ed.), *Issues in computer-adaptive testing of reading proficiency* (Vol. 10, pp. 11–47). Cambridge: Cambridge University Press.
- Grabe, W., & Stoller, F. L. (2002). *Teaching and Researching Reading*. London: Longman.
- Fjeldbraaten, A.-L. (1999). Undervisning i lærings- og studiestrategier i sammenheng med allmennlærerutdanningens pedagogikkundervisning. In I. Bråten & B. S. Olaussen (Eds.), *Strategisk læring* (pp. 122–138). Oslo: Cappelen.
- Hatlevik, I. K. R., & Norgård, J. D. (2001). *Myter og fakta om språk* (5). Oslo: NIFU. Norsk institutt for studier av forskning og utdanning.
- Hellekjær, G. O. (1992). Engelskundervisning og leseforståelse. *Skoleforum*, 2(11/12), 64–67.
- Hellekjær, G. O. (1996). Easy does it: Introducing Pupils to Bilingual Instruction. *Språk og Språkundervisning*, 3, 9–14.
- Hellekjær, G. O. (1998). First Encounters with ESP Textbooks in Higher Education: A Pilot Survey. In L. Lundquist, H. Picht, & J. Qvistgaard (Eds.), *The 11th European Symposium on Language for Special Purposes* (Vol. 1, pp. 886–892). Copenhagen: Copenhagen Business School.
- Hellekjær, G. O., & Hopfenbeck, T. N. (2012). Lesing. In B. W. Svenhard (Ed.), *CLIL: Kombinert fag- og engelskopplæring i videregående skole* (Vol. 28, pp. 84–124). Halden: Fremmedspråksenteret.
- Ibsen, E. B. (2004). *Engelsk i Europa – 2002* (Vol. 2). Oslo: University of Oslo.
- Koda, K. (2005). *Insights into Second Language Reading*. Cambridge: Cambridge University Press.
- Norwegian Ministry of Education and Research [KD]. (1994). *Læreplan for videregående opplæring, Engelsk Studieretningsfag i studieretning for allmenne, økonomiske og administrative fag*. (R94). Oslo: Author.
- Norwegian Ministry of Education and Research [KD]. (2006). *Læreplanverket for Kunnskapsløftet*. Oslo: Author.
- Kjærnsli, M., Lie, S., Olsen, R. V., Roe, A., & Turmoe, A. (2004). *Rett spor eller ville veier?* Oslo: Universitetsforlaget.
- Lie, S., Kjærnsli, M., Roe, A., & Turmoe, A. (2001). *Godt rustet for framtida? Norske 15-åringers kompetanse i lesning og realfag i et internasjonalt perspektiv* (Vol. 4). Oslo: Dept. of Teacher Education and School Development/University of Oslo.

- Oscarson, M. (1997). Self-Assessment of Foreign and Second Language Proficiency. In C. Clapham & D. Corson (Eds.), *Language Testing and Assessment* (Vol. 7, pp. 175–187). Dordrecht/Boston/London: Kluwer Academic Publishers.
- Rayner, K., & Pollatsek, A. (1989). *The Psychology of Reading*. Englewood Cliffs: Prentice-Hall.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston: Houghton Mifflin.
- Stanovich, K. E. (1980). Towards an interactive compensatory model of individual differences in the development of reading fluency. *Reading Research Quarterly*, 16, 32–71.
- UCLES. (2001a). *IELTS Specimen Materials Handbook*. Cambridge: UCLES, The British Council, IPD Education Australia.
- UCLES. (2001b). *IELTS Specimen Materials Test Texts*. Cambridge: UCLES, The British Council, IPD Education Australia.