

Jakonen, T., J. Jalkanen, T. Paakkinen & M. Suni (toim.) 2015. Kielen oppimisen virtauksia. Flows of language learning. AFinLAn vuosikirja 2015. Suomen soveltavan kielitieteen yhdistyksen julkaisuja n:o 73. Jyväskylä. s. 208–225.

Pirjo Pollari

University of Jyväskylä

Can a cheat sheet in an EFL test engage and empower students?

Although occasionally used in language classrooms, cheat-sheet tests have not been explored in foreign or second language education research. This study experimented with cheat-sheet tests in the teaching of EFL in a Finnish upper secondary school. The participants, 101 students, could make a cheat sheet for the grammar part of their English test. A total of 92 students prepared the cheat sheet, nine did not. Students' cheat sheets, test results and comments constituted the data for this study, analysed both qualitatively and quantitatively. The existence of the cheat sheet and its quality (thorough, good or limited) correlated with the grammar test results: students with a thorough cheat sheet scored slightly higher points on average than other groups. Even though the cheat sheet did not markedly improve their test results, the majority of students felt that it had improved their learning and studying. Some students also reported reduced test anxiety.

Keywords: student assessment, testing aids, engagement, empowerment

Asiasanat: oppilasarviointi, arviointimenetelmät, kokeet, voimaantuminen

1 Introduction

Over the past decades, schools, curricula, teaching methods as well as theories of learning have undergone great changes. There is very little research on student assessment in Finland, but student assessment in foreign (FL) or second language (L2) education still appears to be somewhat test-based and limited in scope (Hildén & Härmälä 2015; Tarnanen & Huhta 2011). For instance, despite the requirements of the Finnish national curricula, self- and peer-assessment do not seem to play a significant role in FL/L2 assessment (Tarnanen & Huhta 2011). Yet, teaching and learning cannot really be reformed if assessment methods do not change as well. New avenues should therefore be explored in FL student assessment, both in research and in practice.

This study has a dual aim: it is both a teaching experiment exploring cheat sheets in an English test and a contribution to research on foreign language assessment. After a brief look at the theoretical background, I will introduce the experiment, its participants and the methods used. Then I will present the findings, based primarily on qualitative data. In addition, I will examine quantitatively if the cheat sheets had any measurable impact on students' test results. Finally, I will discuss the findings, their limitations and possible implications.

2 Theoretical background

2.1 The power of assessment

Research in FL/L2 education shows that testing has a significant yet quite complex washback effect (Cheng, Watanabe & Curtis 2004; Hughes 1989; Rea-Dickins & Scott 2007). Although the washback effect is not negative per se, evidence about how tests – high-stakes tests in particular – narrow the curricula into 'teaching to the test' abounds (Rea-Dickins & Scott 2007; Volante 2004). Also students, wishing to succeed, want to study for the test itself, which in turn influences their learning strategies. Many tests still focus on memory and accurate knowledge retention instead of high-order learning and thinking skills such as problem-solving or critical thinking (e.g. Atjonen 2007; Pickford & Brown 2006). So, students often try to memorise the information they think will be tested. This easily leads to superficial rote learning and real conceptual understanding, deep learning, takes a back seat (e.g. Harlen 2012; Volante 2004). Ultimately, passing the exam becomes far more important than learning itself (Harlen 2012).

Furthermore, test anxiety, which is rather common among female students, can weaken memory and knowledge retention and, thus, many students cannot show all that they actually know in test situations (e.g. Hembree 1988). Underperforming in the test can affect their motivation, self-efficacy and self-esteem as learners (Harlen & Deakin Crick 2003). Accordingly, several studies have shown that test anxiety, which is also closely related to foreign language anxiety (e.g. Horwitz 2001, 2010), may affect not only students' test results but also their FL/L2 learning processes, proficiency and motivation (e.g. Aydin 2009; Cheng, Klinger, Fox, Doe, Jin & Wu 2014; Liu & Huang 2011).

Finland has only one national high-stakes examination, the Matriculation Examination. However, the effects of testing on students and their learning are not only limited to high-stakes exams (Harlen & Deakin Crick 2003). Students may feel anxious and powerless in the face of any of the assessment situations that take place dozens of times throughout their school year (Atjonen 2007). Determining students' grades, they, too, have high stakes for students. Furthermore, even though socio-constructivist learning theories – the basis of the Finnish national curricula – emphasise the learner's active role and agency in the learning process (e.g. Tynjälä 1999; von Wright 1993), the test-taker has remained far more often than not an object of assessment, rather than an active agent.

2.2 Cheat sheets in a test?

During the past couple of decades, both teachers and researchers have developed alternative assessment methods that are better aligned with current learning theories. For instance, a cheat-sheet exam, also known as a crib-notes exam, refers to an exam or a test where students can bring into the exam notes they have written themselves for that particular testing situation. Sometimes the notes may be restricted, for instance with regard to their content or size. Some teachers have also insisted on using hand-written notes only.

Although cheat-sheet tests have not been really examined in FL/L2 education research so far, there are some published studies on cheat sheets in other contexts, mainly in psychology and mathematical subjects at the tertiary level. Most of the studies so far have advocated cheat sheets for a variety of reasons. For instance, they have concluded that the engagement in creating a personal cheat sheet – and not only using one in the test – improves studying and learning and thus also performance in the test (Block 2012; de Raadt 2012; Erbe 2007; Larwin 2012; Whitworth 1990). This is attributed to a coding process: when students review, select, organise and rewrite information on their cribs, they process the information more actively and more profoundly than when just trying to memorise it (e.g. Larwin 2012; Whitworth 1990). The improvement

in test results may be rather small (Gharib, Phillips & Mathew 2012) but there are other benefits, for instance decreased test anxiety (Block 2012; Butler & Crouch 2011; Erbe 2007; Whitworth 1990) or simply the fact that students find cheat-sheet exams useful and prefer them over closed-book exams (Block 2012; Erbe 2007; Gharib et al. 2012).

However, some studies have concluded that cheat sheets are not beneficial for learning (Dickson & Bauer 2008; Dickson & Miller 2005; Funk & Dickson 2011) even if they have improved test results and students have found them both helpful and stress-reducing (Dickson & Bauer 2008). Dickson and her colleagues argue that instead of really engaging in studying and learning, the students become dependent on their cribs. To test their dependency hypothesis, Dickson and Bauer (2008) organised a dual test on a course examination of developmental psychology at an American university. First, students had to take an unexpected pre-test without their crib notes and, immediately afterwards, they took the real exam, now with their crib notes. The questions were mostly identical multiple-choice questions. Dickson and Bauer argued that if the reason for an improved test performance lay in the engagement and improved learning, then students who had made the crib notes for the exam should perform just as well with or without the cribs in the actual test situation. As this was not the case (students performed better in the real test with their cribs than in the pre-test), Dickson and Bauer (2008: 117) concluded that “constructing crib sheets did not enhance learning, but use enhanced performance” because students depended on their notes in the exam. In fact, Dickson and Bauer (2008: 117) warned that crib sheets, or “crutches”, actually cripple learning as “students do not learn the course material as well when they expect to use a crib sheet” as they would for a closed-book exam, and advised against using cheat sheets.

To measure the efficacy of learning in another way, Gharib, Phillips and Mathew (2012) gave students surprise post-tests two weeks after the exams. They could not find any significant difference in the retention quiz performance between the students who had taken open-book, closed-book or cheat-sheet exams. Furthermore, they found out that “scores among exam types are positively correlated – students who do well on one exam type tend to do well on the others” (Gharib et al. 2012: 476). They concluded that “all three types of exams are equally effective as teaching tools”, but because of other beneficial factors, they deemed cheat-sheet and open-book exams more learner-friendly than closed-book exams (Gharib et al. 2012: 477).

Although prior research seems to be somewhat conflicting on the benefits of cheat sheets, in their recent meta-analysis Larwin, Gorman and Larwin (2013: 439) found out that “the use of either student-prepared testing aids or open-textbook exams can have a moderate impact on student performance on exams”. Furthermore, on the basis of higher effect sizes for student-prepared testing aids, they concluded as follows:

This outcome suggests some possible additional benefit to students who are required to prepare their own testing aids, thus requiring them to review, organize, and clarify the information on which they are being tested, as has been suggested by earlier research (...). This also suggests a potential benefit in the form of greater student engagement with the course material and information that, as other research has found, can ultimately help students to develop better study strategies and skills that they will incorporate into their other coursework. (Larwin et al. 2013: 439)

Therefore, Larwin et al. (2013) inclined towards favouring cheat-sheet exams.

3 The aim and setting of the study

Some years ago I introduced cheat sheets to some of my EFL courses. In order to examine the cheat-sheet test and its effects more thoroughly, I collected systematic data in 2013 and 2014. My main research interests were to find out how students react to cheat-sheet tests and what kinds of cheat sheets they construct and why. I also wanted to see if cheat sheets affect students' learning, test results as well as their learning and studying experiences. Above all, I wanted to explore if cheat-sheet tests could empower and engage students more in their assessment.

Altogether 101 students (61 females and 40 males, aged 17–18) took part in this study in 2013 and 2014 (47 and 54, respectively). They were on the penultimate compulsory English course (ENAS, the culture course) before the final Matriculation Examination. The cheat sheet was made for the written test, which comprised both grammar and reading comprehension exercises. The grammar exercise was a traditional multiple-choice exercise with 40 items and a maximum of 40 points. Although rather behaviouristic, some of the items required processing two grammatical constructs at once (e.g. articles and capital letters) and, admittedly, the exercise was quite detailed and challenging. The maximum score of the reading comprehension (RC) part was also 40 points.

The contents of the cheat sheet were limited to grammar (articles with proper nouns, punctuation, capital letters, sequence of tenses, conjunctions and linking words, some phrasal verbs). The size of the sheet was restricted (A4 on one side) so that students would have to process and summarise the information they selected. Even though some studies have suggested a link between hand-written notes and better test results (Larwin 2012), the students could prepare their cheat sheets outside lessons as they wanted. My only requirement was that each student made their own cheat sheet, i.e. they were not to copy somebody else's sheet. The cheat-sheet test was not the only assessment method in the course but one among many forms, such as two longer written pieces, an oral presentation, and smaller vocabulary and listening comprehension tests.

After the test, I collected the students' cheat sheets for analysis. Furthermore, using an open-ended questionnaire in Finnish (see Appendix 1), I collected their comments both before and after the test. The final student comments were collected after I had handed students their tests and cheat sheets back. The findings presented in this article are primarily based on the students' cheat sheets as well as their comments.

Inductive qualitative content analysis was used for analysing the students' cheat sheets and comments: after several readings, the cheat sheets as well as the comments were placed into categories emerging from the data (e.g. Elo, Kääriäinen, Kanste, Pölkki, Utriainen & Kyngäs 2014; Tuomi & Sarajärvi 2009). However, not all students answered every questionnaire item, and some answers were also rather vague. When a comment proved difficult to categorise, I consulted a second reader, an experienced educational researcher. Some student comments are used to illustrate the categories in the following text. This way, the reader can evaluate the trustworthiness of the analysis (Elo et al. 2014). The comments are identified by a student number, gender (F/M) and the quality of the cheat sheet (T=thorough, G=good, L=limited and N=no cheat sheet). Originally written in Finnish, I translated not only their meaning but also tried to retain the students' style, grammar and occasional ambiguities. The students' test results and previous grades are also used for additional quantitative analyses.

4 Findings

4.1 The initial reaction, construction and quality of the cheat sheet

When I introduced the idea of a cheat-sheet test to my students, it was a new idea to nearly half of them. In contrast, 15 students said they had used cheat sheets in two or more tests in various subjects, for instance foreign languages and mathematics. At this point, a great majority of the students said they liked the idea:

I was excited! The cheat sheet is familiar to me from the past and I like it a lot. Then you study properly for the test and you also feel more secure when you go to the test. If you get a black-out, you don't need to panic as you can check it from your cheat sheet. (15F, T)

Nice to have some change. A new thing for me. The test didn't stress me so much. (42F, T)

A good thing, we've had one sometime last year as well. Otherwise, it'd be a terrible task to remember all the exceptions. (57M, G)

About 20% of the students had neutral or slightly mixed feelings. Mixed feelings were caused by concern about either the difficulty of the test or the quality of learning:

214 CAN A CHEAT SHEET IN AN EFL TEST ENGAGE AND EMPOWER STUDENTS?

I liked it, I've had one once before. I was a bit worried if the test would be really hard - on the other hand, the cheat sheet relieves anxiety. (01F, T)

Good idea as such because nowadays similar things are used in the working life, for instance. On the other hand, you should know grammar in particular by heart so a cheat sheet for a grammar test may weaken your learning. This test form is new to me. (71M, L)

Wasn't a new thing, and for me, it doesn't really matter what kind of test it'd be. (24M, G)

Initially, five students felt that the cheat-sheet test would be a bad idea:

Cheat-sheet tests belong to the junior high. They are of no use when you are preparing for the Matriculation Exam. (03F, N)

Not a new thing but the cheat-sheet test will be harder, I don't like it myself. (94M, G)

After the test I collected and analysed the cheat sheets. A total of 92 students had made a cheat sheet, nine had not. The cheat sheets were divided into three groups: there were 42 thorough cheat sheets, 44 good ones and six limited cheat sheets that seemed very hastily constructed or consisted of only some short notes. The difference between a good and a thorough cheat sheet was basically in the quality, not in the quantity of information: thorough cheat sheets seemed more processed and organised with colour codes, pictures, the student's own rules or examples, for instance.

Altogether, over half of the female students (55.7%) prepared a thorough cheat sheet compared to 20% of the male students. In other words, 34 thorough cheat sheets were made by girls, eight by boys. Making a thorough cheat sheet therefore seemed to appeal more to the girls than to the boys. Another gender difference appeared among those who prepared the cheat sheet: with boys, the higher the previous grades, the more thorough the cheat sheet, but vice versa with girls (see Table 1).

TABLE 1. Previous English course grades (scale 4–10, 10 being the highest), the cheat sheet and the gender (n=101).

Cheat sheet n all (female / male)	Limited n=6 (2/4)	Good n=44 (23/21)	Thorough n=42 (34/8)	No cheat sheet n=9 (2/7)	All n=101 (60/41)
Mean of all prior course grades: all (female / male)	8.63 (9.25 / 8.31)	8.57 (8.64 / 8.50)	8.65 (8.63 / 8.72)	9.32 (9.68 / 9.21)	8.68 (8.69 / 8.65)
Previous grade: all (female / male)	8.50 (9.00 / 8.25)	8.50 (8.65 / 8.33)	8.43 (8.38 / 8.63)	9.44 (10.00 / 9.29)	8.55 (8.56 / 8.55)

However, nine students, seven boys and two girls, did not prepare a cheat sheet for the test. With one exception, the group was quite homogenous on the basis of their prior grades: the mean of their previous grades was clearly higher than that of those who made the cheat sheet. These students seemingly trusted their skills and preferred to take the test without a cheat sheet:

I don't find it useful in English because my language skills are so good anyway, but it must be helpful when trying to remember small, trivial things. (34F, N)

I wanted to see how well I can do without. (70M, N)

As mentioned earlier, one of the pedagogical premises of the cheat-sheet test is enhanced learning through an engagement in constructing the cheat sheet (e.g. Erbe 2007; Larwin 2012; Whitworth 1990). Also, the limited size of the cheat sheet makes students select material, which necessitates both the self-assessment of their own skills and an evaluation of the relevance of the information (e.g. Whitworth 1990). These ideas of engagement and self-assessment come across clearly in the majority of students' comments when they describe what they wrote on their cheat sheets and why. Some also mentioned their learning styles as a basis for selection:

I wrote nearly all grammar things on the cheat sheet because I revised things that way even if I had known some of the things beforehand. (09F, T)

Mainly I chose things that were difficult for me but also the most important. (08M, G)

For the cheat sheet, I selected things that I thought I might forget or that I didn't master yet so that they would stay better in my memory also for the future. (16F, T)

Some rules, but more examples. I understand things better through examples. (66F, T)

Furthermore, a few students wanted to make sure they would be able to perform well despite their possible test anxiety:

I wrote all I could squeeze in because I was afraid that I might have a 'blackout' in the test situation. (39F, T)

There were, however, some students who did not base their selection on self-assessment but rather more on a presumption of what might be asked in the test. Some students also wanted to take full advantage of the cheat sheet:

I listed all the things that were on the list of the test topics because they will probably be asked. (33M, T)

Almost all the information I could find. If you are allowed to use a cheat sheet, make the most of it then. (60F, T)

In sum, most students liked the idea of a cheat sheet. They had also engaged in preparing their cheat sheets, which, in general, were of good quality: 42 of them were considered thorough, 44 good and six limited.

4.2 The effects of the cheat sheet on learning experiences and results

Immediately after the test, the students were asked if the cheat sheet had been useful or beneficial in the test. Although feeling almost unanimously that the cheat sheet had been helpful in one way or another, the students did not believe its impact on their actual test results would be strong, perhaps a couple of points on average. Overall, the students felt that the cheat sheet had rather helped them to learn better than offered them the right answers in the test, as the following comment shows:

I believe the cheat sheet improved my test results a bit, but not significantly. Or, actually, maybe the cheat sheet improved the result quite a lot. I noticed in the test that I knew the things I had written on my cheat sheet. So, making the cheat sheet had taught me. (82F, T)

Over a third of the students said they had used the cheat sheet mainly for checking some of their answers. Eleven students mentioned that they had used their cheat sheets little or not at all in the test situation even though they had them:

I didn't actually need it more than in a couple of cases where I was wondering if there should be an article or not. (50F, G)

I believe that making the cheat sheet helped my language skills. If I had realised to take it out of my rucksack, I could have checked exercise 1 from it. (06M, G)

Conclusively demonstrating the efficacy of learning and the influence of the cheat sheet on the test results is unfeasible in a real classroom context, since the students cannot take the same test both with and without the cheat sheet. In Dickson and Bauer's (2008) study, college students first had an unexpected pre-test without their crib notes, and then the actual test with their cribs. Using the same research design would not have been possible in this experiment because of the practical time constraints of school life. Furthermore, Dickson and Bauer's (2008) assumption that the slightly better results in the real test with the crib meant dependency on the notes and thus inferior learning is, in my opinion, somewhat fallacious. There is ample evidence that students use different learning strategies when studying for different types of assessment situations (Atjonen

2007; Pickford & Brown 2006). For instance, students have been documented to use more deep learning strategies and skills that lead to better conceptual understanding as well as more self-directed study skills when studying for an open-book exam than for a closed-book exam, the latter evoking more rote learning and memorisation (Block 2012; Boniface 1985; Theophilides & Koutselini 2000). Thus, along the same lines, students who are told that they can use the cheat sheet in the test probably select things differently for the cheat sheet than they would if they knew they cannot use their notes in the actual test situation. For example, if there is something that is difficult to memorise by heart, such as a long list of exceptions, they write them down – as most students did in the present study. Also, if the students recognise their own weaknesses, they write down things they do not master – as, again, many students did in this study. Thus, using the cheat sheet very sensibly for their own needs, they can concentrate on learning and understanding more important concepts. If no cheat sheets are then allowed in the test situation, it handicaps students who have prepared well and rationally – but for a different kind of purpose and test situation.

However, in order to investigate if the cheat sheet had had any measurable effect on test results, I first compared the quality of the cheat sheet and the grammar results (see Table 2). I used the comparison of means and Pearson's correlation coefficient to analyse the test results; t-test, one-way analysis of variance as well as analysis of covariance were used to analyse the statistical significance of the differences of the means (see e.g. Jokivuori & Hietala 2007; Metsämuuronen 2009). On average, the students with thorough cheat sheets scored the highest in the grammar exercise and the six students with limited cheat sheets scored the lowest ($p < .01$, one-way ANOVA).

TABLE 2. The means of the test results, earlier English grades and the quality of the cheat sheet (n=101).

	Limited (n=6)	Good (n=44)	Thorough (n=42)	No cheat sheet (n=9)	All (n=101)
Grammar score A (max. 40p)	29.33	31.39	33.93	33.67	32.52
Reading comp. score B (max. 40p)	31.92	32.50	32.62	36.89	32.91
A-B: the difference	-2.58	-1.11	+1.31	-3.22	-0.38
Total score (max. 80p)	61.25	63.89	66.55	70.56	65.43
Test grade	7.71	8.03	8.36	8.89	8.22
Previous course grade	8.50	8.50	8.43	9.44	8.55
Mean of previous course grades	8.63	8.57	8.65	9.32	8.68

Next, I compared the grammar results with the reading comprehension scores. The underlying assumption was that the reading comprehension results would give an idea of the student's overall language skills and could thus be used as a baseline for the cheat-sheet part of the test. As hypothesised, the students' reading comprehension scores turned out to be in line with the average of their earlier English grades ($r = .83$, $p < .001$). As could be expected on the basis of their previous grades, the students without a cheat sheet scored the highest of all in the RC part (36.89/40p.), but on average 3.22 points less in the grammar section. The students with thorough cheat sheets were the only group that on average had a higher score in grammar than in reading comprehension.

The mean of the previous grades also correlates strongly with the grammar score ($r = .71$, $p < .001$) but the correlation is smaller than with the reading comprehension score. To investigate the effect of the cheat sheet (both its existence and quality) on the grammar score, an analysis of co-variance was run (see e.g. Jokivuori & Hietala 2007). When the effect of the mean of the previous grades was removed, the difference in grammar scores was still statistically significant ($p < .001$, ANCOVA).

With most girls having prepared a thorough cheat sheet, I wanted to see if there were any statistical gender differences (t-test) in the results. As the scatter graph illustrates at the individual student level (see Figure 1), female students tended to score slightly higher in grammar ($m = 33.57$) than in RC ($m = 32.61$). Their grammar scores were also higher than those of the male students ($m = 30.93$, $p < .01$), who, in turn, scored a little higher in RC ($m = 33.35$, $p = ns$).

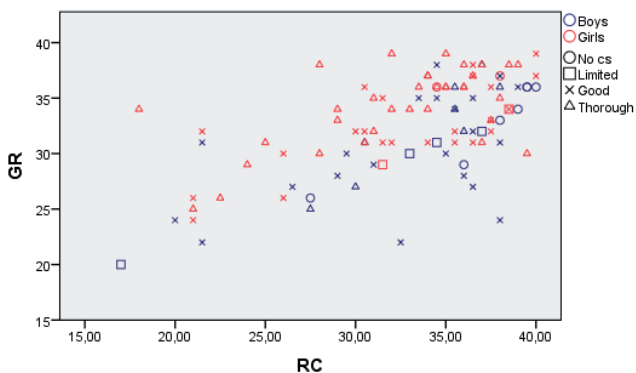


FIGURE 1. Scatter chart of grammar (Y axis) and reading comprehension scores (X axis) of female and male students, displaying the quality of the cheat sheet.

Female students' higher grammar scores further corroborate the effect of the quality of the cheat sheet on grammar results since females also prepared more thorough cheat

sheets than males. All in all, female students did slightly better in this test even though the means of the previous grades of both female and male students were almost identical (see Table).

4.3 The students' final verdict: was it worth it?

When I handed the tests and cheat sheets back, I asked the students for final feedback on the cheat-sheet test: "Now that you have seen your marked and graded test, what do you think of the cheat-sheet test, the cheat sheet itself and its effects? Was it worth it or not? Why? Would you do something differently now?" Out of the 92 students who had prepared a cheat sheet, six said the impact of the cheat sheet had been non-existent or negative, mainly because of the difficulty of the test, lack of time or lack of preparation:

Well, now it seems that it wasn't that useful. Of course, I could have made a bit better cheat sheet but I forgot and made it during the previous break/class so it didn't give me an awful lot of benefit. (40M, L)

The test was really difficult and there was too little time. I felt that the cheat sheet didn't help although I was well prepared otherwise, too. If the test was more difficult because of the cheat sheet, it would have been better to have the test without it. (21F, T)

Two of those six students also considered the cheat sheet detrimental for learning:

I would rather have done an ordinary test even though the grade would probably have been pretty similar. The cheat sheet was helpful in some things because I didn't cram at all, actually. A nice experiment but without a cheat sheet you would learn better. (79F, T)

Still, a great majority of students felt that the cheat sheet had been beneficial because it had enhanced their studying and learning as well as recollection, for instance. Some also said that it had made them prepare for the test better than usually. Fourteen students, all girls, mentioned feeling less insecure or stressed because of the cheat sheet.

A very good thing. I noticed last week that I already knew by heart some pretty difficult phrasal verbs that I had written on my cheat sheet. In the test itself I didn't need the cheat sheet that much but still it was useful in general, e.g. phrasals got into my mind. (22M, G)

The cheat sheet was quite handy because it forces you, in a way, to read and study properly grammar rules that you wouldn't normally bother to study so carefully. (Which probably was the whole point.) Things got into my head. I wouldn't do anything differently. (77F, T)

Yes it was useful. Things went into several boxes in the brain so to speak when you read and wrote at the same time. (47M, T)

It was worth it. Precisely because this test wasn't so stressful as well. (25F, G)

As can be seen from the comments above, preparing the cheat sheet had generally engaged students in the studying and learning process. Nevertheless, would students do something differently now that they had seen the results of their tests? Those 31 students who answered this question mentioned a few changes, for instance investing more effort in the construction of the cheat sheet or making the cheat sheet more condense. Five students mentioned changes that suggest some prior dependency on the cheat sheet: they would either study more or trust themselves more instead of trusting the cheat sheet alone

I wouldn't do anything differently except I would trust my gut feeling more than the cheat sheet. (38F, T)

If I did something differently, then I'd study more and not just trust the cheat sheet. (31F, G)

A few students were also disappointed when they noticed that they had made mistakes in the grammar section despite having had the correct information on their cheat sheets. Had they perhaps not had time to check, or had they not found the information on the cheat sheet – or had they had a false sense of knowing it by heart?

I'd concentrate better in the test because some mistakes were stupid, careless errors. (87F, G)

How about the students who had not made a cheat sheet at all? Would they construct one if given a second chance? Some perhaps would, some would not:

I don't regret my decision. My mistakes were in such small things that I wouldn't have written them on my cheat sheet anyway. (34F, N)

In some small things (like in the article exercise) I would have liked to have had it. (03F, N)

All in all, most students considered the cheat sheet helpful for the learning and studying process as well as for the test situation.

5 Discussion

The aims of the reported experiment were to find out how students react to cheat-sheet tests, what kinds of cheat sheets they construct and why. I also wanted to see if cheat sheets affect students' learning experiences and results. Above all, I wanted to

explore if cheat-sheet tests could give students a more engaged or empowered role in assessment. The qualitative findings showed that a significant majority of the students liked the idea from start to finish. Accordingly, 92 out of 101 students prepared cheat sheets, which in general were of good quality – 44 of them were regarded as good, 42 as thorough – so students clearly invested thought and effort in preparing them. Furthermore, quantitative analysis indicated that a thorough cheat sheet improved their test results a little.

This experiment has some limitations, though. First of all, this was primarily a teaching experiment in order to develop assessment methodology in my own teaching rather than a pure research experiment. The design was therefore not as rigorous as it could have been. Also, the number of students is rather limited for statistical analyses, and as they were not a random sample, the results cannot be generalised. Furthermore, my role as both the teacher and the researcher may have affected some of my decisions, both in teaching and in conducting the study. Some external factors such as poor handwriting may have in some subliminal way influenced the categorisation of the cheat sheets as well.

In addition, we can naturally argue whether the grammar results and those of the rest of the test are comparable and thus, whether they are feasible indicators of either improved test results or improved learning. We can also say that the differences were so small that they are not really significant. Furthermore, we can argue whether the small test result improvements were because of the students' dependency on the cheat sheet in the test situation or because of their engagement in learning while constructing the cheat sheet. However, the evidence clearly shows that female students, who seemed to have invested more in making their cheat sheets, performed better in the grammar exercise.

The students' personal experiences are perhaps the most pertinent issue here. First of all, the majority of the students experienced the cheat sheet as helpful for their learning process. Many students said that making the cheat sheet had improved their learning. Some also said that preparing the cheat sheet made them study better, in a more engaged way. Secondly, most students felt that although the cheat sheet did not increase their test scores much, it helped them in the actual test situation in one way or another: a few mentioned that the cheat sheet decreased their test anxiety and stress, and some said they could check their answers with the help of the cheat sheet. Finally, most of the students liked the cheat sheet, and quite a few would like to use it more often. Although there were a couple of students who may have depended on the cheat sheet, I agree with several researchers who claim that the use of cheat sheets can enhance both performance and learning through increased engagement (e.g. Block 2012; Erbe 2007; Larwin et al. 2013; Whitworth 1990). Even if test results had not improved, I would still

recommend cheat sheets because most students found them helpful. Unlike Dickson and her colleagues (Dickson & Bauer 2008; Dickson & Miller 2005; Funk & Dickson 2011), I believe that learning may manifest itself in many guises, not only in improved test results, and that students' own experiences and reactions are paramount. Positive attitudes, motivation as well as reduced anxiety are key components in learning. They simply cannot be ignored.

Furthermore, the cheat-sheet test empowered students in a very concrete way. First, the students could each decide whether or not to prepare the cheat sheet. Then, they could decide what to include and how. Constructing the cheat sheet developed their self-assessment as well as their learning-to-learn skills, and it introduced a new study method for those who had never written revision sheets. By reducing some students' test anxiety, the cheat sheet allowed them to focus on both studying and taking the test without excessive, disruptive stress. Finally, the students could decide when to use the cheat sheet in the test – a few decided not to consult it at all. So, the students had several opportunities to act as active agents – none of which a closed-book exam allows.

As shown in previous studies, the effect of the cheat sheet on the actual grade in general remained quite small: as Gharib et al. (2012) concluded, students who usually do well tend to do well regardless of the exam type. Thus, students and teachers who worry that cheat sheets might result in everybody getting (too) good grades, regardless of whether they really deserve them or not on the basis of their skills, need not worry. And the two or three disappointed students who had hoped that the cheat sheet would give them an easy escape route in lieu of studying for the test may have learnt more learner responsibility through their disappointment.

In a nutshell, I would argue that, in spite of its clear benefits, a cheat-sheet test is not a panacea, and it should not be used as the only assessment method. Assessment must be versatile and diverse enough to tap into the diverse skills of all students. Thus, further research in FL student assessment, whether dealing with cheat sheets or not, is needed. Yet, the findings of this limited study suggest that a cheat-sheet test is *one* learner-friendly assessment method that most students find beneficial for learning. It also engages and empowers them far more than traditional closed-book tests do. Quite justly, then, I will give the final word to a student whose comment summarises the findings of this study very well:

It was nice to try this. I wouldn't like to have a cheat sheet test every time but it's good every now and then. There was so much stuff that's difficult to learn by heart in a minute. Making the cheat sheet made me study the test area better than I would have studied for an ordinary test. (13F, T)

References

- Atjonen, P. 2007. *Hyvä, paha arviointi*. Helsinki: Tammi.
- Aydın, S. 2009. Test anxiety among foreign language learners: a review of literature. *Journal of Language and Linguistic Studies*, 5 (1), 127–137.
- Block, R. M. 2012. A discussion of the effect of open-book and closed-book exams on student achievement in an introductory statistics course. *PRIMUS*, 22 (3), 228–238.
- Boniface, D. 1985. Candidates' use of notes and textbooks during an open-book examination. *Educational Research*, 27 (3), 201–209.
- Butler, D. & N. Crouch 2011. Student experiences of making and using cheat sheets in mathematical exams. In J. Clark, B. Kissane, J. Mousley, T. Spencer & S. Thornton (eds) *Mathematics: traditions and [new] practices: proceedings of the 34th annual AAMT–MERGA conference*. Adelaide: AAMT and MERGA, 134–141.
- Cheng, L., D. Klinger, J. Fox, C. Doe, Y. Jin & J. Wu 2014. Motivation and test anxiety in test performance across three testing contexts: the CAEL, CET, and GEPT. *TESOL Quarterly*, 48 (2), 300–330.
- Cheng, L., Y. J. Watanabe & A. Curtis 2004. *Washback in language testing: research contexts and methods*. Mahwah, NJ: Lawrence Erlbaum.
- de Raadt, M. 2012. Student created cheat-sheets in examinations: impact on student outcomes. In *Proceedings of the Fourteenth Australasian Computing Education Conference (ACE2012)*, Melbourne: Australian Computer Society Inc., 71–76.
- Dickson, K. L. & J. J. Bauer 2008. Do students learn course material during crib sheet construction? *Teaching of Psychology*, 35 (2), 117–120.
- Dickson, K. L. & M. D. Miller 2005. Authorized crib cards do not improve exam performance. *Teaching of Psychology*, 32 (4), 230–233.
- Elo, S., M. Kääriäinen, O. Kanste, T. Pölkki, K. Utriainen & H. Kyngäs 2014. Qualitative content analysis: a focus on trustworthiness. *SAGE Open*, 4 (1). Available at DOI: 10.1177/2158244014522633.
- Erbe, B. 2007. Reducing test anxiety while increasing learning: the cheat sheet. *College Teaching*, 55 (3), 96–98.
- Funk, S. C. & K. L. Dickson 2011. Crib card use during tests: helpful or a crutch? *Teaching of Psychology*, 38 (2), 114–117.
- Gharib, A., W. Phillips & N. Mathew 2012. Cheat sheet or open-book? A comparison of the effects of exam types on performance, retention, and anxiety. *Psychology Research*, 2 (8), 469–478.
- Harlen, W. 2012. The role of assessment in developing motivation for learning. In J. Gardner (ed.) *Assessment and learning*. London: Sage, 171–183.
- Harlen, W. & R. Deakin Crick 2003. Testing and motivation for learning. *Assessment in Education*, 10 (2), 169–208.
- Hembree, R. 1988. Correlates, causes, effects, and treatment of test anxiety. *Review of Educational Research*, 58 (1), 47–77.
- Hildén, R. & M. Härmälä 2015. *Hyvästä paremmaksi: kehittämisideoita kielten oppimistulosten arviointien osoittamiin haasteisiin*. Koulutuksen seurantaraportit 6. Helsinki: Kansallinen koulutuksen arviointikeskus. Available at http://karvi.fi/app/uploads/2015/03/KARVI_0615.pdf.
- Horwitz, E. K. 2001. Language anxiety and achievement. *Annual Report of Applied Linguistics*, 21, 112–126.
- Horwitz, E. K. 2010. Foreign and second language anxiety. *Language Teaching*, 43 (2), 154–167.
- Hughes, A. 1989. *Testing for language teachers*. Cambridge: Cambridge University Press.

- Jokivuori, P. & R. Hietala 2007. *Määrällisiä tarinoita: monimuuttujamenetelmien käyttö ja tulkinta*. Porvoo: WSOY.
- Larwin, K. 2012. Student prepared testing aids: a low-tech method of encouraging student engagement. *Journal of Instructional Psychology*, 39 (2), 105–111.
- Larwin, K. H., J. Gorman & D. A. Larwin, 2013. Assessing the impact of testing aids on post-secondary student performance: a meta-analytic investigation. *Educational Psychology Review*, 25 (3), 429–443.
- Liu, M. & W. Huang, 2011. An exploration of foreign language anxiety and English learning motivation. *Education Research International*. Available at DOI: 10.1155/2011/493167
- Metsämuuronen, J. 2009. *Tutkimuksen tekemisen perusteet ihmistieteissä: tutkijalaitos*. 4. laitos. Helsinki: International Methelp.
- Pickford, R. & S. Brown 2006. *Assessing skills and practice*. Abingdon: Routledge.
- Rea-Dickins, P. & C. Scott, 2007. Washback from language tests on teaching, learning and policy: evidence from diverse settings. *Assessment in Education: Principles, Policy & Practice*, 14 (1), 1–7.
- Tarnanen, M. & A. Huhta 2011. Foreign language assessment and feedback practices in Finland. In D. Tsagari & I. Csépes (eds) *Classroom-based language assessment. Language testing and evaluation*, Vol. 25. Frankfurt am Main: Peter Lang, 129–146.
- Theophilides, C. & M. Koutselini 2000. Study behavior in the closed-book and the open-book examination: a comparative analysis. *Educational Research and Evaluation*, 6 (4), 379–393.
- Tuomi, J. & A. Sarajärvi 2009. *Laadullinen tutkimus ja sisällönanalyysi*. 6., uud. laitos. Helsinki: Tammi.
- Tynjälä, P. 1999. *Oppiminen tiedon rakentamisena: konstruktivistisen oppimiskäsityksen perusteita*. Helsinki: Kirjayhtymä
- Volante, L. 2004. Teaching to the test: what every educator and policy-maker should know. *Canadian Journal of Educational Administration and Policy*, 35. Available at <http://www.umanitoba.ca/publications/cjeap/articles/volante.html>.
- Whitworth, R. 1990. Using crib notes as a learning device. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, 64 (1), 23–24.
- Wright, J. von. 1993. *Oppimiskäsitysten historiaa ja pedagogisia seurauksia*. Helsinki: Opetushallitus.

APPENDIX.

Nimi/Name:

**Kun opettaja ehdotti lunttilappukoetta, mitä mieltä olit siitä ajatuksena? Miksi?
Oliko se sinulle uusi asia?**

When the teacher suggested a cheat-sheet test to your group, what did you think of the idea? Why?

Was it a new idea to you?

(This question was answered after the initial discussion on a cheat-sheet test in class.)

Mitä asioita kirjoitit lunttilappuusi? Miksi valitsit juuri ne asiat? Jos ET tehnyt lunttia, kerro mikset.

What did you write on your cheat sheet? Why did you choose these things? *If you did NOT make a cheat sheet, tell me why.*

(This question as well as the following two questions were answered immediately after the test.)

**Oliko luntista hyötyä kokeessa? Millaisissa asioissa/tilanteissa? Miksi (ei)?
(Millaisissa tilanteissa olisit kaivannut lunttia?)**

Was the cheat sheet useful in the test? In what kinds of things/situations? Why (not)?
(In what kinds of situations would you have liked to have had a cheat sheet?)

Miten uskot luntin vaikuttaneen koetulokseesi? (Tai sen, ettei sinulla ollut lunttia)

How do you think the cheat sheet affected your test result? *(Or, how did you not having a cheat sheet affect your test result?)*

Nyt kun olet nähnyt kokeesi korjattuna, mitä mieltä olet lunttilappukokeesta, luntista ja sen annista? Kannattiko lunttilappukoe vai ei? Miksi? Tekisitkö nyt jotain toisin?

Now that you have seen your marked and graded test, what do you think of the cheat-sheet test, the cheat sheet itself and its effects? Was it worth it or not? Why? Would you do something differently now?

(This question was answered after the marked tests and the cheat sheets were returned to the students.)