JYX JYVÄSKYLÄN YLIOPISTO UNIVERSITY OF JYVÄSKYLÄ

This is a self-archived version of an original article. This version may differ from the original in pagination and typographic details.

Author(s): Leontjev, Dmitri

Title: Finnish Matriculation Examination, National Curriculum, and teachers' attitudes, perspectives, and practices : when the two assessment cultures meet

Year: 2022

Version: Published version

Copyright: © Kirjoittaja & Suomen soveltavan kielitieteen yhdistys ry, 2022

Rights: In Copyright

Rights url: http://rightsstatements.org/page/InC/1.0/?language=en

Please cite the original version:

Leontjev, D. (2022). Finnish Matriculation Examination, National Curriculum, and teachers' attitudes, perspectives, and practices : when the two assessment cultures meet. In T. Seppälä, S. Lesonen, P. Iikkanen, & S. D'hondt (Eds.), Kieli, muutos ja yhteiskunta (pp. 133-150). Suomen soveltavan kielitieteen yhdistys AFinLA. AFinLA:n vuosikirja, 2022. https://doi.org/10.30661/afinlavk.113656

Seppälä, T., S. Lesonen, P. Iikkanen & S. D'hondt (toim.) 2022.
Kieli, muutos ja yhteiskunta – Language, Change and Society. AFinLAn vuosikirja 2022.
Suomen soveltavan kielitieteen yhdistyksen julkaisuja n:o 79. Jyväskylä. s. 133–150.

Dmitri Leontjev
University of Jyväskylä

Finnish Matriculation Examination, National Curriculum, and teachers' attitudes, perspectives, and practices: When the two assessment cultures meet

The Finnish National Core Curriculum (NCC) and the Matriculation Examination (ME) are likely to shape what happens in classrooms around Finland, including classroom assessment. The present study aimed at exploring English academically oriented upper secondary school English teachers' ($n = 85$) (1) attitudes to changes in the NCC and the ME, (2) their perspectives on these changes, and (3) how their assessment practices were shaped by the NCC and the ME. The data were collected roughly at the time both changes came into force. The data came from a survey and follow-up interviews with four teachers. The results indicated that teachers were quite positive about the changes but anticipated both positive and negative changes in the classroom. Their assessment practices appeared to be shaped more strongly by the ME than the NCC and seemed to impact their attitudes to and perspectives on the two changes.

Keywords: assessment literacy, Matriculation Examination, National Core Curriculum, language assessment
Asiasanat: arviointilukutaito, ylioppilastutkinto, opetussuunnitelman perusteet, kielitaidon arviointi

VERTAISARVIOITU
KOLLEGIALT GRANSKAD
PEER-REVIEWED
www.tsv.fi/tunnus

1 Introduction

Two changes had entered the Finnish education between 2015 and 2019: the only nationwide standardised examination, the Matriculation Examination (ME), was digitised, and the new National Core Curriculum (NCC; FNBE 2016) was published.
The ME is norm-referenced, is taken at the end of the upper-secondary school, and is the only form of high-stakes assessment in Finland, used as an entrance examination by a growing number of higher education institutions. The ME was fully digitised at the end of 2019, with the digitised ME in English as a foreign language (L2), the context of this study, coming in spring 2018. By the time of the study, the Matriculation Examination Board had revealed that the item types in the digitised ME would be expanded to include audio, pictures, and videos (https://www.ylioppilastutkinto.fi/en/), and the new ME format could be practised through the Abitti system (https://www.abitti.fi/).

By the time of the study, the 2015 NCC (FNBE 2016) had been introduced to all stages of education and was being implemented locally. One change in the NCC pertained to assessment:

> The purpose of assessment of learning is to promote the student's learning… Assessment as well as providing feedback during the studies are part of the interaction between a student and a teacher. Feedback as well as self and peer assessment guide the student to define the goals in greater detail and to develop his or her work… The student's learning is assessed during a course (FNBE 2016: 295).

To compare, the 2003 NCC listed the following about assessment:

> The role of assessment of students' learning is to provide students with feedback on their progress and learning results... The purpose of such feedback is to encourage and guide students in their studies… Each course will be assessed upon completion. The role of assessment is to provide students with feedback on how well they have met the objectives of the course and on their progress in that subject (FNBE 2003: 224).

It transpires that the focus of assessment in the 2015 NCC was more on promoting learning, with a stronger recognition of learners' role in assessment than in the 2003 Curriculum. This is evidenced in that assessment of the product and the process of learning are elicited more in the 2003 and 2015 NCC, respectively. Furthermore, feedback in the 2015 NCC is presented as emerging in the interaction between the teacher and learners. Still, both summative (providing evidence for learner achievement at the end of a learning period) and formative (focused on decisions regarding adjusting teaching and learning) functions of assessment (Wiliam & Leahy 2015) are recognised in the 2015 NCC.

The ME and assessment conceptualisation in the NCC are different in that they elicit different assessment cultures (Davison & Leung 2009). The ME elicits assessment

of learning (A*o*L), where summative and formative assessments have different "form and function" (Davison & Leung 2009: 397), and the emphasis is on equality, comparability, and standardisation. The NCC is strongly focused on assessment *for* learning (A*f*L), in which the purpose of all classroom assessments, even initially having the summative function, is to direct learning by suggesting where learners can go next and how (Davison & Leung 2009). The emphasis is on equity and fostering individual developmental trajectories. Whereas A*o*L is rooted in positivist theories, A*f*L is firmly positioned in constructionist view (Lam 2016). Thus, summative/formative and A*o*L/A*f*L are not fully congruent, and the later conceptualisation should be useful when studying teachers' assessment literacy and practices.

What regards assessment practices, teachers in Finland are considered professionals trusted with achieving the NCC goals, so classroom assessment is designed or selected by teachers (Tarnanen & Huhta 2011). Tarnanen and Huhta (2008: 270) argued that "Finland is not a very testing-oriented country." However, overviews of classroom assessment in Finnish academically oriented upper secondary schools (Lukio; e.g., Mäkipää 2021; Välijärvi et al. 2009) show that exams and tests are used the most often.

Recently, there has been more emphasis on assessment promoting learning in educational policies worldwide (Tsagari & Vogt 2017). Simultaneously, the standardisation of assessments to achieve high educational standards increased (Lam 2015). There is also an implied expectation that teachers are responsible for merging the two assessment cultures in the classroom, which they cannot always do (Poehner & Inbar-Lourie 2020). Hence, some challenges teachers in Finland, particularly at the Lukio level, face are similar to other educational contexts.

This study aims to explore Finnish Lukio L2 English teachers' attitudes to and perspectives on (1) the changes in the NCC and ME and (2) the roles of the NCC and the ME in the classroom assessment practices. I define both attitudes and perspectives as referring to mental views and shaping and being shaped by practices. However, *attitudes* involve an affective component, whereas *perspectives* involve an understanding of interrelations between parts forming a larger whole and introspection. I emphasise that I do not compare conceptualisations of language assessment in the NCC and ME. Instead, I explore what they *create* in the classroom.

2 Background and Literature review

2.1 Assessment literacy in practice

According to the comprehensive definition provided by Fulcher (2012), teacher assessment literacy (AL), in addition to understanding assessment concepts and cri-

teria, involves being able to design, conduct, and interpret these assessments and evaluate their impact (see also Looney et al. 2018; Tsagari & Vogt 2017).

Xu and Brown (2016) further argue for expanding the concept of AL to *assessment literacy in practice*, conceptualising AL as several components impacting one another. First, there is *knowledge base*, including knowledge of assessment purposes, methods, and standards and understanding of the role of agents, such as learners in peer- and self-assessment. The uptake of this knowledge is mediated by *conceptions* about assessment, including beliefs about assessment and emotions associated with it. These conceptions are also mediated by *educational contexts*, including educational policies (see Brown et al. 2019). These together mediate assessment *decisions* teachers make in the classroom. Finally, the model includes the teacher's *identity* as an assessor. This model allows to conceptualise teachers' AL as dynamic, changing due to education, experience, and as changes in policy occur, as well as reciprocally related to practices (see also Hill 2017; Inbar-Lourie 2016; Taylor 2013). Indeed, as Coombs et al. (2018) found, early career teachers in the USA first value A*f*L (due to their pre-service training) with the following increase of support of A*o*L (due to the focus on accountability), followed by yet another change towards a more differentiated approach to assessment.

The conceptualisation of AL in this paper is *assessment literary in practice* with the focus on teachers' practices, the role of the ME and the NCC in these, and teachers' attitudes to and perspectives on the changes in the ME and the NCC. Focusing on attitudes, perspectives, and practices encompasses several components of Xu and Brown's (2016) model.

2.2 High stakes assessments and educational policies and teachers' assessment practices

There has been ample research on the impact of high-stakes tests and educational policy on teaching, learning, and assessment. Brown and Remesal (2017) explored Ecuadorian school teachers' conceptions about assessment, discovering that these were largely informed by high-stakes examinations. The education policy that highlighted A*f*L had little influence. Luxia (2005) studied the effect of the high-stakes National Matriculation English Test designed to make teaching and learning communicative in a study of 388 teachers and 986 learners. Examination preparation overpowered the communicative test function. Mäkipää (2021) found that, according to students, Lukio teachers primarily used tests and practice exams, attributing this to the impact of the ME (see also Atjonen et al. 2019; Au 2007).

With regard to teacher reactions to changes in educational policy about assessment, Verhoeven and Devos (2005), in a questionnaire study of 1,274 secondary school teachers in Flanders, found that, even though school policy elicited A*f*L, teachers leaned towards A*o*L. Only 17% of the respondents frequently applied A*f*L

principles in their practices. Similarly, van der Kleij, Cumming, and Looney (2018), studying the change in the educational policy in Australia emphasising A*f*L, found that teachers still primarily relied on A*o*L. Teachers' greater reliance on A*o*L appears to emerge in other educational contexts, e.g., in Greece (Vlachou 2018), Scotland (Hayward & Hedge 2005), and Finland (Hildén & Fröjdendahl 2018; Mäkipää 2021).

3 Methodology

3.1 Research Questions

The aims of the study were realised by finding answers to two research questions:

1. What are upper secondary school L2 English teachers' attitudes to and perspectives on the changes in the Matriculation Examination and the National Core Curriculum?
2. What are the Matriculation Examination and the National Core Curriculum roles in upper secondary school L2 English teachers' assessment practices?

The study had an exploratory nature. However, based on the research outlined in Section 2, the following preliminary hypotheses were formed: (1) there is a range of reactions to the changes due to teachers both realising their benefits and having concerns about accommodating them in the classroom and (2) both the ME and the NCC emerge as strongly contributing to classroom assessment practices, but the role of the ME is stronger.

3.2 Instruments

The main instrument in the study was a questionnaire (see Appendix), having both five-point Likert-scale and open-ended items eliciting:

1. teachers' attitudes to the changes in the ME and the NCC;
2. teachers' perspectives on the changes in the ME and NCC;
3. the role of the ME and the NCC in teachers' assessment practices.

The questionnaire (and the whole survey) was informed by (1) the questionnaire that studied teachers' perspectives on the changes to the Hong Kong Certificate of Education in English in 1996 (Cheng 2005) and (2) the survey of Finnish comprehensive school first and foreign language teachers' assessment practices (Tarnanen & Huhta 2011). To a lesser extent, Wall and Horák's (2006) instrument for exploring teachers' and learners' perspectives on the changes in TOEFL and the instrument for

exploring Finnish teachers' and learners' readiness for the digitised ME (Lakkala & Ilomäki 2013) informed the instrument.

The instrument was first piloted among specialists in language assessment, survey experts, and teachers (n = 8). As a part of the a posteriori validation, an exploratory factor analysis (EFA) was conducted with the data collected in this study (Section 3.3). I studied the items relating to the role of the NCC and ME in teachers' practices. The EFA demonstrated that the model with two factors (role of the NCC and role of the ME in teachers' practices) was consistent with the data, X2(8) = 9.52, p = .301, *RMSEA* = .05, *CFI* = .975, *TLI* = .952, *SRMR* = .08, which added to the validity of the instrument. In this paper, only questionnaire items directly related to the two research questions are discussed (see Appendix).

The semi-structured interview was the other instrument in the study. Most interview themes were the same as those in the questionnaire but also included topics that emerged in the open-ended responses of the interviewed teachers, such as teacher assessment and feedback practices, teachers' understanding of assessment, and merging assessment as elicited in the ME and the NCC.

3.3 Participants, data, and procedures

Lukio L2 English teachers were reached through openly available contacts on school websites. The invitations were also sent to school secretaries and through professional networks. All in all, 692 individual e-mails were sent, though in some cases, it was difficult to determine whether a teacher worked in the Lukio or in the comprehensive school.

The online questionnaire administered in spring 2017 was anonymous unless teachers indicated that they wanted to be interviewed. The teachers were informed about the goals of the study and how the data were to be used. Only the teachers who gave permission to use their data, n = 85, participated in the survey. Twenty-four of them gave their permission to contact them for an interview. In spring of 2016, 504 teachers taught English in Lukio full-time (FNBE 2017), so the sample comprised 12.3% (judging by the number of e-mails sent) to 16.9% of the target population.

Of the background variables, considering Coombs et al.'s (2018) findings, teaching experience should be mentioned. The respondents had, on average, 20.5 years of teaching experience (*min* = 2; *max* = 40); of them, the average of 16 years of teaching at the Lukio (*min* = 1; *max* = 35). Teaching experience did not significantly correlate with any of the studied variables.

The semi-structured interviews took place in late spring–early summer 2017, their lengths ranging from 31 minutes to one hour and fourteen minutes. Four teachers, referred to as T1 to T4, were interviewed. The selection was made such that two teachers indicated in their questionnaire that they were sceptical about the changes in the ME and the NCC and two that they were enthusiastic about them; the

teachers were selected randomly otherwise. This allowed different points of view to be represented in the interviews.

3.4 Analysis

The statistical analysis of the questionnaire was conducted using SPSS version 24. Because the data were on Likert scales and often skewed, non-parametric tests were used: Spearman's correlations, Friedman test, and Wilcoxon signed-rank test. Bonferroni adjustment was used for pairwise comparisons. The open-ended items were content analysed. The grounded coding was completed independently by two researchers (investigator triangulation). After this, the researchers arranged a data session to discuss the coding. They arrived at a joint decision on the few discordant codes (e.g., "improving the exam" *Vs* "versatility of assessment formats") and decided on the inclusion of codes present only in one coder's file. An agreement was reached on all such cases.

The interview data will be used to yield deeper insights into the questionnaire responses (data triangulation). In the paper, the quoted interview excerpts were shortened to save space, the skipped parts indicated with '...'. The results will be interpreted with reference to assessment cultures meeting in the classroom and teachers' assessment literacy in practice.

4 Results

The results will be presented in two sections. First, the teachers' attitudes to and perspectives on the changes in the ME and the NCC will be studied (research question 1). Then, the roles of the ME and the NCC classroom assessment practices, as self-reported by the teachers, will be explored (research question 2).

4.1 Teachers' Attitudes to and Perspectives on the Changes

As a group, the teachers were fairly positive about the changes in the ME and the NCC (Table 1 and Figures 1 and 2; see items 1 and 3 in the Appendix).

TABLE 1. Teachers' attitudes to the changes in the Matriculation Examination and the National Core Curriculum; 1 = sceptical; 5 = enthusiastic.

	Mean	**Median**	**SD**	**Min**	**Max**
Matriculation Examination	3.52	4	1.02	1	5
National Core Curriculum	3.42	3	1.01	1	5

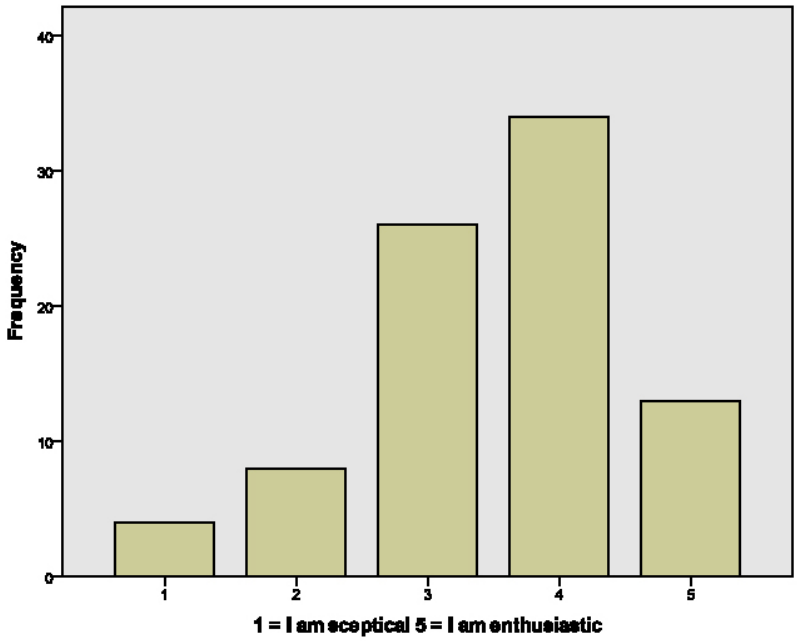

FIGURE 1. Teachers' attitudes to the changes in the Matriculation Examination.

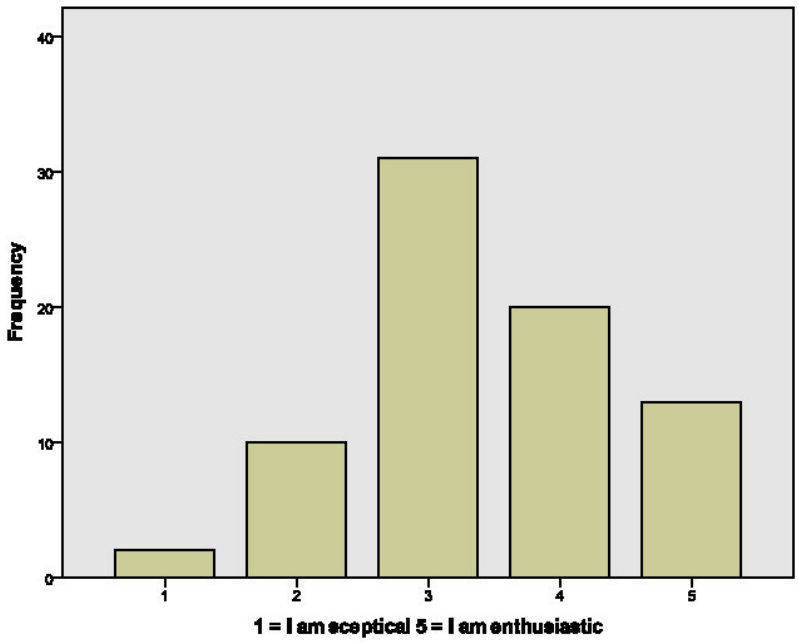

FIGURE 2. Teachers' attitudes to the changes in the National Core Curriculum.

The teachers were slightly more enthusiastic about the digitisation of the ME, but the difference between their attitudes to the two changes was not significant, $Z = -.57$, $p = .571$.

As revealed in the open-ended items (items 2 and 4 in the Appendix), among the positive aspects of the change in the ME were videos as a part of listening ($n = 28$) and a decrease in teachers' workload ($n = 26$). The topics of increased variety of task types ($n = 20$), ME becoming easier ($n = 19$), and authenticity ($n = 16$), too, emerged often.

Among the negative aspects, the overwhelming majority ($n = 50$) mentioned technical or security issues, followed by the lack of computer skills ($n = 14$) and health problems ($n = 11$). Other topics included lack of ME preparation materials ($n = 8$), more mistakes that learners would make due to being less attentive ($n = 4$), and the computerised modality making it difficult to develop speaking examination ($n = 4$).

Nine teachers skipped the item eliciting their attitudes on the changes in the NCC. However, it is difficult to establish relationships between this and other variables due to the small number of the participants who skipped this item. As emerged in the open-ended responses focusing on teachers' perspectives, one reason for teachers' lack of enthusiasm towards the change in the NCC was that some ($n = 8$) thought this was not that much of a change, as exemplified in Excerpt 1.

(1) We already have ongoing assessment and feedback systems in our existing classroom practices… Not that we shouldn't be aware of the role of assessment, only the "new" doesn't seem so new after all.

Others ($n = 5$) expressed worries that they lack expertise in A*f*L or scepticism as to whether other teachers embrace the changes, for example, "I am rather sceptical of how many teachers will dedicate that much time on assessment as they may see it as extra work." Indeed, teachers reported, for example, "Too much assessment during studies will take time from learning." Furthermore, some ($n = 12$) confused the continuous assessment in the NCC with assessment *of* learning, for example, "Testing should be reliable, valid and restricted." During the interview, T2 discussed a similar issue (Excerpt 2).

(2) Many teachers say that "I don't have time for any extra assessment … Do I still need to add more assessment here?" … But then again, if you see it as a separate thing, detached thing, that you have to do some kind of a document or whatever. But seeing it as a natural part of it [teaching and learning].

T2 acknowledged that their colleagues saw continuous assessment as something extra, taking time from learning. At the same time, T2's own understanding was that assessment is an integral part of teaching and learning, which was also mentioned

as a positive aspect in some open-ended survey responses (n = 12). That is, some teachers indeed recognised the focus on A*f*L in the NCC.

4.2 The role of the ME and the NCC in classroom assessment practices

The item *What affects your choice of classroom assessment?* included several aspects, of which I will mainly discuss the role of the NCC, the ME, and teachers' own experience (Table 2). The reason for this was a combination of the feasibility of reporting and the focus of this paper.

TABLE 2. What affects your choice of classroom assessment? 1 = not at all; 5 = a lot.

	Mean	**Mdn**	**SD**	**Min**	**Max**
The National Core Curriculum	3.80	4	0.94	1	5
Local (municipal) and school curriculum	3.58	4	1.06	1	5
Matriculation Examination	3.75	4	0.94	1	5
Your own teaching experience	4.47	5	0.60	3	5
…					
Learners' preferences and expectations	2.99	3	0.76	1	5

The teaching experience was rated significantly higher than the ME and the national and local curricula, as demonstrated by a Friedman test, $X^2(7) = 164.67$, $p < .001$ followed with the pairwise comparisons (here and elsewhere, Bonferroni adjustment made taking into account all of the studied aspects), $Z = -4.76$, $p_{adj} < .001$, $r = -.55$, $Z = -5.32$, $p_{adj} < .001$, $r = -.61$, and $Z = -5.22$, $p_{adj} < .001$, $r = -.60$ respectively, indicating thus a large difference. The pairwise comparisons among the role of curricula and the ME were not statistically significant. Learners' preferences (M = 2.99, Mdn = 3, SD = 0.67) were rated significantly lower than the rest of the variables in Table 2. The pairwise comparison with the role of the local and school curricula, for example, $Z = -3.67$, $p_{adj} = .007$, $r = -.30$, suggested a moderate difference; differences with other variables being large.

There is, thus, somewhat of a contradiction between learners' role in assessment emphasised in the NCC and learners' involvement in classroom assessment. This also emerged in the frequency of self- and peer-assessment (Table 3).

TABLE 3. Frequency of teacher-, self-, and peer-assessment (1 = never; 5 = often).

	Mean	**Mdn**	**SD**	**Min**	**Max**
Teacher	4.91	5	0.33	3	5
Students themselves	3.22	3	0.90	1	5
Peers	2.84	3	0.81	1	5

A Friedman test demonstrated that the difference was significant, $X^2(2) = 119.10$, $p < .001$. The pairwise comparisons demonstrated that significant, and rather large, differences were namely between teacher assessment and self- and peer-assessment, $Z = -7.32$, $p_{adj} < .001$, $r = -.59$ and $Z = -7.63$, $p_{adj} < .001$, $r = -.61$ respectively.

A weak negative correlation was found between the teachers' attitudes towards the changes in the NCC and the self-reported frequency of teacher assessment. Those between the teachers' attitudes and the frequency of self- and peer-assessment were positive and moderate (Table 4).

TABLE 4. Spearman's correlation between teachers' attitudes towards the changes in the NCC and teacher-, self-, and peer-assessment.

	Teacher-assessment	**Self-assessment**	**Peer-assessment**
Teachers' attitudes to the changes in the NCC	$r_s = -.286$ $p = .017$	$r_s = .545$*	$r_s = .446$*

*$p < .001$

Because teachers' assessment practices were correlated to their attitudes to changes to be fully in effect in future, it can be suggested that teachers practising self- and peer-assessment more often were more likely to appreciate the changes in the NCC than those who practised these less.

To illustrate this further, T3, who was rather sceptical about the changes and practised peer- and self-assessment only rarely, reported the following (Excerpt 3):

(3) I'm sceptical because usually, students are not that serious about self-assessment. Almost ninety percent if I ask… them to do it … would answer with usually the best possible mark.

T3's understanding of self-assessment was, thus, informed by A*o*L, as she did not acknowledge the use of self-assessment for developing self-regulated learning, for

example. T1, on the other hand, being enthusiastic about the changes in the NCC, reported the following (Excerpt 4):

(4) [T]hat's why they [the students] have to do it all the time… so they have to think of the reasons, and they have to analyse their skills of learning or if they learn something, … how did you do it and then always why did you succeed or why didn't you succeed.

T1, therefore, recognised that the purpose of self-assessment is, above all, to help learners take charge of their learning. In addition, T1 has also reported the following about peer-assessment, even though they only practised it from time to time (Excerpt 5):

(5) [W]hen they are saying something about their partners' performance… they have to somehow assess the process… Its best purpose is to raise awareness of assessment and the process itself… I'm not like lessening it in any way.

Thus, the self-reported frequency of self- and peer-assessment should not always be equated with less perceived importance.

What regards the types of classroom assessment, tests and essays were reported to be used the most often (Table 5).

TABLE 5. Types of assessments used.

	Mean	**Mdn**	**SD**	**Min**	**Max**
Tests and practice exams	4.19	4	0.81	2	5
Essays and compositions	4.13	4	0.77	2	5
Observation	3.97	4	0.99	1	5
Group work	3.53	4	0.84	1	5
Dialogues	3.05	3	1.19	1	5
Oral tests	2.9	3	0.85	1	5
Portfolios	2.73	3	1.00	1	5

The results of the Friedman test, $X^2(5) = 167.88$, $p < .001$, demonstrated that there was a significant difference among different assessment types. The pairwise comparisons for the frequency of essay assessment and the variables rated lower showed that, except for observations, there were relatively large differences between them (Table 6).

TABLE 6. Pairwise comparisons between the frequency of essays and compositions and lower-rated assessment types*.

	Essays and compositions
Observation	$Z = -1.16$, *n.s.*
Group work	$Z = -4.58^{**}$, $r = -.37$
Dialogues	$Z = -5.18^{**}$, $r = -.42$
Oral tests	$Z = -6.55^{**}$, $r = -.53$
Portfolios	$Z = -7.10^{**}$, $r = -.57$

* Bonferroni adjustment of p-values made, considering all pairwise comparisons across the variables.
$^{**}p_{adj} < .001$

The differences in using tests and practice exams and the rest of the assessment types, with the exception of the non-significant differences with essays and observation, were even larger. Hence, while being generally enthusiastic about changes in the NCC, the teachers resorted to the kinds of assessments used in the ME more often than other assessment types. The following T2's thoughts shed some light onto this (Excerpt 6):

(6) I think this is something that we face in our school because of this problem, these two parts, the matriculation and what we do in the classroom. Like, if we invest time in dialogues and encouragement and motivational talks, like, is it time away from this? Some of our teachers believe that "…Why would you waste time in that? … this is what they are really tested for."

T2 acknowledged the tension that the ME creates in the classroom, suggesting that it can lead to focusing on assessment elicited in the ME. Furthermore, when asked how they could merge the ME and the NCC, T2 emotionally responded, "I wish I knew!" T3's response was similar to T2's.

5 Discussion

The study aimed to explore teachers' attitudes to and perspectives on the changes in the NCC and the ME and the roles the NCC and the ME play in the classroom. The participants felt fairly positive about both changes. Their attitudes, it seems, were shaped by what they perceived the changes to the ME and the NCC would bring,

their classroom practices and experiences with the ME, computerised assessments, and assessment *for* learning.

From the perspective of AL, there were three groups of teachers with regard to their perspectives on the changes in the NCC: (1) thinking it was not much of a change, (2) feeling unprepared for it, and (3) misinterpreting A*f*L as continuous testing. Whether and to what extent teachers' responses in the latter group were related to the power of the ME is more of a speculation, though reports in this study (e.g., Excerpt 2) and the previous research (e.g., Brown & Remesal 2017; Stobart & Eggen 2012) suggest that it might be the case.

Judging by T2's report in Excerpt 6, there could be another possibility for teachers' attitudes and perspectives: acknowledging the value of assessment as integral in the teaching/learning process, they struggled to accommodate both, focusing on preparing their students for the ME (see also Hayward & Hedge 2005; Luxia 2005). This calls for providing support to teachers in the form of in-service training, which is essential when new developments are introduced (Tarnanen & Huhta 2008). Still, judging by T1's and T2's responses, teachers could have had some ideas for merging the two assessment cultures. Further research can focus on how teachers combine A*o*L and A*f*L in the classroom.

Some teachers reported that A*f*L had been a part of their practices, which is more positive than the previous research findings (e.g., Brown & Remesal 2017; Tsagari & Vogt 2017; Verhoeven & Devos 2005). However, generally, teachers marked tests and essays as the most commonly used assessments in the classroom. I should note, though, that it cannot be deduced from teachers' responses whether they also used tests formatively and how they conducted essay assessment; that is, these assessment types could as well be parts of A*f*L culture. Future research could shed more light on this. Teachers resorting to these types of assessment is similar to what other research has found (Mäkipää 2021; Tarnanen & Huhta 2011; Tsagari & Vogt 2017; Van der Kleij, Cumming, & Looney 2018; Välijärvi et al. 2009), suggesting that the ME had more power on teachers' assessment practices than the NCC did (see also Au 2007; Luxia 2005; see Excerpt 6).

The clash between the reported role of the NCC and teachers' self-reported practices was also evident in the lack of involvement of learners in assessment. As emerged in Excerpt 3, some teachers could have used self- and peer-assessment for assessing learning outcomes and, finding that they did not work for this purpose, abandoned them. This suggests that even though self- and peer-assessment are underscored in the NCC, their value may not be recognised by all teachers. Still, the frequency of assessment should not be equated with its usefulness. The area of teacher- *vs* self- and peer-assessment could be further investigated.

Overall, it appears that the impact of the ME was higher than that of the NCC. Furthermore, even those teachers who recognised the value of A*f*L did not necessarily see how the two assessment cultures could be merged in the classroom.

Several limitations should be mentioned to be addressed in future research. One is the size of the sample, so caution should be exercised as to the generalisability of the findings. A further limitation refers to the sampling strategy. While I did checks with regard to how well the sample reflected the population, probability sampling, such as stratified random sampling, would have strengthened the generalisability. I should also note that the teachers' self-reports were not compared with learners' self-reports, though, as far as the types of assessment used in the classroom, the findings are similar to those of Mäkipää (2021), who focused on learners' reports.

6 Conclusion

This study aimed to explore the roles that the National Core Curriculum and the Matriculation Examination play in shaping upper secondary school L2 English teachers' assessment practices and their attitudes to and perspectives on the changes in the NCC and the ME.

Overall, it appears that despite a strong focus on A*f*L in the NCC, teachers in Finland have similar struggles as they do elsewhere, their classroom assessment focusing more on assessment of learning outcomes. This underscores the complex nature of teachers' AL in practice, where knowledge base, conceptions about assessment, and educational policy shape their understanding assessment and assessment practices. Teachers also need support in developing their classroom assessment. As a prerequisite for it, I propose the interplay of teacher experience with high-stakes assessment, teachers' interpretation of the NCC, their conceptions of, and practices in assessment can be further explored. That teachers in Finland are somewhat more likely to embrace A*f*L and find solutions for merging A*f*L and A*o*L than in other educational contexts (cf. Tsagari & Vogt 2017) speaks for eliciting the A*f*L assessment culture in educational policy while giving teachers freedom in implementing curriculum goals their classrooms. More importantly, this suggests collaborative engagements between researchers and teachers with the common goal of merging A*f*L and A*o*L in the classroom, as argued by Poehner and Inbar-Lourie (2020). Such collaborations could focus on aligning the practices proposed by the NCC and tasks and constructs elicited in the ME. After all, the goal of both the NCC and the ME is to contribute to the development of language proficiency, even if they approach this goal from different angles.

Literature

Atjonen, P., Laivamaa, H., Levonen, A., Orell, S., Saari, M., Sulonen, K., ... & Immonen, J. 2019. *"ETTÄ TIETÄÄ MISSÄ ON MENOSSA" Oppimisen ja osaamisen arviointi perusopetuksessa ja lukiokoulutuksessa ["TO KNOW WHERE IT IS GOING" Assessment of learning and competence in primary and secondary education]*. Publications of the Finnish Education Evaluation Centre.

Au, W. 2007. High-stakes testing and curricular control: a qualitative metasynthesis. *Educational Researcher*, *36* (5), 258–267. https://doi.org/10.3102%2F0013189X07306523.

Brown, G. T. L., A. Gebril, & M. P. Michaelides 2019. Teachers' conceptions of assessment: A global phenomenon or a global localism. *Frontiers in Education, 4*. https://doi.org/10.3389/feduc.2019.00016.

Brown, G. T. L., & A. Remesal 2017. Teachers' conceptions of assessment: comparing two inventories with Ecuadorian teachers. *Studies in Educational Evaluation, 55*, 68–74. https://doi.org/10.1016/j.stueduc.2017.07.003.

Cheng, L. 2005. *Changing language teaching through language testing: A washback study*. Cambridge University Press.

Coombs, A., C. DeLuca, D. LaPointe-McEwan, & A. Chalas 2018. Changing approaches to classroom assessment: an empirical study across teacher career stages. *Teaching and Teacher Education, 71*, 134–144. https://doi.org/10.1016/j.tate.2017.12.010.

Davison, C. & C. Leung 2009. Current issues in English language teacher-based assessment. *TESOL Quarterly, 43* (3), 393–415. https://doi.org/10.1002/j.1545-7249.2009.tb00242.x.

FNBE 2003. *National Core Curriculum for upper secondary schools*. Helsinki: National Board of Education.

FNBE 2016 *National Core Curriculum for general upper secondary schools 2015*. Helsinki: National Board of Education.

FNBE 2017. *Opettajat ja rehtorit Suomessa 2016: raportit ja selvitykset 2017:2*. https://www.oph.fi/fi/tilastot-ja-julkaisut/julkaisut/opettajat-ja-rehtorit-suomessa-2016.

Fulcher, G. 2012. Assessment literacy for the language classroom. *Language Assessment Quarterly, 9* (2), 113–132. https://doi.org/10.1080/15434303.2011.642041

Hayward, L. & N. Hedge 2005. Travelling towards change in assessment: policy, practice and research in education. *Assessment in Education: Principles, Policy & Practice, 12*(1), 55–75. https://doi.org/10.1080/0969594042000333913.

Hildén, R. & B. Fröjdendahl 2018. The dawn of assessment literacy – exploring the conceptions of Finnish student teachers in foreign languages. A*pples – Journal of Applied Language Studies, 12* (1), 1–24. https://doi.org/10.17011/apples/urn.201802201542.

Hill, K. 2017. Understanding classroom-based assessment practices: A precondition for teacher assessment literacy. *Papers in Language Testing and Assessment, 6* (1), 1–17.

Inbar-Lourie, O. 2016. Language Assessment Literacy. In Language Testing and Assessment. Cham: Springer International Publishing, 1–14. https://doi.org/10.1007/978-3-319-02326-7_19-1.

Lakkala, M. & L. Ilomäki 2013. *Lukioiden valmiudet siirtyä sähköiseen ylioppilastutkintoon: kahden lukion tapaustutkimus* [Readiness of upper-secondary schools to move to the digital Matriculation Examination: A case study of two upper-secondary schools]. Vantaan kaupunki, Sivistystoimi. http://hdl.handle.net/10138/39406.

Lam, R. 2015. Language assessment training in Hong Kong: Implications for language assessment literacy. *Language Testing, 32* (2), 169–197. https://doi.org/10.1177%2F0265532214554321.

Lam, R. 2016. Assessment as learning: examining a cycle of teaching, learning, and assessment of writing in the portfolio-based classroom. *Studies in Higher Education, 41* (11), 1900–1917. https://doi.org/10.1080/03075079.2014.999317.

Looney, A., J. Cumming, F. van der Kleij, & K. Harris 2018. Reconceptualising the role of teachers as assessors: teacher assessment identity. *Assessment in Education: Principles, Policy & Practice, 25* (5), 442–467. https://doi.org/10.1080/0969594X.2016.1268090.

Luxia, Q. 2005. Stakeholders' conflicting aims undermine the washback function of a high-stakes test. *Language Testing, 22* (2), 142–173. https://doi.org/10.1191%2F0265532205lt300oa.

Mäkipää, T. 2021. *"I have never received feedback from my teachers." Students' perceptions of assessment and feedback practices in Finnish general upper secondary education* (Doctoral Dissertation). Helsingin yliopisto.

Poehner, M. E. & O. Inbar-Lourie (eds) 2020. *Toward a Reconceptualization of Second Language Classroom Assessment*. Springer International Publishing.

Stobart, G. & T. Eggen 2012. High-stakes testing – value, fairness and consequences. *Assessment in Education: Principles, Policy & Practice, 19* (1), 1–6. https://doi.org/10.1080/0969594X.2012.639191.

Tarnanen, M. & A. Huhta 2008. Interaction of language policy and assessment in Finland. *Current Issues in Language Planning, 9* (3), 262–281. https://doi.org/10.1080/14664200802139547.

Tarnanen, M. & A. Huhta 2011. Foreign language assessment and feedback practices in Finland. In D. Tsagari & I. Csépes (eds) *Classroom-based language assessment*. Frankfurt am Main: Peter Lang, 129–146.

Taylor, L. 2013. Communicating the theory, practice and principles of language testing to test stakeholders: Some reflections. *Language Testing, 30* (3), 403–412. https://doi.org/10.1177%2F0265532213480338.

Tsagari, D. & K. Vogt 2017. Assessment literacy of foreign language teachers around Europe: research, challenges and future prospects. *Papers in Language Testing and Assessment, 6* (1), 41–63.

Välijärvi, J., N. Huotari, P. Iivonen, M. Kulp, T. Lehtonen, H. Rönnholm, … S. Ohranen 2009. *Lukiopedagogiikka. Koulutuksen Arviointineuvoston Julkaisuja, 40*. https://karvi.fi/app/uploads/2014/09/KAN_40.pdf.

Van der Kleij, F. M., J. J. Cumming, & A. Looney 2018. Policy expectations and support for teacher formative assessment in Australian education reform. *Assessment in Education: Principles, Policy & Practice, 25* (6), 620–637. https://doi.org/10.1080/0969594X.2017.1374924.

Verhoeven, J. C. & G. Devos 2005. School assessment policy and practice in Belgian secondary education with specific reference to vocational education and training. *Assessment in Education: Principles, Policy & Practice, 12* (3), 255–274. https://doi.org/10.1080/09695940500337231.

Vlachou, M. A. 2018. Classroom assessment practices in middle school science lessons: a study among Greek science teachers. *Cogent Education, 5* (1). https://doi.org/10.1080/2331186X.2018.1455633.

Wall, D. & T. Horák 2006. *The impact of changes in the TOEFL examination on teaching and learning in Central and Eastern Europe: phase 1, the baseline study*. TOEFL Monograph Series, MS – 34. Princeton, NJ: Educational Testing Service. Retrieved from https://www.ets.org/Media/Research/pdf/RR-06-18.pdf.

Wiliam, D. & S. Leahy 2015. *Embedding Formative Assessment: Practical Techniques for K-12 Classrooms*. Learning Sciences International.

Xu, Y. & G. T. L. Brown 2016. Teacher assessment literacy in practice: a reconceptualization. *Teaching and Teacher Education, 58*, 149–162. https://doi.org/10.1016/j.tate.2016.05.010.

Appendix:

Questionnaire items discussed in the study

1. Matriculation Examination is in the transition period to become fully digitised: What do you think about this change as far as English is concerned? 1 = I am sceptical; 5 = I am enthusiastic.
2. What do you think are the positive aspects of this change? What problems (if any) related to this change do you envision?
3. The new National Core Curriculum (opetussuunnitelman perusteet) emphasises assessment for and as learning. What do you think about this change? 1 = I am sceptical; 5 = I am enthusiastic. Skip if you don't have an opinion on this matter.
4. Please elaborate on your above answer (skip if you don't have an opinion on this matter).
5. What affects your choice of classroom assessment? 1 = not at all; 2 = very little; 3 = to some extent; 4 = to a notable extent; 5 = a lot
   - The National Core CurriculumLocal (municipal) and school curriculum
   - Matriculation Examination
   - Your own teaching experience
   - Academic seminars and workshops
   - Students' preferences and expectations
   - Textbooks
   - Discussions with colleagues
   - Anything else that affects your choice of classroom assessment? (please briefly elaborate)
6. Who assesses students' knowledge in your classes? 1 = never; 2 = rarely; 3 = from time to time; 4 = quite often; 5 = often
   - You (teacher assessment)
   - Students themselves (self-assessment)
   - Other students (peer assessment)
7. What types of assessments are used in your classes? 1 = never; 2 = rarely; 3 = from time to time; 4 = quite often; 5 = often
   - tests and practice exams
   - group work (e.g., project work)
   - oral tests (e.g., interviews)
   - dialogues
   - essays and compositions
   - portfolios
   - observation
   - Any other kind of assessment? (please briefly elaborate)