

Students' performance in ethics assignments in the Finnish Matriculation Examination 2017–2021

Mika Perälä & Eero Salmenkivi



Nordidactica

- Journal of Humanities and Social Science Education

2022:2

Nordidactica – Journal of Humanities and Social Science Education

Nordidactica 2022:2

ISSN 2000-9879

The online version of this paper can be found at: www.kau.se/nordidactica

Students' performance in ethics assignments in the Finnish Matriculation Examination 2017–2021

Mika Perälä & Eero Salmenkivi

University of Helsinki

Abstract: What is difficult in ethics teaching for general upper secondary students? Can they achieve as good results in metaethics as in normative ethics? These questions should not be addressed without consideration of the various traditions of ethics teaching. Finnish students complete their studies in general upper secondary school by taking the Matriculation Examination. In recent years, a growing number of students has chosen to take an exam in philosophy but there is no systematic study on how students perform in the philosophy exam. This study is focused on those assignments within philosophy that concern ethics. How do students meet the key goals of ethics education defined in the Finnish National Core Curriculum 2015? The material of the study consists of the curriculum, ethics assignments, grading guidelines and the performance statistics in 2017–2021. The main observation is that the average performance is better in assignments in normative ethics than in metaethics, but in neither area of ethics does the average performance surpass half of the maximum points that is considered the threshold for a satisfactory performance. Implications for ethics teaching? We recommend that normative considerations be combined with metaethical considerations when necessary. However, we issue this recommendation having in mind the teaching of ethics in a philosophy classroom. The question for further research is whether the recommendation can also be reasonably applied to the teaching of ethics in other settings.

KEYWORDS: ETHICS, PHILOSOPHY, MATRICULATION EXAMINATION, CORE CURRICULUM, FINLAND, NORMATIVE ETHICS, METAETHICS, AVERAGE PERFORMANCE, ASSIGNMENT, GRADING GUIDELINES

About the authors: Mika Perälä is Docent in Theoretical Philosophy at the Department of Philosophy, History and Arts Studies, University of Helsinki. His current research focuses on the history of philosophy, especially Aristotelian logic, psychology, and ethics, as well as the philosophy of education, and the teaching of ethics.

Eero Salmenkivi is University Lecturer in Didactics of Philosophy and Culture, Worldview and Ethics, and Docent in Didactics of Philosophy at the Faculty of Educational Sciences, University of Helsinki. His current research focuses on the philosophy, worldview and human rights education, and the question of powerful knowledge in education.

The authors are full members of the Finnish Matriculation Examination Board.

Introduction

In Finland, the main responsibility for teaching ethics in general upper secondary education lies with philosophy, although many other subjects also touch upon some ethical issues. Finnish students complete their studies in general upper secondary school by taking the Matriculation Examination in at least five subjects. In recent years, a growing number of students has chosen to take an exam in philosophy. However, there is no systematic study on how students perform in the philosophy exam in general or in the ethics assignments in particular.

The present study is a first step in acquiring research-based knowledge of students' performance in the philosophy matriculation exam. We confine ourselves to studying those assignments within philosophy that concern ethics. We discuss the question of how students meet the key goals of ethics education that are defined in the Finnish National Core Curriculum for General Upper Secondary Education 2015 (FNCC 2015). The matriculation exams in 2017–2021 are based on this document.¹ Our main observation is that the average performance, according to the matriculation exam statistics, is better in assignments in normative ethics than in metaethics, but in neither area of ethics does the average performance surpass half of the maximum points that is considered a threshold for a satisfactory but not yet good performance.² An average student attains that threshold in only three assignments out of the 17 studied ones. By an average student, we refer to a student who shows average performance in the ethics assignments studied. The average student is merely a statistical construction. To avoid using excessively technical language, we occasionally apply this term in place of the technical term 'average performance'.

As our research material, we use, in addition to FNCC 2015, the assignments and grading guidelines of philosophy matriculation exams during 2017–2021. We also analyse the numeric data about how popular various assignments have been, and how many points an average student has received from them. The assignments and grading guidelines are published online in Finnish and Swedish.³ Our interpretation of the material is partly based on our experience as the chairs of the Philosophy Division in the Matriculation Examination Board.⁴ In this capacity, we have been involved in designing the philosophy assignments and their grading guidelines. We have also been

¹ The assignments up until Spring 2020 are also compatible with FNCC 2003. Even if there are some differences between the two curricula, especially in metaethics, the exam designer team did not consider it necessary to produce alternative assignments for each curriculum during the transition period when the two curricula were both in operation. It is worth noting that FNCC 2003 is more comprehensive with respect to metaethical contents than FNCC 2015. We shall specify the differences in Section 'The Matriculation Examination'.

² The terminology is not uniformly set by the Matriculation Examination Board. In this article, we call an answer that reaches the threshold of 50% of the maximum points satisfactory, and an answer that reaches the 75% threshold good.

³ For the philosophy matriculation exams in Swedish, see <https://svenska.yle.fi/abimix/filosofi>; for the grading guidelines in Swedish, see <https://www.ylioppilastutkinto.fi/sv/studentexamen/beskrivningar-av-goda-svar>.

⁴ Eero Salmenkivi acted as the chair in 2016–2018, Mika Perälä in 2019–2021.

involved in grading students' answers to the assignments. Since the students' results are somewhat disappointing and suggest that the average performance in most of the assignments studied falls short of the threshold of satisfactory performance, we wish to consider some hypotheses as to why this is the case. The following three hypotheses regarding the curriculum and the teaching at school will be discussed: teaching does not keep pace with the changing curriculum, the assignments require more than a student is typically able to learn at school, and the design of the exam reflects an interpretation of the curriculum that is not shared by all teachers and textbooks. In addition to these hypotheses, we make several conjectures as to how the nature of the assignments themselves may affect an average student's performance. We shall analyse the area, structure, material, and the command words of each assignment. Even if we cannot verify or falsify these hypotheses and conjectures in the present study, we hope that our study suggests some directions for further research on the matter.⁵

Philosophy and ethics in general upper secondary education

In Finland, the status of philosophy as a national subject has varied over time. The status of the ethics course has also varied over the years. A brief survey of the recent history helps to understand the context in which ethics is taught. In our view, it also helps to explain in part why an average student, following the FNCC 2015 syllabus, performs slightly better in normative ethics than in metaethics.

Philosophy regained the status of a compulsory national subject in FNCC 1994 (Tomperi 2017). A course in ethics was naturally a part of philosophy, but it was merely optional, which is why very few students completed the course (Turunen et al. 2011, p. 43). The compulsory course in ethics was taught in religious education and its alternative secular subject, which was at that time even called ethics in English translation (in Finnish 'elämäkatsomustieto', in Swedish 'livsåskådningskunskap'; FNCC 1994). The ethics course within philosophy remained optional in FNCC 2003. The approach to teaching philosophy was chiefly academic, while the ethics course aimed at covering the most basic concepts and theories in metaethics.

There was a significant change in teaching ethics when the Finnish Government, in revising the number of lesson hours for general upper secondary school in 2014, decided that the compulsory ethics course should be transferred from religious education to philosophy from August 2016 on. The main argument for this change was that the teaching of ethics is too scattered because it is given either as part of religious education with ten different syllabi for different Christian churches and other religions or as part of the alternative secular subject (see e.g. Salmenkivi et al. 2022). The Minister of

⁵ It would be illuminating to compare the average performance in philosophy with that in fairly similar subjects, such as history, as well as in relatively different subjects, such as biology. However, we have decided not to do so in this study because that would have complicated our research setting considerably. Even so, it can be noted that the average performance is somewhat weaker in philosophy than in history or biology. It is a matter for future research to explain why that is the case.

Education 2013–5, Krista Kiuru, argued that all the Finnish upper secondary school students should have essentially the same basis for ethical knowledge and skills. Hence, ethics should be studied within philosophy, even if various syllabi of religious education and the secular alternative to it still address some ethical issues as well.

As a result of the revised lesson hours, the Finnish National Board of Education issued a new national core curriculum, FNCC 2015, which includes two compulsory courses in philosophy: an introductory course and a course in ethics. Additionally, there are two optional courses: one in social and political philosophy, and another in theoretical philosophy. The change concerned not only the status of the ethics course, but also the approach taken to teaching philosophy in general. The new approach can be characterised as a step in the direction of what is called philosophy for children (Lipman 2003; Juuso 2007; Fisher 2013). This is because FNCC 2015 puts less weight on acquiring comprehensive knowledge of philosophical theories and schools, and instead places greater emphasis on acquiring the ability to conceptualise and address issues in one's own life, school and society at large, including, for example, art and literature.

In the case of ethics, the new approach means that the starting point of teaching is the ethical issues that students face in their own lives, school and society. Teaching is supposed to provide concepts and theories to address these issues. In this approach, then, neither ethical theories nor history of ethics are truly studied unless they help in taking a useful perspective on the issues in question. As a result, the field of ethics is not covered systematically from an academic perspective, but only to the extent that is relevant to the goals of FNCC 2015. Whereas FNCC 2003 includes no less than five core contents in metaethics (quoted in Section 'Material and method' below), FNCC 2015 mentions only two, 'the nature of morality as a system of norms and its distinction from systems based on juridical norms and conventions' and 'moral relativism' (p. 172).⁶ FNCC 2003 comprises no explicit metaethical objective, but FNCC 2015 claims that 'the student perceives the nature of normative statements and their relationship with descriptive statements and is able to justify the conceptions of good and right' and that 'the student is able to justify the binding nature of morality' (p. 171). The changes made to the curriculum are based on the assumption that the two remaining metaethical contents and the two metaethical objectives are the most relevant in conceptualising the ethical issues in a student's daily life.

The question arises, however, whether the changes in the curriculum could explain students' weaker performance in metaethics. As noted, one of our hypotheses is that teaching does not keep pace with the changing curriculum. As a general rule, this hypothesis might have some explanatory value. It may be that many teachers have not yet adjusted their teaching to the new philosophy for children approach required by FNCC 2015,⁷ but we do not as yet have any empirical evidence for this. Even if that

⁶ In citing from FNCC 2003 and 2015, we use the English translations that are published by the Finnish National Board of Education.

⁷ On this point, see especially the analysis of 2018A 2 in Section 'Assignments in metaethics' below.

were the case, we do not think that the matter is so simple. That is because it may be that, if teachers tend to be conservative in matters regarding the curriculum and their teaching, they still put some emphasis on metaethics in the spirit of FNCC 2003. If so, a weaker performance in metaethics cannot be explained by reference to the lack of teachers' emphasis on metaethics. Second, one might think that, regardless of what the two curricula say about metaethics, the assignments in metaethics require more than a student is typically able to learn at school. This is our second hypothesis. Again, we do not have empirical evidence because the matriculation exam results do not tell us what a student is typically able to learn at school: they only tell us about the performance of those who take the philosophy exam, which is around 5% of all those who sit the matriculation exam. Our third hypothesis is this: It may be that the design of the exam reflects an interpretation of the new curriculum that is not shared by all teachers and the authors of textbooks. In other words, it is possible that many students are taught to meet somewhat different objectives than those that the exam designers have in mind. Again, we do not have any empirical evidence for this hypothesis. However, the hypothesis has to be taken into consideration because, first, teachers of philosophy have various philosophical backgrounds, which affect their approach and, second, exam designers have a special role in interpreting the curriculum. We shall describe that role in the following section.

The matriculation examination

The Finnish Act on the Matriculation Examination defines the examination as follows:

A matriculation examination is completed at the end of upper secondary education Students who complete the examination have acquired the knowledge and skills specified under the curriculum for general upper secondary education and attained a level of maturity consistent with the objectives of general upper secondary education. ... Successful completion of the matriculation examination provides general eligibility for further studies in institutions of higher education (Act on the Matriculation Examination, 502/2019, Section 1)

The matriculation examination thus combines the curriculum-based secondary education exit examination (see Bishop 1998; Au 2007) with general eligibility for higher education. Entrance to higher education is very competitive in Finland (Jokila, Haltia & Kosunen 2021, p. 592) and the recent reforms in educational law and administration have increased the importance of the matriculation examination in higher education entrance (OKM 2016, p. 37).

The Finnish matriculation examination 'comprises tests in language and literature in the mother tongue, the second national language, foreign languages, mathematics, humanities and sciences' (Act on the Matriculation Examination, 502/2019, Section 3). Tests are arranged twice a year and each test takes a maximum of six hours unless the student has been given extra time due to a disability. The tests in philosophy and other humanities include six basic assignments (graded on a 0–20-point scale) and three more

demanding ones (graded on a 0–30-point scale). The latter assignments are more demanding in that they may include more extensive material, relate to more than one course in the curriculum, and require the ability to apply concepts and theories in a context that is not familiar. A student is required to complete five assignments of which at most two can yield 30 points. The maximum points in the exam, then, are 120. It is up to the students which assignments they choose to complete. The choice is made during the exam. Prudent students naturally choose those assignments that they feel most competent to tackle.

Since September 2016, the philosophy matriculation exam is implemented digitally. A digital exam makes it possible for an assignment to comprise more extensive material than in the case of traditional paper tests, which were confined to a four-page question booklet. The digital exam, then, may be used to test skills in analysing rather complex material. In the case of ethics, such materials include film clips and excerpts from classical philosophical texts. This diversity of material can make the assignment more interesting to students. In our view, this is an advantage of the digital exam, and it explains in part why the philosophy exam has become increasingly popular in recent years. Figure 1 below shows how many students have taken the philosophy exam during the period of the present study.

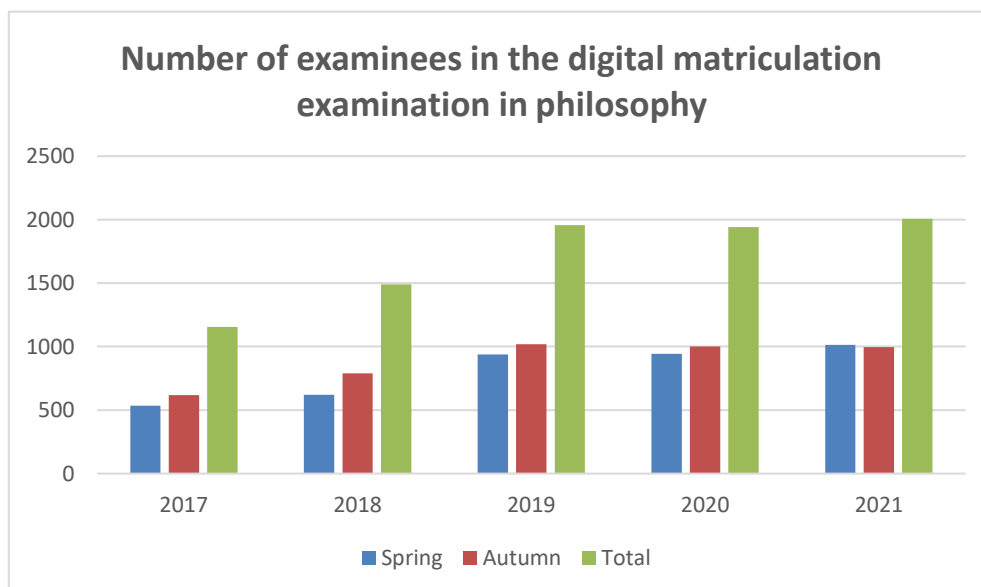


FIGURE 1

Number of examinees in the digital matriculation examination in philosophy

At this point, we should like to make a general comment on the relationship between the matriculation exam and the curriculum. In principle, as implied above, the matriculation exam is supposed to test whether the students meet the goals of the curriculum. However, it has been argued that, since the curriculum defines the competences of each subject on a very general level, the designers of the exam find themselves in a situation in which they, in fact, are the major interpreters of the goals of teaching and learning. According to Löfström et al. (2010, p. 9), '[p]reparing the

exam questions they [exam designers] effectively make authoritative statements about what the core concepts and competences in the subject are'. This argument is based on a study on the situation in social studies during 2000–2010, but in our experience, it also applies to philosophy with certain qualifications during 2017–2021.

In philosophy, the team designing the exam consists of five to seven experts, which is comparable to the team in social studies. Like the designers of the social studies exam, they receive feedback from philosophy teachers after the exam. Even if the feedback plays no official role, it is an important indication of whether the designers of the exam have succeeded in giving a reasonable interpretation of the goals of the curriculum. The feedback also helps to develop the design of the exam. The communication, then, is not only one-way. However, this does not rule out the possibility that the exam designers may interpret the curriculum somewhat differently from the teachers of philosophy and the authors of textbooks. This is partly because the exam designers cannot be school teachers due to the exam secrecy. Their background is mostly in academia. This is evident in some interpretations of the exam designer team. For example, the present team has chosen to use excerpts from the classics of philosophy and contemporary philosophers as the materials for the assignments. Many teachers find excerpts from original texts too difficult for students. The key question, then, is whether their use can be justified by reference to the curriculum. There are reasons for and against. We shall briefly comment on the matter in the case of one assignment in metaethics (2018S 4).

The requirements that the Matriculation Examination Board places on both the exam designing teams and the assignment grading teams have become stricter. This is due to the Act on the Matriculation Examination (502/2019) that gives the examinee or the guardian of the examinee 'the right to obtain information on how the criteria for marking are applied to the candidate's performance of a test' (Section 18). Furthermore, the act decrees that examinees 'must be given an opportunity to see the marking of their written or test performance assessed in other ways [e.g. by oral as well as written tests]' (Section 18).⁸ These rights help the examinee or the guardian to issue a well-informed and considered request for a review of the grading of the test performance (see Section 22 of the act). The new act, then, requires greater transparency in the criteria for grading. As a consequence, the grading guidelines have to be composed with more detail than before. In philosophy, that means, for example, that each assignment is explicitly linked to a certain core content of the curriculum. With this background in mind, it can be concluded that the interpretation of the core contents and objectives of the curriculum is by no means arbitrary, even if they are somewhat ambiguous.

Material and method

The primary material of this study consists of four kinds of documents: (i) the objectives and core contents of the ethics course in FNCC 2015, (ii) 20-point ethics assignments in the philosophy matriculation exam during 2017–2021, (iii) the grading

⁸ For the sake of clarity, we have slightly modified the official translation here.

guidelines for these assignments, and (iv) the statistics about how many students chose to complete the assignments as well as how many points, on average, they received from them. We will limit our study to the ethics assignments that are graded on a scale of 0–20 points, because these assignments do not extend beyond ethics. There are also ethics assignments that are graded on a scale of 0–30 points, but they often contain parts that extend to different areas of philosophy and thereby to different courses. We deal with the exams from Autumn 2017 to Autumn 2021 because the philosophy assignments in these exams comply with FNCC 2015.⁹ In each of these exams, with one exception, there are two 0–20-point ethics assignments.¹⁰ Thus, we analyse altogether 17 assignments.

The assignments analysed are either single-structured or multi-structured. By a single-structured assignment, we refer to an assignment that requires only one answer, an essay, which is graded on a scale of 0–20 points. By a multi-structured assignment, we refer to an assignment that has more than one part and therefore requires more than one answer. Most of the multi-structured assignments that we analyse have two parts. The parts can be graded on different scales. In most of our cases, the scale for each part is 0–10 points, but in one case, the first part is graded on a scale of 0–5 points, and the second on a scale of 0–15 points. All assignments and parts of assignments with a maximum of 10 or more points require an answer that is an essay in form. We have one three-part assignment (4, 6 and 10 points) and one five-part assignment (5 times 4 points).

In interpreting the assignments and their grading, we use the grading guidelines that the Matriculation Examination Board publishes on their website after the exam. Each exam is doubly assessed, first by a teacher and then by an examiner on the Matriculation Examination Board. If the grading varies too much between the teacher and the examiner, another examiner from the Matriculation Examination Board will make a

⁹ They also comply with FNCC 2003 until Spring 2020, but we shall not discuss the differences between the two curricula systematically because that is not relevant for the main arguments of this study. However, we shall make some observations on the differences in metaethics because that may explain the average performance in some assignments. The most recent curriculum is FNCC 2019. However, the first exam that is based on FNCC 2019 will be held in Spring 2023.

¹⁰ The exception is 2020A in which there is only one assignment. On the face of it, assignment 5 in 2020A entitled 'Climate change and global responsibility' appears to be an assignment in ethics because the environment is one of the contents of the ethics course in FNCC 2015. However, we do not categorise this assignment primarily as an ethics assignment because the assignment deals with the distribution of responsibility among individuals, society and the global community, which is a matter of social and political philosophy in FNCC 2015. In particular, the assignment relates to the objectives that a student 'learns to analyse justice, exercise of power, and division of duties in the activities of individuals, communities, and institutions' (FNCC 2015, p. 172) and that a student 'is able to apply what he or she has learned to current societal questions' (p. 172). Furthermore, the assignment relates to the core contents 'political ideals and their implementation: freedom, equality and solidarity; conservatism, liberalism, and socialism; nation states and the global perspective' (p. 172) and 'current questions in social and political philosophy [...] building a sustainable future' (p. 172).

third assessment. Published grading guidelines are meant to assist teachers and examiners in applying equal standards for grading.

In philosophy, the contents of guidelines are complicated because it is possible that even excellent answers vary in many ways. The guidelines tend to be rather long and detailed compared to some other subjects. A guideline consists of general guidance that is basically the same for all exams, and of specific guidance concerning each assignment. The specific guidance contains three parts in the case of each assignment or an independent part of each assignment: i) a summary of the main philosophical contents relevant to the assignment that is meant to give each teacher and examiner an idea of the possible directions of an answer; ii) one to three thresholds of grading, depending on the maximum points; and iii) an account of the core contents to which the assignment relates in FNCC 2015. The thresholds of grading are mostly given in terms of verbs that express the key competences or abilities that a student should show in answering the assignment. In cases in which the threshold defines a performance worth half of the maximum points or more, these verbs match the command verbs by which the assignment is given. Perälä and Salmenkivi (2020) have analysed the command verbs in the philosophy matriculation exam in relation to Anderson and Krathwohl's (2001) version of Bloom's taxonomy.

Even if the grading guidelines do not explicitly link each assignment to one or several objectives of FNCC 2015, we show in this article that this can be done based on the threshold descriptions. According to our interpretation, then, the grading guidelines enable us to link each assignment to both the objectives and core contents of FNCC 2015. In this way, we show what competences each assignment is supposed to test.

Since we focus on the average performance in ethics assignments, we will pay special attention to the threshold that defines a 50% performance in the assignment. A student's answer which reaches the middle threshold, but does not go beyond it, replies to the question or completes the task that is given in the assignment in an appropriate way in terms of both philosophical content and competence. Additionally, the answer has to satisfy the formal requirements that the matriculation exam sets for writing. For example, if the assignment asks the examinee to analyse an argument by Bernard Williams (see 2018S 4.1), the description of the threshold includes the command verb of the assignment: 'The examinee analyses Williams's argument...' (2018S 4.1 guidelines).¹¹ The answer could be better in terms of content and competence, but it shows, first, that the examinee understands the assignment correctly, and second, that he or she is able to answer the assignment adequately in terms of content and competence. A formal requirement is, for instance, that if an essay is required (as is the case when the maximum points are 10 or more), the answer should have the structure of an essay. In this sense, then, the answer is satisfactory. We call an answer that goes beyond this threshold and reaches the next threshold of 75% of the maximum points good. In this case, the threshold description includes the command verb, with some qualification indicating good performance: for example, 'The examinee analyses

¹¹ The English translations from the guidelines are by the authors of this article.

Williams's argument so as to show clearly...' (2018S 4.1 guidelines). There is no further threshold beyond that in the grading guidelines, but the guidelines include a brief description of the merits that an answer that is called 'outstanding' shows. By contrast, an answer that fails to reach the middle threshold is lacking in either or both of the two foregoing respects: it shows a deficient understanding of the assignment, or it is lacking in content or competence, or both. In cases where there is a threshold of 25% of the maximum points, the threshold separates poor performance from unsatisfactory performance: poor performance shows neither proper understanding of the assignment nor is satisfactory in terms of content and competence. The description of the lowest threshold does not typically include the command verb of the assignment: for example, 'The examinee is able to describe the structure of Williams's argument in the main' (2018S 4.1 guidelines). This underlines the fact that the examinee is unable to meet the requirements of the assignment: he or she takes some steps in the right direction, but does not show competence in analysis.

The foregoing considerations about the material indicate that the approach of the present study is analytic and interpretive. In essence, we analyse the numeric data about the average performance and interpret what it tells us about an average student's proficiency in ethics assignments. We concede that our interpretations remain somewhat speculative and constitute at best plausible explanations. However, we hope that our interpretations are nonetheless illuminating as a starting point for further research. Our aim, then, is to do the necessary groundwork for a more extensive study that analyses not only the average performance but also some other relevant numeric data about students' performance. In connection with such a study, it might be interesting, for example, to interview teachers and students on the ways in which metaethical questions are discussed in class (if at all), on whether they apply materials such as those used in matriculation exams, and on what kind of didactic approaches are used in class. Another interesting line of research would be to see if the philosophy textbooks used in schools match the type of questions asked in the matriculation exam.

As mentioned in the introduction of this article, what is special about our study is that we have been involved in producing both the assignments and the grading guidelines. Additionally, we have been involved in grading examinees' assignments as chairs and members of the grading team. This raises the question whether we are biased when conducting proper research into this topic. One might worry, for example, whether we can approach the questions that we raise with a sufficient degree of objectivity. Are we, moreover, at risk of overlooking some alternative explanations for the students' difficulties in the exams, difficulties relating to the nature of the curriculum, or to the wording of the exam questions, or to the grading guidelines. This concern is well justified, and we have kept it in mind in all phases of this study to avoid unwarranted bias as far as possible. That said, we acknowledge that a researcher with a different background might suggest some plausible explanations that we have been unable to discover. However, we should like to emphasise that we do not aim to provide a comprehensive and conclusive interpretation of the material: the suggestions that we make are merely meant to offer a starting point for further research. Moreover, the fact that we have been involved in producing all the material studied can be considered an

advantage rather than a disadvantage for the present study. An informative and effective analysis of the material could not be conducted by a researcher who has little or no experience of grading matriculation exam assignments. Moreover, it is generally acknowledged that it takes several years to acquire the experience that an examiner needs, which is why an easy alternative way of carrying out the present study is not available.

The approach suggested above is based on the assumption that the ethics assignments in the matriculation exam indeed test whether a student achieves the goals of ethics as they are defined in FNCC 2015. We have adduced some considerations in support of this assumption. We have also spelled out our own role in interpreting FNCC 2015. Defending the assumption in more detail, however, is beyond the scope of our article. A more detailed defence would consider the philosophy exam in its entirety. It is worth noting, however, that, if the assumption were not true in general, there would be no point in implementing the philosophy matriculation exam in the first place. Nonetheless, it is possible that some particular assignments, according to certain interpretations, are misconstrued and do not test the goals of the curriculum. However, we do not think that there are such assignments in the material of this study. With these qualifications, we assume that the ethics assignments studied adequately test the goals of FNCC 2015.

According to FNCC 2015, the objective of the ethics course is that the student:

(O1) gets acquainted with the key concepts, questions and theories of philosophical ethics as well as the basics of environmental philosophy

(O2) perceives the nature of normative statements and their relationship with descriptive statements and is able to justify the conceptions of good and right

(O3) is able to structure meaning in his or her life and life choices with the help of philosophical concepts

(O4) is able to justify the binding nature of morality and apply analyses of philosophical concepts and consistent argumentation in morality

(O5) is able to analyse and evaluate actions ethically as well as structure his or her own moral solutions and assessments with the tools of philosophical ethics. (FNCC 2015, p. 171)

In our categorisation, objective O1 is general, that is, applicable to all fields of philosophy, whereas objectives O2 and O4 chiefly relate to metaethics, and objectives O3 and O5 to normative ethics, including what is referred to as applied ethics. Roughly speaking, normative ethics studies the grounds for morally justified action, whereas metaethics explores, for example, the nature of moral goodness in general as well as the possibility of moral truth and moral knowledge. In the main, this categorisation is based on the way in which academic philosophy is divided into different areas. The categorisation does not arise from our research material; it is thus theoretical in methodological terms. It should be noted, though, that objective O3 does not belong to normative ethics as it is usually understood in academia (see e.g. Shafer-Landau 2010). Instead, it belongs to what is called the philosophy of life. Since O3 is confined to a student's own life and his or her choices, it is difficult to assess in the matriculation exam, which is basically an academic test. Note also that the objectives as such do not

reveal how they are to be achieved. However, the general introduction and objectives of teaching philosophy FNCC 2015 clarify the approach to teaching. According to our interpretation, the curriculum suggests an approach that combines the philosophy for children approach with the basics of academic philosophy.

The core contents of the ethics course are the following:

(C1) morality and normative and applied ethics concerning moral issues; the basics of virtue, consequentialist, contractarian, legal, and deontological ethics

(C2) the nature of morality as a system of norms and its distinction from systems based on juridical norms and conventions, moral relativism

(C3) philosophical theories on the goodness and meaningfulness of life as well as a good way of living

(C4) ethics and the moral choices of the individual: interpersonal relationships and life choices

(C5) ethical questions concerning animals and the environment. (FNCC 2015, p. 172)

In our categorisation, contents C1, C4 and C5 pertain to normative ethics, and C2 to metaethics. Content C3 is somewhat difficult to categorise, but it is clearly closer to normative ethics than metaethics.

In our analysis, we will compare the students' ability to answer normative and metaethical assignments. However, it should be noted that the metaethical contents and objectives in FNCC 2015 are more condensed than those described in FNCC 2003. In addition to mentioning metaethics as a core content, the latter curriculum specifies the contents as follows: 'the different grounds for convictions concerning moral values and norms; the relation of morality vis-à-vis justice and religion; the roles of reason and emotion in moral convictions; objectivity and subjectivity in moral values and norms; the questions of rationality of ethical justification and the possibility of ethical truths' (FNCC 2003, pp. 177–178). The metaethical contents were reduced and the concept of metaethics removed when the course became compulsory when FNCC 2015 was adopted. The reason behind the reform was that metaethics was considered very difficult even for those students who chose the optional ethics course in philosophy. It was essential to make the content easier for a more varied student population.¹²

In FNCC 2015, the only metaethical contents are 'the nature of morality as a system of norms and its distinction from systems based on juridical norms and conventions', and 'moral relativism' (FNCC 2015, p. 172). The idea is to teach students the basic nature of morality and discuss the topical but philosophically problematic issue of moral relativism. The main objective related to these contents is that the student 'is able to justify the binding nature of morality and apply analyses of philosophical concepts and consistent argumentation in morality' (FNCC 2015, p. 171). These contents are

¹² One of the authors, Eero Salmenkivi, was mainly responsible for the curriculum reform in philosophy in 2015.

supposed to determine the core performance assessed in the metaethical assignments that we analyse.

Analysis

In the analysis to follow, we divide the assignments into three main categories: assignments in normative ethics, assignments in metaethics, and assignments that include elements of both, i.e., mixed assignments. Within each category, we further divide the assignments into those that are based on some material, for example, a text, and those that contain no material. For the sake of convenience, however, we make observations on the latter only after having considered the assignments according to the main categories. Furthermore, we classify the assignments by reference to the core contents and objectives of the ethics course in FNCC 2015. In considering the reasons behind the average performance in a single assignment, we also make some observations on the structure, command word, and the popularity of the assignment. In our experience, all these factors need to be taken into account in interpreting the statistical data about the average performance in a single assignment. Table 1 presents our analysis in a concise form.

TABLE 1

Analysis of assignments

Assignment	Area	Content	Objective	Examinees	Grading	Material	Structure	Command
2017A 2	NE	1, 3	1, 3	461/619	10.1	N/A	3, 4 + 6 + 10 p.	define, exemplify; discuss; analyse
2017A 4	NE	1, 4	5	409/619	9.7	video	1, 20 p.	analyse & evaluate
2018S 2	NE	1	1	328/621	10.8	N/A	2, 10 + 10 p.	explain
2018S 4	ME	2	2, 4	157/621	8.9	text	2, 10 + 10 p.	analyse, evaluate & explain, argue
2018A 2	ME	2	4	362/790	6.6	N/A	1, 20 p.	select & discuss
2018A 4	NE	5	5	607/790	9.7	two texts	1, 20 p.	evaluate
2019S 2	NE	3	1, 3	427/937	9.5	N/A	1, 20 p.	describe
2019S 4	NE & ME	1, 2, 4	1, 2, 5	707/937	8.3	video	2, 10 + 10 p.	evaluate & discuss
2019A 2	ME	2, 5	1, 2	699/1020	8.5	video	1, 20 p.	discuss
2019A 5	ME & NE	1, 2	1, 4-5	523/1020	9.7	N/A	2, 5 + 15 p.	exemplify, explain & discuss
2020S 2	NE	1	1, 4	349/943	8.0	N/A	2, 10 + 10 p.	explain, exemplify & consider
2020S 5	NE & ME	1, 2	2, 4-5	496/943	7.6	video	1, 20 p.	discuss
2020A 4	ME & NE	1-2	2, 5	505/1000	6.3	figure & video	2, 10 + 10 p.	explain & discuss
2021S 2	ME	2	2	694/1013	11.0	N/A	5, 5 x 4 p.	explain
2021S 4	NE	1	1, 5	729/1013	8.8	video	1, 20 p.	evaluate
2021A 2	NE	1	1, 5	837/995	9.3	N/A	1, 20 p.	discuss
2021A 4	NE	1, 4	1, 5	370/995	7.9	video	1, 20 p.	discuss

Abbreviations in the table and the text

Assignment is given by year followed by Spring (S) or Autumn (A) and by the ordinal number in the exam.

Area comprises normative ethics (NE) and metaethics (ME). Mixed assignments are given as NE & ME.

Content comprises the core contents of FNCC 2015: e.g., 1 in the table and C1 in the main text refer to the first core content of FNCC 2015.

Objective comprises the objectives of FNCC 2015: e.g., 1 in the table and O1 in the main text refer to the first objective of FNCC 2015.

Examinees includes two numbers: first, the number of examinees who completed the assignment in question, and second, the number of examinees who completed the entire exam.

Grading includes the average performance in the assignment in question.

Material refers to the type of material, or lack of material (N/A).

Structure includes two types of information: first, how many parts the assignment has, and second, what the maximum points for each part are.

It is worth noting that when we analyse the grading of assignments, we refer to the statistical average performance. Nevertheless, when we consider the possible reasons for the average performance, the points that we make are partly based on our own experience in grading examinees' answers to assignments. Our observation is that in the material of the present study, the average performance tends to be slightly stronger in assignments in normative ethics than in metaethics. Within each exam, with one exception (2021S), an average student performs better in an assignment in normative ethics than in metaethics. Likewise, within each exam, with no exception, examinees perform better in a mixed assignment than in a metaethics assignment. Figure 2 confirms these observations.

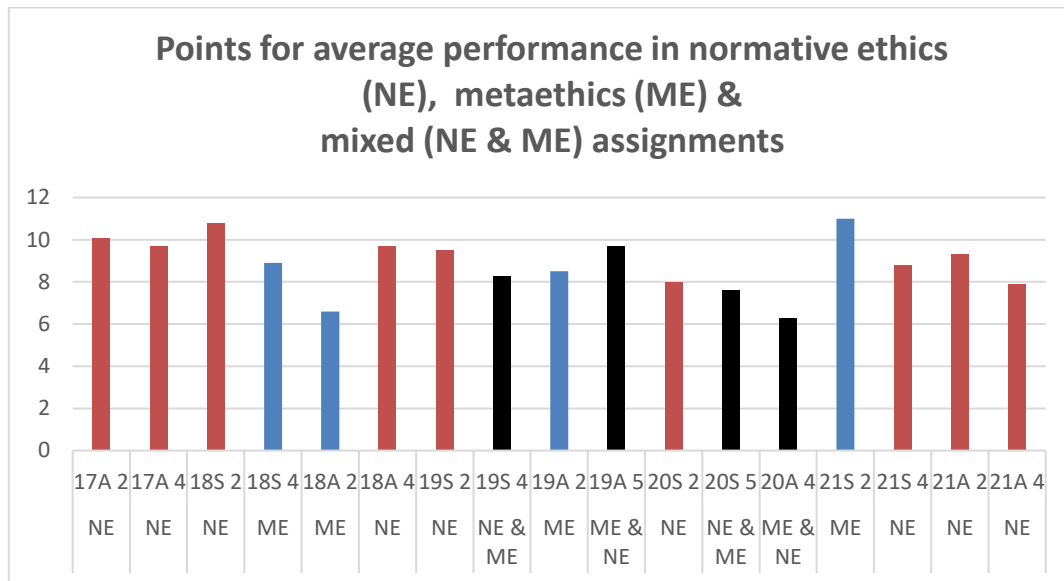


FIGURE 2

Points for average performance in normative ethics (NE), metaethics (ME) & mixed (NE & ME) assignments

In the following subsections, we first discuss assignments in normative ethics, second in metaethics, and third in mixed assignments. Within each section, we group assignments that are related to the same core contents and objectives in FNCC 2015. We conclude the analysis by comparing assignments with and without material.

Assignments in normative ethics

FNCC 2015 suggests that the teaching of ethics in Finnish general upper secondary school focuses on normative ethics. We argue that this is the most important curriculum-based reason why the average performance is stronger in the assignments in normative ethics than in the other assignments. Judging from upper secondary school ethics textbooks, a typical assignment in class is to consider whether a given action or choice is morally right or wrong. Even if normative moral theories such as deontological and consequentialist theories do not directly apply to single actions and choices, students are taught to consider the grounds of single actions and choices based on these theories.

In our material, there are two assignments (2021A 2 and 2021S 4) which test a basic understanding of moral theories (C1, O1) as well as the ability to ‘analyse and evaluate actions’ (O5). The first asks the examinee to discuss whether the actions of Robin Hood are morally justified (2021A 2), and the second asks him or her to judge whether the reasons that a figure called Thanos gives in a clip from the film *Avengers* can be considered utilitarian (2021S 4).¹³ According to the statistics, both assignments are very

¹³ The paraphrase that we give of each assignment in the main body of the article is as close as possible to the original wording in Finnish and Swedish. We give a translation of the assignment in a footnote. For the sake of readability, we have left untranslated or modified some technicalities (e.g. directions regarding the maximum length of an answer and the references to the materials).

popular, but the average performance is not quite satisfactory. In either case, an average student gets less than half of the maximum points: in the first case 9.3, and in the second 8.8. How can we interpret these figures?

Consider the assignment about Robin Hood. In the grading guidelines of this assignment, the middle threshold is given as follows: 'In an answer worth 10 points, the examinee adequately discusses the ethical justification for Robin Hood's action and presents reasons for or against it, applying some moral theory' (2021A 2 guidelines). The account of the core content ends with the following sentences:

It is not the most important matter in an answer to list the various theories in normative ethics and to point out what each of them would suggest in the present case. It is more relevant to discuss consistently Robin Hood's action and to consider grounds for and against the justification of his action. (2021A 2 guidelines)

It is easy to see from this account why an examinee may not reach the middle threshold cited above. The account alludes to two key problems: first, the examinee refers to several moral theories but gives a superficial account of them; secondly, and relatedly, he or she makes a shaky attempt to apply each of these theories to the case in hand. The examinee is unable to score satisfactorily if his or her attempt does not make clear how the theories apply to single cases. In our experience, a very common problem is that moral theories are applied without the necessary qualifications. For example, an examinee may fail to show that the theories are general in nature and that they do not directly determine what is to be done in a particular case. In fact, moral theories suggest principles of principles: they tell us what requirements good moral principles need to satisfy. For example, Immanuel Kant's categorical imperative tells us that a good moral principle must be generalizable. It is only in this way that they may justify a principle for action such as 'Do not steal'. In an average case, an examinee makes passing remarks on how Kantian deontology, Millian utilitarianism, or Aristotelian virtue ethics would address the issue—he or she may even touch upon all these theories—but fails to engage in a proper discussion.

A basic understanding of key concepts and moral theories (C1) is also required in several other assignments in our material. There are two double-structured assignments which do not go beyond content (C1) and the related objective (O1). The first assignment (2020S 2) requires the examinee (i) to explain the difference between act and rule utilitarianism as well as to give an example of the difference, and (ii) to consider how the Finnish moral philosopher Timo Airaksinen criticises utilitarianism in a given

2021A 2. Justification for stealing from the rich. According to English legend, Robin Hood was an outlaw who stole from the rich to give to the poor. Discuss whether the action of Robin Hood is morally justified. (20 p.)

2021S 4. An extreme utilitarian? In the film *Avengers: Infinity War*, directed by Anthony Russo and Joe Russo (2018), a fictitious character named Thanos plans to destroy half of the inhabitants of the universe. His aim is to provide the survivors with better chances to live. [Material: a film clip.] Evaluate whether the reasons that Thanos gives for his action are utilitarian. Make use of the film clip in your answer. (20 p.)

quotation from his textbook *Moraalifilosofia* (1987).¹⁴ The second assignment (2018S 2) asks the examinee to explain the key content of (i) the categorical imperative and (ii) the principle of utility as well as at least one philosophical problem related to each principle.¹⁵ An average student did better in the second assignment (10.8) than in the first (8.0). We suggest that there are two basic reasons for the difference in performance. First, in general, mere explaining based on a knowledge of key contents is cognitively less demanding than explaining that is combined with considering something new to a student. Second, in the present case, an average student could not properly distinguish between act and rule utilitarianism. It is difficult to determine why this is the case, but our conjecture is that in many schools, the teachers do not give sufficient attention to the distinction because, for example, the textbooks that they use are not entirely clear about the distinction. The conjecture is based on the considerations that the examiners of the philosophy exam put forward in a meeting after they had marked the exam in question.

There is a pair of assignments that test basic understanding of key concepts (C1) and competence in 'philosophical theories on the goodness and meaningfulness of life as well as a good way of living' (O3). One of them (2019S 2) is single-structured and asks the examinee to describe two different philosophically motivated views of the good life.¹⁶ The other one (2017A 2) is triple-structured and requires the examinee (i) to define the concept of virtue as well as to give examples of virtues, (ii) discuss briefly whether the most important contemporary virtues are the same as in the past, and (iii) analyse the relationship between virtues and the goal of life.¹⁷ In the first assignment,

¹⁴ 2020S 2. Varieties of utilitarianism. Utilitarianism is an orientation within ethics. There are many different varieties of utilitarianism in philosophical literature. 2.1. Explain what is meant by act utilitarianism and rule utilitarianism, and give an example that sheds light on the difference between the two varieties. (10 p.) 2.2. Philosopher Timo Airaksinen (b. 1947) claims in his treatise *Moraalifilosofia* (1987) that 'the maximisation of the common good (utilitarianism) may apply as a guideline in social politics as long as it is kept in mind that the rights of the minorities are not breached (rights-based ethics)'. He also writes that 'each theory of ethics that aims to analyse rights merely in terms of utility and to account for the validity of rights by reference to their usefulness is artificial'. Consider how Airaksinen criticises utilitarianism in these quotations. (10 p.)

¹⁵ 2018S 2. The principles of normative ethics. The categorical imperative and the principle of utility are well-known principles of ethics. 2.1. Explain the philosophical core content of the categorical imperative and at least one philosophical problem with the categorical imperative. (10 p.) 2.2. Explain the philosophical core content of the principle of utility and at least one philosophical problem with the principle of utility. (10 p.)

¹⁶ 2019S 2. The good life. Philosophers have presented different conceptions of the good life and of the meaning of life. These conceptions differ from each other in various ways. There are disagreements, for example, about the role of reason, will and emotions as well as the relationship between the individual and the community. Behind the disagreements there is often the question about the essence of man. Describe two different philosophically motivated views of the good life. (20 p.)

¹⁷ 2017A 2. Virtues. Virtue ethics flourished from antiquity up to the 18th century, though it has regained popularity in the past half a century. The relatively recent renaissance of virtue ethics is explained by the fact that it considers the conditions of the good life, the goal of life and moral education more comprehensively than other moral theories. 2.1. Define the concept of virtue. Illustrate the matter by giving examples. (4 p.) 2.2. Discuss briefly whether the most important

an average student achieved 9.5 points, in the second 10.1 points. The slight difference between the two may be explained by the fact that students typically perform better in a multi-structured assignment than in a single-structured assignment. That is because, even if they do not show a good understanding, they earn some points in each part of the assignment and thus manage to achieve a decent overall score.

A basic understanding of key concepts (C1) is also required in another pair of single-structured assignments which relate to interpersonal relationships and life choices (C4) and test competence in analysing and evaluating actions (O5). One of the two (2021A 4) asks the examinee to discuss what kinds of moral obligations belong to persons involved in a rape trial that is portrayed in a scene from the series *La otra mirada*.¹⁸ The other one (2017A 4) asks the examinee to analyse the reasons that the vicar Lauri Salpakari, in a Timo Haapasalo film based on the novel *Täällä Pohjantähden alla* (Under the North Star) by Väinö Linna, presents for his decision to expel his tenants from his property that they have taken good care of and even extended, and to evaluate his decision from an ethical point of view.¹⁹ In the first assignment, an average student scored only 7.9 points, in the second 9.7 points.

In the material of the present study, the last assignment in normative ethics belongs to what is called applied ethics (C5) and tests competence in analysing and evaluating actions (O5). The assignment (2018A 4) is single-structured and asks the student to evaluate the reasons for carnivorousism and its limitation to certain species of animals. For this assignment, two texts are given as material: a piece of news that eating cats and dogs is forbidden in Taiwan, and an anti-advertisement that opposes consuming cold cuts by presenting them in a critical fashion.²⁰

contemporary virtues are the same as in the past. (6 p.) 2.3. Analyse the relationship between virtues and the goal of life. (10 p.)

¹⁸ 2021A 4. Moral obligations. The Spanish TV series *La otra mirada*, which is set in 1920s Seville, portrays girls' lives in a boarding school. The series deals with several issues that are related to the rights of women and their societal status. The film clip that is attached [to the assignment] depicts a trial in which one of the school's pupils sues her boyfriend for sexual violence. [Material: a film clip.] Discuss what kinds of moral obligations belong to people involved in the course of action depicted in the film clip. (20 p.)

¹⁹ 2017A 4. An ethical issue in *Täällä Pohjantähden alla*. Before Finnish independence, tenants did not possess land, but it belonged to the landlord. The landlord could regain the land at will from a tenant and even expel him and his family. The clip attached [to the assignment] comes from a Timo Koivusalo film that is based on the novel *Täällä Pohjantähden alla* by Väinö Linna. In the film, vicar Lauri Salpakari faces an ethical problem: should he use his right to regain land from the tenant to meet his obligations towards his family, for example, to secure the dowry of his wife as well as the education of his children, even if the tenant himself has turned peatland into usable land. [Material: a film clip.] Analyse the reasons that bear on Salpakari's considerations and evaluate his decision from an ethical point of view. (20 p.)

²⁰ 2018A 4. Carnivorousism. In Asia, it is a tradition to eat dogs, which is considered inappropriate in Europe. [Materials: a news item and an anti-advertisement]. Evaluate philosophically the justification of carnivorousism and its limitation to certain species of animals. Make use of the news item and the anti-advertisement attached. (20 p.)

Assignments in metaethics

According to FNCC 2015, metaethics is not as central an area of ethics teaching as is normative ethics. However, there are certain contents (C2) and objectives (O2, O4) that are directly related to metaethics. In the material, four assignments can be classified as purely metaethical. Before we go into more detail about the assignment with a text from Bernard Williams, we briefly analyse the three other assignments.

First, there is a multi-structured (i.e., five-part) assignment (2021S 2) that requires the examinee to explain whether a given claim (e.g. 'A person who is older than 12 years is not allowed to cycle on a pavement') is by nature moral or non-moral.²¹ The assignment is related to 'the nature of morality as a system of norms and its distinction from systems based on juridical norms and conventions' (C2) and tests the examinee's ability to perceive 'the nature of normative statements' (O2). The average performance was 11.0 points. This is higher than in any other assignment in metaethics or, for that matter, in assignments in normative ethics or in mixed assignments. This assignment, then, constitutes the exception mentioned above. It does not fit the trend that, within a single exam, the average performance is better in assignments in normative ethics than in metaethics. However, a reasonable explanation can be given for the exception. According to our interpretation, the average performance in this question can be explained by reference to the fact that the assignment tests a very basic understanding. It is also worth noting that the assignment is divided into five parts, which tends to increase the total number of points. Despite the relatively high level of performance, it should not be inferred that an average student has a solid understanding of the distinction between moral and non-moral. When we were grading the examinees' performances, we noticed that in the case of one or two claims all too many students confounded this distinction with the distinction between moral and immoral, which is of course a different matter.

Second, there is a single-structured assignment (2018A 2) that asks the examinee to select a community and discuss what the difference is between morality and conventions in that community.²² The assignment is related to 'the nature of morality as a system of norms and its distinction from systems based on juridical norms and conventions' (C2) and tests the ability to 'justify the binding nature of morality and apply analyses of philosophical concepts and consistent argumentation in morality' (O4). The average performance is the worst in the entire material, namely only 6.6 points. This is not easy

²¹ 2021S 2. Normative claims. Moral claims are often by nature normative. However, not all normative claims are by nature moral. Explain why the following claim is or is not moral in nature. 2.1. A person who is older than 12 years is not allowed to cycle on a pavement. (4 p.) 2.2. A woman who visits a temple should cover her shoulders, and a man should wear long trousers. (4 p.) 2.3. Lying for the sake of personal profit is wrong. (4 p.) 2.4. At least half a kilo of vegetables and fruit should be eaten every day. (4 p.) 2.5. One ought not to cause unnecessary pain or distress to living beings. (4 p.)

²² 2018A 2. Morality and conventions. Both morality and conventions are meant to guide human behaviour. Select some community, big or small—for example, a school, an ideological group or a state—and discuss what the difference is between morality and conventions in that community. (20 p.)

to explain, but our conjecture is the following. According to our interpretation, the assignment better reflects the new philosophy for children approach of FNCC 2015 than the academic approach of FNCC 2003. Since most teachers had not yet adjusted their teaching to match the FNCC 2015 approach, they had not prepared their students to complete assignments in which one should freely choose a case and build one's answer on it. This explains why an average examinee was not well prepared to address the assignment. A further explanation is that, according to our experience, many students have difficulties in distinguishing between morality and other normative systems.

Third, there is a single-structured assignment (2019A 2) that asks the examinee to discuss whether the capacity to understand morality is distinctive of human beings.²³ The student is expected to make use of a video clip from a TED talk by the primatologist Frans de Waal, who argues that in one of his experiments capuchin monkeys show a sense of fairness. In addition to 'the nature of morality' (C2), the assignment pertains to 'ethical questions concerning animals' (C5) and tests not only the ability to 'apply analyses of philosophical concepts and consistent argumentation in morality' (O4) but also the ability to 'justify the conceptions of good and right' (O2). This assignment is clearly very demanding, which explains why the average performance was no more than 8.5 points. Here, again, we discovered that examinees had difficulty in separating morality from closely related phenomena, such as feelings of irritation and outbursts of anger, perhaps largely due to an uncritical attitude towards de Waal's experiment on monkeys.

We then turn to the double-structured assignment (2018S 4), which requires the examinee (i) to analyse an argument by Bernard Williams and evaluate its philosophical significance (10 p.), and (ii) to explain why moral relativism is popular nowadays and argue for or against it (10 p.).²⁴ The assignment is related to 'the nature of morality' and 'moral relativism' (C2) and tests the ability to 'justify the conceptions of good and right' (O2) as well as to 'justify the binding nature of morality and apply analyses of philosophical concepts and consistent argumentation in morality' (O4). Again, an average student did not do particularly well, as he or she earned no more than 8.9 points in total. To comprehend why this is the case, consider the passage from Williams given in the exam:

Let us at this stage of the argument about subjectivism take a brief rest and look round a special view or assemblage of views which has been built on the site of moral disagreements between societies. This is relativism, the

²³ 2019A 2. The morality of animals. Zoologist Frans de Waal (b. 1948) has conducted research into the origins of morality by experimenting on animal behaviour. [Material: a video clip.] Discuss whether the capacity to understand morality is distinctive of human beings. Make use of the video clip in your answer. (20 p.)

²⁴ 2018S 4 Naïve moral relativism. Vulgar or naïve moral relativism, sometimes also called normative relativism, is a surprisingly popular view in everyday thinking. In the text excerpt attached [to the assignment], the English philosopher Bernard Williams (1923–2003) evaluates this view. [Material: a text excerpt.] 4.1. Analyse Williams's argument and evaluate its philosophical significance. (10 p.) 4.2. Explain why relativism is popular nowadays, and present your own considered view for or against relativism. (10 p.)

anthropologist's heresy, possibly the most absurd view to have been advanced even in moral philosophy. In its vulgar and unregenerate form (which I shall consider, since it is both the most distinctive and the most influential form) it consists of three propositions: that 'right' means (can only be coherently understood as meaning) 'right for a given society'; that 'right for a given society' is to be understood in a functionalist sense; and that (therefore) it is wrong for people in one society to condemn, interfere with, etc., the values of another society. [...] Whatever its results, the view is clearly inconsistent, since it makes a claim in its third proposition, about what is right and wrong in one's dealings with other societies, which uses a non-relative sense of 'right' not allowed for in the first proposition. The claim that human sacrifice, for instance, was 'right for' the Ashanti comes to be taken as saying that human sacrifice was right among the Ashanti, and this in turn as saying that human sacrifice among the Ashanti was right; i.e., we had no business to interfere with it. But this last is certainly not the sort of claim allowed by the theory. The most the theory can allow is the claim that it was right for (i.e., functionally valuable for) our society not to interfere with Ashanti society, and, first, this is certainly not all that was meant, and, second, is very dubiously true. (Williams 1993, pp. 20–21; Williams's emphasis)

In this passage, Williams gives a general argument against a certain kind of normative moral relativism and illustrates the matter by giving an example. Williams's argument is based on the idea that moral relativists of this kind are inconsistent in their reasoning because they use the terms 'right' and 'wrong' in two different senses. When they say that human sacrifice was right among the Ashanti, they use the term in a relative sense, i.e. 'right for the Ashanti'. However, when they conclude that it is wrong for us or any other people to condemn the Ashanti for carrying out human sacrifice, they use the term 'wrong' in a non-relative sense, 'wrong for anyone'. To be consistent, relativists should not use the term 'wrong' in a non-relative sense. Therefore, they are not allowed to conclude that we or any other people are wrong to condemn the Ashanti for human sacrifice.

Since the assignment is double-structured, there are different thresholds for each of the two parts. For the first part regarding the structure and philosophical worth of Williams's argument, the middle threshold is the following:

In an answer worth 5 points, the examinee analyses Williams's argument so that the answer shows him or her to understand how the conceptual structure [of Williams's argument] leads to a paradox in naïve moral relativism. Additionally, the answer includes some sort of sensible judgement about the philosophical worth of Williams's argument. (2018S 4.1 guidelines)

For the second part regarding the popularity of moral relativism and the examinee's own argument for or against it, the middle threshold is the following:

In an answer worth 5 points, the examinee adequately explains philosophical, historical or societal reasons for the popularity of relativism. Additionally, he or she presents his or her own considered view for or against relativism. (2018S 4.2 guidelines)

On a first reading, the two thresholds appear to be very demanding. In practice, however, the examinee is able to pass the thresholds by making one sensible point

regarding each of the four tasks. The average performance of 8.9 points in total suggests that this requirement was too hard for many examinees. According to our experience, that is because they face serious difficulties in conducting analysis on a metatheoretical level. This applies not only to analysing conceptual inconsistencies, but also, as seen above, to analysing the relationships between various normative systems, such as morality and conventions, or morality and law.

This is where we should like to return to the question of whether the use of original texts can be justified by reference to FNCC 2015. As we noted, there are reasons for and against. The curriculum can be interpreted in either way. In this connection, we cannot discuss the matter in any detail, but we should like to consider one reason against and another for. One might oppose the use of excerpts from original texts by noting that they are too difficult for most students. The analysis that we have given above might be cited in support of that argument. However, this argument is not compelling. It might be noted that, even if excerpts from original texts are in most cases difficult, they nevertheless help a student to understand how philosophical concepts are applied in their original contexts. This understanding is important because philosophy is not a context-independent discipline: it is inextricably rooted in its own history.

Assignments in normative and metaethics

The assignments that combine normative ethics with metaethics are complex in nature. In most cases, these mixed assignments are at least double-structured, but in cases in which they are only single-structured, they comprise at least two commands, which require the examinee to complete at least two tasks. We make two observations regarding assignments in the same exam: first, that the average performance is stronger in assignments in normative ethics than in mixed assignments, and second, that the average performance is stronger in mixed assignments than in metaethics assignments. Figure 2 confirms the two observations. We suggest that the observations can be explained by reference to the complex structure of mixed assignments.

The material of the present study includes four mixed assignments. Three of them are double-structured (2020A 4; 2019A 5; 2019S 4)^{25, 26, 27} and one is single-structured (2020S 5).²⁸ In none of these assignments did an average student reach 10 points: the highest average mark was 9.7 (2019A 5), and the lowest 6.3 (2020A 4).

For the present purposes, it is not necessary to go into each assignment. Table 1 shows the ways in which we categorise these assignments. As an example, however, we should like to take an assignment that is entitled 'Instincts and morality' (2019S 4), which is combined with a clip from Ruben Östlund's film *Turist*. This assignment requires the examinee (i) to evaluate the protagonist Thomas's action when he flees from an avalanche experiment, leaving his family behind, and (ii) to discuss in general how moral action differs from non-moral action. Discussing the difference between moral action and non-moral action is relevant because Tomas justifies his action by referring to his instinctual reaction to a frightening situation. The assignment is related to three contents: 'normative ethics' (C1), 'the nature of morality as a system of norms' (C2) and 'ethics and the moral choices of the individual' (C4). The assignment tests the student's ability to 'justify the conceptions of good and right' (O2) as well as to 'analyse and evaluate actions ethically' (O5).

An average student scored 8.3 points in this assignment, which is less than the 9.5 points, which is the average performance in the normative ethics assignment (2019S 2)

²⁵ 2020A 4. Self-driving cars and moral decision. In 2018, the Massachusetts Institute of Technology (MIT), USA, published a global research (aka The Moral Machine Experiment) that gave statistical data about the moral choices of people with a view to facilitating societal discussion about the programming of self-driving cars. The research included a test that could be freely taken by anybody on the MIT website. In the test, a subject had to make choices in risky situations in which only bad alternatives were available. [Material: a graph and pictures.] 4.1. Explain why the appended graph that describes the choices that people took cannot show how self-driving cars ought to be programmed. (10 p.) 4.2. Discuss whether some moral theory can offer a principle according to which self-driving cars should be programmed. Make use of the pictures attached [to the assignment]. (10 p.)

²⁶ 2019A 5. Moral dilemma. It is commonplace that one tries to delineate moral principles by means of moral dilemmas. 5.1. Give an example of a moral dilemma and explain why it is a dilemma. (5 p.) 5.2. Discuss different solutions to the dilemma that you gave as an example in 5.1. (15 p.)

²⁷ 2019S 4. Instincts and morality. In the film *Turist*, directed by Ruben Östlund (2014), a family with two children goes on a winter holiday to the Alps. While having a lunchbreak at a slopeside restaurant, they are surprised by an avalanche that in the end proves to be harmless. However, the father Tomas flees, leaving behind his wife Ebba and his children Vera and Harry. In addition to this scene, the film clip attached [to this assignment] portrays Ebba describing the incident, the family friend Mats interpreting Tomas's action as well as Tomas making sense of his feelings. [Material: a film clip.] 4.1. Evaluate Tomas's fleeing from a moral point of view. (10 p.) 4.2. Discuss in general how moral action differs from non-moral action. (10 p.)

²⁸ 2020S 5. Game and morality. The philosopher Albert Camus (1913–1960) once said: 'All that I know most surely about morality and obligations I owe to football.' In a Champions League match between Manchester City and Shakhtar Donetsk at the end of 2018, the referee awarded a penalty to Raheem Sterling of Manchester City when the player tripped over his own feet. After the match, the incident gave rise to much discussion about players' moral obligations. [Material: a video clip.] Discuss what kinds of moral obligations a player has in the situation portrayed in the video clip, and what kind of relationship is here between moral rules and the rules of the game. (20 p.)

in the same exam. The latter assignment is cognitively much easier because it is single-structured and only requires the examinee to describe two different views of good life. This explains why an average student performed better in this assignment even if the performance is not particularly strong in a basic assignment such as this. However, it is remarkable that in the mixed assignment, examinees typically earned most of their points in the first part of the assignment, which concerned normative ethics. The second part of the assignment proved to be very difficult for most of the examinees. This suggests that, even if an examinee fails to acquire points in the metaethical part of the assignment, he or she may succeed in the normative part of the assignment. We argue that this explains why the average performance is stronger in mixed assignments than in metaethics assignments.

Assignments with and without material

We conclude our analysis of the material by comparing and contrasting assignments with and without material. Our observation is that an average student performs slightly better in assignments with no material than in assignments with material. This is most conspicuous in the case of mixed assignments, but it also applies to other cases, though not all (e.g. 2018A 2). The observation, then, does not suggest a general rule. That is why we do not think that our observation could be supported by reference to a single general tendency in students' performance. Each case must be studied separately. In the following, we analyse two cases to show that the matter is indeed rather complex (cf. Puustinen et al. 2020).

Consider the first pair of assignments in normative ethics that we analysed above. The assignment about Robin Hood's actions (2021A 2), which is without material, was judged to be worth 9.3 points on average, whereas the assignment about Thanos's alleged utilitarianism, which had a video clip (2021S 4), was graded as 8.8 on average. One possible reason for the difference in outcome is that the first assignment is cognitively less demanding than the second one, which is why an average student succeeds in reaching a better outcome in the first than in the second. In the first case, it is sufficient that examinees apply their understanding about moral theories to the case in question, whereas in the second case, they are also required to identify and interpret the reasons that Thanos gives for his action in the material.

However, the matter is not so simple. As a possible further reason, we should like to point out that philosophy teachers, as far as we know, have rather little time in the classroom to teach students how to use material in a matriculation exam assignment, or in any assignment. If so, an average student has to rely on the generic skills that he or she is able to acquire in mother tongue and literature classrooms. What we see in grading matriculation exam assignments is that weak students may simply describe the material in their own words, or worse, merely quote the key passages. Students who do this fail to make use of the material for the purposes of the assignment in question, for example, to argue that the considerations put forward in the material constitute a reason for or against a certain claim. With this background in mind, one reason why average students perform worse in an assignment with material may be that they lack, to some extent, an

ability to apply their subject-specific abilities to the material in question. In the case of philosophy, subject-specific abilities include the capacity to raise philosophical questions, to identify and analyse philosophical reasons for conclusions, and to make logical inferences based on these reasons. Even if these abilities are applicable in all philosophical contexts, students may be unable to apply them to particular cases, including the material in question, if they have not exercised the abilities in varied contexts.

Consider two further assignments that are discussed above in connection with assignments in metaethics, and which test partly the same content (C2) and objective (O2) of FNCC 2015: one assignment on normative claims without material (2021S 2) and the other on moral relativism with text material from Williams (2018S 4). The first is given in terms of a single command word, i.e., 'explain', whereas the second comprises no less than four command words: 'analyse', 'evaluate', 'explain' and 'argue'. Based on this observation only, one might suggest, reasonably, that the first assignment is cognitively less demanding than the second, which is why an average student has an increased chance of acquiring higher points in the first (11.0 p.) than in the second assignment (8.9 p.). However, we should like to add that, as shown in Section 'Assignments in metaethics', the second assignment also includes material that is very difficult to interpret. An adequate interpretation of Williams's text requires that a student applies rather extensively the subject-specific abilities mentioned above. In the present case, an average student is unable to do so.

Discussion

The foregoing analysis suggested that, within each exam, the average performance is slightly stronger in assignments in normative ethics than in assignments in metaethics or mixed assignments. As shown above, there are various possible reasons why this is the case. Some of these reasons relate to the curriculum and the teaching at schools, others to the nature of assignments. As an important reason of the former kind, we suggested that, if the teaching of ethics is to comply with FNCC 2015,²⁹ it must place more emphasis on normative ethics than metaethics. As a result, students should get more training in normative ethics than metaethics. However, we have not studied whether there is a similar tendency in the average performance in the case of FNCC 2003, which puts more emphasis on metaethics than FNCC 2015. In any case, this reason should not be overemphasized, because there are also other reasons why students do better in normative ethics.

In our analysis, we focused on those reasons that relate to the nature of assignments: the area, structure, material, and the command words of each assignment. Based on the considerations given, we now suggest two further reasons. First, assignments in metaethics typically require an ability to consider concepts and theories on a level of abstraction that is more general than the level of normative ethics. Second, if there is

²⁹ The new FNCC 2019 is similar in this respect.

some material in an assignment in metaethics, the material is typically more difficult to interpret than the materials in normative ethics. It is worth noting that the two reasons do not depend on any single curriculum, but are more general by nature. Therefore, we assume that they apply equally to FNCC 2003.

The analysis that we gave in the foregoing section also suggests that the average performance is slightly better in assignments with no material than in assignments with material. Again, as we argued, there is more than one reason why this is the case. Firstly, the assignments with no material are cognitively less demanding than the assignments with material, and secondly, philosophy teachers may have only a little time in class to teach students how to use material.

Based on the foregoing considerations, we wish to draw attention to some more general implications. First, even if the students' results are somewhat disappointing, it is important to bear in mind that we have only discussed the average performance, which does not provide a complete account of the distribution of performance. Let us note in passing that, as examiners of the philosophy matriculation exam, we have had the pleasure of reading numerous outstanding answers to each assignment in every exam. Correspondingly, we have taken pains to score poor performance. However, there is a good reason why we have focused on the average performance. Even if a study of the average performance does not provide a complete account of the distribution of individual performances, it nevertheless indicates how a significant number of students perform in an assignment, because performances close to the average are more frequent than performances further from it.³⁰

Secondly, we should like to comment on the relationship between competence in normative ethics and competence in metaethics. One might wonder whether students can show excellence in normative ethics unless they have an in-depth understanding of at least some aspects of metaethics. Consider the assignment with the film clip based on the novel *Täällä Pohjantähden alla* (2017A 4). Even if the assignment belongs to normative ethics, because the examinee is expected to analyse and evaluate the protagonist's reasons for action, an informed analysis and evaluation does not overlook the context for his action. In its context, the protagonist's action complies with the law in the fictitious world of the film and the book that are based on historical reality. In analysing and evaluating the grounds for action, however, it is important to note that acting in accordance with law is not necessarily acting in accordance with morality. A successful examinee thus observes that there is an important distinction between morality and law. This observation is metaethical. Based on this consideration, we suggest in general that, when an issue in normative ethics is related to metaethics, it would be a mistake to promote the competence in normative ethics at the expense of the competence in metaethics. The two are intertwined.

³⁰ The statistical data that we use concerning the individual performances in the philosophy matriculation exam more or less reflect the Gaussian distribution. However, none of the claims that we make about the average performance requires a closer study of the distribution. This is why we do not go into statistical details in this article.

Thirdly, we should like to raise a general question about ethics teaching. If it is agreed that competence in normative ethics should not be promoted at the expense of competence in metaethics, how could teachers of ethics ensure that both competences are attained? There is no easy answer to this question. In general, it can be noted that the teaching of ethics should help students to advance from a basic competence in normative ethics towards a deeper understanding of the nature of morality, such as issues related to moral relativism. We will not make detailed suggestions on how that can be done. However, we wish to recommend that normative considerations be combined with metaethical considerations when necessary.

This brings us to our final point, which concerns directions for further research. In the present article, we have focused on ethics assignments in a rather limited context: within philosophy studies in Finnish general upper secondary school. However, ethics is also taught in other subjects at different levels. Moreover, there is reason to believe that in the future, ethics will be taught even more extensively than before. For example, many researchers on history pedagogics have recently promoted the idea that history cannot be reasonably taught without ethical considerations. The suggestion is that ethics education should be incorporated into history education. There are various ways in which the suggestion could be implemented. It has been suggested, for example, that ethics education requires at least three things: 'moral evaluation of historical actors' conduct', 'the use of historical empathy (perspective-taking)', and 'reflection of the past's moral meaning to the present and the future (historical consciousness)' (Löfström et al. 2021, pp. 240–241). We consider that this suggestion is valuable. We also believe that it is worthwhile to promote integrative approaches in teaching history and philosophy. Like literature and arts, history provides us with ample examples for ethical evaluation. However, we should like to note that great care should be shown in considering how philosophical theories can be used in integrative approaches. Based on the results of the present study, we are concerned that normative ethics is sometimes given priority without the necessary qualifications. This is what Milligan et al. (2018, p. 455) have argued: 'We would argue that normative ethics is the branch most relevant to history education because it could support history students to describe past actions in relation to ethical challenges and deliberate over ethical responses in the present.' We agree that the approaches of normative ethics can be applied to the evaluation of past actions. However, we do not think that normative ethics suffices for this purpose. In evaluating actions and past actions, as given in the above examples, it is crucial to give some consideration to basic issues in metaethics. For example, it is crucial to distinguish between morality and law (see 2017A 4). It is also crucial to address the challenge of moral relativism (see 2018S 4): is it possible that the moral standards were not the same in the past? We recommended above that normative considerations be combined with metaethical considerations when necessary. However, we issued this recommendation having in mind the teaching of ethics in a philosophy classroom. The question that we wish to raise for further research is whether the recommendation can also be reasonably applied to the teaching of ethics in other settings, for example, a history or literature classroom.

This question raises the further question of what constitutes ethics in the first place. Can it be understood differently in different subjects? These are fundamental questions and should not be addressed without serious consideration of the various traditions of ethics teaching.³¹

References

Unpublished sources

Statistical data on the philosophy matriculation exams from Autumn 2017 to Autumn 2021. Matriculation Examination Board. Helsinki, Finland. The data are used under research permission OPH-2327-2020.

Digitally published sources

Philosophy matriculation exams from Autumn 2017 to Autumn 2021. Matriculation Examination Board. Published online in Finnish by Yleisradio and in Swedish by Svenska Yle. In Swedish: <https://svenska.yle.fi/abimix/filosofi> Retrieved 23.1.2022.

Grading guidelines in philosophy matriculation exams from Autumn 2017 to Autumn 2021. Matriculation Examination Board. Published online in Finnish and Swedish by the Matriculation Examination Board. In Swedish: <https://www.ylioppilastutkinto.fi/sv/studentexamen/beskrivningar-av-goda-svar> Retrieved 23.1.2022.

Ylioppilastutkintolautakunta 2021. Digital Matriculation Examination. Published online by the Matriculation Examination Board: <https://www.ylioppilastutkinto.fi/en/matriculation-examination/digital-matriculation-examination> Retrieved 23.1.2022.

Literature

Act on the Matriculation Examination (502/2019). Available in English online: <https://www.finlex.fi/en/laki/kaannokset/2019/en20190502.pdf>. Retrieved 23.1.2022.

Airaksinen, T. (1987). *Moraalifilosofia*. Helsinki: WSOY.

Anderson, L. W. & Krathwohl, D. R. (eds.) (2001). *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. New York: Addison Wesley Longman.

Au, W. (2007). High-Stakes Testing and Curricular Control: A Qualitative Metasynthesis. *Educational Researcher*, 36(5), 258–267.

Bishop, J. H. (1998). The effect of curriculum-based external exit exam systems on student achievement. *Journal of Economic Education*, 29(2), 171–182.

³¹ The authors of this article should like to express their gratitude to the two anonymous reviewers for their criticism and suggestions for improvement.

Fisher, R. (2013). *Teaching Thinking: Philosophical Enquiry in the Classroom*. 4th ed. London: Bloomsbury Academic.

FNCC 1994. *Grunderna för gymnasiet läroplan 1994*. [National Core Curriculum for General Upper Secondary Schools 1994] Helsinki: Finnish National Board of Education.

FNCC 2003. *National Core Curriculum for General Upper Secondary Schools 2003*. Helsinki: Finnish National Board of Education.

FNCC 2015. *National Core Curriculum for General Upper Secondary Schools 2015*. Helsinki: Finnish National Board of Education.

FNCC 2019. *Grunderna för gymnasiet läroplan 2019* [National Core Curriculum for General Upper Secondary Schools 2019]. Helsinki: Finnish National Agency for Education.

Jokila, S., Haltia, N. & Kosunen, S. (2021). Market-Making Practices of Private Tutoring in Finland: Commercialization of Exam Preparation in Admission to Higher Education. *ECNU Review of Education*, 4(3), pp. 590–614.
<https://doi.org/10.1177/2096531120956666>

Juuso, H. (2007). *Child, Philosophy and Education: Discussing the Intellectual Sources of Philosophy for Children*. Oulu: University of Oulu.

Lipman, M. (2003). *Thinking in Education*. 2nd ed. New York: Cambridge University Press.

Löfström, J., Ammert, N., Edling, S. & Sharp, H. (2021). Advances in Ethics Education in the History Classroom: After Intersections of Moral and Historical Consciousness. *International Journal of Ethics Education*, 6, 239–252.
<https://doi.org/10.1007/s40889-020-00116-w>

Löfström, J., Virta, A. & van den Berg, M. (2010). Who Actually Sets the Criteria for Social Studies Literacy? The National Core Curricula and the Matriculation Examination as Guidelines for Social Studies Teaching in Finland in the 2000's. *Journal of Social Science Education*, 9(4), 6–14. <https://doi.org/10.2390/jsse-v9-i4-1140>

Milligan, A., Gibson, L. & Peck, C. (2018). Enriching Ethical Judgment in Historical Education. *Theory & Research in Social Education*, 48(3), 449–479.
<https://doi.org/10.1080/00933104.2017.1389665>

OKM 2016. *Valmiina valintoihin. Ylioppilastutkinnon parempi hyödyntäminen korkeakoulujen opiskelijavalinnoissa*. Opetus- ja kulttuuriministeriön julkaisuja 2016: 37. Helsinki: Opetus- ja kulttuuriministeriö.

Perälä, M. & Salmenkivi, E. (2020). Käskeysanat ja ajattelun taidot filosofian ylioppilaskokeessa. *niin & näin*, 106(3), 120–124. <https://netn.fi/artikkeli/kaskeysanat-ja-ajattelun-taidot-filosofian-ylioppilaskokeessa>

Puustinen, M., Paldanius, H. & Luukka, M.-R. (2020). Sisältötiedon toistamista vai aineiston analyysia? Tiedonalakohtaiset tekstitaidot historian ylioppilaskokeen tehtävänannoissa ja pisteytysohjeissa. *Kasvatus & Aika*, 14(2), 9–33. <https://doi.org/10.33350/ka.84579>

Salmenkivi, E., Kasa, T., Putkonen, N., & Kallioniemi, A. (2022). Human Rights and Children's Rights in Worldview Education in Finland. *Human Rights Education Review*, 5(1), 47–69. <https://doi.org/10.7577/hrer.4456>

Shafer-Landau, R. (2010). *The Fundamentals of Ethics*. Oxford: Oxford University Press.

Tomperi, T. (2017). *Filosofianopetus ja pedagoginen filosofia. Filosofia oppiaineena ja kasvatuksena*. Tampere: niin & näin.

Turunen, H., Herajärvi, S., Kupiainen, S., Pirkkalainen, L., Syyrakki, S., Virtanen, V., Öhman, T. Knubb-Manninen, G., Mehtäläinen, J. & Ohranen, S. (2011). *Lukiokoulutuksen opetussuunnitelman perusteiden ja tuntijaon toimivuuden arviointi*. Koulutuksen arviointineuvoston julkaisuja 55. Helsinki: Koulutuksen arviointineuvosto.

Williams, B. (1993 [1972]). *Morality: An Introduction to Ethics*. Cambridge: Cambridge University Press.