

**This is an electronic reprint of the original article.
This reprint *may differ* from the original in pagination and typographic detail.**

Author(s): Utriainen, Jukka; Marttunen, Miika; Kallio, Eeva; Tynjälä, Päivi

Title: University Applicants' Critical Thinking Skills: The Case of the Finnish Educational Sciences

Year: 2017

Version:

Please cite the original version:

Utriainen, J., Marttunen, M., Kallio, E., & Tynjälä, P. (2017). University Applicants' Critical Thinking Skills: The Case of the Finnish Educational Sciences. *Scandinavian Journal of Educational Research*, 61 (6), 629-649.
doi:10.1080/00313831.2016.1173092

All material supplied via JYX is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the repository collections is not permitted, except that material may be duplicated by you for your research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone who is not an authorised user.

University Applicants' Critical Thinking Skills: the Case of the Finnish Educational Sciences

Corresponding author:

Jukka Utriainen, University of Jyväskylä, Department of Education. P.O. Box 35, FI-40014
University of Jyväskylä, Finland. Tel. +358408054283. E-mail: jukka.t.utriainen@jyu.fi

Co-authors:

Miika Marttunen, University of Jyväskylä, Department of Education. Email:
miika.marttunen@jyu.fi

Eeva Kallio, University of Jyväskylä, Finnish Institute of Educational Research. Email:
eeva.k.kallio@jyu.fi

Päivi Tynjälä, University of Jyväskylä, Finnish Institute of Educational Research. Email:
paivi.tynjala@jyu.fi

ABSTRACT

This study investigates the quality of the critical thinking skills of applicants (n=77) seeking entry to the faculty of educational sciences in a Finnish university and how these skills are associated with the applicant's age, previous higher education experience and matriculation and entrance examination scores. The data consist of the applicants' responses to problem-solving tasks and their matriculation and entrance examination scores. Critical thinking skills were measured with comparison and argumentation tasks. The results indicate that comparison of the texts and analysis of the arguments they contained were more difficult tasks than putting forward arguments both for and against of one's personal standpoint. In addition, previous experience of higher education predicted participants' comparison skills and their matriculation examination grades predicted their argumentation skills. The feasibility of using critical comparison tasks in the entrance examination tests is discussed.

Keywords: critical thinking, argumentation, entrance examination, academic achievement, higher education

Critical thinking and well-developed argumentation skills are commonly seen as important from the civic participation perspective (Arum & Roksa, 2011; Kuhn, 2005; Williams, Foster & Krohn, 2008), and are also regarded as key learning outcomes by higher education institutions (Halpern, 2001). Hence it is not surprising that university students' critical thinking skills have been widely studied (e.g., reviews by Pascarella & Terenzini, 2005; Niu, Behar-Horenstein & Garvan, 2013), and also that various definitions of critical thinking have been proposed. For example, Elder and Paul (2010) define critical thinking as a process of assessing and analyzing one's own thinking with the aim of improving it, whereas Ennis (1993) conceptualizes critical thinking as reasonable reflective thinking, aiming at deciding what to believe or not. These definitions highlight the importance of knowledge and its justification in conceptions of critical thinking. Ennis (2008) has labeled views of this kind as epistemic approaches to the conception of critical thinking, since they share the purpose of finding the truth or at least the closest approximation to it. In contrast to this common epistemic element is the claim that there is no unanimous definition of critical thinking (Flores, Matkin, Burbach, Quinn & Harding, 2012). Nevertheless, three common dimensions for critical thinking have been suggested, namely the ability to perform rational and reasonable thinking, the ability to see alternative viewpoints, and the ability to reflect on one's own thinking and its quality (Niu et al., 2013; Flores et al., 2012).

Like the definitions of critical thinking, the research results on learning critical thinking skills as an outcome of university studies are also conflicting. In some studies, participation in higher education has been found to develop certain elements of students' critical thinking, that is, the ability to approach problems from various viewpoints (Kim, Sharma, Land & Furlong, 2013), and to evaluate knowledge and make informed judgments (Wass, Harland & Mercer, 2011). Similarly, Pascarella and Terenzini (2005) found that senior college students had moderately higher critical thinking skills than freshmen, while, in contrast, Arum and Roksa (2011) detected only a small increase in students' critical thinking skills after three semesters of college studies.

In similar vein, research results on the relationship between critical thinking skills and academic achievement in the university are contradictory. For example, higher critical thinking skills have been found to be associated with higher academic achievement in some studies (Badcock, Pattison & Harris, 2010; Gadzella, Ginther & Bryant, 1997; Stark, 2012; Stone, Davidson, Evans & Hansen, 2001; Williams, Oliver, Allin, Winn & Booher, 2003), but not all. For example, Lobb, Wilkin, McCaffrey, Wilson and Bentley (2006) found no predictive relation between university students' critical thinking skills and their first-year academic achievement.

From the perspective of university pedagogy, fostering the learning of critical thinking is of crucial importance from the outset of studies (Bovill, Bulley & Morss, 2011), as university freshmen face an abundance of new information, conceptions and theories, that is often contradictory in nature. In order to comprehend and make sense of this conflicting knowledge, students need to apply critical thinking skills. However, little is known about the critical thinking skills of new students or university applicants. Knowledge about the level of applicants' critical thinking skills, in particular, is almost nonexistent, although it could provide important information from the point of view of developing university admission procedures, such as entrance examinations, and university teaching. Therefore, the purpose of the present study is to examine critical thinking skills of university applicants.

Interconnections of critical thinking and epistemic beliefs

The various definitions of critical thinking suggest, at least in the context of higher education, that critical thinking is linked to epistemic beliefs, that is, students' understanding of the nature of knowledge (Hofer, 2001). A recent study on Finnish third-year bio-science students' epistemological beliefs supports the idea that epistemic beliefs and critical thinking are interconnected, as the students applied critical thinking in order to strengthen their thinking in general (Hyytinen, Holma, Toom, Shavelson & Lindblom-Ylänne, 2014). There is also evidence that the more education adults have received, the more probable it is that they will perceive knowledge as complex and constantly in flux (Schommer, 1998). Further, this belief in the complexity and perpetual transformation of knowledge increases the individual's willingness to take multiple perspectives and alter their thinking (Schommer-Aikins & Hutter, 2002). In brief, well-developed epistemic beliefs are related to important elements of critical thinking, namely to the ability to see alternative viewpoints and to reflect on one's own thinking.

Owing to the epistemic aspect of critical thinking, it is not surprising that cognitive developmental models have been suggested as a method of gaining insight into the progression of critical thinking skills (Kuhn, 1999). According to Kuhn (1999), these models provide a roadmap of development from simple to more complex skills. Perry (1970) was one of the first researchers to investigate the intellectual development of university students. His model describes the development of assumptions about knowledge (epistemic beliefs) through three levels, starting from Dualistic (absolute) thinking, through Relativistic thinking to Committed thinking (Perry, 1970).

Baxter Magolda (2004) applied Perry's method and found that the majority of college freshmen held absolute epistemic beliefs while a minority held a transitional way of knowing, in which certain areas of knowing include relativistic elements. By their senior year, the majority of students had adopted a transitional way of knowing. More recently, Finnish first-year students of education showed a similar transition from dualistic to more relativistic views after only three weeks' participation in an innovative research methods course (Lahtinen & Pehkonen, 2012). In contrast, in another Finnish study, third year bio-science students held neither an absolutistic nor a relativistic epistemic viewpoint, but a fallibilist view of knowledge, in which they believed that it is possible to find measures for assessing and comparing conceptions, theories and beliefs (Hyytinen et al., 2014). In addition to education, also age and epistemic beliefs are linked together. Older students rely less on experts or authorities as sources of knowledge than younger students, and believe less in simple knowledge, and are more likely to believe that knowledge is uncertain (Bendixen, Schraw & Dunkle, 1998; Chan, 2003; Walker et al., 2009).

It has been suggested that many of the models describing the development of epistemic beliefs since the pioneering work of Perry share a common feature, that is, thinking moves towards more integrative forms (Kallio, 2011). In this connection, King and Kitchener's (1994) Reflective Judgment Model (RJM) is described in more detail below, first because it illustrates the integration of uncertain elements in knowing and, second, because it is also the theoretical foundation for the model that we applied in our data analysis (described later in this paper). The Reflective Judgment Model describes the progression of epistemic beliefs, namely how an individual's assumptions about knowledge are related to the justifications about that individual offers for his or her solutions to ill-structured problems (King & Kitchener, 2004). An ill-structured problem does not have a simple correct or certain solution (King & Kitchener, 1994), and hence requires the individual to consider the advantages and disadvantages of alternative possible solutions (Jonassen, 1997; 2000). Epistemic beliefs are argued to be central to solving ill-structured problems, as they affect how individuals select and defend their solution options (King & Kitchener, 1994).

The Reflective Judgment Model comprises seven stages that are grouped into three broader levels: prereflective thinking (Stages 1-3), quasi-reflective thinking (Stages 4-5), and reflective thinking (Stages 6-7) (King & Kitchener, 2004 p. 6). First, in prereflective thinking, knowledge is understood to be certain and gained through one's own perceptions or from a specific authority; hence, all problems are treated with high degree of certainty and ill- and well-defined problems are treated similarly (King & Kitchener, 1994). This absolute approach to knowledge makes

elementary critical thinking skills possible, because claims are not accepted as such, but they are evaluated by regard to this certain knowledge through cognitive operations (Kuhn, 1999).

Second, the quasi-reflective level is characterized by individuals' realization that, because ill-structured problems have numerous possible solutions which can be justified by various types of evidence, knowledge is uncertain and therefore evidence plays a central role in the construction of knowledge (King & Kitchener, 1994; 2004). The recognition that there are various types of evidence and solutions to problems is also essential in critical thinking, as it involves the skills of interpreting and evaluating information from various sources (Fisher, 2007). Last, on the reflective level of thinking, the evaluation of information becomes more sophisticated, as the person understands that knowledge is context-dependent, and that its quality can be assessed by the information at hand, and hence some knowledge can be better justified than some other knowledge (King & Kitchener, 1994; 2004). Similarly, determining the most reasonable approximation to the truth is the purpose of critical thinking (Ennis, 2008).

Use of the Reflective Judgment Model has been successfully extended to the assessment of college students' critical thinking in essay assignments (Wolcott & Lynch, 1997). Subsequently, Lynch and Wolcott (2001) applied the Reflective Judgment Model and the dynamic skill theory (Fischer & Bidell, 2006) in developing their Steps for Better Thinking Model for assessing ill-structured problem-solving performance, especially in relation to learning outcomes in higher education (Wolcott, 2006; 2011).

The Steps for Better Thinking Model consists of four hierarchical levels (steps) describing the progression of ill-structured problem solving skills (Wolcott, 2011, p. 76): In step 1 (Identifying), the individual recognizes the absence of a single solution to a problem, and identifies the relevant and yet uncertain information that is needed to solve the problem. Step 2 (Exploring) is characterized by the exploration of interpretations of and connections between pieces of information that are needed to solve the problem. Evidence is interpreted from multiple viewpoints to examine the assumptions and reasoning underlying these viewpoints. Next, in step 3 (Prioritizing), the alternative solutions to the problem are prioritized and the soundest option is chosen by applying guidelines developed after a comprehensive analysis and evaluation of the most important factors related to the solution. Lastly, in step 4 (Envisioning), the individual's problem-solving skills are integrated into an on-going process which aims to generate and use information to develop a better solution to the problem. In consequence, the limitations of a chosen solution are acknowledged and explained.

Critical thinking and argumentation

The interconnected and developmental nature of critical thinking and epistemic beliefs are evident in argumentation, which can be defined as "the process of developing arguments, the exchange of views, the seeking and provision of good evidence to support claims and propositions" (Andrews, 2005, p. 110). Typically, providing good evidence requires that one evaluates various types of information, which in turn demonstrates reflective thinking and good problem-solving skills. Argumentation is particularly valuable in higher education, where it works as a mediating element connecting abstract and critical thinking on a more practical level, such as in argumentative writing (Andrews, 2010).

Wu (2006) argues that undergraduate essay assignments often require students to analyze and evaluate content knowledge, to choose their position in debates in their discipline and argue accordingly. Hence, argumentation is a key requirement of essay writing by undergraduate students (Wingate, 2012). Previous research results also suggests that students with better argument analysis skills seem to produce better arguments and counterarguments in their essays (Preiss, Castillo, Flotts & San Martin, 2013). In general, however, it has been demonstrated that university students have problems in achieving mastery in argumentation skills (Marttunen, 1997), and that they lack sufficient understanding of what argumentative essay writing requires (Wingate, 2012).

Bensley (2011) argues that both lack of agreement on what specific skills should be included in critical thinking skills and vague definitions of these skills have hindered research on critical thinking. In light of Bensley's critique, *comparison* and *argumentation* skills are defined in this study as the key critical thinking skills. In focusing on these two aspects of critical thinking skills, we are aware that this leaves out other important elements of critical thinking such as dispositions and attitudes (see Butler & Halpern, 2011) and reflective dimensions. However, the two foci of critical thinking selected in this study can be argued for, first, by noting that both argumentation and argumentative reasoning skills are at the core of critical thinking (Kuhn, 1991). Second, comparison skills have been specified as one of the concrete operationalizations of analytical skills (Facione, 1990; Kuhn, 2005). Moreover, Halonen (1995) explicitly names comparison and contrasting skills as key critical thinking skills alongside analysis and prediction skills. In addition, it has been suggested that novice higher education students lack both the will and effort needed to compare set texts (Andrews, Torgerson, Beng-Huat, 2010).

Critical thinking and academic achievement

An important argument for developing critical thinking assessments in university admissions is that measures of a broad range of domain-general abilities could serve as additional predictors of students' academic achievement, alongside aptitude tests and their high school grade point average (HSGPA) (Stemler, 2012). The predictive values of aptitude tests and HSGPA for academic achievement in universities have been well verified in the United States (e.g., Geiser & Santelices, 2007; Schmitt et al., 2009; Sackett, Kuncel, Arneson, Cooper, & Waters, 2009). In Finland, while universities can decide their student admission criteria independently, matriculation examination grades and/or entrance examinations are the most commonly used. However, studies on the predictive value of the Finnish university admissions criteria are scarce. The few existing studies suggest that matriculation and entrance examinations predict pace of study at least in the social sciences and engineering (Häkkinen, 2004), and in medical school (Lindblom-Ylänne, Lonka & Leskinen, 1996; 1999).

The matriculation examination evaluates knowledge of the upper secondary school curriculum and students have to take a minimum of four subjects (The Finnish Matriculation Examination, 2014). The mother tongue test is mandatory to all and it consists of two parts. The first part aims to measure candidates' textual skills (e.g., analysis skills) and the second part education and thinking on a more general level. The remaining three tests have to be selected from the following subjects: the second national language, a foreign language, mathematics, and one from the general studies test battery in sciences and humanities (The Finnish Matriculation Examination, 2014). Previous research results suggest that success in the matriculation examination is mostly explained by previous academic achievement and parental level of education (Häkkinen, Kirjavainen & Uusitalo, 2003; Kirjavainen, 2012).

In Finland, some studies have focused on developing entrance examination tests that measure a broader range of cognitive skills. For example, in the 1990s the examinations for entry to the University of Helsinki medical school introduced new essay tasks measuring synthesis writing and critical thinking (Lindblom-Ylänne et al. 1996; 1999). Nonetheless, matriculation exam grades and traditional multiple-choice science tests were the strongest predictors of academic achievement, even if the synthesis writing task predicted study pace and the critical thinking task predicted academic achievement (Lindblom-Ylänne et al. 1996; 1999).

The Finnish government has urged universities to streamline admissions by making greater use of matriculation examinations as admissions criteria and by increased cooperation in setting entrance examinations (Ahola & Kokko, 2001). As a result, most Finnish educational faculties have started to use a jointly organized entrance examination. Applicants are required to read a book (about 180 pages) consisting of scholarly articles, and to answer multiple-choice questions designed to measure memorization, comprehension and the ability to utilize knowledge drawn from the articles (Malinen, Väisänen & Savolainen, 2012). However, the entrance examination measures, to some extent at least, the same abilities as the matriculation exam, as a positive correlation has been found between the initial points given on the basis of the applicants' matriculation examination grades and entrance examination scores (Räihä, 2010).

The previous research results, thus, suggest that investigation of the quality of educational science applicants' critical thinking skills and the associations these skills have with applicants' entrance examination scores and matriculation examination grades would have benefits for the further development of the entrance examination. It is these questions that we tackle in this study.

The purpose of the study and research questions

New university students encounter an abundance of conflicting information from the outset of their studies. Universities should, therefore, select students with a good ability to deal with information of this kind, that is, students who possess at least moderate critical thinking skills. In particular, focusing on the development of critical thinking in teacher training has been proposed to increase the use of critical thinking in the society as a whole (Williams, 2005). This further suggests that investigation of the critical thinking skills of applicants to teacher education programs, and to programs of educational sciences in general, is important. Also Stemler (2012) has suggested that university admission procedures could benefit from the inclusion of a component assessing critical thinking. However, little is known about the quality of the critical thinking skills of university applicants and how critical thinking could best be taken into account in screening for university admission. In Finland, the recent introduction of a new kind of entrance examination in the educational sciences has raised the question of the relationship between the matriculation examination grade and entrance examination score and the possible associations of these measures with measures of critical thinking. Therefore, the purpose of this study is to clarify the associations between critical thinking skills of educational science applicants and the currently used admission criteria. Because previous studies suggest that epistemic beliefs become more sophisticated with

age and are also affected by having a university education, we also investigate whether age or previous experience of university studies are related to these applicants' critical thinking skills.

The following research questions were addressed:

1. What is the level of educational science applicants' critical thinking skills with regard to comparison and argumentation skills?
2. How are applicants' matriculation examination score, entrance examination score, age and prior higher education experience associated with their critical thinking skills?
3. Do applicants' matriculation examination score, entrance examination score, age and prior higher education experience predict their critical thinking skills?

Method

Participants

The target group of this study consists of persons who had applied to study education (including programs in teacher education, early childhood education, special education, adult education, and general educational sciences) in one Finnish university in the year 2012 and who could be reached by their email address from the applicant registry (N=3 660). Our original plan was to integrate the data collection with the national entrance examination for admission to the educational sciences. However, for administrative reasons this was not possible, and thus we had to search for volunteers and construct a separate test for this group. Personal invitations to participate in the study were sent by email to the applicants after the entrance examination (high-stakes test). Because it was expected that recruiting participants after a demanding entrance examination would not be easy, a draw for a mobile phone was used as an additional incentive to participate in the study. After the second invitation, a small sample (n=77) consisting of 67 (87 %) women and 10 (13 %) men enrolled as participants in the study.

These voluntary participants ranged in age from 18 to 47 ($M=23$ $SD=6.3$ years). Most of them (84 %) had completed their upper secondary school education, a few (8 %) had a bachelor's degree, 4 % had completed a master's degree, and 4 % of participants did not provide any information about their educational background. The majority of the participants (55 %) had previously taken some higher education courses (e.g., courses at the Open University or the University of Applied Sciences); the remainder (45 %) had no previous experience of higher

education studies. The mean number of self-reported study credits (ECTS) accumulated by those who had previously participated in higher education studies was 62.4 ($SD = 100.2$). It was not possible to obtain information on how many of the participants were admitted to the university, but we assume that the data includes both admitted and rejected applicants.

Data collection

On a broad level, two main methods for measuring critical thinking can be distinguished. In addition to general critical thinking tests measuring a broad range of skills by using multiple choice items, another approach is to use essays or other forms of constructed measures (Butler & Halpern, 2011; Niu et al., 2013). In this study, the latter approach was selected as argumentative writing tasks, in particular, require the use of many critical thinking skills in the same task.

An online learning environment was used for the data collection. The invitation sent to the applicants included individual usernames and passwords for the online learning environment. After having logged into the environment, the participants read an introduction page which contained general information about the critical thinking tasks and the learning environment. Next, the participants read more specific instructions on how to use the environment, followed by a consent form. The participants were then asked to read three texts, each of which was an abridged version of a longer source text. All three abridged texts discussed class size and its effect on learning and academic achievement. The first source text was from the journal of the Finnish teachers' union, the second was a letter to the editor of a national newspaper from a professor of special education, and the third was from an article on the effects of class size on academic achievement published in a scholarly journal. The source texts were purposefully chosen to include partially contradictory information and conflicting viewpoints on the topic. The total length of the three abridged texts was 9 pages (2 819 words).

After reading these texts, the participants completed two ill-structured source-based writing tasks, designed to measure critical thinking, by writing short essays. Task 1 (Comparison task) was “*Compare the views presented in the three texts on the effects of class size on learning and academic achievement*”. Task 2 (Argumentation task) was “*On the basis of what is said in the texts, present and justify your own viewpoint on whether class sizes should be reduced or not*”. After finishing these two tasks, the participants answered a short feedback questionnaire enquiring about their approach to learning and their experiences related to performing the tasks, and asking for some general background information.

Data of the study

The data consisted of the participants' writing task essays, their entrance examination scores, and their matriculation examination grades. The essays, were written in Finnish, were saved automatically to a file in the online environment. The writing task data comprised 77 essays for task 1 (mean length 243 words, $SD = 119$) and 76 essays for task 2 (mean length 179 words, $SD = 89$). The participants' matriculation and entrance examination results were retrieved from the university's student services office.

Data analysis

A review of the research literature clearly demonstrated that critical thinking is entwined with epistemic beliefs. Consequently, developmental models of epistemic beliefs could be used to construct criteria for categorizing critical thinking skills into levels of quality. In this study, the Steps for Better Thinking model (Wolcott, 2011) was applied, both because it is theoretically grounded on such a developmental model and because it has also been previously applied in evaluating higher education students' critical thinking in several disciplines (e.g., Puteh & Hamid, 2014; Sungur, Tekkaya & Geban, 2006; Zygmunt & Karen, 2006). Argumentative essay writing tasks were also applied, as well-developed argumentation can be seen to require similarly developed epistemic beliefs. When analyzing the quality of the applicants' critical thinking skills, a whole essay was used as the unit of analysis. Four analytical categories (scored from 0 to 3) were used with all the critical thinking variables. The analytical framework for critical thinking skills was based on Wolcott's (2011) Steps for Better Thinking model, where 1 point corresponds to step 1, 2 points to step 2 and 3 points to step 3.

Comparison task

In the comparison task (task 1), the participants were asked to compare the information and viewpoints presented in the various source texts. The following variables were used to analyze the quality of the participants' critical thinking skills: 1) *Comparison of texts*, 2) *Conclusion based on the comparison*, and 3) *Number of viewpoints used in the comparison*.

The variable *Comparison of texts* measures how well the viewpoints, opinions and their justifications contained in the source texts were related to each other in the essay. It also measures how well similarities and/or differences between the texts were recognized and explicated. No

points were given if information from the texts was presented in a disorganized manner or if the texts were not referred to at all; 1 point was given if information from the texts was merely described and presented one text at a time; 2 points were given if information and viewpoints from the different texts were compared; and 3 points were given if the information and viewpoints extracted from the texts were compared in relation to some broader perspective or context.

The variable *Conclusion based on the comparison* measured the quality of a summary included in the answer that was based on comparison of the texts. In the lowest category (0 points), the essay ended with no conclusion; 1 point was given if the essay included a conclusion indicating either some similarities or differences between the texts; 2 points were given if the essay included a conclusion in which different viewpoints drawn from the texts were identified, named and related together; and 3 points were given if the essay included a conclusion in which the different viewpoints extracted from the texts were related together as a part of some broader context or perspective.

Last, the variable *Number of viewpoints from the texts used in the comparison* described how well the essay covered the topics relevant to the task. This quality was measured by counting how many of the viewpoints contained in the texts the participants used in their answers. In the lowest category (0 points), the answer included either no viewpoints or only 1 viewpoint; 1 point was given if either 2 or 3 viewpoints were used in the comparison; 2 points were given if the comparison included either 4 or 5 viewpoints; and for 3 points the comparison needed to include from 6 to 8 viewpoints. Examples of the analysis of the variables are given in Table 1 below. The comments in brackets are made by the authors.

Table 1. Examples of the analysis of the Comparison variables.

Variable	Scoring example	Interpretation
Comparison of texts	0 points: S 70 <i>"The extent to which the amount of pupils affects school success is probably school-specific. Especially in terms of how individual teaching can be arranged in the school. Everyone agrees that teachers are heavily burdened and they have too many irons in the fire all the time."</i>	Participant ignores comparison, indicated by the absence of information from the texts.
	1 point: S 1 <i>"In the first text it came out that shy students suffer when compared to louder and more restless students. This occurs when the teacher's attention is taken up by these louder students, leaving these shy students more in the background. ... The second text went over a project which aimed to reduce class sizes [misunderstanding of the texts contents]. The association between class size and positive learning outcomes is not linear, but undulating. ... The third text compared grade point averages between boys and girls in classes of different sizes. ..."</i>	In this excerpt, information from the texts is presented on the basis of individual texts. No comparison between the texts has been made.
	2 points: S 48 <i>"Teachers think that class sizes should be smaller, so that they could give enough attention to everyone and that everyone would get teaching that matches their own level and the attention they need. However, according to the results in the second text, [in fact, this result is from the third text teachers wouldn't be able to change their teaching styles even if class sizes were reduced."</i>	Participant compares the content of text 1 about the benefits of smaller class size on teachers' teaching with the content of another text on the stability of teachers' teaching styles in classes varying in size.
	3 points: S 98 <i>"According to the texts by Saloviita [text 2] and Nissilä [text 1] both the highest educational administration and teachers see the reduction of class sizes as a significant approach to promoting the quality of teaching."</i>	Information from texts one and two are integrated under a broader concept (Quality of teaching) presented in text two.
Conclusion based on the comparison	1 point: S 88 <i>"The three texts gave very different reasons for reducing class sizes. The first text dealt with this issue mostly through teachers' informal knowledge, without presenting any facts, whereas in the second text a professor of special education put forward his opinion. In the third text the research results were contradictory."</i>	Participant concludes that the texts gave very different reasons for reducing class sizes and describes those standpoints, but does not comment on the relationship between the texts
	2 points: S 2 <i>"The studies clearly don't support smaller class sizes. However, teachers who work in these environments report experiencing smaller classes as better learning environments, and also as affecting their own work."</i>	In this excerpt the participant explicates the relationship between different perspectives about the effects of class size together with comparison of the research results and teachers' experiences.
	3 points: S 12 <i>"It can be concluded from all the three texts that small class sizes don't guarantee better academic achievement. Student misbehavior and the wellbeing of the teacher can be more easily influenced in smaller classes. Bigger schools and larger class sizes have more problems, but according to the research results, they don't seem to have an effect on students' grade point average."</i>	The effects of class size on academic achievement and problems in the classroom are integrated and followed by the conclusion that they do not seem to have a combined effect.

Note: S = Subject

Argumentation task

In the argumentation task (task 2), the participants were asked to analyze and apply the information given in the three texts in presenting and justifying their own viewpoint. The quality of their argumentation was analyzed with the following variables: 4) *Analysis of arguments in texts* 5) *Arguments for and against*, 6) *Number of viewpoints used in argumentation*.

The variable *Analysis of arguments in texts* focused on the examination of the arguments and their inter-relationships present in the texts (see Facione, 1990; Fisher, 2007). The focus was on how well a participant was able to name and state the relationship between the arguments included in the texts and whether this relationship was extended to some broader or new context. In the lowest category (0 points), either none of the arguments in the texts were analyzed or the analysis was unclear; 1 point was given if the arguments in the texts were only described and presented individually; 2 points were given if the relationship between the arguments present in the texts were made explicit; and 3 points were given if arguments from two or more text were analyzed as a part of some broader or new context.

The variable *Arguments for and against* measured how well the participants had put forward both arguments supporting their own standpoint, and counterarguments, and whether the argument and counterargument were integrated through rebuttal. Arguments are usually generated more in favor of a position a person supports than arguments countering it (Stein & Bernas, 1999). Expanding one's argumentation to include views counter to one's own demonstrates an understanding of alternatives, which is also considered an important element in high-level critical thinking (Fisher, 2007). Integrative rebuttal, which requires integration of an argument and counterargument by forming a rebuttal to the counterargument (Kuhn, 1991), is the most complex form of argumentation analyzed in this study. No points were given if no relevant arguments were presented in the answer; 1 point was given if only arguments supporting one's own standpoint were put forward; 2 points were given if the essay included both arguments supporting one's own stand and counterarguments against it; and 3 points were given if the essay included an argument for one's own standpoint, a counterargument against it, and also an integrative rebuttal connecting the argument and the counterargument. Successful integrative rebuttal is usually achieved in two principal ways: a) demonstrating a counterargument to be invalid and restoring the validity of the original argument, or b) by comparing the original theory (argument) against the alternative theory (counterargument) and arguing that the original theory is the more valid one (Kuhn, 1991).

Number of viewpoints used in argumentation was determined by counting the number of the main viewpoints from the texts that the participant had used. The variable was scored in the same way as the variable *Number of viewpoints used in the comparison* task: in the lowest category (0 points) either none or only 1 viewpoint was used; 1 point was given if 2 to 3 viewpoints were used; 2 points were given if from 4 to 5 viewpoints were used; and 3 points were given if from 6 to 8 viewpoints were used. Examples of the scoring of the argumentation variables are given in Table 2.

Table 2. Examples of the analysis of the argumentation variables.

Variable	Scoring example	Interpretation
Analysis of arguments in texts	0 points S 70 : <i>"I am a professional school assistant. It seems crazy how much worry is caused by teachers' workload and class sizes, while at the same time school assistants are suffering from unemployment or are working as substitute teachers."</i>	The participant ignores the information in the texts. Instead, she describes the situation from the point of view of her own occupation as a school assistant.
	1 point S 24 : <i>"It's difficult for a teacher to notice whether everyone has learned new things, if class sizes are enormous, as was stated in the first text Individual teaching and attention given to students remains too limited."</i>	This excerpt illustrates how the arguments in the texts are listed but not analyzed.
	2 points S 11 : <i>"However, one clearly notices that even if students' success isn't badly disturbed in large-sized classes (rather the other way round), it is quite possible that teachers' ability to cope with the situation can be very limited."</i>	Participant compares the effects that large class sizes were claimed to have and analyzes arguments in the texts that make their relationship explicit.
	3 points S 99 : <i>"There are two noteworthy issues facing each other: teachers' coping and looking after students' wellbeing, and academic achievement. Which has greater importance? ... As a teacher, the most important person for me in the classroom is the student. I want to support his/her education and learning. Even if recent Finnish research indicates that smaller class sizes have led to weaker academic achievement, I think that interaction between teacher and student plays a greater role in this."</i>	The participant integrates an argument about the class size effect on learning and an argument about interaction between a teacher and a student together under the broader concept of supporting a student's education and learning.
Arguments for and against	0 points S 85 : <i>"The texts caused me to think quite contradictory thoughts. I was a pupil in a small rural school myself and at that time small class sizes were a commonplace thing. However, nowadays the amount of students in these so called school centers has prompted concern. Teachers don't have time to take the necessary precautions for the good of the children, as was noted in the first text"</i>	The participant has presented no arguments relevant to the task at hand.
	1 point S 6 : <i>"In my opinion class sizes should be reduced to a maximum of 20 students. I believe that reducing class sizes gives more benefits than disadvantages. ... The teacher can take the quieter students and those with special needs better into account, when there are fewer students in the classroom."</i>	Only arguments supporting the participants' own standpoint are given.
	2 points S 98 : <i>"A survey conducted with teachers indicates that they do not see anything good in large class sizes, which can be expected. However, according to the research results presented in the texts the effect of class size is not so significant, especially if they are already reasonably sized."</i>	This excerpt shows how an argument that supports smaller class sizes (teachers didn't find anything good to say about them) and a counterargument against it (class size effects are not significant) were linked together.
	3 points S 54 : <i>"I believe that a class size of about 15–20 students would be ideal. I think that in a class this size a teacher could also give individual guidance to students and there could also be a couple of so called special needs students in the classroom. It was stated in the introduction to the third texts that teachers wouldn't change their teaching styles if class sizes were significantly reduced. I think that after the teacher has got used to this new class size and its new possibilities for herself, she would develop her teaching methods in a way that is more beneficial for the class."</i>	First, the participant gives an argument supporting his/her own viewpoint (teachers would have more interaction with students), then counterarguments against it (Teachers wouldn't change their teaching styles), and finally a rebuttal that integrates the argument and the counterarguments together (teachers would develop new teaching methods after they have got used to new class sizes).

In order to determine the reliability of the analysis 15 essays from each task (20 % of all the essays) were analyzed by two judges. Disagreements between the judges were discussed. The inter-rater reliabilities (Cohen's Kappa) for the variables of the comparison task varied from .81 to 1.0, and for the argumentation task from .70 to 1.0.

Formation of sum and background variables

Two sum variables were formed to describe the participants' overall performance in both the Comparison and the Argumentation task. The sum variables were formed by summing the values of the original three critical thinking variables for the comparison and argumentation tasks separately, and then by dividing the sum values by three to recreate the original scale (scale from 0 to 3). The sum variables were named as follows: *Comparison, sum total* (Cronbach's α .72) and *Argumentation, sum total* (Cronbach's α .80).

To describe the participants' overall academic achievement in upper secondary school, the variable *Matriculation exam, sum score* (minimum 8, maximum 28 points) was calculated by taking the best grades awarded for the tests on the four mandatory subjects in the matriculation examination. Matriculation examinations are evaluated on the following 7-point scale (percentage of all students with each grade in parentheses): Improbatur, failed (0 points, 5 %); approbatur (2 points, 11 %), lubenter approbatur (3 points, 20 %), cum laude approbatur (4 points, 24 %), magna cum laude approbatur (5 points, 20 %), eximia cum laude approbatur (6 points, 15 %) and laudatur (7 points, 5 %) (The Finnish Matriculation Examination, 2014). We adopted the original scale used to evaluate acceptable tests (from 2 to 7 points). The variable *Entrance exam score* indicated the score a participant achieved in the entrance examination test (minimum -147.25 and maximum 199).

The variable *age* described the participant's age in years, and variable *HE credits* indicated the higher education course credits accumulated by the participants. The variables *essay, 1 length* and *essay 2, length* were formed by counting the number of words in the participants' essays.

Statistical analysis

Descriptive statistical analysis was used to describe the quality of critical thinking skills (research question 1), and bivariate statistical analysis (Pearson correlation coefficient) was used to examine the associations between the participants' critical thinking skills, matriculation and entrance examination scores and background variables (research question 2). Furthermore, in order to test

to what extent the participants' matriculation and entrance examinations, age and higher education experience predicted their critical thinking skills (research question 3), multivariable regression analysis (linear regression) was used. Statistical analyses were conducted with IBM SPSS Statistics 20. The variables used in the statistical analyses are summarized in Table 3.

Table 3. Research questions, variables, and related statistical tests.

Research question	Variables	Statistical tests
1. What is the level of educational science applicants' critical thinking skills with regard to comparison and argumentation skills?	<ol style="list-style-type: none"> 1. Comparison of the texts 2. Conclusion based on the comparison 3. Number of viewpoints used in the comparison 4. Comparison, sum total 5. Analysis of arguments in texts 6. Arguments for and against 7. Number of viewpoints used in argumentation 8. Argumentation, sum total 	Descriptive statistics
2. How are applicants' matriculation and entrance examination scores, age and prior higher education experience associated with their critical thinking skills?	<ol style="list-style-type: none"> 1. Comparison of the texts 2. Conclusion based on the comparison 3. Number of viewpoints used in the comparison 4. Comparison, sum total 5. Essay 1, length 6. Analysis of arguments in texts 7. Arguments for and against 8. Number of viewpoints used in argumentation 9. Argumentation, sum total 10. Essay 2, length 11. Matriculation exam, sum score 12. Entrance exam score 13. Age 14. HE credits 	Pearson correlation coefficient
3. Do applicants' matriculation and entrance examination scores, age and prior higher education experience predict their critical thinking skills?	<ol style="list-style-type: none"> 1. Comparison, sum total 2. Argumentation, sum total 3. Matriculation exam, sum score 4. Entrance exam score 5. HE credits. 	Linear regression analysis

Results

Level of the participants' critical thinking skills

The results showed that, on average, the applicants' performance was around the middle of the critical thinking skills scale (min 0 and max 3), as can be seen from the total mean scores (see Table 4) for the Comparison (1.43) and Argumentation (1.52) variables. Furthermore, their level of thinking in the use of the number of relevant viewpoints used in comparison in their essays was twice as high as their level in presenting conclusions on their comparison (1.90 vs. 0.88). Likewise, the applicants' performance on the analysis of the texts' arguments was about half a level lower than their performance in presenting arguments both for and against their own standpoint (1.39 vs. 1.96).

Table 4. Means and standard deviations of the variables.

	Variable	M	SD
Comparison variables	Comparison of texts	1.51	0.91
	Conclusion based on the comparison	0.88	1.04
	Number of viewpoints used in the comparison	1.90	0.95
	Comparison, sum total	1.43	0.77
Argumentation variables	Analysis of arguments in texts	1.39	0.78
	Arguments for and against	1.96	0.84
	Number of viewpoints used in argumentation	1.21	0.77
	Argumentation, sum total	1.52	0.67
Background variables	Matriculation exam, sum score	18.6	3.86
	Entrance exam score	107.42	40.10
	Age	23.1	6.33
	HE credits	62.44	100.20
Descriptive variables	Essay 1, length	243.13	119.26
	Essay 2, length	178.83	88.00

The results showed that about 40 % (Figure 1) of the participants were able effectively to compare the information in the texts (scored 2 points), and that about one-tenth extended their comparison into some broader context (scored 3 points). Conversely, more than half of the participants included no concluding summary in their essays (scored 0 points), and about one-tenth concluded their essay by describing only some features of the texts (scored 1 point). Further, about one-third of the participants explicitly drew attention to any relationship between two or three of the texts in their conclusions (scored 2 points) and only a few integrated the content of the texts into some broader context in their conclusions (scored 3 points).

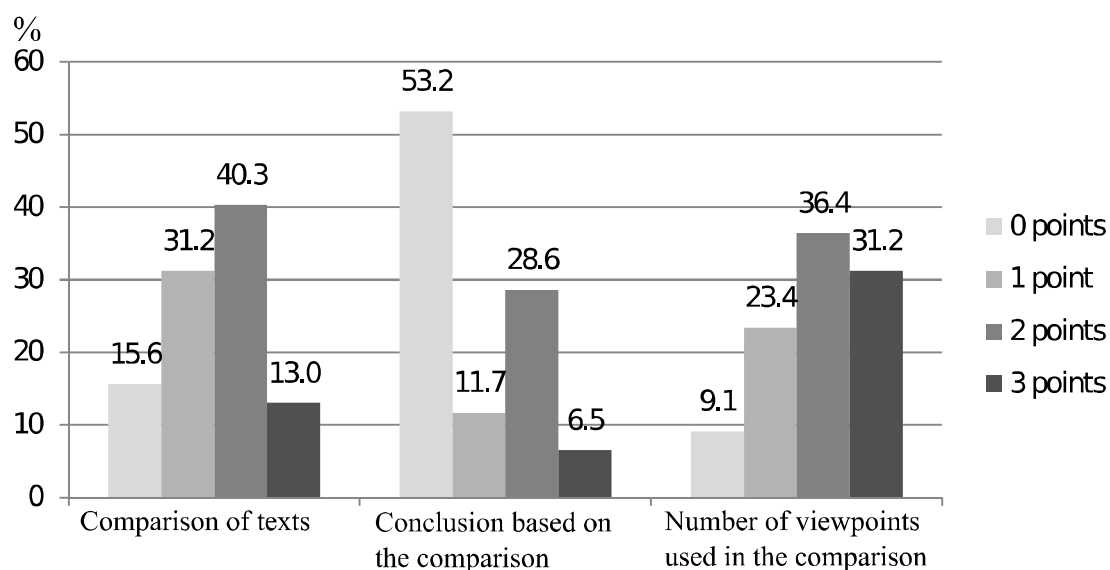


Figure 1. Quality of the participants' comparison skills.

In the Argumentation task almost two-thirds of the participants (Figure 2) inadequately analyzed the arguments in the texts merely describing or completely ignoring arguments (scored 1 and 0 points) in their essays. However, the participants performed better when the task required the use of versatile argumentative perspectives, as almost three out of four successfully put forward both arguments for and counterarguments against their own viewpoint in their essays (scored 2 or 3 points). More importantly, a quarter of them reached the highest level, successfully integrating the argument and the counterargument through rebuttals.

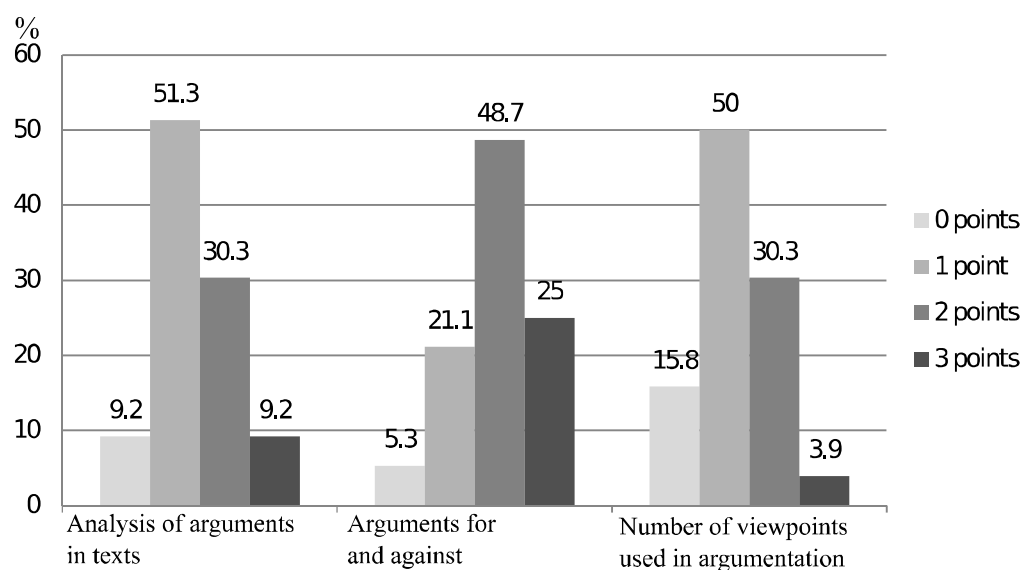


Figure 2. Quality of the participants' argumentation skills.

Associations of critical thinking skills with matriculation and entrance examination scores and background variables

The correlations shown in Table 5 indicate that the better the matriculation exam grades and the higher the entrance examination scores a participant had, the higher also were his or her overall comparison skills. Likewise, the correlations indicated a similar positive association between the participants' matriculation and entrance examination scores and their overall argumentation skills: persons who performed better in both the matriculation and entrance examinations also possessed better argumentation skills. In general, the correlations presented in Table 5 indicate that argumentation skills (Variables 5 to 8) showed stronger associations with the matriculation exam grades and entrance examination test scores than comparison skills (Variables 1 to 4).

A closer investigation of the results for the comparison skills showed some low positive correlations between the variables of interest: the variable Entrance exam score correlated with the variable Comparison of texts (.26), and those applicants who succeeded better in the matriculation examination also had higher scores in the variable Conclusion based on the comparison (.24). The associations of the matriculation and entrance examinations with the skill to put forward arguments both for and against one's own standpoint were moderately positive (.34 and .40, respectively).

Age was not associated with any of the critical thinking skills variables. On the contrary, the more higher education experience the participant had, the higher were his or her comparison and argument analysis skills. Likewise, the length of the participants' response showed strong

positive correlations with the points given for both the comparison sum total (.69) and argumentation sum total (.52). The greater the word count of the participants' essays was, the higher also were their critical thinking skills scores.

Table 5. Inter-correlations of the study variables.

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Comparison of texts	1												
2. Conclusion based on the comparison	.54**	1											
3. Number of viewpoints used in the comparison	.52**	.33**	1										
4. Comparison, sum total	.84**	.79**	.76**	1									
5. Essay 1, length	.52**	.48**	.66**	.69**	1								
6. Analysis of arguments in texts	.23*	.23*	.24*	.29*	.16	1							
7. Arguments for and against	.31**	.22	.26*	.33**	.16	.60**	1						
8. Number of viewpoints used in argumentation	.39**	.24*	.37**	.41**	.15	.59**	.54**	1					
9. Argumentation, sum total	.37**	.27*	.34**	.41**	.18	.86**	.85**	.83**	1				
10. Essay 2, length	.35**	.24*	.25*	.35**	.37**	.48**	.39**	.45**	.52**	1			
11. Matriculation exam, sum score	.19	.24*	.14	.24*	.13	.41**	.34**	.56**	.52**	.15	1		
12. Entrance exam score	.26*	.24*	.21	.30*	.10	.40**	.40**	.31*	.40**	.23	.43**	1	
13. Age	.01	-.07	.11	.02	.06	.14	.00	-.06	.03	.02	-.04	.18	1
14. HE credits	.26*	.12	.28*	.27*	.26*	.29*	.23	.12	.26*	.19	.11	.18	.67**

Note₁: * = $p < .05$; ** = $p < .01$

Note₂: In the statistical analysis, owing to missing background information, the numbers of participants were as follows: matriculation exam grades $n=67$, entrance examination results $n=68$, age $n=74$, and HE credits $n=72$.

Matriculation examinations and entrance examinations and higher education experience as predictors of critical thinking skills

Multiple linear regression analysis was conducted to investigate whether success in the matriculation examinations and entrance examinations, or previous higher education experience predicted the participants' overall comparison and argumentation skills. Two separate analyses were conducted so that the variables *Matriculation exam, sum score*, *Entrance exam score* and *HE credits* were entered as predictors, first for the variable *Comparison, sum total* and then for the variable *Argumentation, sum total*. Age was omitted from the analysis, as it did not show significant correlations with the dependent variables.

The first regression analysis indicated that previous higher education experience ($\beta = .29$, $p = .02$) significantly predicted overall comparison skills, but that matriculation examination grades ($\beta = .14$, $p = .32$) and entrance examination scores ($\beta = .14$, $p = .31$) did not. The regression model explained 11 % of the variance of the variable *Comparison, sum total* ($R^2 = .11$, $F(3, 57) = 3.52$, $p = .02$).

In the second regression analysis, the participants' matriculation examination grades ($\beta = .41$, $p = .001$) were the only significant predictor of their overall argumentation skills, while entrance examination sum scores ($\beta = .20$, $p = .10$) and previous higher education experience ($\beta = .21$, $p = .07$) were non-significant. This model explained 31 % of the variance of overall argumentation skills ($R^2 = .31$, $F(3, 56) = 9.95$, $p < .001$).

Discussion

Because critical thinking seems to be intertwined with epistemological beliefs (Hyytinen et al. 2014; Schommer-Aikins & Hutter, 2002), we first discuss the results of the quality of the critical thinking skills of our sample of applicants for entry to a university faculty of educational sciences in light of the Reflective Judgment Model (King & Kitchener, 1994) and other previous studies on the epistemological aspects of critical thinking. Typically, college students' epistemic beliefs can be located on the quasi-reflective level, in which multiple viewpoints and information related to ill-defined problems are recognized (King & Kitchener, 1994). In this study, most of the participants (in this study university applicants, not yet students) were able to identify the relevant viewpoints and arguments presented on various aspects of class size in the texts which might be considered to reflect the quasi-reflective level of epistemic beliefs. Hence, the participants' performance could be interpreted to correspond to the relativist stance identified in Finnish first-

year students in the educational sciences (Lahtinen & Pehkonen, 2012), or the quasi-reflective stance found among undergraduate students of accounting (Wolcott & Lynch, 1997). However, it must be borne in mind that the majority of the present participants were not completely unfamiliar with higher education and thus might have had previous experience in critical thinking tasks and essay writing.

Despite the view that epistemic beliefs and critical thinking are closely intertwined, the utilization of alternative viewpoints cannot be taken for granted. Undergraduate essay writing often requires analysis and evaluation of information (Wu, 2006). However, university students are claimed to lack both the motivation and effort to compare information when they write essays (Andrews et al., 2010). Such a lack of comparison skills was also demonstrated in this study, as half of the participants merely described and listed the content of the texts or even ignored it. On the other hand, the other half of the participants successfully compared the texts, which rather runs counter to the findings by Andrews et al. (2010). The limitations noted in the participants' comparison skills could be explained by the task assignment, which did not explain how the comparison should be performed. Alternatively, the result could indicate limited knowledge of what an effective comparison requires, that is, the ability to recognize and point out the similarities and/or differences between texts discussing the same topic. Either way, the finding indicates that it might be beneficial to focus on teaching critical evaluation and comparison skills as early as in upper secondary school as well as in the initial stages of higher education studies.

The quality of the participants' argumentation skills was also split in the same way as their comparison skills. Previous research has indicated that mastering argumentation skills is difficult for university students (Marttunen, 1997), and that low achieving students in argumentative essay writing have problems in being critical of and/or analyzing their source materials (Wingate, 2012). The present findings, showing that a large proportion of the participants were struggling in their analysis of the texts are in line with these earlier results. In contrast to their argument analysis skills, one-half of the participants demonstrated at least an elementary understanding of argumentation (see, Wingate, 2012), as they presented both arguments supporting their personal stand and counterarguments against it. In light of the fact that a quarter of the participants presented integrative rebuttals, the quality of their argumentation skills broadly resembles results that was previously reported among college students (Kuhn, 1991).

The associations of critical thinking skills with matriculation and entrance examination results and background variables

Earlier studies indicate that the more education one receives, the more sophisticated one's epistemological beliefs become (Baxter Magolda, 2004; Schommer-Aikins, 1998), a pattern also echoed in the development of critical thinking (Arum & Roksa, 2011). The present results seem to support these findings, as the participants who had more higher education experience also had higher critical thinking skills. Similarly, our findings parallel earlier results, according to which critical thinking skills are positively associated with higher academic achievement (Badcock et al., 2010; Stone et al., 2001). The results indicated that success in both the matriculation and entrance examination was associated with higher critical thinking scores. These results differ from those reported for Finnish medical school entrance examinations, where tasks requiring critical thinking were not associated with matriculation examination grades (Lindblom-Ylänne et al., 1996).

Besides education, also students' age have been found to be associated with their epistemic beliefs, so that older students tend to believe less on authoritative sources of knowledge and perceive knowledge more complex (Bendixen, Schraw & Dunkle, 1998; Chan, 2003; Walker et al., 2009), which may also reflect better performance in critical thinking. Interestingly, in this study, age was not associated with the quality of the participants' critical thinking skills. However, there was rather a narrow variation in the participants' age, which may have caused the lack of effect of age. In addition, in this study critical thinking skills were measured by using argumentative source based writing tasks, which measure more skills that require academic training than age-related skills. Therefore, these results need to be interpreted with caution.

Although matriculation and entrance examination outcomes and previous higher education experience were positively associated with the participants' critical thinking skills, these variables predicted their critical thinking skills inconsistently: the matriculation examination was the only significant predictor of an individual's argumentation skills, while previous higher education experience was the only significant predictor of an individual's comparison skills. These findings suggest that the Finnish matriculation examination is a valuable criterion in selecting students, and that university education has successfully developed students' critical thinking skills.

Implications of the study

Our results suggest that to some extent the matriculation examinations measure students' critical thinking, particularly argumentation skills. Furthermore, when taking into account the previous

findings on the predictive validity of matriculation examinations for university studies (Häkkinen, 2004; Lindblom-Ylänne et al. 1996; 1999), the demand by the Finnish Ministry of Education that the matriculation examination be more widely used as an admission criterion may be warranted (Ahola & Kokko, 2001; Ministry of Education, 2011).

On the other hand, it has been suggested that university admissions benefit from the use of broader cognitive measures alongside previous academic achievement (Stemler, 2012). Our results indicated that the critical comparison of information requires skills different from those evaluated in the matriculation examinations or in the entrance examination for degree programs in the educational sciences. Therefore, the possibility to include critical comparison tasks in, for example, the entrance examination used in the educational sciences, should be further examined. For instance, tasks requiring comparing and contrasting texts dealing with recently publicly debated topics in education could be suitable for this purpose. In previous studies, critical thinking skills have also predicted academic achievement (e.g., Gadzella et al., 1997; Lindblom-Ylänne et al., 1999 Stark, 2012; Williams et al., 2003), which further suggests that critical thinking skills should be evaluated more explicitly in entrance examinations.

From the perspective of university education, the applicability of the Steps for Better Thinking model (Wolcott, 2011), demonstrated in this study, suggests that it could be considered as an assessment and teaching tool from the outset in study programs in the educational sciences. Course work built on this kind of developmental model could help students to better understand how to properly carry out tasks requiring comparison and analysis of different viewpoints and arguments. University students' active engagement in learning critical thinking skills is important, especially since the development of these skills during university studies has been found to be mediocre or limited (Pascarella & Terenzini, 2005).

Limitations of the Study

First, the number of participants ($n=77$) was small, and therefore the sample cannot be considered representative of the population of applicants for entry into educational science faculties, which limits the generalizability of our findings. Moreover, linear regression analysis was used, despite the recommendation of a minimum sample size of 100 participants (Berk, 2004). However, as the other assumptions for regression analysis were met, we felt justified in using it, while acknowledging the limitations of having a small number of participants. Thus, the results of this study should be interpreted with caution.

Second, as the participants had to perform the tasks independently without training or live instructions on how to perform the tasks, they might have incurred additional cognitive load, thereby impeding their performance. However, the possibility to do the tasks on a computer could also be less demanding on them, as writing with a computer enables flexible and versatile editing of answers. Lastly, participants' prior knowledge or attitudes towards the topics that the texts discussed was not controlled, which could have had an effect on the results of the study.

Conclusions

Despite the limitations of this study, the results yield a valuable insight into an under-studied area: the quality of the critical thinking skills of educational science applicants. However, it is problematic to give any recommendations on whether the use of matriculation examination grades or alternatively the use of entrance examination tests should be emphasized. On the one hand, placing more weight on the matriculation examination grade would impede applicants who have not been successful in upper secondary school. On the other hand, making more use of more entrance examination tests, which would measure critical thinking more explicitly, could benefit older applicants who might already have experience in higher education studies. To conclude, emphasizing one admission criterion at the expense of other possible criteria may favor certain applicants and discriminate against others. Therefore developing university admissions is a complex problem for which simple or straightforward recommendations cannot easily be issued.

For the development of admission criteria, it would be important in future studies to examine how educational science applicants' critical thinking skills are associated with success in being offered a place and how these skills predict their academic achievement and pace of study in the university. In order to develop and analyze the validity and reliability of the critical thinking measures, critical thinking could be investigated in relation to performance in standardized critical thinking tests. Finally, future studies exploring the quality of educational science applicants' critical thinking would benefit from the use of broader critical thinking measures which would also include, e.g., dispositions and attitudes towards critical thinking.

References:

- Ahola, S., & Kokko, A. (2001). Finding the Best Possible Students: student selection and its problems in the field of business. *Journal of Higher Education Policy & Management*, 23(2), 191–203. doi:10.1080/13600800120088689
- Andrews, R. (2005). Models of argumentation in educational discourse. *Text*, 25(1), 107–127. doi:10.1515/text.2005.25.1.107
- Andrews, A. (2010). *Argumentation in higher education: improving practise through theory and research*. New York, NY: Routledge.
- Andrews, A., Torgerson, C., & Beng-Huat, S. (2010). Discipline-specific skills in argumentation. In A. Andrews, *Argumentation in higher education: improving practise through theory and research* (pp. 54–80). New York, NY: Routledge.
- Arum, R., & Roksa, J. (2011). *Academically Adrift: limited learning on college campuses*. [ebrary ProQuest Reader]. Retrieved from <http://site.ebrary.com/lib/jyvaskyla/reader.action?docID=10443372>
- Badcock, P. B. T., Pattison, P. E., & Harris, K. (2010). Developing generic skills through university study: a study of arts, science and engineering in Australia. *Higher Education*, 60(4), 441–458. doi:10.1007/s10734-010-9308-8
- Baxter Magolda, M. B. (2004). Evolution of a constructivist conceptualization of epistemological reflection. *Educational Psychologist*, 39(1), 31–42. doi: 10.1207/s15326985ep3901_4
- Bendixen, L. D., Schraw, G., & Dunkle, M. E. (1998). Epistemic beliefs and moral reasoning. *The Journal of Psychology*, 132(2), 187–200. doi:10.1080/00223989809599158
- Bensley, D. A. (2011). Rules for reasoning revisited: toward a scientific conception of critical thinking. In C. P. Horvath, & J. M. Forte (Eds.), *Education in a competitive and globalizing world: critical thinking* (pp. 1–36). [ebrary ProQuest Reader]. Retrieved from <http://site.ebrary.com/lib/jyvaskyla/reader.action?docID=10686269>
- Berk, R. A. (2004). Statistical inference for simple linear regression. *In advanced quantitative techniques in the social sciences series, 11: Regression analysis: a constructive critique*. (pp. 39–81). Thousand Oaks, CA: SAGE. doi:10.4135/9781483348834.n4
- Bovill, C., Bulley, C. J., & Morss, K. (2011). Engaging and empowering first-year students through curriculum design: perspectives from the literature. *Teaching in Higher Education*, 16(2), 197–209.

- Butler, H., & Halpern, D. (2011). *Critical thinking*. Oxford Bibliographies Online: Psychology. doi:10.1093/obo/9780199828340-0019
- Chan, K. W. (2003). Hong Kong teacher education students epistemological beliefs and approaches to learning. *Research in Education*, 69(1), 36–50. doi:10.7227/RIE.69.4
- Elder, L., & Paul, R. (2010). Critical thinking: competency standards essential for the cultivation of intellectual skills, part 1. *Journal of Developmental Education*, 34(2), 38–39. Retrieved from <http://search.proquest.com/docview/1237487185/fulltextPDF?accountid=11774>
- Ennis, R. H. (1993). Critical thinking assessment. *Theory into Practise*, 32(3), 179–186. Retrieved from <http://www.jstor.org/stable/1476699>
- Ennis, R. H. (2008). Nationwide testing of critical thinking for higher education: vigilance required. *Teaching Philosophy*, 31(1), 1–26. doi:10.5840/teachphil20083111
- Facione, P. A. (1990). *Critical thinking: a statement of expert consensus for purposes of educational assessment and instruction*. American Philosophical Association. (ERIC Document Reproduction Services no. ED 315 423).
- Fischer, K. W., & Bidell, T. R. (2006). Dynamic development of action and thought. In W. Damon, & R. M. Lerner (Eds.), *Theoretical models of human development. Handbook of child psychology*. (pp. 313–399) New York, NY: John Wiley & Sons.
- Fisher, A. (2007). *Critical thinking. An introduction*. Cambridge: Cambridge University Press.
- Flores, K. L., Matkin, K., S., Burbach, M., E., Quinn, G., E., & Harding, H. (2012). Deficient critical thinking skills among college graduates: implications for leadership. *Educational Philosophy and Theory*, 44(2), 212–230. doi:10.1111/j.1469-5812.2010.00672.x
- Gadzella, B. M., Ginther, D. W., & Bryant, G. W. (1997). Prediction of performance in an academic course by scores on measure of learning style and critical thinking. *Psychological Report*, 81(2), 595–602. doi:10.2466/pr0.1997.81.2.595
- Geiser, S., & Santelices, M., V. (2007). *Validity of high-school grades in predicting student success beyond the freshman year: high-school record vs. standardized tests as indicator of four-year college outcomes*. Center for Studies in Higher Education, UC Berkeley. Retrieved from <http://escholarship.org/uc/item/7306z0zf>
- Halonen, J. S. (1995). Demystifying critical thinking. *Teaching of Psychology*, 22(1), 75–81. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=afh&AN=9504201003&site=ehost-live>

- Halpern, D. F. (2001). Assessing the effectiveness of critical thinking instruction. *The Journal of General Education*, 50(4), 270–286. Retrieved from <http://www.jstor.org/stable/27797889>
- Hyytinen, H., Holma, K., Toom, A., Shavelson, R. J., & Lindblom-Ylänne, S. (2014). The complex relationship between students' critical thinking and epistemological beliefs in the context of problem solving. *Frontline Learning Research*, 2(5), 1–25. doi: 10.14786/flr.v2i4.124
- Hofer, B. K. (2001). Personal epistemology research: Implications for learning and teaching. *Educational Psychology Review*, 13(4), 353–383.
- Häkkinen, I. (2004). *Do University Entrance Exams Predict Academic Achievement?* Working paper, 2004:16, Department of Economics, Uppsala University. Retrieved from <http://hdl.handle.net/10419/82773>
- Häkkinen, I., Kirjavainen, T., & Uusitalo, R. (2003). School resources and student achievement revisited: new evidence from panel data. *Economics of Education Review*, 22(3), 329–335. doi:doi:10.1016/S0272-7757(02)00060-2
- Jonassen, D. H. (1997). Instructional design models for well-structured and ill-structured problem-solving learning outcomes. *Educational Technology Research and Development*, 45(1), 65–94. doi:10.1007/BF02299613
- Jonassen, D. (2000). Toward a design theory of problem solving. *Educational Technology Research and Development*, 48(4), 63–85. doi:10.1007/BF02300500
- Kallio, E. (2011). Integrative thinking is the key: an evaluation of current research into the development of adult thinking. *Theory & Psychology*, 21(6), 785–801. doi:10.1177/0959354310388344
- Kim, K., Sharma, P., Land, S. M., & Furlong, K. P. (2013). Effects of active learning on enhancing student critical thinking in an undergraduate general science course. *Innovative Higher Education*, 38(3), 223–235. doi: 10.1007/s10755-012-9236-x
- King, P. M., & Kitchener, K. S. (1994). *Developing reflective judgment: understanding and promoting intellectual growth and critical thinking in adolescents and adults*. San Francisco, CA: Jossey-Bass.
- King, P. M., & Kitchener, K. S. (2004). Reflective judgment: theory and research on the development of epistemic assumptions through adulthood. *Educational Psychologist*, 39(1), 5–18. doi:10.1207/s15326985ep3901_2
- Kirjavainen, T. (2012). Efficiency of Finnish general upper secondary schools: an application of stochastic frontier analysis with panel data. *Education Economics*, 20(4), 343–364. doi:10.1080/09645292.2010.510862

- Kuhn, D. (1991). *The skills of argument*. Cambridge, England: Cambridge University Press.
- Kuhn, D. (1999). A developmental model of critical thinking. *Educational Researcher*, 28(2), 16–25. doi:10.3102/0013189X028002016
- Kuhn, D. (2005). Educating for thinking. [ebrary ProQuest Reader]. Retrieved from <http://site.ebrary.com/lib/jyvaskyla/detail.action?docID=10318431>
- Lahtinen, A. M., & Pehkonen, L. (2013). 'Seeing things in a new light': conditions for changes in the epistemological beliefs of university students. *Journal of Further and Higher Education*, 37(3), 397–415. doi: 10.1080/0309877X.2011.645458
- Lindblom-Ylänne, S., Lonka, K., & Leskinen, E. (1996). Selecting students for medical school: what predicts success during basic science studies? A cognitive approach. *Higher Education*, 31(4), 507–527. Retrieved from <http://www.jstor.org/stable/3447823>
- Lindblom-Ylänne, S., Lonka, K., & Leskinen, E. (1999). On the predictive value of entry-level skills for successful studying in medical school. *Higher Education*, 37(3), 239–258. Retrieved from <http://www.jstor.org/stable/3447993>
- Lobb, W. B., Wilkin, N. E., McCaffrey III, D. J., Wilson, M. C., & Bentley, J. P. (2006). The predictive utility of nontraditional test scores for first-year pharmacy student academic performance. *American Journal of Pharmaceutical Education*, 70(6), 6–128. doi:10.5688/aj7006128
- Lynch, C. L., & Wolcott, S. K. (2001). Helping your students develop critical thinking skills. *IDEA paper* No. 37. Manhattan, KS: IDEA Center. Retrieved from http://ideaedu.org/sites/default/files/IDEA_Paper_37.pdf
- Marttunen, M. (1997). *Studying argumentation in higher education by electronic mail*. Jyväskylä Studies in Education, Psychology and Social Research 127. University of Jyväskylä, Jyväskylä.
- Malinen, O. P., Väisänen, P., & Savolainen, H. (2012). Teacher education in Finland: a review of a national effort for preparing teachers for the future. *Curriculum Journal*, 23(4), 567–584. doi: 10.1080/09585176.2012.731011
- Ministry of Education and Culture (2011). Annual Report 2010. Publications 2011:24. Helsinki
Downloaded from www.minedu.fi/export/sites/default/OPM/Julkaisut/2011/liitteet/OKM24.pdf?lang=en
- Niu, L., Behar-Horenstein, L. S., & Garvan, C. W. (2013). Do instructional interventions influence college students' critical thinking skills? A meta-analysis. *Educational Research Review*, 9, 114–128. doi:10.1016/j.edurev.2012.12.002

- Pascarella, E. T., & Terenzini, P. T. (2005). *How college affects students. Volume 2. A third decade of research*. San Francisco, CA: Jossey-Bass.
- Perry, W. G. (1970). *Forms of intellectual and ethical development in the college years: a scheme*. New York, NY: Holt, Rinehart and Winston.
- Preiss, D. D., Castillo, J. C., Flotts, P. & San Martino, E. (2013). Assessment of argumentative writing and critical thinking in higher education: educational correlates and gender differences. *Learning and Individual Differences*, 28, 193–203. doi:10.1016/j.lindif.2013.06.004
- Puteh, M. S., & Hamid, F. A. (2014). A Test on Critical Thinking Level of Graduating Bachelor of Accounting Students: Malaysian Evidence. *Procedia-Social and Behavioral Sciences*, 116, 2794–2798. doi:10.1016/j.sbspro.2014.01.658
- Räihä, P. (2010). Vakava-hankeesta ei tullutkaan uuden ylioppilaan pelastajaa [The Vakava project did not become the savior of the new upper secondary school graduate]. *Kasvatus*, 41(3), 213–225.
- Sackett, R. P., Kuncel, N. R., Arneson, J. J., Cooper, S. R., & Waters, S. D. (2009). Does socioeconomic status explain the relationship between admissions tests and post-secondary academic performance? *Psychological Bulletin*, 135(1), 1–22. doi:10.1037/a0013978
- Schommer, M. (1998). The influence of age and education on epistemological beliefs. *British Journal of Educational Psychology*, 68(4), 551–562. doi: 10.1111/j.2044-8279.1998.tb01311.x
- Schommer-Aikins, M., & Hutter, R. (2002). Epistemological beliefs and thinking about everyday controversial issues. *The Journal of Psychology*, 136(1), 5–20. doi: 10.1080/00223980209604134
- Schmitt, N., Keeney, J., Oswald, F. L., Pleskac, T. J., Billington, A. Q., Sinha, R., Zorzie, M. (2009). Prediction of 4-year college student performance using cognitive and noncognitive predictors and the impact on demographic status of admitted students. *Journal of Applied Psychology*, 94(6), 1479–1497. Retrieved from http://ovidsp.uk.ovid.com/sp-3.13.1a/ovidweb.cgi?&S=NHNIPDFLLKHFDIMFFNLKGFAGHPAAAA00&Link+Set=S.sh.22.24%7c12%7csL_10
- Stark, E. (2012). Enhancing and assessing critical thinking in a psychological research methods course. *Teaching of Psychology*, 39(2), 107–112. doi:10.1177/0098628312437725
- Stein, N. L., & Bernas, R. (1999). The early emergence of argumentative knowledge and skill. In J. Andriessen, & P. Coirier (Eds.), *Foundations of argumentative text processing* (pp. 97–116). Amsterdam: Amsterdam University Press.
- Stemler, S. E. (2012). What should university admissions tests predict? *Educational Psychologist*, 47(1), 5–17. doi:10.1080/00461520.2011.611444

- Stone, C., Davidson, L., Evans, J., & Hansen, M. (2001). Validity evidence for using a general critical thinking test to measure nursing students' critical thinking. *Holistic Nursing Practice, 15*(4), 65–74. Retrieved from http://ovidsp.ovid.com/sp-3.13.1a/ovidweb.cgi?&S=PJNNPDDLEFHFDFICLFLNKDDOFOBBEAA00&Link+Set=S.sh.22.24%7c10%7csL_10
- Sungur, S., Tekkaya, C., & Geban, Ö. (2006). Improving achievement through problem-based learning. *Journal of Biological Education, 40*(4), 155–160. doi:10.1080/00219266.2006.9656037
- The Finnish Matriculation Examination. (<http://www.ylioppilastutkinto.fi/fi/english>).
- VAKAVA. The national selection cooperation network in the field of education. (<http://www.helsinki.fi/vakava/english/examination.htm>).
- Walker, S., Brownlee, J., Lennox, S., Exley, B., Howells, K., & Cocker, F. (2009). Understanding first year university students: Personal epistemology and learning. *Teaching Education, 20*(3), 243–256. doi:10.1080/10476210802559350
- Wass, R., Harland, T., & Mercer, A. (2011). Scaffolding critical thinking in the zone of proximal development. *Higher Education Research & Development, 30*(3), 317–328.
- Williams, R. L. (2005). Targeting critical thinking within teacher education: The potential impact on society. *The Teacher Educator, 40*(3), 163–187. doi:10.1080/08878730509555359
- Williams, R. L., Foster, L. N., & Krohn, K. R. (2008). Relationship of patriotism measures to critical thinking and emphasis on civil liberties versus national security. *Analyses of Social Issues and Public Policy, 8*(1), 139–156. doi:10.1111/j.1530-2415.2008.00165.x
- Williams, R. L., Oliver, R., Allin, J. L., Winn, B., & Booher, C. S. (2003). Psychological critical thinking as a course predictor and outcome variable. *Teaching of Psychology, 30*(3), 220–223. doi:10.1207/S15328023TOP3003_04
- Wingate, U. (2012). 'Argument!' helping students understand what essay writing is about. *Journal of English for Academic Purposes, 11*(2), 145–154. doi: 10.1016/j.jeap.2011.11.001
- Wolcott, S. K. (2006). *One Size Doesn't Fit All: Create a Customized Critical Thinking Rubric*. Unpublished manuscript.
- Wolcott, S. K. (2011). Professional attributes: Teaching the fine arts of being a professional accountant. In E. M. Wiecek & G. Beal (Eds.), *Leveraging change – the new pillars of accounting education*. (pp. 68–86). Toronto: The Canadian Institute of Chartered Accountants and University of Toronto.
- Wolcott, S. K., & Lynch, C. L. (1997). Critical thinking in the accounting classroom: a reflective judgment developmental process. *Accounting education: A journal of theory, practice and research, 2*(1),

59–78. Retrieved from
[http://search.ebscohost.com/login.aspx?direct=true&db=bsh&bquery=\(SO+\(Accounting+education\)\)AND\(DT+1997\)AND\(TI+critical+thinking+%26quot%3bin%26quot%3b+the+accounting+classroom\)&type=1&site=ehost-live](http://search.ebscohost.com/login.aspx?direct=true&db=bsh&bquery=(SO+(Accounting+education))AND(DT+1997)AND(TI+critical+thinking+%26quot%3bin%26quot%3b+the+accounting+classroom)&type=1&site=ehost-live)

Wu, S. M. (2006). Creating a contrastive rhetorical stance: investigating the strategy of problematization in students' argumentation. *RELC Journal*, 37(3), 3–353. doi: 10.1177/0033688206071316

Zygmunt, D. M., & Karen, M. S. (2006). Writing across the nursing curriculum project. *Annual Review of Nursing Education*, 4, 275–290. Retrieved from
<http://search.proquest.com/docview/216908887?accountid=11774>