Development and Application of Assessment Standards to Advanced Written Assignments

Antti Miihkinen* and **Tuija Virtanen****

*Aalto University, Finland; **University of Helsinki, Finland

September 2017

Authors' contact information:

Antti Miihkinen (Correspondence Author)

Aalto University School of Business (Department of Accounting)

P.O. Box 21220, FIN-00076 Aalto

Tel. +358 50 5469573

Email: antti.miihkinen@aalto.fi; antti.miihkinen@alumni.ufl.edu

Website: https://people.aalto.fi/new/antti.miihkinen

Note: Part of this work was done during my post-doc work at the University of Florida (Fisher School of Accounting)

Tuija Virtanen

University Lecturer (University of Helsinki, Faculty of Agriculture and Forestry,

Department of Economics and Management)

P.O. Box 27, FI-00014 University of Helsinki

Tel. +358 40 7242917

E-mail: tuija.virtanen@helsinki.fi

Development and Application of Assessment Standards to Advanced Written Assignments

ABSTRACT   *This study describes the results of a project that focused on developing an assessment rubric to be used as the assessment criteria for the written thesis of accounting majors and the quality of the coursework during the seminar. We used descriptive analysis and the survey method to collect information for the development work and to examine the effect of the rubric on learning. We find that the rubric has a positive effect on students' understanding, self-assessment, confidence, and integration. We contribute to the extant literature by adding to prior work that has examined factors that can improve students' learning outcomes. By synthesizing theories on approaches to learning and self-regulation, and combining them with literature on self-efficacy and social/academic integration, we bring conceptual clarity to the elements of learning in a course, which consist of written assignments and the accompanying group work. The paper demonstrates a way to help university students to learn via explicit assessment rubrics, and thus offers novel ideas for accounting educators.*

1.   Introduction

Assessment is a popular topic in pedagogical discussions. However, it could be emphasized more in university teaching (Lindblom-Ylänne & Nevgi, 2009). According to earlier studies, assessment practices have a strong impact on learning, and students often choose their approach to learning on the basis of assessment methods (e.g., Marton & Säljö, 1976; Thomas & Bain, 1984; Scouller, 1998). In the 2000s, Assurance of Learning (AOL) accreditation standards have been increasingly emphasized, especially in evaluations of business schools. The basic idea of the AOL standards is that properly set learning outcomes and assessment methods lead to better learning outcomes (AACSB White Paper 2007).[1] At their best, the AOL standards can constitute an internal working method at the university level that guides students' learning, but also the basis for external accreditations.

This study is motivated by the importance of assessment as a determinant of learning outcomes (e.g., Biggs, 1996; Duff & McKinstry, 2007; Kostons, van Cog & Paas, 2012). Accounting education literature does not frequently discuss assessment. There is only scant evidence on how students' learning can be supported in advanced writing assignments and accompaning seminar work. In addition, although the education literature demonstrates several factors that correlate with students' learning outcomes (Richardson, Abraham, & Bond, 2012), it lacks evidence regarding how rubrics as a self-assessment tool can be utilized to improve their capabilities for learning, following the predictions of self-regulation theory (e.g., Vermunt, 1998; Puustinen & Pulkkinen, 2001; Kostons et al., 2012). Moreover, evidence on the assurance of learning on courses designed around written assignments would be useful for many because written assignments are widely used in Anglo-Saxon countries and improvements in writing skills are very welcome (Riley & Simons, 2013; Dale-Jones, Hancock & Willey, 2013). Finally, many European undergraduate students write a mandatory bachelor's thesis, which highlights the importance of developing assessments in these courses.

We aim at bringing more evidence to the assessment of *advanced written assignments* and the *accompanying group collaboration* in accounting courses at the undergraduate and graduate levels.[2] Advanced written assignments are demanding because usually they require levels of understanding that occur at the highest levels of the structure of the observed learning outcome (SOLO) taxonomy (e.g., create, analyze, criticize, theorize; see Biggs & Collis, 1982). Consequently, these assignments require students to employ deep or strategic learning approaches, but at the same time may cause anxiety, lack of motivation, and procrastination for students if they do not know what they are expected to do in order to obtain certain grades. We predict that by explicitly showing students the expected quality criteria for the assignment and the seminar work, we can affect their learning positively by facilitating optimal learning strategies and by increasing their feeling of self-efficacy. An optimal learning strategy is needed because of the time constraints that usually prevail in these courses. Moreover, explicit assessment standards should help to enhance teaching and learning in the course because the constructive alignment between the course elements increases (Biggs, 1996). As suggested in the assessment literature (Moss, 1992; Biggs, 1995),

---

[1] AACSB = The Association to Advance Collegiate Schools of Business
[2] We define advanced written assignment to mean assignments that are written after the basic studies at the undergraduate level (e.g., electives in accounting) or which are written at any level of graduate studies. We acknowledge that, in many universities, undergraduate students do not have advanced studies (intermediate studies being the highest category). However, we use 'advanced' to signal that we are examining written assignments that require and develop deeper understandings to create new content from students' existing knowledge following the principle of constructive learning.

constructive alignment in course elements supports students' endeavors to achieve the high levels of understanding specified at the expected learning outcomes of the course. *We study this issue by describing the development and implementation of an assessment rubric for bachelor's thesis seminars in accounting at the Aalto University School of Business to ensure the quality of students' learning.*[3]

Dawson (2017) defines a rubric as a tool used in the process of assessing student work that usually includes Popham's (1997) three essential features: evaluative criteria, quality definitions for those criteria at particular levels, and a scoring strategy. We use descriptive analysis and the survey method to justify the need for, and utility of, the development work for the rubric. Moreover, we obtain evidence regarding the advantages and disadvantages of the rubric from semi-structured feedback interviews with the students. Through its AACSB accreditation, the Aalto University School of Business is committed to the AOL standards, and improvements in assurance of learning are needed at the course and curriculum levels.[4] By developing assessment rubrics, we open the 'black box' of assessment to the students and explicitly show them what is required. An assessment rubric is a tool that helps to understand assessment methods (i.e., written assignment and seminar work). Although the assessment methods (see Stevenson, Ferguson, & Power, 2014, 438) remain the same, the rubric provides students with early evidence on assessment which gives them the opportunity to understand the intended learning outcomes early and to focus on doing the things that will direct them to learning in a restricted time (usually one semester) frame. We also develop explicit quality criteria for the seminar work because prior literature recommends that the focus in teaching should not only be on learning outcomes, but also on the learning process (Lindblom-Ylänne & Nevgi, 2009).

The theoretical background of the paper derives from the constructivist learning theory in which learning is considered to be context-specific and dependent on the activity of the student in her/his learning environment (e.g., Biggs, 1996; Tynjälä, 1998; Biggs, 2003). We add to previous Presage-Process-Product (3P) models (Biggs, 1993; Duff & McKinstry, 2007; Lucas & Mladenovic, 2014) and introduce a modified model of learning tailored for advanced written assignments. This model (see section 2.3) synthesizes several theories that are linked to students' learning in context. Consequently, we suggest that learning in courses that employ advanced written assignments and accompanying group work is dependent on the adopted approach to learning (e.g., Lucas & Mladenovic, 2004; Duff & McKinstry, 2007) and the quality of the student's self-assessment (e.g., Vermunt, 1998; Kostons et al., 2012). Moreover, feeling of self-efficacy (Bandura, 1977; Bandura, 1997) and social/academic integration (Richardson et al., 2012) are influenced by the current learning context. These factors also impact students' perceptions of task requirements, which determine the adopted approaches to learning.

The results provide evidence on the usefulness of the assessment rubric as a tool for assessing and improving students' learning. The statistical analysis of the survey data provides evidence that, on average, students obtain good grades, which suggests that they are learning deeply (Biggs,1989; Ramburuth & Mladenovic, 2004). Grades are statistically significantly and positively related to changes in understanding and confidence during the

---

[3] The format of a bachelor's seminar is clearly different from other courses at the bachelor's level. The role of students' independent work is highlighted. In addition, particularly for thesis-writing seminars, acting as an opponent to other students' theses and active participation in class are emphasized.

[4] In 2007, the Aalto University School of Business was granted an AACSB accreditation as the first business school in the Nordic countries. In 2012, the accreditation was renewed. The accreditation process will subsequently be renewed every five years if the required criteria are met. Aalto University School of Business is one of the six schools of Aalto University that was created as a combination of the Helsinki School of Economics, the Helsinki School of Technology, and the Helsinki School of Art in 2010.

course. We demonstrate that those students who benefited more from the rubric had a larger increase in understanding and confidence, which constitutes evidence on the usefulness of the rubric in improving learning. Increase in understanding is also closely linked to deep learning (Duff & McKinstry, 2007) and increase in confidence is linked to self-efficacy theory and flow-feeling (Bandura, 1997; Krapp, 2002; Csikszentmihalyi, 2008).

Furthermore, the qualitative analysis of the survey data provides support for the utility of the rubric in every suggested dimension in the theoretical framework: student's approaches to learning, student's self-assessment, self-efficacy, and social/academic integration. The rubric can help students to learn more deeply or more strategically, for example, by increasing their understanding or helping them in time management (Duff & McKinstry, 2007). These learning approaches can also intertwine (Biggs, 1989; Biggs, 1993; Beattie, Collins & McInnes, 1997). Students may also feel that the rubric helps them to self-assess their work in line with self-regulation theory (e.g., Puustinen & Pulkkinen, 2001). Furthermore, the rubric can increase students' self-confidence and motivation, and help them to integrate with their peers in the course. Finally, with the rubric, it may be easier for the students to understand the role of the process in the coursework.

This paper contributes to the literature by demonstrating how an assessment rubric can be employed to support students learning in courses incorporating advanced written assignments and accompanying group work (Daff, de Lange, & Jackling, 2012; Dale-Jones et al., 2013; Riley & Simons, 2013). The rubric improves the learning context and thereby affects students' perceptions of task requirements as suggested in constructivist learning theory (e.g., Biggs, 1996; Tynjälä, 1998; Duff & McKinstry, 2007). We also add to the literature, which suggests that effective self-regulated learning requires that students can accurately assess their own performance (Vermunt, 1998; Puustinen & Pulkkinen, 2001; Kostons et al., 2012). Finally, we contribute to the literature on self-efficacy (Bandura, 1977; Bandura, 1997) and social/academic integration (Tinto, 1975; Richardson et al., 2012) by suggesting that these determinants of learning are important in courses relying on advanced written assignments and accompanying group collaboration.

Our constructivist model of learning for advanced written assignments and accompanying group work combines the above mentioned theories into a framework that can be used in future development work. It provides further insights into the various factors that together influence students' learning outcomes (Richardson et al., 2012) and also adds to the extant assessment literature by focusing on accounting courses (e.g., Panadero & Jonsson, 2013; Panadero & Romero, 2014). An important practical contribution of the paper is its detailed description of the development of the rubric, which should encourage and help accounting educators to implement assessment rubrics in their coursework.

We provide the theoretical underpinnings related to assessment of learning and learning theories in the next section. The next section also introduces the theoretical framework of the paper. Section 3 presents the context of the study, and in Section 4 we provide the results of the descriptive study that justify the development of the rubric. Section 5 describes the development work, gives the empirical results regarding the implementation of the rubric, and also provides a discussion on the usefulness of the rubric. We present the summary and concluding remarks in Section 6.

2.   Assessment of Learning and Learning Theories

2.1   *Assessment in Higher Education*

Before making decisions on assessments, it is important to determine learning objectives. Bloom's (1957) taxonomy divides learning objectives into three domains: cognitive, affective, and psychomotor.[5] According to Bloom, students must always learn the basic concepts at the lower level before they can proceed to the upper level. The higher the student proceeds in Bloom's taxonomy, the better (s)he will understand the issues being taught. In the first level (Knowledge), the student can perceive and remember information. In the last level (Evaluation), the student can assess theories, compare ideas, and assess associations and/or causal relationships. Advanced written assignments (e.g., preparing a thesis) require that students have the ability to rise to the highest level in the taxonomy. This, together with Allan's (1996) assertion that assessment is difficult in universities, underlines the importance of finding the optimal ways to assess undergraduate and graduate students' learning in their thesis seminar courses.

Another important framework developed for assessment is the structure of the observed learning outcome (SOLO) taxonomy (Biggs & Collis, 1982). SOLO taxonomy directs assessment towards the quality of the learning process. Consequently, it is not concerned with the number of right answers to certain questions. It is more important that the student has a high-quality learning process. In the highest levels of the SOLO taxonomy, the student is expected to be able to set research hypotheses, theorize, and reflect on previously learned issues. Moreover, Biggs (1996, 2003) has developed the concept of 'constructive alignment.' Accordingly, teachers first think about the learning objectives, and then adjust their teaching and assessment methods to fit those objectives.

The assessment literature partly builds on the literature on different approaches to learning (see section 2.2), which suggests that students choose their learning approaches based on those criteria that are utilized in the evaluation of the assignment. The need for assurance of learning is recognized in the literature on higher education (French, Summers, Kinash, Lawson, Taylor, Herbert, Fallshaw, & Hall, 2014). Although assessment is a core research area in higher education, there is surprisingly little discussion of assessment, in general, and self-assessment tools, specifically, in accounting education journals. Rubrics constitute one way to apply assurance of learning in practice at the accounting course level and assist students to monitor and self-evaluate their learning. The ability to reflect against the rubric increases students' possibilities to make a distinction between what they have done correctly and where they need improvement. Hence, students are more capable of regulating their learning when assessment rubrics are available (Panadero & Romero, 2014). Rubrics can also shape the perception of task requirements, and thereby affect the adopted approach to learning. This is important because the continuum of different alternative approaches to learning is wide (e.g., Vermunt, 1998; Lucas & Mladenovic, 2004).

Assessment at the course level should be aligned with 'Assurance of Learning' (AOL) at the curriculum level. AOL has been part of the AACSB accreditation standards since 2003. AOL standards evaluate how well the school accomplishes the educational aims at the core of its activities. The 2003 standards emphasize direct outcomes assessment. AACSB expects accredited institutions to formulate specific learning goals and conduct appropriate direct assessment of learning to improve curricula when deficiencies or opportunities for

---

[5] The modified version of Bloom's taxonomy was introduced in 2001 (see Anderson & Krathwohl, 2001).

improvement are identified. The standards support two principles: accountability and continuous improvement. The end goal is continuous improvement in learning and achievement (AOL Handbook, 2011).

Martell (2007) provides examples of how assessment results can be used to improve curricula. Her results suggest that curriculum alignment is not, in and of itself, AOL. One step in the AOL process is to align the curriculum with the learning goals. This is typically done with a matrix, with courses along the rows and learning goals across the columns. However, developing these matrices does not assess student learning. Martell (2007) also points out that the main point of the AOL standards is continuous improvement of curricula to ensure learning, not just completing the assessment and meeting the standards.

Previous literature on assurance of learning demonstrates that assessment is challenging (e.g., Pringle & Michel, 2007; Scherer, Javalgi, Bryant, & Tukel, 2005). Assurance of learning procedures at the university level are often driven by a need to show external accreditors that students learn. However, these procedures will only guide and direct students' learning if they concretely affect students' learning at the grass-roots level. Proper assessment can support students' learning, which is a multifaceted phenomenon.

2.2 *Learning Theories*

Approaches to learning paradigms have had a major impact on how we understand students' learning (e.g., Marton & Säljö, 1976; Ramsden, 1979; Lucas & Mladenovic, 2004; Duff & McKinstry, 2007). Accordingly, students may adopt either a deep, surface, or strategic approach to learning. The characteristics of these different approaches are summarized in Duff (2014, 172-173). *Deep approach* means that the student searches for meaning in the topic that (s)he is studying and tries to create links to other experiences and ideas. A deeply learning student wants to relate new ideas to previous knowledge, and relate evidence to conclusions. (S)he is also interested in the logic of the argument. All of these skills are needed in advanced written assignments.[6] *Surface learning* means that the student is rote-learning and memorizing in isolation of other ideas. Previous accounting literature suggests that it may be too simplistic to make a dichotomy between these two forms of learning (Thomas & Bain, 1984; Beattie et al., 1997). *Strategic learning* is characterized by a motivation to achieve high grades. A strategically-oriented student makes decisions on studying based on this objective. In the strategic approach, the student is interested in the cues about assessment criteria, and wants to organize time and distribute effort to achieve the greatest effect (Duff, 2014; Lucas & Mladenovic, 2004).[7]

Popular frameworks that model students learning build on Biggs (1993) 3P model of classroom learning in which a student progresses from presage to process to product and each component of the model interacts with all other components before forming a system in the equilibrium. In the model, student presage factors (e.g., prior knowledge, abilities, preferred ways of learning) and the teaching context (e.g., curriculum, teaching method, assessment) affect the task processing of the student which determines the learning outcomes. Biggs (1993) describes the system as being an educational swamp which in the tertiary context comprises several nested micro-systems (e.g., the student system, the classroom system).

---

[6] Aggregate final grade is one measure for academic performance, and its positive association with deep learning has been discussed and documented in the existing accounting education literature (Biggs, 1989; Ramburuth & Mladenovic, 2004).

[7] Earlier work uses the terms strategic and achieving approach interchangeably (Lucas and Mladenovic, 2004).

Changes in the system, which either lead to new equilibrium or reversal to the old status quo, may be challenging and make the work of innovators more difficult.

In this systemic approach, teaching, assessment, and student perceptions are expected to strike a balance that supports learning. A student's approach to learning is seen as relational and thus deep approach is comprehended as a way of describing how (s)he relates to a task (Ramsden, 1987). The following work has made minor revisions to the original model but collectively, in these 3P models learning can be affected by the context and assessment is one item of the learning context (see Duff & McKinstry, 2007; Lucas & Mladenovic, 2014). These models represent the constructivist learning theories (e.g., Biggs, 1996; Tynjälä, 1998; Biggs, 2003). Accordingly, the student is an active operator in her/his learning environment, and constructive alignment between teaching and assessment is important (Biggs, 1996).

Constructivist learning theory assumes that a student builds new knowledge on previous knowledge by reflecting new information against her/his prior knowledge and experiences, just as in deep learning. Although the university cannot influence individual factors, it can influence factors relating to the learning environment at the university, curriculum, and teacher level (e.g., Hall, Ramsay, & Raven, 2004). Interestingly, Lizzio, Wilson, and Simons (2002) found that undergraduate students' perceptions of their current learning environment were a stronger predictor of learning outcomes at the university than was prior scholastic achievement.

Constructive learning theories are only one way to structure learning. Another important theory is *self-regulation theory*. It suggests that students have to be able to accurately assess their own performance to be able to achieve effective self-regulated learning (e.g., Vermunt, 1998; Puustinen & Pulkkinen, 2001; Kostons et al., 2012). Puustinen and Pulkkinen (2001) characterize self-regulated learning as an intermediate construct describing the ways in which individuals regulate their own cognitive processes within an educational setting. Vermunt (1998) finds that students' self-regulation of learning is a much better predictor of constructive processing strategies than external regulation. This is also closely linked to the motivation of using rubrics in education, especially rubrics that are used for formative assessment purposes (see Panadero & Jonsson, 2013).

Student learning in higher education can be assessed in two ways: formative and summative assessment (Lindblom-Ylänne & Nevgi, 2009). Prior literature suggests that formative assessment methods are more effective than summative assessment methods (e.g., Hoge & Coladarci 1989; Black & Wiliam, 1998; Nicol & Macfarlane-Dick, 2006). Formative assessment can be characterized as contributing to the objectives of learning and guidance. Under formative assessment, students receive feedback on their learning process and, at the same time, teachers get feedback on how they could improve their teaching methods. In contrast, summative assessment emphasizes evaluation. The focus of assessment is on students' learning outcomes. Course grades are often given on the basis of summative assessment (Biggs & Tang, 2007; Brown, Bull, & Pendlebury, 1997; Light & Cox, 2001). Consequently, assessment rubrics that are given to students at the beginning of the written assignments are supposed to be utilized for formative purposes. Ideally, the rubric should help students in self-assessment during the whole study process, and provide them explicit guidance for where to benchmark their own performance. However, Panadero and Romero (2014) point out that rubrics are not always used for formative assessment purposes.

Kostons et al. (2012) discuss the earlier evidence, which shows that students have difficulties with proper self-assessment and task selection. Their results confirm that self-assessment and task selection skills help in self-regulated learning, and that students benefit from training in these skills. There are studies which show that rubrics assist learning and improve performance if they are used long enough (Panadero & Jonsson, 2013), and if the rubric follows most of the recommendations set for self-assessment as specified in Andrade

and Valtcheva (2009) following Goodrich's (1996) suggestions. Panadero and Romero (2014) compare the effects of situations in which self-assessment is done against the rubric and in which the self-assessment tool is missing. They find that the rubric group uses more developed learning strategies, and achieves better performance and accuracy. Interestingly, they also demonstrate that the rubric group has more problems handling stress and a higher level of performance/avoidance self-regulation, and both findings are contrary to the hypothesis. In contrast, many other studies demonstrate that the use of rubrics helps students to feel more secure and also have a negative effect on anxiety (Andrade & Du, 2005; Reynolds-Keefer, 2010). Panadero and Romero (2014) explain these counterintuitive findings by the time pressures that the participants felt. They allocate only one hour for the learning task, whereas previous studies have preferred entire semester-long learning tasks. Thus, the importance of the careful design of the use of rubrics is one take-away from their study. Jonsson and Svingby (2007) conclude in their review article that rubrics have the potential to promote learning and/or refine teaching, mainly because rubrics make expectations and criteria explicit, which also improves feedback and self-assessment.

Students' learning can also be influenced by their feelings. In the 3P model, student factors include affect, which means the experience of feeling or emotion. One essential feeling in higher education is the feeling of self-efficacy. *Self-efficacy theory* suggests that students who believe in their own skills and capabilities in academic studies also perform better than those who have a low level of self-efficacy (Bandura, 1977; Bandura, 1997). Self-efficacy consists of academic self-efficacy and performance self-efficacy (see Richardson et al. 2012, p.356).[8] Academic self-efficacy relates to new occasions where one's own capabilities have to be assessed based on one's own skills (Zimmerman, Bandura, & Martinez-Pons, 1992). Self-efficacy is also related to context. It is not a general feeling, but may vary depending on the performed action based on earlier experiences of the individual (Zimmerman, 2000). The ratio between the skills that the person perceives herself to have and the challenge that the person is facing determine how she feels during the task. If the person has a high self-efficacy and the task at hand is challenging, (s)he can achieve a flow-feeling where the sense of time and place disappears. On the contrary, it is easy to deduce that if the student has a low feeling of self-efficacy, the same challenging task can make her/him anxious (Csikszentmihalyi, 1994; Csikszentmihalyi, 2008). Flow in learning causes positive emotional experiences. These experiences are connected with intrinsic motivation, and interest-triggered learning usually leads to a higher degree of deep learning (Krapp, 2002).

An additional element which has an impact on learning at the university level is the psychosocial context (Richardson et al., 2012). Especially, we suggest that an assessment rubric for the written assignment and seminar work may help students to achieve a higher degree of integration with the supervisor and the other seminar participants. Tinto (1975, 1982) suggests that if the student does not achieve *social or academic integration* during her studies, it will increase the risk that (s)he will drop-out from the program. In Tinto's educational persistence model (1975, p. 95), peer-group interactions and faculty interactions affect both social and academic integration. This model suggests that university systems interact with student characteristics and experiences, and thus shape the degree of interaction that students can achieve within social and academic systems. Optimal interaction will

---

[8] Academic self-efficacy means general perceptions of academic capability. An example of a representative statement is "I have a great deal of control over my academic performance in my courses". Performance self-efficacy means perceptions of academic performance capability. An example of a representative statement is "What is the highest GPA that you feel completely certain you can attain?" (Richardson et al., 2012).

improve students' social, academic and institutional integration, which supports goal commitment, persistence, and academic achievement (Richardson et al., 2012).[9]

Finally, the accounting education literature demonstrates several other determinants of student performance, in addition to those discussed above (Koh, 2014). One of them deserves further attention, i.e., procrastination. Ackerman and Gross (2005) define procrastination as the delay of a task or assignment that is under one's control and that needs to be done within an expected time frame. There are various documented reasons for procrastination, such as lack of intrinsic motivation and problems in self-regulation (Senecal, Koestner, & Vallerand, 1995), fear of negative evaluation (Saddler & Buley, 1999), locus of control (Janssen & Carton, 1999), and low self-efficacy (Haycock, McCarthy, & Skay, 1998). In an accounting context, procrastination is documented to be associated with lower task performance, even after controlling for student quality (Rotenstein, Davis & Tatum, 2009). Interestingly, Janssen and Carton (1999) demonstrate that students with an internal locus of control completed assignments sooner than their fellows with an external locus of control.[10] Collectively, following previous literature, it is expected that assessment rubrics can prevent students' procrastination in various ways, and more specifically, procrastination is also linked to self-regulation and self-efficacy, which we discussed in the preceding paragraphs.

2.3 *Theoretical Framework: Constructivist Model of Learning for Advanced Written Assignments and the Accompanying Group Work*

Figure 1 shows the theoretical framework of the study. It is based on Biggs (1993) and Ramsden's (2003) constructivists models of student learning which are synthesized in the recent literature (Duff & McKinstry, 2007, p. 187; Lucas & Mladenovic, 2014, p. 127). The 3P model consists of presage, process and product phases, and assessment plays an integral role at the presage phase. The 3P model is a useful framework for reflecting teaching and learning because it stresses the importance of achieving constructive alignment between the three phases (Biggs, 2003; Lucas & Mladenovic, 2014).

We modify the model to fit it with the context of written assignment and the accompanying group work. We place assessment methods and tools in the core of the current learning context because assessment is the topic of the paper. We make assessment rubrics an explicit subcategory of assessment and suggest that they can assist learning in several ways. First, they shape students' perception of task requirements as part of the current learning context. Second, they interact with student factors and thus can increase feelings of self-efficacy and improve students' social and academic integration. Following the previous 3P models, we link student factors to students' perceptions of task requirements.[11] Students'

---

[9] The three forms of integration are defined in Richardson et al. (2012, p.358). Social integration means perceived social integration and the ability to relate to other students. An example of a representative statement is "I find it easy to get to know other people". Academic integration means perceived support from professors. An example of a representative statement is "Professors take a personal interest in helping me with my work". Institutional integration means commitment to the institution. An example of a representative statement is "I am confident that I made the right decision in choosing to attend this university".

[10] Internal versus external control refers to the degree to which people expect that a reinforcement or an outcome of their behavior is contingent on their own behavior or personal characteristics versus the degree to which people expect that the reinforcement or outcome is a function of chance, luck, or fate, is under the control of powerful others, or is simply unpredictable (Rotter, 1989).

[11] We acknowledge that the psychosocial context (Richardson et al., 2012) may affect learning outcomes also through an alternative path. For example, the student may achieve a flow-feeling, which directly improves the learning process without affecting her/his perception of task requirements. Moreover, we acknowledge that our model aggregates students' feelings and does not make a distinction between past feelings and those which occur

perceptions of task requirements affect their approaches to learning which explain their learning outcomes. Third, assessment rubrics can help students to self-assess their work in line with the theory on self-regulation, and thereby improve their learning process, which leads to better learning outcomes.

We predict that the rubric helps students to learn because it aligns the assessment with the expected learning outcomes of the bachelor's thesis seminar course. By following the principle of constructive alignment (Biggs, 2003; Lucas & Mladenovic, 2004), the rubric assists students in learning by making the learning outcomes more concrete. The positive effects of increased clarity in assessment can occur either through an improved learning approach, enhanced self-assessment capabilities, or increased feelings of self-efficacy. All of these elements are also linked to increased motivation to engage in the writing task, and with the rubric the probabilities are higher that the student may achieve a flow-feeling in writing. Moreover, the rubric may assist students to integrate socially and academically at the university, which also motivates them to learn more. The accompanying group work is important for the social integration and group dynamics in the course.

(Figure 1 about here)

3.   The Context of the Study

3.1   *The Bachelor's Program at the Aalto University School of Business*

In the Finnish system, university studies consist of a three-year bachelor's degree program and a two-year master's degree program. In this system, students are required to write a thesis at the undergraduate level, a tradition widely adopted in European universities. In contrast, in English-speaking countries, students rarely write a thesis at the undergraduate level. Despite this difference, Riley and Simons (2013) demonstrate that, in the U.S. and Australia, accounting curricula include several types of writing assignments (e.g., Laufer & Crosser, 1990; English, Bonanno, Ihnatko, Webb, & Jones, 1999; Riordan, Riordan, & Sullivan, 2000; Ashbaugh, Johnstone, & Warfield, 2002).

In Finland, most of the students continue to the master's degree program in which they have to write a master's thesis. The bachelor's thesis thus trains students for the more demanding thesis at the master's level. In addition to the transferred knowledge, the bachelor's program teaches students writing skills and generic skills (e.g., interaction and presentation skills), which all constitute important assets in their working life. Assessment rubrics are one way to help students learn these skills. The need to develop these skills is highlighted in the recent accounting education literature (e.g., Daff et al., 2012; Dale-Jones et al., 2013; Turner & Baskerville, 2013). It shows that approaches to improve students' writing and seminar skills are very welcome regardless of the educational system.

Our development project was carried out at the Aalto University School of Business in the Department of Accounting. The accounting specialization in the bachelor's program gives students the ability to produce, analyze, interpret, and communicate accounting information

---

in the context of the writing course. For the sake of brevity, we do not add these additional dimensions in the model.

and to use this information to assist the decision-making of internal management and external stakeholders. In general, the program offers students opportunities to develop analytical thinking, a scientific approach to solving problems, teamwork skills, an ethical understanding of business, and language and communication skills.[12]

3.2 *Bachelor's Thesis Seminar*

The bachelor's thesis seminar is one course included in the specialization studies of the bachelor's program. It is a compulsory 12-credit course typically taken in the fifth or sixth semester. The syllabus of the bachelor's thesis course can be found in Appendix A. The thesis can be exclusively a literature review or can also include an empirical part. The bachelor's thesis is limited to 20-25 pages (excluding the references and appendices).

The main learning outcome of the course is to engage students in performing independent accounting research. Passing the course requires students to write a thesis, act as an opponent for other students' theses, take notes in the seminar, and actively participate in the sessions. The written thesis must be presented to other group members in seminar sessions held towards the end of the term. 70% of the course grade is based on the thesis and 30% on the process. The process consists of the quality of the opponent work (15%) and course activity (15%). The maximum grade for the course is 5.

The course introduces students to different research topics and methods and, even more importantly, it provides them with a good setting for developing research skills. At their best, bachelor's thesis seminars are dynamic sessions in which students participate in discussions and enjoy an academic atmosphere. At their worst, they may be inflexible sessions in which students do not have the motivation and/or courage to get involved. Participating in the course is expected to increase students' feeling of self-efficacy, and thereby decrease their fears about doing research (see Bandura, 1977). The bachelor's thesis seminar differs from the previous courses in the program because students receive more responsibility and feel free to start examining what they are really interested in, as can be seen from the examples of thesis topics in Appendix B.

4. Justification for the Development Work

Before developing the assessment rubrics, we utilized surveys to examine students' opinions of the course. The surveys were intended to justify the need for assessment rubrics and to help better understand the research context. Students' motivation, confidence, and opinions of the existing assessment criteria (pre-rubric era) were first examined via closed-ended questions (LIKERT-scale range 1-5). After this, they were asked to answer certain open-ended questions that were analyzed by following the principles of content analysis. The survey-method was chosen as the research method because we wanted to collect comparable data (between groups and questionnaires) from a sample of students (see Cooper & Schindler, 2008, p. 215). We collected the survey data at the beginning and at the end of the bachelor's thesis seminar in spring 2012. Detailed descriptions of the starting and ending questionnaires are available from the authors upon request.

---

[12] Further information on the Department of Accounting can be found at the following address: http://accounting.aalto.fi/en/

4.1   *Summary of the Survey Results*

The starting questionnaire demonstrated that, on average, students had enough motivation and self-confidence to start writing their bachelor's thesis. Most of the students understood fairly well that the process was also being assessed. Open questions revealed that students considered that obtaining support from their fellow students and supervisor during the course was important, as can be concluded from the following comments:

> *'It depends very much on the group members how active and useful discussions we will have on the thesis topics.'*

> *'Other students give peer group support and important feedback when they act as opponents. Sometimes I become blind to my own work and my fellow students see better what should be done.'*

> *'The supervisor has an important role in guiding during the process. I can ask guidance and opinions from my supervisor regarding my thesis topic, which helps to increase the quality of the work.'*

We also asked for each student's own perception about her or his role as a seminar participant. We found that many students had a clear view of their role, but some of them felt that there were many challenges concerning their role, as can be seen from the following comments:

> *'My own role is to present my thesis for the opponents and the audience and act as the 'specialist in that area'. On the other hand, it is important to participate in the discussion when other course participants present their work.'*

> *'I think of it as a challenge that I should defend my work and criticize other students' work because I am not that good at answering and accepting the critique.'*

The ending questionnaire revealed that students understood the grading principles of the bachelor's thesis seminar very well. This finding implies that most of the students understood that 30% of the course grade comes from the process (i.e., opponent work and course activity) and 70% from the written study. However, the survey provided evidence that students had problems understanding the grading principles of the bachelor's thesis.

The ending questionnaire also helped us to understand the role that a bachelor's thesis seminar can have in motivating students, as can be seen from the following quote:

> *'In the third year, my motivation started to reappear and it increased a lot as I started to write a bachelor's thesis. I found a very interesting subject, and I think that I would like to work with it also in the future. This has given me totally new motivation for the future master's studies.'*

Descriptive statistics and correlation analysis have a communicative role in this paper. Table 1 gives descriptive statistics for the examined survey variables that we collected in 2012-2014. The sample consists of 64 students who returned both the starting and ending questionnaires.[13]

The mean grade of the bachelor's thesis seminar course (*BS_Grade*) is 4.141. The mean, median and mode values of the variable are very similar, which constitutes evidence that the variable is normally distributed. The maximum grade is 5.000 and the minimum grade is 2.000, which suggests that, on average, students perform well on the course. 65.60% of the observations are men (*Male*), which provides evidence that there is some male dominance in the sample. Usually, the course assists students to increase understanding of academic research, as can be seen from the mean value 0.539 of *Ch_understanding*. Moreover, we find evidence that, on average, students' motivation to write a bachelor's thesis in the beginning of the course is higher than their motivation to write a master's thesis in the end of the course (mean -0.211 for *Ch_motivation*). On the contrary, students' self-confidence develops positively, as can be seen from the mean value of 0.273 for *Ch_confidence*.[14] There is lot of variation in the changes of students' understanding, motivation and confidence, as can be seen from the high coefficient of variation values. The mean (median) values for *Pre_understanding*, *Pre_motivation*, and *Pre_confidence* are 3.68 (4.000), 3.820 (4.000), 3.359 (3.000), respectively. However, there is also variation between the students because all of these variables have the maximum value of 5, and the minimum value is 2.000 for understanding and confidence, and 1.000 for motivation. This finding signals that there were differences in the students' understanding, motivation, and self-confidence to start writing the thesis at the beginning of the course.

(Table 1 about here)

Table 2 gives the correlation matrix. *BS_grade* is positively and significantly correlated with *Ch_understanding* and *Ch_confidence*. This finding suggests that course grade is associated with similar constructs (e.g., learning) that also determine changes in understanding and confidence. Students who feel more improvement in the capability of doing research after the course have obtained higher grades. Interestingly, male students had higher confidence to start writing the bachelor's thesis, as can be seen from the positive and significant correlation coefficient between *Male* and *Pre_confidence*. We also find evidence for the benefit of the course for increasing understanding of academic research. Especially, those students who have a low initial understanding, motivation, and confidence demonstrate learning new on the course (*Ch_understanding* negatively and significantly correlated with *Pre_understanding*, *Pre_motivation* and *Pre_confidence*). Low levels of initial understanding, motivation, and confidence are also negatively and significantly correlated with changes in

---

[13] We used descriptive analysis to justify the need to develop the assessment rubric. However, in Table 1, we report statistics which also use available data for understanding, motivation, and confidence in the time period after the implementation of the rubric to increase the sample size.

[14] In the end of the bachelor's thesis seminar course, we asked for undergraduate students' motivation and self-confidence regarding the writing of the master's thesis in the future. We acknowledge that the motivation and confidence towards the writing of the bachelor's thesis is not totally the same as motivation and confidence towards the writing of the master's thesis. However, comparing these measures provides a good surrogate for the development of motivation and confidence during the bachelor's thesis seminar course.

motivation and confidence. Change in motivation is positively correlated with change in self-confidence. Moreover, starting motivation and confidence are also strongly correlated. The mutual correlations of the covariates are low enough to suggest that multicollinearity threat should not be severe in the multivariate analysis. The highest correlation coefficient (Pearson) is -0.617, which is documented between *Ch_understanding* and *Pre_understanding*.

(Table 2 about here)

Finally, we analyze the determinants of students' performance in the bachelor's thesis seminar from a wider perspective. Student performance is expected to be partly driven by student factors in addition to the current learning context, as suggested in the constructivist model of learning (see Ramsden, 2003; Duff & McKinstry, 2007; Lucas & Mladenovic, 2014). We attempt to examine some of those factors by using the archival study register data of the Aalto University School of Business. The final sample includes 204 observations. We regress the grade of the bachelor's thesis seminar on the following variables: grade in the Principles of Accounting course, gender, age, and the student's basic studies grade point average.

The results (untabulated) provide evidence that the explanatory power of the model is 15.80%. Age is negatively and significantly associated with the grade, and the student's basic studies grade point average is positively and significantly correlated with the grade. Moreover, historically, male students' grades have been significantly lower. For this study, the most important information is that 84.20% of the bachelor's thesis seminar grade can be explained by other factors. It is very probable that a student's grade in the Principles of Accounting course and the student's basic studies grade point average correlate with individual factors that explain learning outcomes. This finding implies that, on average, students' grades in the bachelor's thesis seminar are not only determined by 'ex-ante factors' that lead to good learning outcomes despite the learning environment. Thus, the results demonstrate that much room exists for affecting students' learning by using a proper assessment rubric to help them learn.

5.   The Development of the Assessment Rubric

5.1   *The Structure of the Rubric*

When we started to develop the rubric, we had two objectives. First, we wanted to develop a framework that guides students in their work, but that also leaves some 'power' to the instructor to adjust grading if needed. Second, we wanted to integrate the process view into the rubric and explicitly show what a high-quality process would mean.

The development project started with discussions between the authors. Both of them have much experience teaching bachelor's thesis seminars. One has supervised more than 400 theses focusing on management-accounting-related topics. The other has focused on financial-accounting-related topics and guided approximately 80 students' work. Consequently, considerable experience could be used in the development work. Our experience helped us to distinguish the challenging issues for students who are writing their

bachelor's thesis. The authors have also evaluated each other's teaching in the bachelor's thesis seminars, which has increased their understanding of the interaction between the students and instructor during the seminar sessions. They have also discussed the key criteria that they apply when they assess a bachelor's thesis and the quality of the seminar process. Moreover, they have gathered tacit knowledge from discussions with other teachers on the assessment of the bachelor's thesis seminar course. Because one of the authors specialized in qualitative research in management accounting and the other in quantitative research in financial accounting, they were able to take both of these research methodologies into account in the analyses of the survey evidence.

During the discussions, the main criteria that should be included in the rubric were considered. The authors agreed that the following principles should be kept in mind when formulating the rubric. First, it should be easy for the undergraduate students to understand. It should not include any 'boiler plate' text, which merely provides general information. The definitions of the quality criteria should be clear and should provide concrete examples of those issues that should be taken into account in the thesis. Second, the rubric should demonstrate that the quality of the thesis originates from the students' performance in several areas, all of which will be assessed. For example, preparing an empirical statistical study in the thesis is not a reason to put less effort into the literature review. An empirical study will only mean there will be less room for the literature review, which underlines the importance of the paper's being well organized.

Third, the rubric should be a good tool for the students' self-assessment and the instructors' course grading. There should still be flexibility in the rubric that makes it possible to take into account different circumstances and different groups. The most important function of the rubric should be to demonstrate to students the elements of high-quality coursework, and in this way to increase students' motivation and feeling of self-efficacy. Fourth, criteria for the study process should also be included in the standard to provide students with detailed examples of what differentiates excellent from inadequate work.

After developing the first version of the rubric, the authors analyzed its content critically and decided on the needed changes. There were several development rounds before the final rubric was ready. In the first round, the authors prepared a working version of the rubric and had an in-depth discussion on its content. The rubric was revised based on the feedback of the authors, and these revisions continued until the authors did not have any more ideas to improve the content of the rubric. The development work for the rubric was part of the advanced pedagogical studies that prepared the authors to obtain the university teacher degree. Consequently, they received feedback from their supervisors during the project, which also had some influence on the formulation of the assessment criteria. Appendix C and D illustrate the assessment rubric for the bachelor's thesis seminar. We decided to divide the assessment criteria into seven themes:

1. Research question
2. Structure of the study
3. Literature review
4. Empirical analysis
5. Conclusions
6. Fine-tuning and style
7. Process

We developed the assessment criteria for each theme for the following grades: inadequate work (2), fairly good work (3), good work (4), and excellent work (5). We did not want to include the lowest level (1) in the rubric because the statistics documented that students very seldom receive the lowest grade in the course. By including only four levels in the rubric, we hoped to increase its readability and clarity. We believe that four levels make the rubric more understandable for the students.[15]

5.2   *The Content of the Assessment Criteria*

The above-mentioned themes were those that we considered to be the most important for assessing the quality of the written thesis and process during the seminars. For example, motivating and formulating the research question and explicitly demonstrating the need for the research (research gap) is one of the most challenging tasks for students. Moreover, a balanced structure is important; the literature review and, if included, the empirical analysis should both receive enough space in the thesis and the paper should not exceed the length limits. One essential assessment criterion for us is the quality of the literature review. First, various high-quality references from journals are appreciated. Second, because one of the purposes of university studies is to teach students academic critical thinking, we consider dialogue between the references to be an important factor that should be included in the rubric. By showing the ability to take into account different views presented in the previous literature, the student can demonstrate that (s)he understands 'the big picture' behind her/his research question.

We discussed the assessment of the empirical analysis based on our methodological specializations. Furthermore, credible empirical analysis and interesting results are valuable. One common problem from our experience is that students do not link the results of the study to the research question and existing theory and/or previous empirical findings. Therefore, we included this as an assessment component under the 'Conclusions' criterion. Moreover, excellent work always also includes the researcher's 'own voice,' which assists the reader to understand how the author interprets the literature and the results.

Fine-tuning and style are important factors, but often underappreciated among the students. How you present the paper to the reader counts greatly, as every scholar is aware. For example, by using metatext, the author can make the reading a pleasant and a somewhat exciting experience for the reader. An excellent bachelor's thesis is also clearly written and consistent. In addition, carefully prepared self-explanatory tables are important and increase the readability of the study.

Appendix D gives detailed assessment criteria for the process. The process criteria have been divided into two components: opponent work and activity. It demonstrates very clearly what is needed for good course activity. If the student participates actively in discussion, (s)he will also be awarded even though the quality of the comments could be improved. In contrast, passive students who do not contribute to classroom discussions will not be rewarded. Furthermore, the rubric very clearly described that high-quality opponent work includes careful preparation, effective cooperation with the second opponent, and the right balance between fundamental issues and minor details. The rubric is supposed to help students in their work. In line with the theoretical model of the paper, we propose that the rubric can assist students to optimize their learning approach or to self-assess their

---

[15] For the sake of brevity, we give the final rubric in Appendix C and D. The final rubric is the improved version of the rubric that we tested with the students (please see section 5.7).

performance. It can also increase the self-efficacy and motivation, and improve students' social and academic integration.

5.3   *The Implementation of the Rubric in 2013 and 2014 Bachelor's Thesis Seminars*

We implemented the rubric in 2013 and 2014 during the spring and fall semesters. In the spring semester 2013, we only had a survey at the end of the course. In other semesters, we conducted the survey both at the beginning and at the end of the course. Moreover, in 2013, semi-structured discussions with the group members were incorporated in the feedback sessions after the seminar presentations, in addition to the surveys. This way, we were able to obtain more insights into the usefulness of the new rubric. The feedback on the impact of the rubric on learning was encouraging. We examined students' views on the rubric via six closed-ended questions. The results provided evidence that, on average, the rubric was a useful tool during the bachelor's thesis seminar, as illustrated in Figures 2 and 3. A majority of the students gave score 4 (scale 1-5) for the effect of the rubric on learning motivation. Moreover, students also considered the rubric to be very useful in supporting their working (scores 3 and 4 were the most common answers).

(Figures 2 and 3 about here)

Table 3 reports all of the questions and their summary statistics for the utility of the rubric at the beginning and at the end of the course.[16] The number of observations is 35 (in all statements, except one where it is 34). In the ending survey, the highest mean value (3.886) is reported for question 6. Accordingly, many students felt that the rubric assisted them to understand the components of the course grade. For most of the statements, the mean and mode value is 4.000, the 25th percentile value is 3.000, and the 75th percentile value is 4.000. The maximum value is 5.000 for all statements in both the starting and ending questionnaires. The minimum value is between 1.000-3.000 for the starting questionnaire and 1.000 for the ending questionnaire.

The standard deviation is also higher for all statements in the ending questionnaire than in the starting questionnaire. This finding suggests that, in the beginning of the course, the students were more unanimous in thinking that they would benefit more from the rubric than in the end of the course. At that point, they had not yet utilized the rubric, and perhaps had somewhat vague expectations about it. During the course, they realized more in detail whether or not the rubric is beneficial for them, and the answers regarding the benefit of the rubric were more realistic. This can be seen from the lower mean values for the statements in the ending questionnaire than in the starting questionnaire. Moreover, some students explicitly mentioned in the survey that they did not use the rubric. For example, one of them forgot to use the rubric and gave grade 1 for every statement in the ending questionnaire. However, the answers of the ending questionnaire reveal that, for the majority of the students,

---

[16] In the first survey in spring 2013, the phrasing of the questions varied slightly compared to the following surveys (starting from fall 2013), although the meaning of the questions was largely the same. In addition, this survey was conducted only at the end of the course. In Table 3, we omit those students, and only use participants who handed in both the starting and ending questionnaires in fall 2013, spring 2014, and fall 2014. The mean values for the six questions regarding the usefulness of the rubric in spring 2013 are 4.67, 4.33, 4.44, 4.22, 3.89, and 4.67, respectively. The number of observations is 9.

the rubric was a useful tool during the course. *Rubric_benefit* (see Table 1) is the average score of the answers relating to the usefulness of the rubric in the ending questionnaire. The mean (median) is 3.417 (4.000), and the 25th (75th) percentile value is 3.000 (4.000).

(Table 3 about here)

We next conduct multivariate analyses for the determinants of course grade and report the results in Table 4. In the first model, we regress *BS_grade* on *Rubric_benefit* and the control variables. The number of observations in the model is 35, the model F-value is 2.820, and the adjusted R-square is 30.00%. The regression coefficient of *Rubric_benefit* is positive and non-significant (P-value 0.256 in a two-sided test). Taking into account the small sample size and the threat of type-two errors, it is however worthwhile to note that the sign of the coefficient is positive and we are not too far from the 10% significance level (in a one sided-test) with the small sample. *Ch_understanding*, *Ch_confidence*, and *Pre-confidence* are positively and significantly associated with *BS_grade*. In line with the results of the correlation analysis, this finding suggests that those students who develop more in the course (in terms of improvements in understanding and confidence) obtain higher grades. Starting confidence is a good surrogate for student factors that are expected to affect learning outcomes in the constructivist model of learning, and therefore, the significance of the variable is not surprising.

In the second model, we regress *Ch_understanding* on *Rubric_benefit* and the relevant covariates. The model is significant, and the adjusted R-square is 25.60%. *Pre_understanding* is negatively and significantly associated with changes in understanding. Thus, the result that was documented in the correlation analysis holds after controlling other potential drivers of improved understanding. After this, we regress *Ch_confidence* on *Rubric_benefit* and the control variables. The explanatory power of the significant model is 48.90%. Interestingly, *Rubric_benefit* is positively and significantly related to improvements in confidence (P-value 0.048). However, if we control changes in understanding and motivation (model 4), the positive regression coefficient of the rubric benefit is non-significant. The small sample size together with some multicollinearity between the covariates may explain why we cannot document the significance of the rubric in increasing confidence in model 4.[17] As can be seen from the standardized regression coefficients, the relative impact of the rubric on changes in confidence is also smaller than the impact of *Ch_motivation* and students factors (*Pre_understanding*, *Pre_motivation*, *Pre_confidence*).

(Table 4 about here)

Finally, we examine if the rubric has an indirect influence on course grade via its positive effect on understanding and confidence, which were the most significant determinants of course grades in the regression analysis. We compute two variables to measure if the students obtained a significant treatment from the rubric. *High_usefulness1* obtains value 1 if the score of Rubric_benefit is higher than the median value. *High_usefulness2* obtains value 1 if the score of Rubric_benefit is equal to, or higher than, the 75th percentile value. We then employ propensity score matching and try to find student pairs

---

[17] The mean VIF-value of the regressors is 1.61, and all VIF-values are below 2. Thus, multicollinearity probably has only a small effect on the results.

that are close to each other regarding the covariates (Male, Pre_understanding, Pre_motivation, Pre_confidence), but only the other received the treatment.[18] We obtained evidence that those students who benefited from the rubric (treatment: *High_usefulness1*) had significantly higher *Ch_understanding*. The average treatment value on treated (ATT) is 0.667, and for the control group the mean value is 0.144. The difference is also significant (t-value 2.583). If we use *High_usefulness2* as a treatment variable, the equivalent mean values are 0.800 and 0.317, respectively. However, the significance is slightly lower (t-value 1.743).

For change in confidence (treatment: *High_usefulness1*), we document the mean value 0.633 for the treated, and the mean value 0.400 for the control group. For the second treatment, the gap between the treatment group and control group values is narrower. These differences are not significant, however. Interestingly, if we run a t-test for the two independent samples (treatment: *High_usefulness1*), the mean value of *Ch_confidence* is 0.633 for the treated and 0.100 for the control group, and this difference is also statistically significant (t-value 2.000). This finding suggests that, in terms of confidence, some covariates that impact the usefulness of the rubric may be omitted from the first stage matching model, and therefore the unmatched test works better. To summarize, the propensity score matching results supplement the regression results, and provide confirming evidence that the rubric has an indirect effect on course grades through changes in understanding in the bachelor's thesis seminar course. For confidence, we obtain some evidence on the indirect effect if we utilize an unmatched sample.

5.4 *The Effect of the Rubric on the Approaches to Learning*

Several interesting viewpoints on the advantages of the rubric were presented in the group discussions and in students' answers to the open-ended questions of the surveys.[19] The results of the qualitative analysis gave confirming evidence that, for many students, the rubric was beneficial in improving the quality of their coursework. Next, we discuss our findings regarding the following main reasons why the rubric is expected to be helpful for the students: an optimal approach to learning, self-assessment, self-efficacy, and integration.

We first use Duff and McKinstry's (2007) classification for the characteristics of *deep learning* to analyze students' answers on the rubric. The survey evidence demonstrates that the rubric increases *intention to understand* and assists students to *use evidence and logic to understand content*. The following quotes illustrate these views:

> *'It helps in understanding which items are important. Then, it is easier to focus on these items, which also leads to better learning outcomes.'*

> *'Similarly to the thesis, I also understand the process better via the grading principles.'*

> *'It offers a platform that can be used to assess the quality of my own work. It also helps in understanding the content and presentation technique requirements of the*

---

[18] By using propensity score matching, we can control the observed student heterogeneity and make some suggestions regarding causality.

[19] Open discussions on the advantages of the assessment rubric were conducted on April 2013 and December 2013. They were semi-structured and lasted approximately 15 minutes.

bachelor's thesis. Perhaps it helps understanding scientific research from a wider perspective.'

'The rubric was clear and defined the grading principles accurately. This helped in understanding the objectives and supported learning.'

'It also helps in focusing on the weaknesses. It also clearly explains the requirements for a certain grade, which helps in setting goals.'

'Information on assessment criteria helped me to understand what the thesis should include. It motivated to focus on certain elements in the thesis. In addition, clear assessment criteria also impacted me by encouraging me to focus on the opponent work and course activity.'

Another dimension of deep learning is students' ability *to relate and organize ideas and concepts* (Duff & McKinstry, 2007). The survey evidence suggests that the rubric was useful from this perspective, as well.

'With the rubric, I can better set the pace for the working process when I know to what parts of the work it is encouraged to put emphasis.'

'The working process became easier when I was able to use a framework to follow. With the rubric, it was easy to split the work into smaller items which were much easier to control than focusing all of the time on the whole work.'

'It helped to perceive the content and more specific elements of the bachelor's thesis. Otherwise, I would not necessarily have taken into account all those elements in the writing of the thesis.'

'When I was the opponent, I could compare the theses to the rubric, which made it easier to find good/bad issues in the work and to justify my own feedback.'

'It helped in focusing on high-quality outcomes. Information on the grading principles assists in structuring the thesis. The informed required level brings motivation when you know what is required for good work.'

The rubric was also beneficial in making students learn more deeply because it increased interest and motivation for the work by assisting students to *take an active interest in the subject* (Duff & McKinstry, 2007). The following quote illustrates this:

'Personally, the information that I got in advance motivated me a lot. I am more motivated when I know how to achieve a goal, and this increased motivation correlates strongly with the learning outcomes, as well.'

Another main approach to learning that is documented in the education literature is the *strategic approach*. In Duff and McKinstry (2007), it is characterized as the following: intention to obtain the highest possible grades, willingness to work hard to excel, effective organization of studying, and effective time management. Based on the survey findings, we can conclude that some students were using the rubric for strategic purposes:

'With the rubric, I know what sections of the thesis I need to put specific emphasis on to get a good grade.'

'It gives me motivation to be active in the seminar meetings because it impacts quite a lot to the grade.'

'The rubric assisted me in putting emphasis on the right things, and in this way, I managed my time more effectively.'

'I focused a lot on acting as an opponent because it was emphasized in the assessment.'

Finally, as suggested in Duff and McKinstry (2007), surface learning might be indicative of students' anxiety about coping with the demands of the course, propensity to memorize information needed for assessments, difficulties in making sense, and problems relating concepts. We obtained very little evidence that the rubric was utilized for surface learning purposes. That would be very difficult in courses involving advanced written assignments. However, some students seemed to be reluctant to use the rubric for various reasons, as can be seen from these quotes:

'I mostly read the instructions. I did not read the rubric, properly speaking, at all because the grade and its determinants do not matter to me. I just tried to do what I was able to do.'

'Less, most probably I will not suddenly become more talkative or start working in a more organized way. On the other hand, as I said, maybe the grade is not that important. Maybe the rubric just gives some guidance for how I can try to work.'

5.5   *The Effect of the Rubric on Self-assessment, Self-efficacy, and Integration*

Next, we reflect the survey results against self-regulation theory and the rubric's suggested improvements in self-assessment. We obtained substantial evidence that supports this argument.

'During the writing of the thesis, I can regularly check if the thesis fills, for example, the requirements that are set for grade 4.'

'The grading principles help in concentrating and putting emphasis on the relevant things that are evaluated. These principles bring transparency to the assessment, which makes self-assessment of my work easier.'

'I wish that I get from the rubric guidance and tips for how to write a good thesis and to keep this guidance in my mind throughout the entire writing process.'

'Through the rubric, I will focus more on the different sections of assessment and to make a laudable and rightly structured thesis where all sections are taken into account.'

20

'It helps me to assess my own work and perceive what kinds of elements should be taken into account in the writing of the thesis. Moreover, I understand better that course grade is not only dependent on the bachelor's thesis, but also that the quality of the process influences it.'

Regarding the self-efficacy, we obtained support for the prediction that the rubric is helpful because it increases the self-efficacy and motivation of the students. It may increase or keep the motivation to work because it explicitly shows what the students are expected to do. Thus, it may also reduce the number of dropouts in the course. The following quotes reflect this effect:

'I knew better what is required, so I was able to focus on the right things. Lack of information could have led me in a wrong direction, and then I would not have learned the right things. Lack of information would have also discouraged me.'

'I know what is required for good work, which increases my self-confidence.'

'Personally the information that I received in advance motivated me a lot. I am more motivated when I know how to achieve a goal, and this increased motivation correlates strongly with the learning outcomes, as well.'

'It gives me some perspective on the length and quality requirements of the thesis. It is important that also some more detailed issues, such as the quality of the tables, are stated. Prior knowledge regarding the assessment makes it easier to plan your own working load and try harder. Encouragement for course activity motivates me to participate in discussions.'

'It is easier to see the essential points of the thesis and understand how it should look. On the other hand, it also increases pressure when you see all those requirements and feel that your work will not necessarily satisfy them. However, I believe that my learning results may improve and my motivation will increase because I know the requirements.'

The survey evidence also indicated that the rubric was useful in improving social and/or academic integration:

'Information encourages preparation for the opponent work in seminars very well and perhaps to put more emphasis on the most useful issues that can be taken into discussions to help others.'

'It spurred me to put effort into the opponent work. So, you really thought what kind of feedback would be useful and constructive. As an opponent, you were able to compare the theses to the rubric, which made it easier to find good/bad issues in the work and to justify your own feedback.'

'It did not affect writing the thesis, but increased my active participation in class.'

'You can better prepare for the seminar, and that adds to your own and others´ learning.'

'It motivates me to prepare for the opponent work with care and to take part in discussions.'

5.6   *Other Evidence on the Usefulness of the Rubric*

As demonstrated in the preceding sections, the rubric assisted students in various ways. In addition, we obtained evidence that the students used the rubric both in the starting and fine-tuning phases of the bachelor's thesis:

'Very useful, especially in the beginning of the course.'

'With the rubric, it was easy to start working on the course because you understood all those elements which are included in the assessment of the course.'

'With the rubric, my working process was clearer and more determined – especially at the latter part of the course.'

'I knew what was required for the grade that I wanted. I tried to use this information in the fine-tuning phase of the bachelor's thesis.'

'It will surely make it easier to prepare the thesis. Especially in the fine-tuning phase of the manuscript, it is important to know the requirements of different elements. I think that it will increase my motivation to improve and finalize the thesis.'

It was noted that it would be beneficial if the instructor reminded students of the rubric during the course.

'The instructor should probably remind students to use the rubric also during the latter part of the course.'

Those students who did not benefit from the rubric based on certain criteria considered it to be more helpful based on other criteria, as can be seen from the following quote:

'It did not affect my working and motivation very much, but provided me with the structure of the assessment anyway.'

This finding suggests that there is not any one right way to use the rubric, and the students benefit from it in different ways.

Finally, we complement the survey evidence with evidence from two semi-structured interviews which were held after the spring 2013 and fall 2013 courses. In these feedback discussions, the students pointed out that it is important that the instructor follows the themes of the rubric in her/his comments. There was also some consensus that the rubric is just a tool. Emphasizing it too much may reduce students' eagerness to do their first academic study. Hence, its use should not be overemphasized.

5.7   *Epilogue: The Final Rubric*

The rubric is supposed to facilitate students learning. However, to be credible, the teachers also have to utilize the rubric in a consistent manner. After developing the rubric, we looked back and discussed how the rubric could be improved by fine-tuning it and by norming it to be more widely applicable for assessment.[20] We reviewed the rubric that we employed for the survey years. After this, we wanted to revise it to its final form and demonstrate how it could be normed for wider use. Thus, we used our long experience as thesis supervisors and the received feedback and experiences from the rubric as a benchmark and discussed how we could improve it further. We discussed scoring based on different assessment criteria. We decided to quantify the rubric more and make it more specific and explicit, because we could guide students more clearly in this way. The revisions were as follows.

First, we increased the clarity of the rubric by making clearer distinctions between different grade categories for each valued criteria. Second, we increased the concreteness of the framework by using exact numbers to show the number of needed pages of written text for each grade, and to specify the number of needed references for the literature review. Third, we decided to weight the assessment criteria to signal students the slight differences in their importance. We decided that the research question, literature review and conclusions obtain a little bit more weight (20%) in the rubric. This way, we can demonstrate their importance for the students. Structure and empirical analyses get 15% of weight and the fine-tuning and style count for 10%. The process is equal to 30% of the course grade and both of its dimensions (opponent work and activity) count for 15%. The grading principle would be to give the grade based on these weights for the dimensions.

Finally, we chose two papers from the 2012-2014 courses (one from both supervisors' groups) and used these as practice papers that we scored independently. After this, we compared the assessment scores in every dimension and had a profound discussion on the results. In the assessment of the first paper, the other scholar gave 4.6 points for the thesis and the other 4.2. After the discussion, we decided that practical motivation should be considered as a value-adding, but not necessary, element for grade 5 in the assessment of the research question. This is so because sometimes scientific motivation dominates and practical motivation is not that easy to find. We also noticed that it would be easier to give a 10% weight to both components of the literature review (used literature; critical thinking of the literature review) because the student does not necessarily get the same grade from these elements. In the assessment of the second paper, the other scholar gave 3.75 points and the other 4.25. Discussion of this grading made us add the comment, on over 25 page-long papers, to the criterion 'structure of the study'. We agreed that if the thesis is longer than 25 pages, it should not be repetitive. Otherwise it has a negative impact on the grade. Other differences in grading were also discussed against these practice papers. For example, we discussed the importance of linking the results to the research question and previous literature. We inferred that one good approach would be to have a separate section called 'interpretation of the results', and this could be followed by the 'conclusions' section where the research is summarized and conclusions are given at the meta-level. Furthermore, we decided to add two elements in the overall assessment of the thesis which give the teacher some latitude to adjust the grade. Both the independence of the student and the novelty of the research question have positive effects on the grade.

---

[20] We obtained information on the norming of the rubric from the following site: http://www.teachingmatters.org/toolkit/norm-setting-protocol

An interesting remark is that both of us gave a lower grade for the thesis than originally after the course. We concluded that quantifying the assessment rubric and discussion with a peer to norm the rubric gave us confidence to give lower grades for specific assessment criterion when needed. Another interesting finding was that both of us gave higher scores to that thesis which was out of her/his area of expertise (i.e., financial accounting or management accounting).

The final rubric is given in Appendix C and D. For example, let us assume that the grades for the six dimensions of the thesis are 5 (research question), 4 (structure of the study), 5 (used literature), 4 (critical thinking of the literature review), 5 (empirical analysis), 4 (conclusions), and 3 (fine-tuning and style). In this case, the grade of the thesis would be as follows: $0.20*5 + 0.15*4 + 0.10*5 + 0.10*4 + 0.15*5 + 0.20*4 + 0.10*3 = 4.35$. Then let us assume that the student got 4 from the opponent work and 3 from activity. This would give the following points: $0.5*4 + 0.5*3 = 3.50$. The final grade from the bachelor's thesis seminar course would then be the following $0.7*4.35 + 0.3*3.50 = 4.095$. This would be rounded to 4.

5.8 *Discussion of the Usefulness of the Rubric for Learning*

With the rubric, we can increase transparency and show students the essential quality dimensions of the written thesis. Furthermore, we can set the standard for the oral seminar work and thereby create room for more dynamic and interactive sessions. We cannot assert that the rubric helps some students to learn more deeply or more strategically. Students probably use the rubric in heterogeneous ways. Some students can utilize it strategically to focus their efforts on the right things. This is beneficial because the writing time for the thesis is usually from six to eight weeks. Hence, a too-deep learning approach is not optimal either, if it takes excessive time, and increases the risk of dropping-out from the course because of procrastination. Finally, we admit that some students do not necessarily need the rubric which, however, does not reduce its merits for the majority of the students. In our opinion, the course is difficult, if not impossible, to pass by surface learning because memorizing does not help in the writing of the thesis. However, we recognize the possibility that some students may employ the rubric only superficially.

Moreover, a student's approach to learning is not absolutely rooted to deep, strategic or surface categories, but can include elements from various categories (Thomas & Bain, 1984; Beattie et al., 1997). Although the student would follow a strategic approach and get better grades, her/his approach may also relate to deep learning if (s)he engages with the task at hand (Biggs, 1989; Biggs, 1993). According to Biggs (1993), a student's approach to learning is relational (Ramsden, 1987) and therefore, a deep approach is seen as a way of describing how a student relates to a task. Teaching a student a bundle of deep strategies for future use without linking them to a context is not recommended (Biggs, 1993). Biggs (1989) suggests that teaching should encourage deep and achieving approaches to learning.

Those students who were more grade-oriented were probably able to put their effort in the essential things. It is difficult to believe that their learning results would have been any better without the rubric. Those students who were seeking deeper meaning got structure from the rubric. For some deeply-processing students, the problem may be that they cannot cope with the time frame given for the written assignment. If they try to write their assignment without proper guidelines, the required level of performance may become too abstract and students may start suffering from procrastination and a loss of motivation. Moreover, the rubric was not only useful for the written assignment. It also gave the structure for the classroom learning. By being explicit about what is required, we were able to set the standard

for the oral seminar work. This way, we created room for more dynamic and interactive seminar sessions.

The quantitative analyses demonstrate that the rubric improves students' learning. On average, the course grades were high, which is a coarse surrogate indicating that students learned to do research (Biggs, 1989, Ramburuth & Mladenovic, 2004). Moreover, we find that the rubric helps students to increase their understanding of research, and that increase in understanding is positively related to course grades. This is an important finding because understanding is linked to deep learning (Duff & McKinstry, 2007). It seems evident that, for many students, the rubric is a tool which helps them to comprehend what research is about. This interpretation is reinforced by the qualitative analyses of the survey answers. The results indicate several positive opinions regarding the usefulness of the rubric in aiding students. For example, the rubric assists them in understanding which items are important, and focusing on these items leads to better learning outcomes.

Our constructivist model of learning for advanced written assignments and the accompanying group work takes a broader view of learning theory. We added student's self-assessment, self-efficacy, and integration in the model as additional determinants of students' learning. The survey evidence demonstrates that the rubric was a beneficial self-assessment tool. The grading principles assisted in concentrating on the relevant factors that are evaluated. These principles brought transparency to the assessment, which made self-assessment of one's own work easier. Self-regulation theory suggests that the quality of self-assessment is an important determinant of learning (e.g., Vermunt, 1998; Kostons et al., 2012). Our results support this argument in a writing assignment context.

The importance of the self-efficacy deserves recognition. Our result suggests that the confidence in the beginning of the course and the change in the confidence during the course are positively and significantly associated with course grades. The quantitative analyses also give moderately significant evidence on the positive effect of the rubric on change in confidence. Especially those students who have a low starting confidence may benefit from the rubric if it increases their feeling of self-efficacy. The unmatched results provide evidence that those students who benefited more from the rubric had a significantly higher change in confidence (0.633 vs. 0.100 in tests for high_usefulness1). The qualitative results are in line with this finding because we received many comments which highlighted the benefit of the rubric for increasing the self-efficacy. According to the students' answers, they knew better what is required and they were able to focus on the right things. Lack of information would have also discouraged them. The effect of the rubric on self-efficacy in the course is in line with prior self-efficacy literature in general (Banduras, 1977; Banduras, 1997) and specifically with those views which suggest that self-efficacy is related to flow-feeling, positive emotional experiences, and a higher degree of deep learning (Csikszentmihalyi, 1994; Krapp, 2002; Csikszentmihalyi, 2008).

Following Tinto's (1975, 1982) suggestions, it is expected that a higher degree of integration during the writing course reduces the risk of drop-outs from the course. We find evidence that the rubric may be beneficial in enhancing peer-group and faculty interaction, which influences social and academic integration in the course. The assessment rubric encouraged strong preparation for the opponent work in seminars and putting emphasis on the most useful issues that can be taken into discussions to help others. Integration is beneficial for the students because it supports goal commitment, persistence, and academic achievement (Richardson et al., 2012). Moreover, rubrics can also be utilized to improve opponents' peer evaluation. Finally, rubrics are a powerful way for the instructor to justify to students their final grades from the thesis seminar. This can be done by using a written thesis assessment form in which the instructor explicitly shows how the student performed in each of the quality criteria.

One good question for further developments work is how quantified the rubric should be. We made it explicit and specific, and gave the range of the needed number of pages and references in the thesis. Transparency should enhance the guidance effect. However, it may also make the rubric too grade-oriented, which starts reducing the joy of learning. In addition, the teacher may lose some flexibility in her/his grading. We believe that the value weights for the specific assessment criteria are a good way to show students the order of importance of different themes.

Furthermore, the inclusion of 'process' in the rubric could be criticized by researchers who argue that students' temperament affects their activity (Mullola, Hintsanen, Jokela, Lipsanen, Alatupa, Ravala, & Keltikangas-Järvinen, 2014). These researchers assert that activity in the class should not be evaluated as part of the course grade because it is so tightly tied to a student's inherent temperament. However, in our opinion, the bachelor's thesis seminar is an excellent venue for teaching those communication and interaction skills that are also needed at work, especially in those jobs that newly graduated business students obtain. We admit that, for some students, those issues that are evaluated in the process are inherently easier. Nonetheless, we also argue that explicit assessment rubrics may be very helpful for those students for whom interaction is more difficult. Clear guidance is a good way to decrease the mystery behind the course grades and to level the playing field, despite the temperament of the student.

The rubric also provides a tool to improve the curricula for the universities. Often, assurance of learning is an item that has to be demonstrated for AACSB evaluators. However, we still lack tools that help students in learning during the course. In our opinion, rubrics are an effective way to help students in courses that require a lot of independent work (e.g., advanced written assignments) and accompanying group work (e.g., interactive discussions). We see that the majority of the students liked the rubric because it increases clarity around the often 'scared' research paper that most of the students are preparing for the first time in their lives. Based on the findings of this study and the experiences that we received from the development work, we suggest that transparency in course assessment is one effective way to assist students to optimize their learning approach, help them in self-assessment, and boost their self-efficacy and integration in courses incorporating advanced written assignments and the accompanying group collaboration.

6.   Summary and Conclusions

This study describes the results of a project that focused on developing and implementing an assessment rubric for the bachelor's thesis and the accompanying group work to ensure the quality of students' learning. We utilize the constructivist model of learning and modify it for written assignments and the accompanying group work by adding student's self-assessment, self-efficacy, and integration in the model (modified from Ramsden, 2003; Duff & McKinstry, 2007; Lucas & Mladenovic, 2014). All of these elements together with students' approach to learning are expected to influence learning outcomes.

The development project was carried out at the Aalto University School of Business in the Department of Accounting. We used surveys and descriptive analyses to collect data on students' feelings and perceptions before the development work. These analyses showed that there is a need to obtain more detailed assessment rubrics for the course to ensure that students learn. The final assessment rubric consists of seven main assessment criteria. One novel characteristic of the rubric is that it also takes into account the quality of the process, i.e., the quality of the opponent work and course activity. The process perspective in the

rubric is in line with formative assessment of students' learning that takes place during the entire learning process (see e.g., Lindblom-Ylänne & Nevgi, 2009; Turner & Baskerville, 2013). The rubric was implemented after the development in 2013 and 2014, and feedback on its usefulness was collected via surveys and analyzed both quantitatively and qualitatively. In addition, group discussions were used to collect additional feedback. The final rubric was formulated by following the principles of norming.

The evidence that we have shown in this paper constitutes firm support for the argument that the rubric is useful for learning in courses which incorporate advanced written assignments and accompanying group work. The statistical analysis of the survey data provided evidence that, on average, students get good grades, which suggests that they are learning deeply (Biggs, 1989). Grades are statistically significantly and positively related to changes in understanding and confidence during the course. We also find that those students who benefited more from the rubric gained a statistically significantly higher increase in understanding. In addition, we obtained some evidence that the rubric was useful in increasing the self-confidence of the students. Increase in understanding is closely linked to deep learning (Duff & McKinstry, 2007), and increase in confidence is linked to self-efficacy theory and flow-feeling (Bandura, 1997; Krapp, 2002; Csikszentmihalyi, 2008).

The qualitative analysis of the survey data gave support for the utility of the rubric in all suggested dimensions of the theoretical framework: student's approaches to learning, student's self-assessment, self-efficacy, and social/academic integration. We received several comments that the rubric assisted students to learn more deeply or more strategically (Duff & McKinstry, 2007). For example, students felt that the rubric helped them understand what they were expected to do or aided them in time management. It is important to note that these learning approaches are not necessarily totally exclusive (Biggs, 1989; Biggs, 1993; Beattie et al., 1997). Students also felt that the rubric helped them to self-assess their work in line with self-regulation theory (e.g., Puustinen & Pulkkinen, 2001). The rubric also increased students' self-confidence and motivation, and helped them to integrate with their peers and the teacher in the writing course. In addition, students felt that the rubric assisted them to understand the role of the process in the coursework.

Students utilized the rubric both in the starting and fine-tuning phases of the bachelor's thesis. Students did not always benefit from the rubric in certain areas, but still considered it to be useful in some other areas. In the future, the rubric can be used as a guideline and an assessment tool in bachelor's thesis seminars. Both instructors and students can use it. It will be especially useful for students' self-evaluation of the quality of their written thesis and the accompanying coursework. Moreover, instructors can use it to justify their grading against a quantified framework.

This paper contributes to prior literature by demonstrating how a constructivist-learning-approach-based tool (i.e., the assessment rubric) can be employed to enhance students' learning in courses incorporating advanced written assignments and accompanying group work (Daff et al., 2012; Dale-Jones et al., 2013; Riley & Simons, 2013). The influence can be seen through the optimized approach to learning (e.g., Lucas & Mladenovic, 2004; Duff & McKinstry, 2007), improved self-assessment (e.g., Vermunt, 1998; Puustinen & Pulkkinen, 2001; Kostons et al., 2012), enhanced self-efficacy (e.g., Bandura, 1979; Bandura, 1997), and through better social/academic integration in the course (Tinto, 1975; Richardson et al., 2012). In this way, we provide new insights into the various factors that affect students' learning outcomes from written assignments and also add to the existing assessment literature (e.g., Richardson et al., 2012; Panadero & Jonsson, 2013).

This paper provides evidence on assessment in one top European business school. Hence, this paper also makes a practical contribution to accounting educators who wonder how they can ensure that students learn in thesis seminars or other courses that involve

advanced written assignments. The findings and implications of the study could be applied to improve the curriculum and assessment methods in other business schools, as well. European business schools usually require bachelor's theses at the undergraduate level, which makes these assessment rubrics applicable in those schools. U.S. and Australian business schools, which do not usually include bachelor's theses in their curriculum, may also benefit from the results. For example, they can consider developing assessment rubrics for some other courses at the undergraduate level. In addition, assessment rubrics could also be used to demonstrate the key quality criterion of theses at the master's and Ph.D. levels. It is important to remember that assessment influences learning only if it guides students proactively during the course. Keeping this principle in mind increases the applicability of the rubrics across universities and countries to improve students' writing and the accompanying group work.

One limitation of the study is that we do not have comparable data where, in the same course, the same supervisors taught some students by using the rubric and some students without it. However, we were able to make a decent control group by using students' opinions on the usefulness of the rubric. Another limitation of the study is the small sample size. Therefore, the results may not be generalizable to all courses that require independent work, writing, and/or active participation in seminars. We understand that learning contexts may differ by country and across universities. However, our evidence has a clear theoretical foundation that should increase the external validity of the results. Moreover, we have quantified the rubric and described its norming. This should make it easier to follow our development example and to norm a rubric for a specific context. Through its AACSB accreditation, the Aalto University School of Business is committed to the assurance of learning (AOL) framework. Our findings suggest that the AOL standards should not only be the benchmark for external accreditations, but increasingly a motivator for developing internal assessment methods that can be used to guide students to better learning outcomes. In future studies, it would be interesting to study the implementation and the effects of assessment rubrics on students' learning in other business schools. Their influence on learning in other accounting courses could also be examined.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

AACSB White Paper No. 3 issued by: AACSB International Accreditation Coordinating Committee, AACSB International Accreditation Quality Committee, 20 November 2007.

Aalto University School of Business, Department of Accounting, http://accounting.aalto.fi/en/ (accessed 1.4.2017)

Ackerman, D.S., & Gross, B.L. (2005). My instructor made me do it: Task characteristics of procrastination. *Journal of Marketing Education*, 27 (1), 5–13.

Allan, J. (1996). Learning outcomes in higher education. *Studies in Higher Education,* 21(1), 93–108.

Anderson, L.W. & Krathwohl, D.R. (2001). A taxonomy for learning, teaching, and assessing: a revision of Bloom's taxonomy of educational objectives. New York: Longman.

Andrade, H.L., & Du, Y. (2005). Student perspectives on rubric-referenced assessment. *Practical Assessment, Research and Evaluation*, 10(3), 1–11.

Andrade, H., & Valtcheva, A. (2009). Promoting learning and achievement through self-assessment. *Theory Into Practice*, 48, 12–19.

AOL Handbook. (2011) (version18.5.2011). *Assurance of learning for the curricula of the degree programs in economics and business administration*. Aalto University School of Economics.

Ashbaugh, H., Johnstone, K. M., & Warfield, T. D. (2002). Outcome assessment of a writing-skill improvement initiative: Results and methodological implications. *Issues in Accounting Education,* 17(2), 123–148.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review,* 84(2), 191–215.

Bandura, A. (1997). *Self-efficacy. The exercise of control.* New York: Freeman and Company.

Beattie, V., Collins, B., & McInnes, B. (1997). Deep and surface learning: A simple or simplistic dichotomy? *Accounting Education*, 6(1), 1–12.

Biggs, J.B. (1989). Approaches to the Enhancement of Tertiary Teaching. Higher Education Research & Development, 8(1), 7-25.

Biggs, J.B. (1993). From theory to practice: a cognitive systems approach. *Higher Education Research and Development*, 12(1), 73–85.

Biggs, J.B. (1995). Assessing for learning: some dimensions underlying new approaches to educational assessment. *Alberta Journal of Educational Research*, 41, 1-18.

Biggs, J. (1996). Enhancing teaching through constructive alignment. *Higher Education,* 32(3), 347–364.

Biggs, J. (2003). *Teaching for quality learning at university* (2nd ed.). Buckingham: Society for Research into Higher Education, Open University Press.

Biggs, J., & Collis, K. (1982). *Evaluating the quality of learning: The SOLO taxonomy.* New York: Academic Press.

Biggs, J., & Tang, C. (2007). *Teaching for quality learning at university* (3rd ed.). Suffolk, UK: Society for Research into Higher Education & Open University Press.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy and Practice*, 5(1), 7–73.

Bloom, B.M. (1957). *Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook I: The Cognitive Domain.* New York: David McKay Co Inc.

Brown, G., Bull, J., & Pendlebury, M. (1997). *Assessing students learning in higher education.* London: Routledge.

Cooper, D. R., & Schindler, P. S. (2008). *Business Research methods* (10th ed.). Singapore: McGraw – Hill International edition.

Csikszentmihalyi, M. (1994). *The evolving self: A psychology for the third millennium*. New York: Harper Perennial.

Csikszentmihalyi, M. (2008). *The psychology of optimal experience*. New York: Harper Perennial.

Daff, L., de Lange, P., & Jackling, B.(2012). A comparison of generic skills and emotional intelligence in accounting education. *Issues in Accounting Education,* 27(3), 627-645.

Dale-Jones, G., Hancock P., & Willey, K. (2013). Accounting students in an Australian university improve their writing: but how did it happen? *Accounting Education, An International journal*, 22(6), 544-562.

Dawson, P. (2017). Assessment rubrics: towards clearer and more replicable design, research and practice. *Assessment & Evaluation in Higher Education,* 42(3), 347-360.

Duff, A. (2014). *Learning styles and approaches.* Chapter 8 in The Routledge Company to Accounting Education (Edited by Richard M.S. Wilson), 163–188.

Duff, A., & McKinstry, S. (2007). Students' approaches to learning. *Issues in Accounting Education*, 22(2), 183–214.

English, L., Bonanno H., Ihnatko T., Webb C., & Jones, J. (1999). Learning through writing in a first-year accounting course. *Journal of Accounting Education,* 17 (2/3), 221–254.

French, E., Summers, J., Kinash, S., Lawson, R., Taylor, T., Herbert, J., Fallshaw, E., & Hall, C. (2014). The practice of quality in assuring learning in higher education. *Quality in Higher Education*, 20(1), 24–43.

Goodrich, H. (1996). *Student self-assessment: At the intersection of metacognition and authentic assessment*. Unpublished doctoral dissertation, Harvard University, Cambridge, MA.

Hall, M., Ramsay, A., & Raven, J. (2004). Changing the learning environment to promote deep learning approaches in first-year accounting students. *Accounting Education, An International journal*, 13(4), 489–505.

Haycock, L.A., McCarthy, P., & Skay, C.L. (1998). Procrastination in college students: the role of self-efficacy and anxiety. *Journal of Counseling & Development*, 76 (summer), 317–324.

Hoge, R.D., & Coladarci, T. (1989). Teacher-based judgments of academic achievement: A review of literature. *Review of Educational Research,* 59(3), 297–313.

Janssen, T., & Carton, J.S. (1999). The effects of locus of control and task difficulty on procrastination. *The Journal of Genetic Psychology*, 160(4), 436–442.

Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review*, 2, 130–144.

Koh, H.C. (2014). *Determinants of students' performance in accounting programmes.* Chapter 20 in the Routledge Companion to Accounting Education (Edt. Richard M.S. Wilson), 449–468.

Kostons, D., van Gog, T. & Paas, F. (2012). Training self-assessment and task-selection skills: A cognitive approach to improve self-regulated learning. *Learning and Instruction*, 22, 121–132.

Krapp, A. (2002). Structural and dynamic aspects of interest development: Theoretical considerations from an ontogenetic perspective. *Learning and Instruction*, 12, 383–409.

Laufer, D., & Crosser, R. (1990). The ''writing-across-the-curriculum'' concept in accounting and tax courses. *Journal of Education for Business,* 66(2), 83–87.

Light, G., & Cox, R. (2001). *Learning & teaching in higher education: The reflective professional*. London: Paul Chapman Publishing.

Lindblom-Ylänne, S., & Nevgi, A. (eds.). (2009). *Yliopisto-opettajan käsikirja.* Helsinki: WSOYpro Oy. (Translation: The handbook of the university teacher)

Lizzio, A., Wilson, K., & Simons R. (2002). University students' perceptions of the learning environment and academic outcomes: implications for theory and practice. *Studies in Higher Education,* 27(1), 27–52.

Lucas, U., & Mladenovic, R. (2004). Approaches to learning in accounting education. *Accounting Education*, 13(4), 399–407.

Lucas, U., & Mladenovic, R. (2014). *Perceptions of accounting*. Chapter 6 in the Routledge Companion to Accounting Education (Edt. Richard M.S. Wilson), 125–144.

Martell, K. (2007). Assessing student learning: are business schools making the grade. *Journal of Education for Business* (March/April), 189–195.

Marton, F., & Säljö, R. (1976). On qualitative differences in learning–II Outcome as a function of the learner's conception of the task. *British Journal of Educational Psychology,* 46(2), 115–127.

Moss, P.A. (1992). Shifting conceptions of validity in educational measurement: implications for performance assessment. *Review of Educational Research* 62, 229–258.

Mullola, S., Hintsanen, M., Jokela, M., Lipsanen, J., Alatupa, S., Ravala, N., & Keltikangas-Järvinen, L. (2014). Associations between teacher-rated versus self-rated student temperament and school achievement. *Scandinavian Journal of Educational Research*, 58(2), 147-172.

Nicol, D.J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education,* 31(2), 199–218.

Panadero, E. & Jonsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited: A review. *Educational Research Review*, 9, 129–144.

Panadero, E., & Romero, M. (2014). To rubric or not to rubric? The effects of self-assessment on self-regulation, performance and self-efficacy. *Assessment in Education: Principles, Policy & Practice,* 21(2), 133-148.

Popham, J. (1997). What's wrong and what's right with rubrics. *Educational Leadership*, 55, 72-75.

Pringle, C., & Michel, M. (2007). Assessment practices in AACSB-accredited business schools. *Journal of Education for Business* (March/April), 202–211.

Puustinen, M., & Pulkkinen, L. (2001). Models of self-regulated learning: a review. *Scandinavian Journal of Educational Research*, 45(3), 269-286.

Ramburuth, P. & Mladenovic, R. (2004). Exploring the relationship between students' orientations to learning, the structure of students' learning outcomes and subsequent academic performance. *Accounting Education*, 13(4), 507–527.

Ramsden, P. (1979). Student learning and perceptions of the academic environment, Higher Education, 8, 411–427.

Ramsden, P. (1987). Improving teaching and learning in higher education: The case for a relational perspective. *Studies in Higher Education*, 12(3), 275-286.

Ramsden, P. (2003). Learning to teach in higher education, 2nd edn, London: Routledge.

Reynolds-Keefer, L. (2010). Rubric-referenced assessment in teacher preparation: An opportunity to learn by using. *Practical Assessment, Research and Evaluation*, 15(8), 1–9.

Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological Bulletin,* 138(2), 353–387.

Riley, T.J., & Simons, K.A. (2013). Writing in the accounting curriculum: A review of the literature with conclusions for implementation and future research. *Issues in Accounting Education,* 28(4), 823–871.

Riordan, D. A., Riordan, M. P., & Sullivan, M.C. (2000). Writing across the accounting curriculum: An experiment. *Business Communication Quarterly,* 63(3), 49–59.

Rotenstein, A., Davis, H.Z., & Tatum, L. (2009). Early birds versus just-in-timers: the effect of procrastination on academic performance of accounting students. *Journal of Accounting Education*, 27, 223–232.

Rotter, J.B. (1989). Internal versus external control of reinforcement. A case history of a variable. *American Psychologist*, 45(4), 489–493.

Saddler, C.D., & Buley, J. (1999). Predictors of academic procrastination in college students. *Psychological Reports*, 84, 686–688.

Scherer, R.F., Javalgi, R.G., Bryant, M., & Tukel, O. (2005). Challenges of AACSB International Accreditation for Business Schools in the United States and Europe. *Thunderbird International Business Review,* 47(6), 651–669.

Scouller, K. (1998). The influence of assessment method on student's learning approaches. Multiple choice question examination versus assignment essay. *Higher Education,* 35, 453-472.

Senecal, C., Koestner, R., & Vallerand, R.J. (1995). Self-regulation and academic procrastination. *The Journal of Social Psychology*, 135(5), 607–619.

Stevenson, L.A., Ferguson, J. & Power, D.M. (2014). The use of teaching resources in accounting education. Chapter 19 in the Routledge Companion to Accounting Education (Edt. Richard M.S. Wilson), 420–446.

Thomas, P., & Bain, J. (1984). Contextual dependence of learning approaches: the effects of assessments. *Human Learning,* 3, 227–240.

Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research*, 45(1), 89–125.

Tinto, V. (1982). Limits of theory and practice in student attrition. *The Journal of Higher Education*, 53 (6), 687–700.

Turner, M., & Baskerville, R. (2013). The Experience of Deep Learning by Accounting Students. *Accounting Education, An International journal*, 22(6), 582–604.

Tynjälä, P. (1998). Traditional studying for examination versus constructivist learning tasks: Do learning outcomes differ? *Studies in Higher Education,* 23(2), 173–189.

Vermunt, J.D. (1998). The regulation of constructive learning processes. *British Journal of Educational Psychology*, 68, 149–171.

Zimmerman, B.J. (2000). Self-efficacy: An essential motive to learn. *Contemporary Educational Psychology*, 25, 82–91.

Zimmerman, B.J., Bandura, A. & Martinez-Pons, M. (1992). Self-motivation for academic attainment: The role of self-efficacy beliefs and personal goal setting. *American Educational Research Journal*, 29 (3), 663–676.

Table 1. Descriptive statistics of the variables

	Obs	Mode	Mean	Median	Lower quartile	Upper quartile	Std.dev.	Coef.var.	Max	Min
BS_grade	64.000	4.000	4.141	4.000	4.000	5.000	0.774	0.19	5.000	2.000
Rubric_benefit	35.000	3.667	3.417	3.667	3.000	4.000	0.824	0.24	4.667	1.000
Male	64.000	1.000	0.656	1.000	0.000	1.000	0.479	0.73	1.000	0.000
Ch_understanding	64.000	0.000	0.539	0.250	0.000	1.000	0.638	1.18	2.000	-1.000
Ch_motivation	64.000	0.000	-0.211	0.000	-1.000	0.000	0.760	3.60	1.000	-3.000
Ch_confidence	64.000	0.000	0.273	0.000	0.000	1.000	0.811	2.97	2.000	-2.000
Pre_understanding	64.000	4.000	3.680	4.000	3.000	4.000	0.613	0.17	5.000	2.000
Pre_motivation	64.000	4.000	3.820	4.000	3.750	4.000	0.747	0.20	5.000	1.000
Pre_confidence	64.000	3.000	3.359	3.000	3.000	4.000	0.804	0.24	5.000	2.000

This table reports the summary statistics for the examined variables computed from the surveys in 2012-2014. *BS_grade* is the grade of the bachelor thesis seminar course. *Rubric_benefit* is the average score of the answers relating to the usefulness of the rubric in the ending questionnaire. This variable relates to the time after the development of the rubric, but is reported in descriptive statistics for the sake of brevity. *Male* is an indicator variable for gender which obtains value 1 if the student is a male, otherwise 0. *Ch_understanding* describes how much change there was in the student's understanding of academic research during the course. *Ch_motivation* describes how much change there was in the student's motivation during the course. *Ch_confidence* describes how much change there was in the student's self-confidence during the course. *Pre_understanding* describes the student's understanding of the academic research in the beginning of the course. *Pre_motivation* describes the student's motivation for the writing of the bachelor's thesis in the beginning of the course. *Pre_confidence* describes the student's confidence for the writing of the bachelor's thesis in the beginning of the course. The descriptive statistics are computed from those observations which returned both the ending and starting questionnaires (Nobs = 64).

Table 2. Correlation matrix

	1	2	3	4	5	6	7	8	9
1. BS_grade		0.210	0.175	*0.310*	0.024	*0.254*	-0.054	0.099	0.096
2. Rubric_benefit	0.250		-0.134	0.037	0.222	<u>0.284</u>	0.211	0.000	-0.014
3. Male	0.186	-0.109		-0.137	-0.093	-0.142	0.133	-0.153	<u>0.244</u>
4. Ch_understanding	**0.350**	0.147	-0.130		**0.353**	**0.324**	**-0.617**	*-0.260*	*-0.291*
5. Ch_motivation	0.034	0.130	-0.093	*0.308*		**0.481**	*-0.266*	**-0.452**	-0.160
6. Ch_confidence	<u>0.235</u>	0.266	-0.092	*0.306*	**0.413**		-0.045	-0.153	**-0.530**
7. Pre_understanding	-0.060	0.164	0.145	**-0.593**	<u>-0.228</u>	-0.017		0.184	<u>0.237</u>
8. Pre_motivation	0.153	0.124	-0.156	<u>-0.213</u>	**-0.413**	-0.132	0.137		**0.453**
9. Pre_confidence	0.149	-0.009	<u>0.245</u>	*-0.279*	-0.165	**-0.542**	0.194	**0.479**	

This table reports the Spearman rank correlation coefficients (below the diagonal) and Pearson correlation coefficients (above the diagonal) for the examined variables computed from the surveys in 2012-2014. The correlation coefficients are computed from those observations which returned both the ending and starting questionnaires (Nobs = 64) for all variables except Rubric_benefit. Correlation coefficients between Rubric_benefit and other variables are computed from those 35 observations that returned the ending questionnaire after the implementation of the rubric in 2013-2014. Correlation coefficients significant at 1% or better are **bolded**, at 5% or better are given in *italics*, and at 10% or better are <u>underlined</u>.

Table 3. Descriptive statistics for the closed-ended questions of the surveys examining students' opinions on the usefulness of the rubric

Questions: Fall 2013, Spring 2014 and Fall 2014	Nobs	Mode	Mean	T-test	Median	Wilcoxon	Q1	Q3	StDev	Max	Min
1a. The assessment criteria of the rubric influence my learning outcomes regarding the written thesis.	35	4.000	3.829		4.000		3.000	4.000	0.707	5.000	3.000
1b. The assessment criteria of the rubric influenced my learning outcomes regarding the written thesis.	35	4.000	3.457	0.026	4.000	0.025	3.000	4.000	0.919	5.000	1.000
2a. The assessment criteria of the rubric influence my learning motivation regarding writing the thesis.	35	4.000	3.686		4.000		3.000	4.000	0.900	5.000	1.000
2b. The assessment criteria of the rubric influenced my learning motivation regarding writing the thesis.	35	4.000	3.371	0.054	4.000	0.070	3.000	4.000	0.942	5.000	1.000
3a. The assessment criteria of the rubric influence my learning outcomes regarding the process of the bachelor's thesis seminar course.	35	4.000	3.714		4.000		3.000	4.000	0.710	5.000	2.000
3b. The assessment criteria of the rubric influenced my learning outcomes regarding the process of the bachelor's thesis seminar course.	35	4.000	3.457	0.059	4.000	0.067	3.000	4.000	0.852	5.000	1.000
4a. The assessment criteria of the rubric influence my learning motivation regarding the process of the bachelor's thesis seminar course.	34	3.000	3.706		4.000		3.000	4.000	0.799	5.000	2.000
4b. The assessment criteria of the rubric influenced my learning motivation regarding the process of the bachelor's thesis seminar course.	35	4.000	3.371	0.062	3.000	0.102	3.000	4.000	0.942	5.000	1.000
5a. How much do I need the rubric to support my work?	35	4.000	3.457		4.000		3.000	4.000	1.010	5.000	1.000
5b. How much did I need the rubric to support my work?	35	3.000	2.957	0.031	3.000	0.054	2.000	4.000	1.032	5.000	1.000
6a. How much does the rubric help me to understand the components of the course grade?	35	4.000	4.029		4.000		4.000	5.000	0.857	5.000	2.000
6b. How much did the rubric help me to understand the components of the course grade?	35	4.000	3.886	0.500	4.000	0.513	3.000	5.000	1.078	5.000	1.000

This table describes the descriptive statistics of the answers that students gave in the starting and ending surveys after the implementation of the rubric in fall 2013, spring 2014 and fall 2014. All participants answered both the starting and ending questionnaires (Nobs = 35), and the same questions were asked in both cases. The only difference was that in the starting questionnaire the questions concerned the future and in the ending questionnaire the questions concerned the past. In the table, question a relates to the starting questionnaire, and question b to the ending questionnaire. The number of observations is 34 for question 4a because one student left that specific question unanswered, all of the other questions were answered by that student. The t-test describes the significance of the paired t-test, and Wilcoxon gives the significance of the matched-pairs signed-ranks test.

Table 4. Regression results

Dependent variable	BS_grade			Ch_understanding			Ch_confidence			Ch_confidence		
Independent variables	Coef.	Stand. coef.	P-value	Coef.	Stand. coef.	P-value	Coef.	Stand. coef.	P-value	Coef.	Stand. coef.	P-value
Rubric_benefit	0.118	0.124	0.256	0.107	0.145	0.373	**0.230**	**0.233**	0.048	0.120	0.122	0.230
Male	0.316	0.198	0.319	-0.074	-0.060	0.729	-0.026	-0.015	0.911	0.289	0.174	0.106
Ch_understanding	**0.629**	**0.486**	0.004	----	----	----	----	----	----	0.382	0.284	0.113
Ch_motivation	-0.257	-0.203	0.446	----	----	----	----	----	----	**0.469**	**0.356**	0.014
Ch_confidence	**0.500**	**0.521**	0.041	----	----	----	----	----	----	----	----	----
Pre_understanding	-0.021	-0.014	0.934	**-0.645**	**-0.564**	0.000	0.279	0.181	0.135	**0.616**	**0.400**	0.022
Pre_motivation	-0.071	-0.068	0.692	-0.042	-0.052	0.723	0.193	0.178	0.239	**0.450**	**0.415**	0.009
Pre_confidence	**0.369**	**0.394**	0.070	-0.136	-0.188	0.254	**-0.693**	**-0.709**	0.000	**-0.742**	**-0.760**	0.000
Intercept	2.042		0.080	3.214		0.005	0.114		0.930	-1.870		0.219
Model F-value (prob)	2.820 (.022)			3.340 (.000)			7.510 (.000)			8.770 (.000)		
Adjusted R^2	0.300			0.256			0.489			0.615		
Nobs	35			35			35			35		

This table reports the regression results for the determinants of course grades, understanding, and confidence. *Rubric_benefit* is the average score of the answers relating to the usefulness of the rubric in the ending questionnaire. Other variables are defined in Table 1. Regression coefficients significant (two-tailed) at 10% or better are **bolded**. P-values are computed from the heteroscedasticity corrected standard errors. Stand.coef. is the standardized regression coefficient.

Table 5. Propensity score matched results on the usefulness of the rubric in improving understanding and confidence

Computation of propensity scores					
Treatment variable = High_usefulness1, High_usefulness2			Covariate balance		
Covariates	Coef.	P-value	Mean_treated	Mean_Control	P-value
Male	0.120	0.803	0.600	0.733	0.456
Pre_understanding	0.307	0.463	3.733	4.100	0.088
Pre_motivation	0.368	0.294	3.800	3.867	0.799
Pre_confidence	-0.445	0.155	3.133	3.533	0.176
intercept	-1.286	0.527			

Matching based on propensity scores (treated observations = 15; control observations = 20)					
Dependent variable = Ch_understanding					
		Treated	Controls	Std.error	T-value
Unmatched	High_usefulness1	0.667	0.475	0.207	0.930
	High_usefulness2	0.800	0.460	0.221	1.540
Treatment (ATT)	High_usefulness1	0.667	0.144	0.210	2.583
	High_usefulness2	0.800	0.317	0.278	1.743

Computation of propensity scores					
Treatment variable = High_usefulness1, High_usefulness2			Covariate balance		
Covariates	Treated	P-value	Mean_treated	Mean_Control	P-value
Male	0.120	0.803	0.600	0.733	0.456
Pre_understanding	0.307	0.463	3.733	4.100	0.088
Pre_motivation	0.368	0.294	3.800	3.867	0.799
Pre_confidence	-0.445	0.155	3.133	3.533	0.176
intercept	-1.286	0.527			

Matching based on propensity scores (treated observations = 10; control observations = 25)					
Dependent variable = Ch_confidence					
		Treated	Controls	Std.error	T-value
Unmatched	High_usefulness1	0.633	0.100	0.266	2.000
	High_usefulness2	0.450	0.280	0.307	0.550
Treatment (ATT)	High_usefulness1	0.633	0.400	0.341	0.677
	High_usefulness2	0.450	0.300	0.418	0.360

This table reports the results of the tests for the mean. Unmatched results provide the results of the t-test for two independent samples without matching, and treatment (ATT) results give the results of the t-test after propensity score matching. Propensity scores are obtained from the regression where the treatment variable is regressed on the covariates (Male, Pre_understanding, Pre_motivation, Pre_confidence). The dependent variables are those that were the significant determinants of course grade in Table 4: Ch_understanding and Ch_confidence. *Ch_understanding* describes how much change there was in the student's understanding of academic research during the course. *Ch_confidence* describes how much change there was in the student's self-confidence during the course. The treatment variables are as follows. *High_usefulness1* obtains value 1 if the score of Rubric_benefit is higher than the median value, otherwise 0. *High_usefulness2* obtains value 1 if the score of Rubric_benefit is equal or higher than the 75th percentile value, otherwise 0. Rubric_benefit is the average score of the answers relating to the usefulness of the rubric in the ending questionnaire. Covariate balance describes how balanced the treated and control groups were after the propensity score matching in every covariate category. The treatment (ATT) results are the mean results of three independent propensity score matchings after random sampling of the observations. The covariates are defined in Table 1.
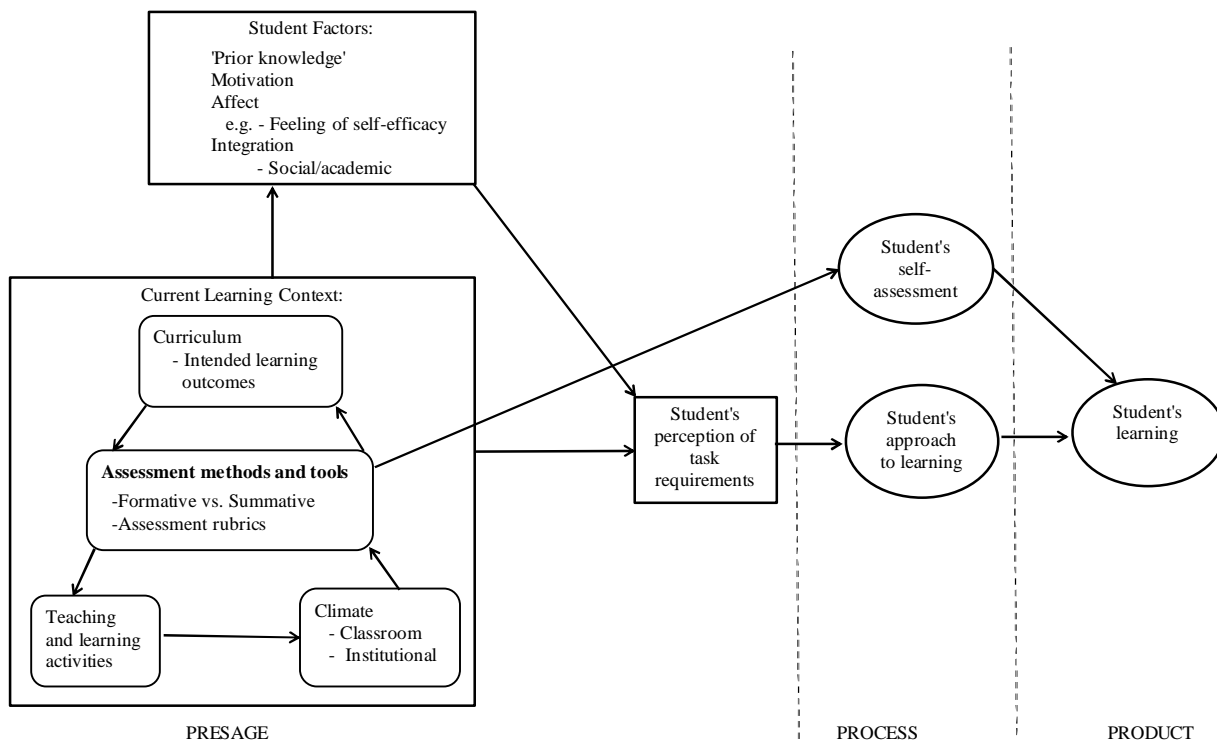
Figure 1. Constructivist model of learning for advanced written assignments and the accompanying group work (modified from Biggs, 1993; Ramsden, 2003; Duff & McKinstry, 2007; Lucas & Mladenovic, 2014)
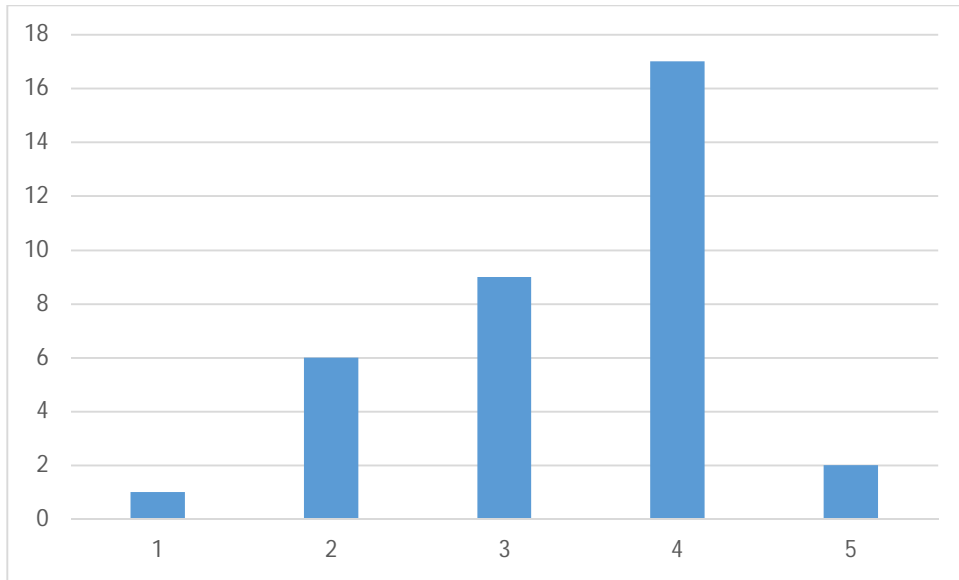
Figure 2. The assessment criteria of the rubric influenced my learning motivation regarding writing the thesis (scale 1-5, Nobs=35)
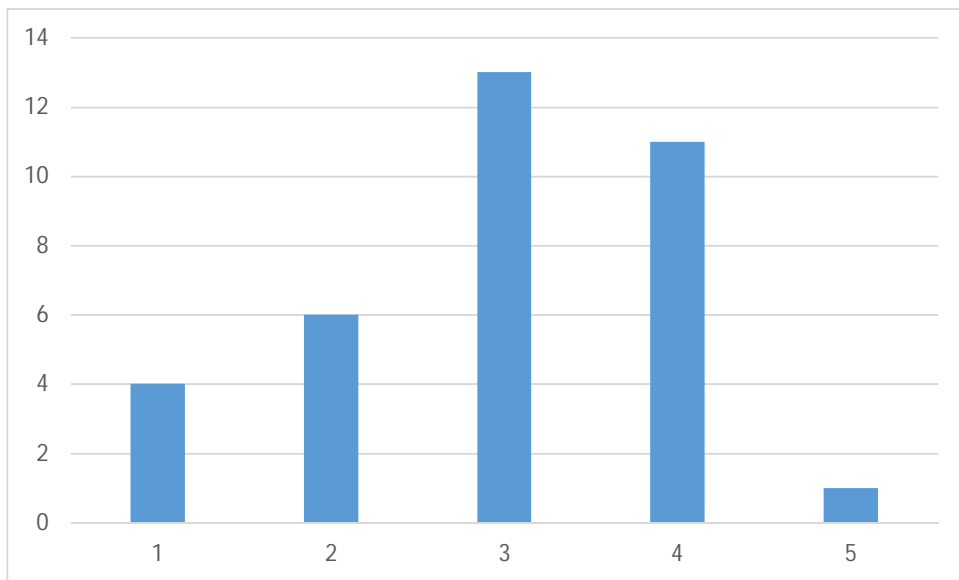

Figure 3. How much did I need the rubric to support my working? (scale 1-5, Nobs=35, one observation gave answer 3.5 and that was rounded up to 4)

Appendix A. Course syllabus of the bachelor's thesis seminar course

22C99901 Bachelor's Thesis (10 cr)

Status of the Course:

Bachelor's program, compulsory course in the accounting specialization area.

Level of the Course:

Intermediate studies

Teaching Period:

1-11, 111-V (autumn 2013, spring 2014)

Workload:

-Classroom hours, 28 h

-Class preparation, 56 h

-Thesis writing, 232 h

-Maturity exam, 4 h

Learning Outcomes:

The student understands the main theories, research problems and research methods in his/her chosen topic area, and he/she has capabilities to conduct independent scientific research in accounting. The student can also apply the theoretical knowledge, as well as analytical and problem-solving skills, to practical decision-making problems.

Maturity tests for the bachelor's degree will be arranged at the end of the seminar. In the test, the student writes an essay based on the Bachelor's Thesis in his/her first language, and shows that the main content of the thesis has been mastered.

Content:

Participation in the seminar, writing the thesis, maturity test

Assessment Methods and Criteria:

1. Submitting and presenting a bachelor's thesis (70% of the grade)
2. Acting as an opponent and taking minutes in the seminar (15% of the grade)
3. Active participation in the seminar (15% of the course grade)

4. Maturity test (Pass/Fail).

The seminar will be arranged in several groups in management accounting and financial accounting.

Prerequisites:

Introduction to Accounting, Introduction to Finance, Kirjanpito (translation: Bookkeeping), and Johdon laskentatoimi I (translation: Management Accounting I). In addition, it is recommended to have completed Johdon laskentatoimi II (translation: Management Accounting II) or Tuloslaskenta (translation: Financial Accounting) according to the topic of the thesis.

Evaluation:

0-5

Registration:

To register for seminar groups, please fill out an electronic form available at the web site of the course. Registration for groups in the fall term ends on May 31 and for the groups in the spring term on November 30.

Language of Instruction:

Finnish.

22C99902 Bachelor's Thesis Seminar (2 cr)

Content:

See 22C99901 Bachelor's Thesis

Evaluation:

Pass/Fail.

Appendix B. Examples of bachelor's thesis topics

The determinants of the quantity and quality of IFRS 8 based segment reporting.

The impact of the capital structure on the market value of the firm. Environmental reporting in developed and developing economies. Comparison of Finnish- and Russian-listed firms.

The possibilities for using cash flow information and empirical evidence on its use from investors' perspective.

Investor relations professionals and voluntary disclosure in Finnish-listed firms. Review of disclosures on future prospects.

The valuation of start-up companies. Comparison of the methods, problems, and viewpoints from the field.

Earnings management via the manipulation of discretionary accruals. Empirical evidence on the impact of firm size on manipulation in Finnish-listed firms.

The valuation of small- and medium-sized companies: special issues, methods, and problems from the valuator's perspective.

The determinants of audit fees in the listed companies of the Helsinki stock exchange.

Uncertainty of international investment decisions.

Risk management in light of the ERM frameworks: case television production process.

Basel III capital adequacy framework's impact on banks' financial solidity.

The role of sustainability in management accounting.

The value of corporate responsibility reporting in management decision-making – Is there any?

The contemporary CFO's role and the importance of actual accounting skills in the profession: Case Nomovok, Ltd.

Developing the scorecard after the implementation – Challenges and new frameworks.

Risk management of a Finnish ESCO company's energy efficiency investment in China: analysis of various types of risks.

Appendix C. Assessment rubric for accounting students' bachelor thesis seminar

Assessment criterion	Inadequate work (2)	Fairly good work (3)	Good work (4)	Excellent work (5)
1. Research question (20%)	The research question is only slightly motivated.	Research question is moderately motivated.	The research question is considerably and credibly motivated from a scientific and practical perspective. The research gap is considerably clearly formulated.	The research question is extremely well-linked to the existing literature. The student explicitly states the existing research gap that (s)he is attempting to fill. Scientific motivation can be supplemented with practical motivation if it is relevant for the subject.
2. Structure of the study (15%)	There are significant deficiencies in the structure of the study. Typical sections of the bachelor's thesis are not reported. For example, there is not a summary and conclusions section. The length of the paper is between 15-17 pages of written text. If the thesis is longer than 25 pages, it should not be repetitive.	The structure of the bachelor's thesis is moderately formulated. The length of the paper is between 18-19 pages of written text. If the thesis is longer than 25 pages, it should not be repetitive.	The structure of the bachelor's thesis is considerably formulated. The literature review and empirical part are in balance. The length of the paper is between 20-25 pages of written text. If the thesis is longer than 25 pages, it should not be repetitive.	The structure of the bachelor's thesis is extremely well-formulated. The literature review and empirical part are in balance. The length of the paper is between 20-25 pages of written text. If the thesis is longer than 25 pages, it should not be repetitive.
3. Literature review a. Used literature (10%)	The reference list is significantly inadequate: there are altogether 0-14 references in the study and 0-1 references to scientific journal articles.	The reference list is inadequate: there are altogether 15-19 references in the study and 2-7 references to scientific journal articles. The references are moderately relevant to the studied subject.	Various high-quality references have been used in a balanced way. The author succeeds in referring to the references as instructed. There are altogether 20-29 references in the study and 8-15 references to scientific journal articles. The reference list includes both old and new studies. The references are relevant to the studied subject.	Various high-quality references have been used in a balanced way. The author succeeds in referring to the references as instructed. There are altogether more than 30 references in the study and more than 16 references to scientific journal articles. The reference list includes both old and new studies. The reference list includes scientific articles, but also recent working papers in the area. The references are relevant to the studied subject.
b. Critical thinking of the literature review (10%)	There is not any dialogue between the references or other proofs of critical thinking.	The literature review rests too much on a few references. There is only slight evidence for dialogue between the references or other proofs of critical thinking.	The literature review uses several references in a balanced way, and there is moderate evidence for dialogue between the references or other proofs of critical thinking.	The literature review is a versatile synthesis of the aspects of different references. Dialogue between the references is on an excellent level, and the end-result provides a comprehensive view of the subject. There are considerable proofs of other forms of critical thinking in the literature review, which reflect a deep understanding of the subject.
4. Empirical analysis (15%) a. Statistical research	Testing is not based on any research hypothesis. The research method is not appropriate, and it is not described in the study. Empirical analyses are improperly conducted, and the research results are not credible.	Testing is not based on any research hypothesis. The research method is appropriate for the problem, but it is described unclearly. Empirical analyses are moderately well-conducted, and the research results are moderately credible.	Testing is based on a research hypothesis or hypotheses. The research method is appropriate for solving the research problem, and it is described clearly in the study. Empirical analyses are considerably well-conducted, and the research results are considerably credible.	Testing is based on a research hypothesis or hypotheses. The research problem is appropriate for solving the research problem, and it is described clearly in the study. Empirical analyses are extremely well-conducted, and the research results are considerably credible. The research problem is analyzed from different aspects, i.e., certain robustness checks are performed after the main analyses.
b. Qualitative research	The research method is not appropriate for solving the research problem, and it is not described and justified in the study. The research results are not reported in a credible way.	The research method is appropriate for the research problem, but it is justified and described unclearly. Empirical material is inadequately described and analyzed, and the results are reported in a moderately credible way.	The used research method is appropriate for solving the research problem, and it is justified and described in the study. Empirical material is considerably well-described and analyzed. The research results are reported in a considerably credible way.	The used research method is appropriate for solving the problem, and it is justified and described in the study. Empirical material is extremely well-described and analyzed. The research results are reported in an extremely credible way and reflected against the theoretical framework.

Assessment criterion	Inadequate work (2)	Fairly good work (3)	Good work (4)	Excellent work (5)
5. Conclusions (20%)	There are not discussions or conclusions between the literature review and the findings. The researcher's own interpretations cannot be determined from the study.	The findings of the literature review or empirical analysis are not interpreted. There is only a slight amount of discussion and conclusions in the study.	The researcher succeeds in linking the results of the literature review or empirical analysis to the research question and previous literature. There is a moderate amount of discussion and conclusions in the study. The researcher's own voice' can be heard, but there could still be slightly more interpretation.	The researcher succeeds excellently in linking the results of the literature review or empirical analysis to the research question and previous literature. There is a considerable amount of discussion and conclusions in the study. The researcher's 'own voice' can be heard from the study extremely clearly.
6. Fine-tuning and style (10%)	The study is inconsistent and contains several errors, which hamper reading. There are errors in the text and reference list, and a lack of metatext. The tables and figures are also incomplete.	The study contains some errors in the text and reference list, and some lack of metatext. The tables and figures are moderately well-formulated.	The study contains only a few errors in the text and reference list, or a slight lack of metatext, which however does not disturb reading. The tables and figures are considerably well-formulated.	The study is clearly written and consistent. The text contains sufficient metatext, and the reader can feel the 'tension', which maintains her/his interest. The tables are uniform, precisely done, and self-explanatory.
7. Process	Opponent work is inadequate, and the student does not participate in discussions with other group members.	Opponent work is moderately well-executed, but the student does not participate in discussions with other group members.	Opponent work is considerably well-executed, and the student sometimes participates in discussions with other group members.	Opponent work is excellent, and the student is an active group member in discussions.

Fundamental errors which have negative effect on the course grade are the following:
- Being late, relative to the agreed-upon schedules
- Fundamental factual errors in the study, which are beyond the scope of the rubric
- Misbehaviour towards other group members during the bachelor's thesis seminar course

The independence of the student has a positive effect on the grade.
The novelty of the research question has a positive effect on the grade.

Appendix D. Detailed assessment criteria for the quality of the process

Assessment criterion	Inadequate work (2)	Fairly good work (3)	Good work (4)	Excellent work (5)
7. Process (60%) a. Opponent work	Opponent's preparation is insufficient, and (s)he has not read the study. Comments are more like statements and not very constructive. Opponent's discussion is too short, and the supervisor has to help the opponent during her/his work. Opponent does not activate the audience to discussion. Opponent focuses on minor details without a more comprehensive understanding of the most fundamental issues for developing the study.	Opponent is not adequately prepared for the work. Comments are more like statements and not very constructive. However, opponent succeeds in her/his role without the supervisor's help, although the balance between fundamental issues and minor details could be better during her/his opponent work. Opponent moderately activates the audience to discussion.	Opponent is prepared for the work in a considerable way. (S)he gives constructive feedback to the presenter, and the fundamental issues and minor details are in a good balance in the discussion. Discussion and cooperation with the second opponent is considerably fluent. However, opponent's discussion may be slightly too long or short. Opponent considerably activates the audience to discussion.	Opponent is prepared for the work in an excellent way. (S)he gives constructive feedback to the presenter, and also has the courage to ask questions of the presenter. The fundamental issues and minor details are in a good balance in the discussion. Discussion and cooperation with the second opponent is extremely fluent. The length of the discussion is as instructed. Opponent is excellent in activating the audience to discussion.
7. Process (40 %) b. Activity	The student is passive and does not give any audience comments during the bachelor's thesis seminar course.	The student's activity could be better. (S)he gives a couple of comments during the bachelor's thesis seminar course.	The student participates actively in discussions. The quality of the comments is considerably good, but with a more active approach the student can help the presenter to a greater extent.	The student participates actively in discussions. The comments are constructive and extremely high-quality, which indicates that the student has prepared to provide comments for the study. The presenter derives a lot of benefit from the comments.