The art of assessment - How to utilise multiple-choice within the field of law

Berit Kjosnes
Abstract
The purpose of this essay then is to gain insight into the utilisation of MCQ’s within the field of law at Swedish upper secondary schools and in this question what kind of knowledge requirements can be tested with MCQ. Also, the difference in test results between non-MCQs and MCQ was analyses in large and when it came to gender. Here a MCQ test was evaluated based upon the knowledge requirements while quantitative date gathered from a utilisation of MCQ within the field of employment law was analysed. It was found that it should be possible to utilise MCQ within the field of law. With regards to difference in results between non-MCQ and the MCQ under scrutiny it was found that high performing students scored above one grade lower on MCQ than the average of three non-MCQ’s in other subjects. Low performing students experienced a little improvement in their results. There was a slightly difference when it came to score in between gender. Although the average female score on all tests where higher than their male counterpart, females scored a little lower on their first MCQ test. It was felt that the scope of this research is too small to allow any conclusions to be drawn.

Key word: MCQ, Multiple Choice, Objective Assessment, Select Response, Assessment.
# Table of contents

1. Introduction 4  
1.1 Purpose and research questions 5  
1.1.1 Purpose 5  
1.1.2 Research questions 5  

2. Background 6  
2.1 The legal frame of assessment 6  
2.1.1 General applicable law 6  
2.1.2 Regulations applicable at Swedish upper secondary schools 7  
2.1.3 Knowledge requirements applicable within the field of law in general and civil law in particular 8  
2.2 Objective and transparent assessment 10  
2.2.1 Objective assessment in large and MCQs in particular 10  
2.2.3 Transparency in assessment 11  
2.2.4 Summary 12  
2.3 How to use MCQs? 13  
2.3.1 Limitations regarding authentic assessment 13  
2.3.2 Sharp versus fuzzy criterions 13  
2.3.3 Deep versus surface level knowledge 13  
2.3.4 Gaining pass through guesswork 14  
2.3.5 Gender 15  
2.3.6 Ease of use 15  
2.3.7 Summary 15  
2.4 A call for stability in assessment 16  

3. Method 17  
3.1 Guidelines regarding construction 17  
3.1 Method of data analysis 19  
3.3 Method discussion 20  
3.3 Delimitations 21  

4. Results and analysis 22  
4.1 Knowledge requirements that might be tested by MCQ’s 22  
4.1.1 Results 22  
4.1.2 Analysis 22  
4.2 Difference in results between MCQ and non-MCQ assessment 23  
4.2.1 Results 23  
4.2.2 Analysis 23  
4.3 Difference in results between sexes 24  
4.3.1 Results 24  
4.3.2 Analysis 24  

5. Conclusion and discussion 26  
5.1 Conclusion 26  
5.1.1 What kind of knowledge requirements can be tested by MCQ’s? 26  
5.1.2 What kind of difference are there between MCQ’s and non-MCQ’s when it comes to results? 26  
5.1.3 What kind of difference in results are there between MCQ’s and non-MCQ’s when it comes to gender? 26  
5.2 Discussion 27  
5.2 Reflections 27  

Bibliography 28
1. Introduction

In the Swedish magazine Lärarnas Tidning the seventh of April 2016 one found that three secondary school students submitted an identical school assignment to different teachers. The same assignment was assessed as following: One girl got B, a second, C and the boy received D.

Merriam-Webster define assessment as “the act of making a judgment about something”. While what can be judged might vary, the focus of this paper is to address assessment as a mean of passing judgement within the field of education.

Throughout our life we are assessed several times. Assessment impact us. It is personal, it influences not only our identity, but also how and what we are to learn. Through assessment societies are formed and replicates (Boud, 2000). Through assessment we are taught what is acceptable or not, what to focus on, and what to leave out. Thus, assessment becomes a powerful tool in the process of socialization. A technique, specific to a certain age, employed to govern people to obtain that people behave and believe in a certain way (Foucault, 1983). Assessment then becomes an important tool of power which impact and forms the people it governs.

The act of assessment impacts student’s perception of their own ability to meet their own future learning needs (Boud, 2000). Therefore, among others, Boud argue that a sustainable assessment process must prompt considerations beyond the act of assessment. A given tool of assessment then must be constructed with impact and implications in mind.

A sustainable tool of judgement must not only provide students with equal opportunity but it must also judge the same performance equally. As a new teacher, I have often heard the necessity of giving students equal opportunity to show their knowledge. However, I question if the need of equal judgment of same performance might be an even bigger problem.

Today students from different upper secondary schools in Sweden compete about what some might consider to be few available spaces at very attractive educations. These spaces are occupied either based upon the Swedish Scholastic Aptitude Test or grades. If these spaces are filled with students coming from schools or teachers that apply different criterions when grading both student and university will suffer.

The students in question having their performance upgraded might overestimate their ability to deal with the new study situation and even develop a perception of teachers as being in error or unqualified. This while the university will have to deal with students not able to pass fundamental examinations, complaining and unsatisfied students that require a lowering of the overall assessment level, teachers that are required to teach at a lower level or teachers that rate same performance differently.

Students receiving lower grade on same performance not only loose out of their opportunity to someone less fitted to study but suffer regarding the perception of own ability. This while the University loose students better prepared for higher education.
Perceiving then that assessment build societies the message given through arbitrary assessment might tremendous. Therefore, there is a prevalent need to investigate whether objective methods of assessment can be implemented.

The most prevalent form of assessment utilised today within the field of objective testing are multiple choice tests (MCQ)s. MSQ tests belong to the category of selected-response tests (SR)s and can be machine scored or scored with a key. (Rodriguez & Haladyna, 2013) A MCQ item typically consist of a question or statement with several possible fixed answers.

Seeing then the need of implementation of objective assessment methods at Swedish upper secondary school the objective of this essay is to investigate whether MCQ’s can be a suitable tool of objective judgement within the field of law.

1.1 Purpose and research questions

1.1.1 Purpose

The purpose of this essay then is to gain insight into the utilisation of MCQ’s within the field of law at Swedish upper secondary schools.

1.1.2 Research questions

1) What kind of knowledge requirements can be tested with MCQs?
2) What kind of difference are there when it comes to results between MCQ’s and non MCQ?
3) What kind of difference in results are there between MCQ's and non-MCQ's when it comes to gender?
2. Background

To understand the utilisation of MCQ’s this chapter address general and specific applicable legal requirements and with this Objectivity and transparency. limitations and possibilities of MCQ’s and with that impact on gender and seeing the need of sustainable assessment and holistic thinking also the issue of stability in assessment has been addressed.

2.1 The legal frame of assessment

According to the Instrument of government chapter 1, art. 1 and 2 the Swedish government are to exercise its power under law and with respect for the equal worth of all. This chapter will therefore address regulatory documents that influence assessment in Swedish upper secondary schools. Firstly, general applicable laws are addressed, then regulations affecting assessment at upper secondary schools in general and lastly regulations affecting assessment within the field of law at upper secondary school in Sweden.

2.1.1 General applicable law

According to the Instrument of government chapter 1, art. 9 it applies that:

*Courts of law, administrative authorities and others performing tasks within the public administration shall have regard in their work to the equality of all before the law and shall observe objectivity and impartiality.*

The question of equality has been addressed within the field of education among else by. Gripps & Murphy (Gripps & Murphy, 1994). Here one finds that equity in assessment does not imply equity of outcome nor that all students share the same experience but rather that assessment practice and interpretation of results are fair and just for all groups. This perspective seems to agree with the Swedish national board of education here one finds that equality in education necessitate that one pay attention to the students’ different conditions, needs as well as level of knowledge (SKOLFS2017:16). The general recommendations (SKOLFS2011:149) Art. 3 translate the principle as the need to utilise appropriate methods that provide students with good opportunities to demonstrate their knowledge in different ways. This while the same rules of assessment shall apply to all students. Giving that all students performing at the same level shall receive the same grade. (Prop.1973:90 bil. 3s. 170 f.)

The principle of equality, objectivity and impartiality applies to all that perform tasks within the field of public administration (Prop. 2009/10:80, page 116). In Prop. 2009/10:80, page 247, one finds that the rule applies even if the person performing the tasks are not be employed by a public agency if the tasks he or she performs can be perceived as an execution of public administration. Then it does not matter if a teacher is employed in a public school or private school. A teacher then cannot afford to be a kind grader without endangering the equal rights of other students.

Assessment, as such, are only to be based upon what objectively matters in the given circumstances, that is the legal frame and the course requirements put forth in 2.1.3.
Assessment within the sector of education is also regulated by the 1986 Administrative Procedure Act. Under the general requirements with regards to handling of matters, section 7 one finds that one of the goals with the Act is to make sure that:

_Each matter to which a person is a party shall be handled as simply, rapidly and economically as is possible without jeopardizing legal security._

In section 20, one finds that the reason for a given decision should be put forth in an easy and understandable way.\(^1\) Here one also find that if reasons are not provided in the decision itself, the authority should upon request of a party put forth the reasons after on. In section 21 one finds that decision-making authorities are compelled to inform for example student of the decision unless this is obviously unnecessary. Notification can be made orally, by letter or in some other way unless the party request that it is made in writing. Therefore, it follows that the decision made must be specific enough that it is possible to appeal it. That is while it is possible to appeal a vague decision the reasons of the same might not be specific enough for a given appeal to address.

### 2.1.2 Regulations applicable at Swedish upper secondary schools

Students at Swedish upper secondary schools have a right to receive a grade in accordance with the Swedish Education Act (2010:800) art. 13. In Art. 22-24 one finds that a teacher must provide the student with a grade between A-E at the end of each course the student finalised and F if he or she fails. Here also one finds that a given grade are to be based upon the knowledge requirements within a given course.

The knowledge requirements are to be specified to reveal what kind of knowledge are needed for each grade according to the School Ordinance (2010:2039) art. 2.

The purpose of the grade is to express the extent to which an individual student has met the national knowledge requirements for a given course and the objectives of the upper secondary school according to Swedish National Agency for Education, curriculum for upper secondary school (SKOLFS 2011:144 appendix 2.5). Here one also finds that to fulfil the objective of secondary schools a teacher among else are required to report to the student on which ground grading takes place. Students on their hand should be able to assess their study results but also the need for development in relation to the requirements of their education (SKOLFS 2011:144, appendix 2.5).

---

\(^1\) Section 20 applies only on the final decision (grade). Furthermore a teacher only have to provide the reasons that induced that same final decision see prop. 1971:30 part 2 page. 491-494 and Hellner, Trygve and Boo, Malmqvist, 2010, page. 231–243.

**Berit Kjosnes**  
HT 2017  
Examensarbete, 15 hp  
Examensarbete med ämnesdidaktisk inriktning, VAL, PE248HT17, 15 hp, page 7
2.1.3 Knowledge requirements applicable within the field of law in general and civil law in particular

If a student chooses to study law at a Swedish upper secondary school he or she will find that in law one deal with the Swedish legal system, that is laws regarding governing Sweden but also laws regulating companies, organisations and citizens (SKOLFS 2010:14). The aim is to provide students with both an understanding of the importance of legal order in a democratic society and how this order is influenced by international law (SKOLFS 2010:14). At the same time, the aim is also to develop students' ability to analyse and judge legal problems through the utilisation of legal method and the ability to use legal sources (SKOLFS 2010:14).

In the national knowledge requirement for law one finds that assessment tasks must be designed in such a way that these enable students to demonstrate relevant outcome in each subject (SKOLFS 2010:90). Within the field of law, one differentiates between public and civil law. While public law in large address how public powers are distributed and its relationship with international law, civil law regulates relations between people, people and businesses as well as in between businesses.

For civil law, the relevant outcome in question has already been defined by the Swedish National Agency for Education (SKOLFS 2010:90). Here one finds that the aim of the study among else is:

- to develop knowledge of the Swedish legal system and an understanding of how applicable laws regulate private, professional and business life (SKOLFS 2010:90).

A mere understanding is not enough seeing that students are to develop the ability to express themselves in a correct and precise way. They are to learn to utilise legal language through study of court cases and legal literature. They are to take on legal language as their own language and utilise this language both in speech and language. They are to learn practical legal problem solving and develop their analytic skill through practical work with legal case histories. It is also worth to notice that students should also learn to refer to and interpret relevant legal sources (SKOLFS 2010:90).

A teacher is bound by the Swedish National Agency for Education to assess a student based upon the following knowledge requirements: The ability to account for, the given field, the concepts, to solve problems, to reason regarding the effect of legislation, to explain legal documents, to draw up legal documents, to apply a legal philosophical and an ethical perspective and lastly to assess the own ability (SKOLFS 2010:90), Art. 3).

This study will only address the first four of these eight areas of assessment namely I and II the ability to account for a field of study and concepts, III- the ability to solve problems and IV- to reason regarding the effect of legislation. This seeing the ability to explain and draw up legal documents require authentic situations (2.3.1), the ability to apply ethical perspectives seem to be a fuzzy criterion (2.3.2) and the ability to assess own abilities has not been tested by MCQ. The extend one can test own ability by MCQ remains unexplored to me. See also chapter 2.3.7.
I- To gain an E a student must be able to give an account in basic terms of limited parts of the legal areas that constitute civil law. To reach a C, the students account of parts of the legal areas sorting to civil law while an A require not only a detailed account of comprehensive parts of civil law but also that this account must be balanced.

II- In addition to the account the student at E level must be able to in basic terms define the meaning of legal concept and utilise these with some certainty in relevant contexts. At C level, the student must give detailed account of the meaning of legal concept and utilise these with some certainty in relevant contexts while an A student are required to give account in detail and in a balanced way of the meaning of legal concepts while using them with certainty in relevant contexts.

III- At E level a student, based upon simple case descriptions, should be able to solve with satisfactory results specific problems with the help of legal sources. This, while at C level the student should base upon case descriptions should be able to identify and in detail describe problems in different private legal areas and solve these with satisfactory results by choosing, interpreting and applying relevant legal sources. The student must also provide the reasons for solutions with well-grounded argument. The requirements of a student performing on an A level are more extensive. This student must base upon complex case descriptions identify and describe in a detailed and balanced way problems in different areas within civil law, and solve these with good results by selecting, interpreting and applying relevant legal sources. The student must also give the reasons for their proposals for solutions in well-grounded and balanced arguments.

IV- In addition, at E level students apply simple reasoning to the effects of legislation, and draw simple conclusions. This while students at C must apply well-grounded reasoning about the effects of legislation, and draw well-grounded conclusions. To get an A students, apply well-grounded and balanced reasoning about the effects of legislation, and draw well-grounded and balanced conclusions.

For a teacher to provide the student with D or B he or she must fulfil the knowledge requirements for E respective C in full plus most of the requirements for C respective A according to the School Ordinance Art. 2. These requirements should be identified and specified to such an extent that a teacher can analyse which parts of the overarching knowledge requirements the student's knowledge correspond to (SKOLFS2011:149) Art 3.
2.2 Objective and transparent assessment

In legal literature one finds that a judge must act strictly impartially when deciding procedural and criminal issues. Giving that all judgments, decisions and assessments are to be objective. (Heuman, Lars, 2004/5, p. 214) While a teacher cannot claim to hold the office of a judge he or she still pass on judgement on students’ performance through the process of assessment. These judgement impacts the student’s perception of themselves and their future.

According to Heuman the concept of objectivity may apply to different intellectual activities (Heuman, 214). Heuman argue that one science cannot disregard any of the requirements of objectivity unless one would want the results of a given investigation to be based upon personal preferences without even the individual in question being aware of this.

In this context, the focus needs to be on the thoughts behind the design of assessment tasks, the actual assessment activity and finally the evaluation of a student's performance. Seeing that the goals has already been provided the focus would be on how to measure a given students' performance.

2.2.1 Objective assessment in large and MCQs in particular

Objectivity has been addressed within the literature of education. Here one finds that Rodrigues and Haladyna makes a difference between tests that can be objectively respective subjectively scored (Rodriguez & Haladyna, 2013). The extent to which tests are to be perceived as objective versus subjective is usually a function of the type of scoring the test require. For a test to be perceived as strictly objective the correct responses must be known and responses scored by a key or machine scored. (Rodriguez & Haladyna, 2013).

A subjectively scored tests on the other hand will often but not always require human scoring or a complex scoring guide as a rubric. In these tests outcome differs in that answers might be partly correct or there might exist several possibilities regarding answers.

And still the picture here is not one of black and white seeing that among subjective scored items one also finds short-answer constructed-response (CR) items and even essay that also might be objectively scored with simple too complex to complex computer scoring engines. (Rodriguez & Haladyna, 2013). In this class of CR items one finds fill-in the blank, grid-in-response, short answer formats, the cloze and some experiments or tasks that results in specific products. This while more subjective scored items include demonstrations, essays, debates, interviews and so on. (Rodriguez & Haladyna, 2013).

Among objectively scored tests one finds selected- response (SR) items. In a test with SR format examinees are required to choose his or hers answer to a question or a statement from a listing of several possible answers. (Downing, 2006, Rodriguez & Haladyna, 2013)
The way CR test items then differ from strictly objective test items in that in the first situation the student provide or construct the word while in the latter situation the options are provided directly in the actual test item. The student then is to choose which option represent the correct answer.

Within SR one can choose between among else conventional multiple-choice tests (MCQs), true or false tests, multiple true-false. Multiple choice tests normally describe tests where one can choose one answer between different answers. In true or false tests, a person can choose between two answers where one is true and the other is false. In multiple true or false tests, a person need to evaluate each sentence to see if the sentence in question is true or false.

Thus, said any tests, whether that be a machine scored tests or so called subjective assessment, require human hands. Thus, while a test can be objectively scored one must still question whether the test ensure objective measurement of a given relevant criteria. After all the goals provided have been expressed by man, interpreted by man and translated into questions with given answers by human minds.

2.2.3 Transparency in assessment

Transparency is a vehicle that ensure objectivity and with that enable democracy. Teachers are to provide students with information regarding not only their progress but also the basis on which grades are awarded (SKOLFS 2011:144, 2:5). Seeing that to not be able to comprehend nor reach learning goals can have a negative impact on learning in more than one way (Wiiland, Tove, 2005). Transparency becomes important.

Here transparency can be perceived as student awareness of the purpose of the assessment and assessment criteria (Wiggins, Grant, 1998, p. 250). Through transparency the rules of the game become visible for both students and teachers. Then all parts know what apply. (Jönsson, Anders, 2010). Transparency is necessary for a student to interpret the feedback given and to be able to work with self-assessment in a way which reinforce understanding the assessment criteria. (Jönsson, Anders, 2010). Transparency empower students and increase their level of performance. (Jönsson, Anders, 2010). Rubrics seem to be a necessary tool in assessment to increase transparency however research has found that students have a hard time understanding academic language used to express assessment criteria (Jönsson, Anders, 2013, p. 34-35, 2014 and Higgins, Richard, Hartley, Peter & Skelton, Alan).

The issue of not understanding assessment criteria also becomes visible when it comes to feedback. Feedback has been defined by Ramaprasad, 1983, p 4, as the:

"information about the gap between the actual level and the reference level of a system parameter which is used to alter the gap in some way."
To trigger a desired change in a student he or she must understand not only the goal or the reference level being aimed for but also be able to compare the actual or current level of performance with this standard and engage in appropriate action which leads to close that gap (Boud, 2015; Sadler, 1989). Here MCQ's with its clear-cut form, already specified answers and objective scoring might prove to be a teacher's friend.

MCQs might be used as a tool to increase transparency by clarifying basic concepts, direct students focus to the principles which form the very base of the legal language and stimulate the development of an academic language (Gibbons, 2013, p. 224-225) This seeing that Rød, Eiklund & Fjær, Olav found that MCQs could be utilised as a tool to increase the level of reflection and the encourage students to read continually throughout the semester (Rød & Fjær, 2010). One might argue that this could be due to an increased transparency seeing that students that read more during the course should be more empowered than students that do not read.

2.2.4 Summary

Within assessment one differs between subjective assessment, objective assessment that can be subjective scored and strictly objective assessment techniques such as MCQ. And still while an assessment tool can be objectively scored it does not apply that it is objective. This seeing the questions must be based upon predetermined knowledge requirements and a student should score the same at different tests. A well written MCQ is clear cut and might prove itself a tool to increase transparency in assessment. It is easy to provide model answers and some research has shown that students are encouraged to read continually throughout the semester when MCQ has been utilised.
2.3 How to use MCQs?

The aim of this chapter is to provide a peek into some limitations and possibilities with MCQ's in assessment.

2.3.1 Limitations regarding authentic assessment

Jönsson makes a difference between authentic assessment and MCQs. Authentic assessment is utilised to assess the student’s ability to perform, demonstrate, and act in real life situations (Jönsson, Anders, 2013, p. 44-45). Performance and ability require authentic situations and happens in real time in this situation the very act is what is being assessed (Jönsson, Anders, 2013, p. 44). However, while MCQ's cannot be utilised as a tool to observe real time performance one might be able to use it when assessing reflections regarding own behaviour or what kind of considerations that led to a given performance.

Rodriguez & Haladyna questions whether it is possible to utilise objective scoring in assessments where students should demonstrate whether they have mastered specific skills and competencies by performing or producing something (Rodriguez & Haladyna, 2013). In law, this situation could apply to legal problem solving. Here the students are presented with a problem, to score higher the student need to identify problems in the problem and present a balanced solution.

2.3.2 Sharp versus fuzzy criterions

Fuzzy criteria are often found among else within the social sciences and require qualitative judgement while sharp criteria enable assessment tasks which might be solved in a standardised way therefore can be subject to immediate feedback (Sadler, 1989, p. 119-144). In the field of sociology few professors utilise multiple-choice tests. This seeing that it would not make sense in that students would not be able to demonstrate mastery of the learning goals (Clark and Filinson, 2011). Instead the ultimate products which is the centrepiece of assessment in sociology is a seminar paper designed so that the students can demonstrate cumulative mastery of a variety of skills. Thus, it might be that MCQs in certain situations cannot replace other types of examination.

2.3.3 Deep versus surface level knowledge

Subjective assessments like essay examinations are easy to construct, provide a teacher with access to deep thoughts behind the answer, and a possibility to explore the breath of student knowledge (Epstein et. al., 2002). MCQ's has been criticised for not being able test higher order cognitive skills in the same way as laboratory write-ups and essays (Fellenz, 2004 and Scouller, 1998). Scouller found that MCQ's has been perceived by students and teachers alike as primary accessing fact recall (Scouller, 1998).
Students were more likely to employ surface learning approaches when they were preparing for their MCQ examinations and that the students perceived the latter as assessing lower levels of cognitive processing (Scouller, 1998).

According to Bloom’s Taxonomy fact recall seems to sort to under level I; Knowledge. This seeing that fact recall give a student a possibility to exhibit memory recall of previously learned material and do not require much cognitive processing (See Bloom's Taxonomy Cheat Sheet by Kugelman, Francie). However, when Rodriguez did a meta-analysis with the purpose of investigating difference between results between constructed responses and selected response items he found that these produced nearly equivalent scores (Rodriguez, 2003).

Palmer and Dewitt argue that the problem of MCQ's is a construction problem as such and not a problem with the tool (Palmer and Devitt, 2007). They are supported by Downing who argue that most of criticism can be explained by poorly written SR items and is not attributable to the SR form per se. (Downing, 2002b) This might be why Haladyna, Downing and Rodriguez call for better guidelines regarding the construction of MCQs (Haladyna, Downing and Rodriguez, 2002, p. 311).

Palmer and Devitt state that a well written MCQ’s address even the highest order of cognitive skills (Palmer and Devitt, 2007). This wile Downing, 2002a perceive objective testing in the select-response (SR) item format to be the best choice for test developers that are interested in measuring cognitive achievements or abilities in an effective and efficient way.

In fact, it can be argued that one of the main benefits with MCQ's is that a teacher can test large domain of knowledge sampled at multiple cognitive levels (Wendler, Cathy LW and Michael E. Walker, in Handbook of Test Development, 2006). This might be why MCQs to a large extend are utilised in testing complex knowledge within the field of medicine.

One way to test higher order skills seems to be having a series of linking questions surrounding a scenario, topic, item or material (Gwinnett, 2017, p. 282). This however seem to contradict Rodriguez and Haladyna, 2013 who emphasis that item contents are to be kept independent of each other.

2.3.4 Gaining pass through guesswork

MCQs has been criticised for allowing participants to pass through guesswork. (Fellenz, 2004 and Scouller, 1998). The danger of gaining a pass through guesswork seems to be prevalent but this seems to have more to-do with test construction than the tool.

This seeing that regarding the possibility of gaining a pass-through guesswork Gwinnett in her chapter addressing design and implementation of MCQs in Forensic Science Assessment states that MCQ’s with fewer options, like true or false, have greater effect of chance and guessing upon the results (Gwinnett, Claire, 2017, p. 276).
2.3.5 Gender

One should also consider whether the usefulness of MCQs depend not only on how tests are constructed but also how they are coupled with other forms of learning activity. Including a variety of assessment techniques might for example be a good way to reduce bias associated with certain techniques in assessment (Röd & Fjär, 2010).

This seeing that a tendency among girls to perform poorer on MCQs while boys have poorer results on essays (Murphy, 1982).

2.3.6 Ease of use

From the teacher’s perspective, one might find that one of the main benefits of MCQ’s is that one can examine large groups of students in an objective way and provide fast feedback (Nicol, 2007, Gwinnett, 2017). The main challenges here might be the construction of MCQ’s. This seeing that while essay questions are easy to construct the construction of MCQ’s questions are time consuming.

From a student perspective MCQs are relatively easy to score, minimise subjectivity and the student can receive his or her feedback within short time (Epstein et. al., 2002).

While delayed feedback might serve to maintain student’s attention to the task under certain circumstances for example with children that are not ready to receive the answer (Aiken, 1968). In courses like civil law where one module follows another module and the modules as such not necessary directly are connected. Then it might apply that some have found that delayed feedback might hinder learning (Kulik & Kulik, 1988; Epstein, 2002).

Students seems to like fast feedback and find model answers beneficial (Huxham, 2007; Nicols, 2007). A model answer is a 100% answer. The thought behind model answers are that if students are not exposed to examples of what constitutes good work they are likely to adopt bad practices and engage in negative modelling (Boud, 2000).

2.3.7 Summary

Seeing that research has revealed that there where nearly no difference in results between CR and SR test it applies that the likelihood of earning a pass through guesswork’s should be low. It seems plausible that a well-designed MCQ can address even the highest order of cognitive skills and prove to be one of the best possibilities for a test developer aiming to measure cognitive achievements in an effective and efficient way and thus fulfill the requirement of the law. One of the main benefits with MCQ’s are that one can test large domain of knowledge sampled at multiple cognitive levels. One way of doing this might be to utilise linked questions.

Even if a MCQ is well written it does not seem to be an authentic assessment tool in that it seems difficult to assess a student’s ability to perform, demonstrate, and act in real life situations.
With regards to assessment of law, a student cannot verbally argue nor draw up a legal document. The ability to assess own knowledge might be possible but remain unexplored.

On the other hand, with a well written MCQ one should be able to assess the ability to account for a given field of study, concepts, the ability to solve problems when this do not require authentic assessment, and the ability to reason regarding effect of legal clauses.

However, when situations demand fuzzy criterions as when there exist more than one solution MCQ finds its limitation. On the other hand, if there are two solutions and both can be clear cut then MCQ’s can be applied. Giving that while MCQ’s can be utilised as a tool of assessment when it comes to solving problems and reasoning regarding legislation it would be suitable to supplement MCQ’s with another assessment tool.

MCQ can successfully be utilised in large classes with many students. Although demanding to construct it provide a teacher with a powerful tool that enable fast feedback and high levels of transparency. The value of fast feedback increases in courses where several subjects are examined over a shorter period. MCQ is easy to score for students.

With MCQ’s gender might be an issue seeing that girls tend to perform poorer on MCQs than boys. One way to deal with this might be to standardise the method of assessment.

Seeing the need to consider construction issues when it comes to MCQ’s these has in large been addressed in the method chapter.

2.4 A call for stability in assessment

While students are to be able to show their skills in multiple ways a problem arise if the method of assessment must be learnt (SKOLFS 2011:149). It might be that a variation of tests might however counteract our very aim.

According to Boud one of the” implicit aims of higher education is to enable students to become better judges of their own work.” (Boud, Lawson, Romy & Thomson, 2013).

This seeing that self-evaluation enchants student achievement. (Boud, 2000, p 157). Now while students seem to develop their capability to judge own performance during the course and over years, also this capability must be learned (Falchikov, N., and Boud, D,1989).

While familiarity in assessment type seem to accelerate student’s ability to make accurate judgments about their own effort, disruptive assessment patterns, that is change of assessment type and or assessment criteria, together with the introduction of new subject matters increase cognitive load and leads to delay in student’s development of self-assessment expertise (Boud, David; Lawson, Romy & Thompson, Darrall G., 2016). Thus, while variation might be necessary from time to time it might also create confusion among students in that these must learn how to respond in different ways, thus adding to their cognitive load and negatively impact student achievement.
3. Method

First it seems necessary to address the issues in 2.3.3 (see also 2.3.7) regarding construction of MCQs. Thereafter one finds the method of analysis and the method discussion.

3.1 Guidelines regarding construction

Rodriguez and Haladyna, 2013 emphasis that most of MCQ construction guidelines are based upon principles of good writing, lessons learned from practicing and logical reasoning. These authors provide us with the following fundamental principles regarding what to think about when to construct MCQs:

1) Each item is to be formatted vertically instead of horizontally.
2) Each item should be edited and proof read
3) The location of the right answer should be varied.
4) Options should be placed in logical or numerical order.
5) One is to avoid utilisation of wording as none of above, all above or I don’t know.
6) Test important content. Avoid opinions unless these are qualified.
7) Three options are often sufficient.
8) One should utilise students’ typical errors in writing distractors. All distractors should be plausible.
9) No humorous options.
10) Also, it seems important to not give clues regarding the right answer.
11) Word stems and options positively, avoid negatives such as not or except.
12) One should not utilise trick items.
13) New material should be introduced to elicit higher level thinking.
14) Make sure that the linguistic complexity for the group are appropriate.
15) Only one of the options should have the right answer.
16) Each item is to be based upon one type of content and cognitive demand
17) The amount of reading in reach item are to be minimised.
18) Items content are to be kept independent of one another
19) Utilise no specific nor overly general contents.
20) State the central idea in the stem very clearly and concisely. Repetitious wording is to be avoided.
21) The contents of options are to be kept independent. Options should not overlap each other.

Among these guidelines provided by Rodriguez and Haladyna, number 1-17 was taken into consideration at construction of the MCQ test in Appendix II as following:

All items have been formatted vertically. The text was proofread mainly by me. Seeing that I am Norwegian I often ask other teachers to read through the text. Most likely this happened also here. Thus, said teachers have limited time and I often find that students find error both with mine text but also what has been proof read by others or even taken from a book.

The location of the right answer varied and the options where placed in numerical order. Expressions like none of above, all above or I do not know where not utilised. Content where evaluated and only what was deemed as important was tested.

The aim has been in this test to not air or promote opinions. In twelve of fourteen MCQ question more than three options where utilised, of these four and five options where utilised five times each. Only in two instances where more than five options utilised.

Berit Kjosnes
HT 2017
Examensarbete, 15 hp
Examensarbete med ämnesdidaktisk inriktning, VAL, PE248HT17, 15 hp, page 17
The thought behind this was to make it difficult for students gaining a pass through guesswork as well as activate what Bloom might have called higher levels of thinking in that it requires not only and interpretation and understanding, but also activate analysis and synthesis (Bloom’s Taxonomy Cheat Sheet by Kugelman, Francie).

This seeing that meanings where not only meanings rephrased but when doing the exercise, the student had to identify the correct information and remove other plausible solutions.

Only plausible distractors have been utilised. Distractors have been constructed with the aim to mimic plausible student errors. The aim has been to only utilise plausible options and none trick items. None of the options have a humorous content. Options where constructed with the aim of not giving clues regarding the right answer. No sentence containing the correct answer where shorter, respective longer than the other. The wording of the sentences was made in a way that ensured similarity while enabling differentiation. This again to stimulate higher level thinking as already mentioned above.

None of the stems have negative wording. Questions are instead open to enable considerations regarding which answer could be the right ones. Normally I would introduce new material in one of the questions to move the students forward and increase their vision. Here this is in the last one. The linguistic complexity for the group was taken into consideration during construction in that the text only contain known words and words introduced during the given course. The challenge here is for the students that have problem with an academic language. However, there might be a benefit as well in that the words are already expressed in a MCQ test.

In four of the fourteen MCQ questions more than one option contained a right answer. In three of these the pupil would lose point if he or she picked a wrong answer. The number of possible answers were not given. The purpose was again to ensure the absence of guesswork in providing an answer while stimulating deeper analysis. The parts were rephrased and together the parts when combined made up a holistic whole.

While the aim was to base each item upon one type of content and cognitive demand. Reflecting upon the feedback given regarding question number one this was obviously not the case. Here the pupil where to provide an answer to a dual question, identifying both the gain and loss.

Although one has tried to minimise the amount of reading to reach each item this proved not to easy, se for example Q12. This applies if one believes in problem based learning. More exercise might solve this issue.

In addition to what has been said each question was weighted but this was not visible for the students. I now feel that the weighting should have been visible.

Guidelines 18-21 was not taken into consideration. This because of ignorance at the time of construction.

My experience with MCQ’s has been that the tool has been criticised by some students for eliciting fact recall while inducing anger when it proved not enough for students to prepare for the test in a way that enabled surface learning.

Berit Kjosnes
HT 2017
Examensarbete, 15 hp
Examensarbete med ämnesdidaktisk inriktning, VAL, PE248HT17, 15 hp, page 18
3.1 Method of data analysis

The data in this essay was collected during my spring 2017 teaching appointment in civil law. In this it represents a convenience sample (May, Tim, 1997, page 161). From January to June 31 students in the same class went through seven examinations, one of these was a home assignment with max grade pass or more than pass. Two examinations were too small to take into consideration as these consisted of material equal to one to three pages. In one of these the students could only get pass and more than pass. The other where downgraded with regards to grade value at the of the year when the material was weighted. Four examinations where traditional summative assessments made at the school in form of a traditional written test at the end of a given course. The three first examinations were made with open ended questions. The subjects examined where rent and tenancy, property rights, consumers rights and contract law. MCQ was utilised in the last traditional examination. This was an examination of employment law at the end of the year. Students were given a possibility to complement their grade with an additional work in the working environment act. At the end of the year there where one student with a grade lower than C.

To prepare the students for MCQ, Socrative had been utilised as a formative assessment tool. The test was printed from Socrative. Socrative provide an online classroom facility and enable construction of MCQ’s.

I felt that quantitative method of research was needed in that the questions of this study required numeric in order to compare data (Repstad, Pål, 1993, p. 14).

Here each grade was converted to a number. E and E+ = 1; E/D and D/E = 1,5; D-, D and D+ = 2; D/C and C/D = 2.5; C-, C and C+ = 3; C/B and B/C = 3.5; B-, B and B+ = 4; B/A and A/B = 4.5; A-, A and A+ = 5. The average mean based upon the three traditional examinations where then calculated and compared with the mean of the MCQ assessment (Byström, Jan 1992, page; Befring, Edvard, 1994, page 108-109). The mean of each grading assessment was taken into consideration during the analysis. There was a difference between means in the non-MCQ assessment. Seeing that this difference did not change the outcome on p. 22, this difference has not been discussed in this study.
3.3 Method discussion

All assessment instruments should be able to withstand the scrutiny of content and construct validity, reliability with more. (Palmer and Devitt, 2007). For MCQ's to benefit from a high reliability, and validity, questions must be relevant, linked and unambiguous (Palmer and Devitt, 2007). Collin found that if MCQ questions were "drawn from a representative sample of content areas that constitute predetermined learning outcomes" these would have a high degree of validity (Collins, 2006). Predetermined knowledge requirements then increase validity. In this context, the most important matter in that assessment was identified and the test was developed based upon the knowledge criterions and with this in aim.

Thus, said it is felt that when considering the construction of the actual MCQ test it seems wrong for the author of a test to evaluate his or her own work. The danger of creating a situation of self-confirming bias is overwhelming.

looking in the mirror it is easy to see that an evaluation of design in large and whether the actual knowledge requirements have been fulfilled this should have been done by another teacher in the same subject.

When it comes to clarity and understanding the answers. The test should have been evaluated by student's that are representative for the population being tested (Befring, Edvard, 1994, page 41-42). Here this population consist of higher secondary level students about 16 to 17 years old that study civil law. The sample population then should have had the same knowledge experience as the students taking the test. If using a representative population in another class there is a challenge in that because of the time it takes to construct a MCQ test a teacher do not want to take the risk that a test that represent so much work becomes public knowledge.

The second challenge is to find a population that are representative. If one asks the actual students that have taken the test these might rate the test more negatively based upon their lower score.

Comparing the actual MCQ with the criterions might suffer from bias independent of whatever done.

Reliability can increase if a given student score the same at different tests. In this context, the given knowledge requirements where only tested by the utilisation of MCQ unless students requested a different test. This happens once and the outcome did not change. The number of students that requested another test was too small for a conclusion to be drawn. Another problem is that all tests might have some type of flaws and students do not necessary score the same pending the day, the situation and the type of test.

Quantitative research methods are often utilised on larger groups. The mean as a tool for analysis was felt appropriate seeing the size of the group. It is not possible draw general conclusions based upon such a small sample in a local research.
A scientist is bound to make ethical considerations. In this instance, the weight of new knowledge has been balanced against the need of protection of the individuals. Due to the nature of this study, in that it consists of only quantitative data, that names have been replaced by a number and the place of study has not been revealed, the possibility of improper insight into individual privacy is limited (Vetenskapsrådet, 2002). The student was not informed and could not opt out. Seeing the nature of the research it was felt that this was not needed. However, if someone wanted to consider the analysis and control the numbers this might prove a benefit. This seeing that statistics can have errors in coverage and errors in content, in coverage when a person is not counted or counted more than once, in content when information is not correctly reported or tabulated (Frankfort-Nachmias, Chava et al., 2000, page 279, 282-283 and May, Tim, 1997, page 161).

3.3 Delimitations

The aim of this study is only to discuss MCQs in assessment. Legal rules applicable to assessment will therefore be addressed with only this aim. The aim has been to gain an overview of applicable rules and not to conduct a deep study. Therefore, case law and doctrine has not been included. Praxis has not been addressed. The reader will find that also here is a limitation in that issues in relation to handicap, second language and so on are not discussed in this study. Neither are special support nor special education addressed. Issues regarding claims or proceedings are not addressed.

Questions regarding a teacher right to hold back information when this can jeopardise legal security regarding assessment will not be addressed in this essay. One example on situations which actualise this problematic are when for example not all students have taken a given test or if the test is to be reused. Thus, issues regarding information and secrecy is not discussed. Issues regarding whom to grade are not addressed. Feedback has only been addressed in relation to MCQs. The issue of continuously feedback has not been addressed. Neither is the issue of cross study evaluation or considering all available information when it comes to assessment. In this essay MCQs are addressed as an example on strictly objective assessment. In this study, one only finds one analysis of one MCQ within the field of employment law.
4.0 Results and analysis

This chapter start out by evaluating what kind of knowledge requirement that might be tested by MCQ's, the difference in a result between three non-MCQ assessments and one MCQ assessment.

4.1 Knowledge requirements that might be tested by MCQ's

4.1.1 Results

Within employment law the MCQ test under scrutiny was designed with the aim to test the knowledge criterion put forth in 2.1.3 as following (see Appendix I for Q):

I  Q1, Q2, Q3, Q4, Q5, Q6, Q7, Q8, Q9, Q10, Q11, Q12, Q13, Q14
II Q1, Q2, Q3, Q4, Q5, Q6, Q7, Q8, Q9, Q10, Q11, Q12, Q13, Q14
III Q6, Q7, Q8, Q9, Q10, Q12, Q14
IV Q5, Q6, Q9, Q11, Q12, Q13, Q14

To pass the I criterium the student needed to get 50% of the answers correct, to gain a D the student needed 60% correct, for C 70%, B 80% and A 90%. The questions had been weighted but this was not visible for the students. Here the teacher must provide balanced solutions for a student to gain A.

4.1.2 Analysis

It seems to be possible to test student’s ability to account for employment law (Criterion I) In large the same applies for legal concepts (Criterion II). Here one finds that different parts of a given concept can be tested in different questions. Seeing that a student is to verbally utilise a concept these can only in part be tested by MCQs. Regarding criterion III, problem solving, it seems possible to address this criterium also through MCQ and still there are limited possibilities to explore different angles. Here as in criterion I, the teacher must provide the balanced solution but taken the need to argument it seems beneficial to complement MCQs with some problem-solving exercises where the students can argue for a given solution. The same applies regarding criterion IV.

Regarding criterium IV, the teacher should provide the possible interpretations and again also the balanced reasoning that is needed. Thus, increasing the workload on the teacher.
4.2 Difference in results between MCQ and non-MCQ assessment

4.2.1 Results

In table I, appendix II, one finds the grading in four subjects, namely rent, tenancy and property rights (bo), consumers rights (ko), contract law (av) and employment law. The class consisted of 31 students, fifteen males and sixteen females. The average results in the three non-MCQ tests was compared with the mean in the MCQ test in the appendix. The school was situated in an above average socioeconomic area. Students travel from afar to attend this school.

4.2.2 Analysis

The average unadjusted grade of the students for the first tree non-MCQ examinations was 2.9 while the adjusted grade came out as 3. That is equal to C and C-. Here an unadjusted score represents the actual score while the adjusted score represents retake of exam at the end of the year or if in the case of failure some connection with the test.

The unadjusted MCQ grade came out as 2.3 while the adjusted as 2.4. That is equal to D+ and D/C. Although there is a variation in between the mean of each non-MCQ assessment it seems to be clear that the pupil scored lower at the MCQ assessment than the three non-MCQ assessments. Thus, said one of the student (number. 30) wanted to retake the examination with open ended questions and where given the possibility. The score did not change.

In average then the students scored 0.6 lower on the MCQ than the average of the three non-MCQ tests. This result did not change even if the student where given a possibility to improve their results with a new exam.

Student no. 7, 9, 12, 16 and 30 had scored B or higher on previous traditional tests. Except for student no. 30 then who went through some difficult time at the test the average mean of these high performing students was 4.4 or B/A pre-to taking this test. The average score of these high performing students on the MCQ test where 3.25 or C+. The difference in test outcome between the three non-MCQs and the MCQ was 1.15 or a little more than one grade.

Student number 22, 24 and 25 where lower performing students with an average unadjusted mean predating the MCQ test of 1.9 or a D-. At the MCQ test the average mean increased with 0.3 to 2.2 points or a D+. 
4.3 Difference in results between sexes

<table>
<thead>
<tr>
<th>Males</th>
<th>Mean Bo-ko-av</th>
<th>Mean un-improved</th>
<th>MCQ</th>
<th>Mean un-improved</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2,5</td>
<td>2,5</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>3,5</td>
<td>3,5</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>3,17</td>
<td>3,17</td>
<td>3,5</td>
<td>3,5</td>
</tr>
<tr>
<td>4</td>
<td>3,5</td>
<td>3,5</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>3,33</td>
<td>3,33</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>10</td>
<td>2,33</td>
<td>2,33</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>14</td>
<td>1,33</td>
<td>1,33</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>15</td>
<td>3,33</td>
<td>3,33</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>21</td>
<td>2,83</td>
<td>2,83</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>24</td>
<td>2,33</td>
<td>1,83</td>
<td>3,5</td>
<td>3,5</td>
</tr>
<tr>
<td>25</td>
<td>1,83</td>
<td>1,83</td>
<td>2,5</td>
<td>0</td>
</tr>
<tr>
<td>26</td>
<td>3</td>
<td>3</td>
<td>3,5</td>
<td>3,5</td>
</tr>
<tr>
<td>27</td>
<td>3</td>
<td>3</td>
<td>3,5</td>
<td>3,5</td>
</tr>
<tr>
<td>28</td>
<td>2,83</td>
<td>2,83</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>29</td>
<td>2,67</td>
<td>2,67</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>41,48</td>
<td>38,48</td>
<td>33,5</td>
<td>31</td>
<td></td>
</tr>
</tbody>
</table>

Table II

4.3.1 Results

Fifteen male students and sixteen females have got their average results in the three non-MCQ tests compared with the mean in the MCQ test in the table above.

4.3.2 Analysis

With regards to the differences between male and females. The average score for a male on the three first assessments was 2.8 adjusted and 2.6 unadjusted. That is a grade estimated to C- and C/D. For MCQ's the equivalent score was 2.2 and 2.1 (table 2 - total/ 15). That is D and D+. The average male then scored slightly lower than the class mean at both non-MCQ and the MCQ examination.

The average female score of the same tree non-MCQs adjusted was 3.3 (class total 93.94 (Table I) - males 41.48 = females 52.46/16 persons) and 3.2 unadjusted (89.94 (Table I) -38.48=51.46/16). Equalling the grades C+. The average female score at the MCQ test was about 2.6 (75(Table I)-33.5=41.5) both adjusted and non-adjusted scores (72.5-31=41.5/16).

Berit Kjosnes
Ht 2017
Examensarbete, 15 hp
Examensarbete med ämnesdidaktisk inriktning, VAL, PE248HT17, 15 hp, page 24
That is C/D. The average female score then was higher on both non-MCQ tests and the MCQ test than the male score.

The score difference between females and males was a little higher for females. 0.6 (0.7) for females respective 0.6 (0.5) for males. Indicating that on first try female students scored lower than their male counterparts on the MCQ.
5. Conclusion and discussion

5.1 Conclusion

5.1.1 What kind of knowledge requirements can be tested by MCQ’s?

MCQ can successfully be utilised in large classes with many students. Although demanding to construct it provide a teacher with a powerful tool that enable fast feedback and high levels of transparency. The value of fast feedback increases in courses where several subjects are examined over a shorter period. MCQ is easy to score for students.

While it seems plausible that a well-designed MCQ can prove to be an effective and efficient tool utilised to test large domain of knowledge sampled at multiple cognitive levels, and in that, even the highest order of cognitive skills this depends in large on the construction of the same (2.3.3 and 2.3.7).

As put forth by Jönsson, Rodriguez & Haladyna there are some limitations when it came to creating an authentic situation where students could verbally express themselves utilising legal language, solve problems and reason regarding legal clauses (2.3.1).

5.1.2 What kind of difference are there between MCQ’s and non-MCQ’s when it comes to results?

In his meta-analysis Rodriguez found that the difference between constructed response tests and select response test where minimal (2.3.3 versus 2.3.7). The expectation with regards to the actual MCQ test was that the difference between non-MCQ and the MCQ test would be minimal. However, in my test students, in average, scored 0,6 lower on the MCQ compared with the average of the tree non-MCQ tests. 0,6 is about half a grade.

I found that the discrepancy between the MCQ and the non-MCQ tests for high performing students was higher than for low performing students. High performing students scored above one grade lower on the MCQ than what they had done in average on the non-MCQ tests. In addition to what has been said already a possible explanation can be an overestimation of students’ abilities. Low performing students on the other hand increased their result with 0.3 points or nearly half a grade.

5.1.3 What kind of difference in results are there between MCQ’s and non-MCQ’s when it comes to gender?

Seeing there seem to be a tendency among girls to perform poorer on MCQ’s (Murphy, 1982 (2.3.5 and 2.3.7) I expected a discrepancy based upon gender.

In this research, the average female scored higher than the average male independent whether assessment was done by utilisation of MCQ or non-MCQ. However, the research revealed that there was a difference in score between male and females when it came to non-MCQ test or the MCQ. The difference for females was 0.2 higher than their male counterparts on first trial. Giving that the MCQ test lowered females score slightly more than their male counterparts but only on the first trial.
While considering whether this is a gender issue one must take into account what has already been said in 5.1.2 seeing that the average mean of male students was slightly lower than the class mean at both non-MCQ and the MCQ examination. At second try females scored the same as their male counterpart.

5.2 Discussion

With regards to testing of knowledge requirements. While it appeared to be possible to test students’ ability to account for employment law and concepts even at higher levels one must make sure that the test measures what it is supposed to measure (4.1.1 and 4.1.2). The ability to measure should have been evaluated by another set of eyes than mine.

A possible solution regarding concept would be to create an authentic problem-solving exercise. One example is a court trial session where the students can act at the end of the year. While the ideal MCQ got three options the test utilised had in several instances more than three options and might have in some instances proved confusing (2.3.7). Another problem addressed in 3.2.3 is the problem of a researcher evaluating his/her own tool. The probability of self-confirming bias is overwhelming. The results then should be received.

Difference in scoring: This difference in scoring can be explained both by errors in MCQ construction (3.2.1), unfamiliarity with the test instrument (2.4) and the small sample. After all it was just 31 students and the test was only done only once.

And again, adding to what has been said the improvement of low grade students results on MCQs could also be caused by guesswork, or the possibility that students with few words find a way to express themselves through MCQ’s. It is also possible that MCQ could prove to be a valuable tool for improving student performance in line with Rød & Fjär in that low performing students read more throughout the course (2.2.3). Anyhow based upon what has been said the test need to be replicated and flaws corrected to provide an answer.

Gender: One explanation to the difference between the first and the second test might be that be unfamiliarity with the method or language in the tool of assessment. Here a standardisation of the assessment tool might be beneficial. It is felt that the difference in score is too small for a conclusion to be drawn.

5.2 Reflections

Law and theory seems to call for objectivity in assessment. How to acquire this objectivity with equality in mind might not be the easiest question to answer. It is however felt that by standardising the tool of assessment on a national level one might increase objectivity and impartiality while improving student performance by easing the cognitive load. Here the utilisation of MCQs might prove to be, in part at least, one of the answers. Adding then that MCQ are provide specific answers and thus increase transparency the perception of fair treatment might increase. Then it might apply as it is written, same rules of assessment apply to all.
Bibliography

Aiken, Edwin G. (1968). Delayed feedback effects on learning and retention of morse code symbols, Psychological Reports, 23, 723-730.


Byström, Jan (1992), Grundkurs i Statistik, natur och kultur, Falköping: Gumnessons Tryckeri AB.


Foucault, Michel (1983). Foucault replies to questions from the audience and Berkeley’s History Department in 1983, Foucault’s lecture, Berkeley


Heuman, Lars (2004/5). Domarens och åklagarens skyldighet att vara objektiva, Juridisk Tidskrift s. 42, nr. 1.


Repstad, Pål (1993). Närhet och distans, kvalitative metoder i samhällsvetenskap. Lund: Studentlitte-
ratur.


Sandgren, Claes (2011). Rättsvetenskap för uppsatsförfattare, Ämne, material, metod och argumenta-
tion, Stockholm: Nordstedts Juridik.

Scouller, Karen (1998). The Influence of Assessment Method on Students’ Learning Approaches: Multi-

Hellner, Trygve and Malmqvist (2010), Boo: Förvaltningslagen: med kommentarer, Stockholm: Nordstedt Juridik AB

Vetenskapsrådet (2002). Forskningsetiska principer inom humanistisk - samhällsvetenskaplig forsk-

Wendler, Cathy LW & Michael E. Walker (2006), "Practical Issues in Designing and Maintaining Mul-
tiple Test Forms for Large-Scale Programs", in Handbook of Test Development ed. Steven M. Dow-

Wiidland, Tove (2005). Examinationen som vägvisare. Högskolestudenter upplevelse av examinatio-


**Legal sources:**


The Administrative Procedure Act (1986:23)

The 1986 Administrative Procedure Act, available online in english at: https://www.kriminalvarden.se/globalassets/om_oss/lagar/forvaltningslagen-engelska.pdf

Berit Kjosnes
HT 2017
Examensarbete, 15 hp
Examensarbete med ämnesdidaktisk inriktning, VAL, PE248HT17, 15 hp, page 30
The Swedish Education Act (2010:800)
The School Ordinance (2010:2039)
Prop. 2009/10:80 En reformerad grundlag
Prop 1978:90 bil. 3s. Kungl. Maj:ts proposition med förslag till ny regeringsform och ny riksdagsordning m. m.
SKOLFS2017:16 Förordning om ändring i förordningen (SKOLFS 2011:144) om läroplan för gymnasieskolan
SKOLFS2011:149 Skolverkets allmänna råd om planering och genomförande av undervisning
SKOLFS2011:144 Skolverkets föreskrifter om läroplan för gymnasieskolan
SKOLFS 2010:90 Skolverkets föreskrifter om ämnesplan i Juridik
SKOLFS 2010:14 Förordning om examensmål för gymnasieskolans nationella program

Links:
http://lararnastidning.se/larare-satte-tre-olika-betyg-pa-samma-uppgift/Lärarnas tiding accessed 31/1
http://www.bloomstaxonomy.org/Blooms%20Taxonomy%20questions.pdf Bloom’s Taxonomy accessed 31/1

A Sin rätt till semesterlön!
B Sin rätt att sluta tidigt
C Pensionsgrundad inkomst
D Rätten att vara hemma med sjukt barn
E Rätten att ta semester!
F Skydd enligt försäkringskassans arbetsavtal!
G Skydd enligt LAS
H Rätten till semesteranställning
I Rätten till arbete vid tillfällig behov
J Rätten att vare hemma med barn utan lön!
K Hon behöver inte betala skatt!
L Rätten till betalt skönhetsvård på arbetstid!

2. Vad är ett kollektivavtal?

A Ett kollektivavtal är ett kollektivt avtal anställda emellan
B Ett kollektivavtal är ett avtal som kollektivtrafiken gör med sina passagerare
C Ett kollektivavtal är ett avtal som olika fackliga organisationer gör mellan sig
D Ett kollektivavtal är ett avtal mellan en grupp av människor och en eller flera arbetsgivare
E Ett kollektiv avtal är ett avtal mellan en människa och en arbetsgivare


A Vem som helst!
B Enskilda arbetstagare som är med i ett fackförbund.
C Enskilda företag
D En grupp av arbetstagare organiserat i ett fackförbund.
E En grupp av företag
4. Vem berörs av ett kollektivavtal?
   A All anställda berörs direkte!
   B Enbart dessa som är med i ett fackförbund berörs.
   C Arbetstagare som är med i fackförbundet i första hand men också andra berörs indirekta.
   D Enskilda arbetstagare som inte är med i fackförbundet i första hand men också andra berörs indirekta.
   E Alla företag!

5. I §11 Medbestämmande lagen (MBL) läser man om att man måste förhandla i vissa situationer. Man säger att man har förhandlingsskyldighet. Vad menas med detta?
   A Att man har rätt att bestämma hur arbetsgivaren får driva sin verksamhet!
   B Att arbetsgivaren måste göra som facket säger!
   C Att den enskilda arbetstagare har rätt att få sin röst hörd. Därför är arbetsgivaren skyldig att lyssna.
   D Att det enskilda företaget har rätt att säga till vidrörande facket!
   E Att arbetsgivaren är skyldig att höra på vad facket säger men måste inte göra som dessa vill!

6. Har ett företag rätt att ersätta en anställd med konsulter?
   A Ja, företaget får ersätta en anställd med konsulenter om dessa blir billigare facket kan inte göra något enligt MBL 38-40§§.
   B Ja, i princip har företaget rätt att ersätta en anställd med konsulter, men den fackliga organisationen kan använda sin vetorätt enligt MBL 38-40§§
   C Nej, företaget får inte ersätta anställda med konsulenter då detta inte är sjysst mot de anställda enligt MBL 38-40§§

7. Som en huvudregel gäller en anställning...
   A Tillsvidare
   B En bestämd tid
   C Vid tillfälligt behov
   D Som en provanställning
   E En viss säsong

8. Linus undrar lite på vad som gäller vidrörande uppsägning. Kan du hjälpa honom?
   A En uppsägning får vare muntlig
   B En uppsägning kan, men behöver inte vare, skriftlig
   C En uppsägning måste vare skriftlig
   D En uppsägning ska vare båda muntlig och skriftlig
9. Ett företag med 25 anställda måste säga upp tre anställda och undrar nu på vilken som blir kvar (enligt LAS) av följande:
   A MILZOVITS, som arbetat 21 år på företaget och som fyllt 50 år men som fortfarande talar knagglig svenska
   B VESNA, som är 23 år och som arbetat 2 år på företaget och till vilken chefen har ett mycket gott öga.
   C STJAKOVITS, som är 30 år och som arbetat på företaget i 8 år och som är en verklig klippa och som anses oumbärlig.
   D PUTTIS som har vissa alkoholproblem och som nyligen fyllde 55 år och har arbetat 28 år på företaget

    A En kille som har ett alkoholproblem
    B En tjej som kommit för sent på arbete tre gånger i rad
    C En kille som har tankat sin privata bil och betalat med företagskortet
    D En gammal och sjuk tant som har varit "för länge" på företaget
    E En kille som lånade en cykel utan att fråga om tillåtelse
    F En tjej som "lånat" lite från dagskassan

11. Vem avgör ytterst om arbetsbrist föreligger?
    A Den anställda
    B Facket
    C Arbetsgivaren
    D Arbetsgivarens organisation

    A Ja, man kan ta upp fallet i Arbetsdomstolen
    B Ja, man kan ta upp fallet i tingsrätten
    C Ja, man kan be sitt ombud (advokat) ta upp fallet i Arbetsdomstolen
    D Nej, loppet är kört!

13. Hur lång tid har arbetsgivaren till att ta ställning till om en uppsägning av personliga skäl ska ske eller ej? (från kännedom om händelse till varsel sker)

A Sture uppmanas att gå med i fackrörelsen Syndikaterna men vägrar. Sture mobbas och känner sig tvunget att gå med för att vara kvar på arbetsplatsen.

B Lindas chef har uttalat sig negativt gällande facket i massmedia

C En arbetsgivare utfärdade ett arbetsintyg till en anställd med en skrivning om att den anställda var mycket lojal gentemot sin fackliga organisation

D En arbetsgivare frågade vid anställningsintervjun om ansökaren var med i facket. Den sökande svarade ja och fick då reda på att någon anställning inte var aktuell.
<table>
<thead>
<tr>
<th>Student no.</th>
<th>Bostad</th>
<th>Konsument</th>
<th>Avtals</th>
<th>Arbetsrätt</th>
<th>MCQ</th>
<th>Mean</th>
<th>Bo-ko-av</th>
<th>Mean</th>
<th>Improved MCQ</th>
<th>Mean</th>
<th>Improved MCQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>D</td>
<td>2,0</td>
<td>C/D</td>
<td>2,5</td>
<td>C-</td>
<td>3</td>
<td>E</td>
<td>1</td>
<td>1</td>
<td>2,0</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>2,0</td>
<td>C/D</td>
<td>2,5</td>
<td>C-</td>
<td>3</td>
<td>3</td>
<td>C/D</td>
<td>2,5</td>
<td>3,5</td>
<td>3,5</td>
<td>3,5</td>
</tr>
<tr>
<td>3</td>
<td>2,5</td>
<td>C-</td>
<td>3</td>
<td>C/D</td>
<td>3,33</td>
<td>3,33</td>
<td>C/D</td>
<td>3,33</td>
<td>3,5</td>
<td>3,5</td>
<td>3,5</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>C/D</td>
<td>3,5</td>
<td>B+</td>
<td>4</td>
<td>4</td>
<td>B</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>C/D</td>
<td>3,5</td>
<td>B+</td>
<td>4</td>
<td>4</td>
<td>B</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>6</td>
<td>2,5</td>
<td>C/D</td>
<td>3,5</td>
<td>B+</td>
<td>4</td>
<td>4</td>
<td>B</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
<td>C/D</td>
<td>3,5</td>
<td>B+</td>
<td>4</td>
<td>4</td>
<td>B</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>8</td>
<td>3</td>
<td>C/D</td>
<td>3,5</td>
<td>B+</td>
<td>4</td>
<td>4</td>
<td>B</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>9</td>
<td>2,5</td>
<td>C/D</td>
<td>3,5</td>
<td>B+</td>
<td>4</td>
<td>4</td>
<td>B</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>10</td>
<td>2</td>
<td>C/D</td>
<td>3,5</td>
<td>B+</td>
<td>4</td>
<td>4</td>
<td>B</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>11</td>
<td>2,5</td>
<td>C/D</td>
<td>3,5</td>
<td>B+</td>
<td>4</td>
<td>4</td>
<td>B</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>12</td>
<td>2</td>
<td>C/D</td>
<td>3,5</td>
<td>B+</td>
<td>4</td>
<td>4</td>
<td>B</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>13</td>
<td>2,5</td>
<td>C/D</td>
<td>3,5</td>
<td>B+</td>
<td>4</td>
<td>4</td>
<td>B</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>14</td>
<td>2</td>
<td>C/D</td>
<td>3,5</td>
<td>B+</td>
<td>4</td>
<td>4</td>
<td>B</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>15</td>
<td>2,5</td>
<td>C/D</td>
<td>3,5</td>
<td>B+</td>
<td>4</td>
<td>4</td>
<td>B</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>16</td>
<td>2</td>
<td>C/D</td>
<td>3,5</td>
<td>B+</td>
<td>4</td>
<td>4</td>
<td>B</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>17</td>
<td>2,5</td>
<td>C/D</td>
<td>3,5</td>
<td>B+</td>
<td>4</td>
<td>4</td>
<td>B</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>18</td>
<td>2</td>
<td>C/D</td>
<td>3,5</td>
<td>B+</td>
<td>4</td>
<td>4</td>
<td>B</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>19</td>
<td>2,5</td>
<td>C/D</td>
<td>3,5</td>
<td>B+</td>
<td>4</td>
<td>4</td>
<td>B</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>20</td>
<td>2</td>
<td>C/D</td>
<td>3,5</td>
<td>B+</td>
<td>4</td>
<td>4</td>
<td>B</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>21</td>
<td>2,5</td>
<td>C/D</td>
<td>3,5</td>
<td>B+</td>
<td>4</td>
<td>4</td>
<td>B</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>22</td>
<td>2</td>
<td>C/D</td>
<td>3,5</td>
<td>B+</td>
<td>4</td>
<td>4</td>
<td>B</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>23</td>
<td>2,5</td>
<td>C/D</td>
<td>3,5</td>
<td>B+</td>
<td>4</td>
<td>4</td>
<td>B</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>24</td>
<td>2</td>
<td>C/D</td>
<td>3,5</td>
<td>B+</td>
<td>4</td>
<td>4</td>
<td>B</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>25</td>
<td>2,5</td>
<td>C/D</td>
<td>3,5</td>
<td>B+</td>
<td>4</td>
<td>4</td>
<td>B</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>26</td>
<td>2</td>
<td>C/D</td>
<td>3,5</td>
<td>B+</td>
<td>4</td>
<td>4</td>
<td>B</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>27</td>
<td>2,5</td>
<td>C/D</td>
<td>3,5</td>
<td>B+</td>
<td>4</td>
<td>4</td>
<td>B</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>28</td>
<td>2</td>
<td>C/D</td>
<td>3,5</td>
<td>B+</td>
<td>4</td>
<td>4</td>
<td>B</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>29</td>
<td>2,5</td>
<td>C/D</td>
<td>3,5</td>
<td>B+</td>
<td>4</td>
<td>4</td>
<td>B</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>30</td>
<td>2</td>
<td>C/D</td>
<td>3,5</td>
<td>B+</td>
<td>4</td>
<td>4</td>
<td>B</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>31</td>
<td>2,5</td>
<td>C/D</td>
<td>3,5</td>
<td>B+</td>
<td>4</td>
<td>4</td>
<td>B</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

**Table 1**

Appendix II