Challenges of Assessment in Ethics – Teachers' reflections when assessing National Tests

*Christina Osbeck, Olof Franck, Annika Lilja
and Annika Lindskog*

The aim of this article is to identify and discuss challenges in assessing pupils' knowledge in Ethics. The background of the study is the development of the knowledge field Ethics, a part of Religious Education (RE) in Sweden which in the sixties went from being a Christian school subject to a pluralistic and non-confessional one. The knowledge field Ethics is, in this school context, marked by vagueness, due partly to its indistinct frames and partly to a duality in the aims represented in the directives of the curriculum. *Methodologically*, data was produced through a think-aloud study where six teachers assessed pupil-responses in Ethics within the National Test for RE. The teachers' ongoing assessments were audio recorded, transcribed and analysed in a qualitative text analysis that focused on different kinds of challenges that the teachers experienced. Since this kind of study has not been conducted before, the analyses were explorative and, rather than being theory guided, were guided by the aim and the research question of the study. The *findings* show three groups of challenges related to the assessment processes. The first group of challenges concerns interpretations of the assessment instructions, the second regards competing ideas about what students should be given credit for during assessment in Ethics and the third is about being positioned between the pupils' need for good grades and the task of differentiating between students. The *discussion* distinguishes between general challenges in assessment and challenges related to Ethics and its partly unclear character in the RE school context. In such a situation when the construction of a knowledge field is vague, the influence of national tests can be assumed to be considerable and the need for research regarding different conceptions of ethical competence is therefore urgent.

CHRISTINA OSBECK, OLOF FRANCK, ANNIKA LILJA
AND ANNIKA LINDSKOG

Christina Osbeck, Associate Professor, University of Gothenburg
christina.osbeck@gu.se
Olof Franck, Associate Professor, University of Gothenburg
olof.franck@gu.se
Annika Lilja, Senior Lecture, University of Gothenburg
annika.lilja@ped.gu.se
Annika Lindskog, PhD Student, University of Gothenburg
annika.lindskog@ped.gu.se

Introduction

By tradition Religious Education (RE), in a Swedish context, involves teaching in Ethics. This tradition can be traced back to those days when this subject´s predecessor, the school subject Christianity, was taught in school, resting upon a confessionally defined ground. Christianity was, in an era and in a society where Christian perceptions of life and values were explicit and supposed to mark identity as well as community, thought to be the most important school subject, making sure that children and young people were educated to be prepared for living in a spirit of Christian faith and Christian morals (Selander, 1993; Algotsson, 1975).

Since the middle of the 1960s when the subject of Christianity was replaced by RE (see, for example, Osbeck, 2013), Ethics has continued to be a central strand within the subject. Of course, RE is now shaped and practised within a non-confessional context, but this fact has not broken the connection built between, on the one hand, religious and existential dimensions and, on the other, ethical and moral ones (Olivestam, 2012). Furthermore, the subject seems to rest upon two keystones that, although perhaps in a somewhat different way, were also salient within the teaching in Christianity in older days.

First, relating to the task regarding the conducting and moulding of the fundamental democratic values, especially highlighted in their present form since the 1990s (Osbeck, 2004), one could, within RE

as well as in the educational context as a whole, identify what has become known as *values education*, a label suggested as an umbrella term including moral education, character education, civic education and citizenship education (Taylor, 1994; Thornberg, 2004). Secondly, in relation to this kind of education, one could identify what may be called *ethics education*. This second kind could be conceptually apprehended in terms of being formal and school based rather than informal, explicit rather than implicit (ibid).

In relation to the subjects of both RE and of Christianity, it would be possible to identify two kinds of education, one relating to transmitting values and one relating to knowledge development. The borders do not seem to be strict or clear-cut. At the core of values education are educational ideals regarding personal development in children and youngsters. Certainly this task, within the confessional subject, was performed and developed with reference to at least partly different premises than those from which the task of transmitting values is carried out today (Selander, 1993). While the ideal distinguished by the subject of Christianity seems to have been defined with regard to a life lived by a spiritually conscious person, practising what was assumed to be Christian ethics, the corresponding ideal that was presumed to mark RE could be characterized in terms of a democratic approach, applied within a pluralistic national and global society, focusing on a deliberate reflection upon and practising of universal human rights, rejecting various kinds of discrimination and harassment. The moral and character education, however, as well as the civic and citizenship one, seems to share evaluative visions of making judgments about what are to be counted as right actions and a good life. In both cases an ethical code or a moral canon works as a point of reference, in relation to which issues concerning how to reason and how to act are discussed and assessed (Franck, 2013a).

With regard to ethics education, similarities may also, consequently, be observed. Here, however, one can identify differences that may have relevance to the questions raised and treated in this article. While an education within the subject of Christianity seems to have revolved around the children´s and young people's learning of fundamental moral rules and concepts, signifying and expressing what was taken to be ethics founded in Christian tradition and belief, today´s focus in RE is not anchored in a confessional canon of presumably knowable

facts. Rather ethics education is directed towards the knowledge requirements stated in the syllabus, highlighting skills of reasoning in moral matters using a variety of ethical concepts (Franck, 2013b). The meaning of *ethical competence* seems to have shifted from referring to an understanding and an application of moral rules and concepts, defined with reference to Christian fundamentals, to an interpretation according to which such a competence is shown and expressed by the use of the skill of ethical *reasoning* involving a relevant and correct use of selected ethical concepts. In the first case, ethics education is aimed at and believed to lead to knowledge about moral codes *that are regarded as right and praiseworthy*. In the second case, such an education is aimed at and believed to lead to knowledge about moral codes in relation to which *no definite opinion is prioritized as universally preferable or tenable for moral choice and action.*

This latter approach, practised and developed within RE, seems to imply a rather open-ended field regarding how to define what kind of skills and what kind of knowledge are to be focused on in order to carry out ethics education, involving assessment in relation to the knowledge requirements stated in the syllabus. A variety of conceptions of "ethical competence" and "ethical knowledge" may here come to mind – and with them a range of challenges mirroring different approaches to, and within, relevant assessment procedures (Tillson, 2011). Research into these conceptions, and the challenges relating to them, would presumably contribute to the shaping of knowledge about ethics education in theory and practice, as well as to a development of strategies for handling challenges facing teachers in RE.

Questions of ethics have, however, been researched comparatively little in the field of RE. The National Tests in RE, involving items relating to ethics education, may provide information that can be used in relation to the aim of making the previously mentioned challenges visible, creating space for *analysis and discussion.*

The empirical study presented in this article is thought to contribute to this process. By an analysis of the ongoing assessments of the teachers involved in the study, space is made for both identifying challenges in the assessment of pupils´ knowledge in Ethics, and relating these challenges to the teaching context characterized by the indistinct frames of the school subject Ethics, and the duality apparent in the aims of the curriculum.

Aim

The aim of this article is to identify and discuss challenges in assessing pupils' knowledge in Ethics. A description of analytical perspectives, shaped with reference to curricular considerations, in relation to such challenges, has to take into account the fact that ethics education is obligatory in compulsory school in Sweden. Ethics is a field of reflection, argumentation and analysis with a focus upon issues and areas belonging to existential and moral dimensions of life, and it does not seem possible, at least not at first sight, to place it among subjects that make use of standard methods of assessment and grading. At the same time, the curriculum for compulsory school contains specific and seemingly strict knowledge requirements in order that pupils´ skills in moral reasoning and their knowledge about ethical concepts may be assessed in ways that are similar to ways of assessing more concrete forms of knowledge regarding, for example, world religions. The demand for teachers to make use of these requirements when assessing pupils' statements seems to create challenges which could perhaps be difficult to trace and identify.

In order to investigate such challenges and make them explicit, the authors have performed a think-aloud study where some teachers evaluate certain tasks belonging to Ethics within the National Tests in RE that were given in the spring of 2013.

The research question in focus for the following analysis can be formulated in this way: Which challenges can be identified when six teachers assess pupils' responses in Ethics in the National Tests in RE?

Previous research

Research about assessment in school is a rather wide field. In this study, attention is drawn towards assessment research that focuses on instructions for assessment, and challenges that can arise when using pre-set criteria as well as when assessing specifically within RE.

Challenges concerning pre-set criteria and instructions for assessment

Regarding grading, the first distinction that has to be made is between holistic grading and analytic grading (Sadler, 2009). In holistic grading

the assessor builds up a complex mental response to different parts of a student's work. In analytic grading the teacher makes separate qualitative judgements on each of the pre-set criteria. The teacher determines the criterion prior to marking. According to Sadler (ibid.) the idea of focusing on criteria specifically related to quality was first proposed in the 1920s.

The aim of analytic assessment is to achieve certain practical ends, for example, to compare the quality of a student's work with fixed criteria, and this is educationally more defensible than making comparisons with how other pupils perform. Another advantage is that explicit criteria enable students to understand the assessment process. A problem that comes with analytic marking is, however, that it is not possible to take into account all the necessary nuances of expert judgement in the same way as in a holistic grading process. In this sense, analytic marking is deficient (ibid.). Sadler (1987) also notes that the meaning of verbal descriptions is always, to some degree, vague and fuzzy. The criteria need to be interpreted (Connolly, Klenowski & Wyatt-Smith, 2012). These interpretations inevitably differ and according to Smaill (2011), a number of conditions must be fulfilled to help the teachers to develop a common understanding of the pre-set criteria. Assessors need concrete examples demonstrating the levels of achievement, and assessors also need opportunities to participate in a social moderation process. A social moderation process involves a comparison and alignment of participants' judgement of pupils' work in relation to the pre-set criteria. During the social moderation process the teachers develop a shared understanding of the qualitative descriptors (ibid.).

Instructions for assessment constitute an important part of the national tests in Sweden. They express an interpretation of the knowledge requirements, developed to support the teachers' assessments. The instructions can thereby be experienced as a disruption to intuitive ways of working with the pupils (Gipps et al., 1996). However, when studying the role of assessment criteria during teachers' collaborative assessment of pupils' portfolios, van der Schaaf et al. (2011) found that without assessment criteria, the teachers based their judgement more on their personal opinion and less on the evidence found in the pupils´

work. But when the teachers used assessment criteria, they based their judgements significantly less on personal characteristics of pupils. A conclusion in this study is that the quality of the judgement process is more construct valid when teachers use the assessment criteria.

It has also been observed that ways of working with assessment and making professional judgements are dependent on the subject in question (Wyatt-Smith & Klenowski, 2012). According to Wyatt-Smith and Klenowski, teachers of mathematics tended to give emphasis to stated standards. They regarded assessment as an objective process and were part of an assessment-as-measurement tradition. Teachers of English, on the other hand, had a more reflective and holistic approach to assessment that was often intuitive and usually non-numerical. Both approaches were found to be problematic in some ways. The teachers who saw assessment as an objective process devalued their own first-hand assessment knowledge as subjective and not good enough, and the teachers who assessed in a more holistic way lacked confidence in their self-knowledge as assessors. Based on their findings and Sadler's criteria, which have enabled individual assessors to change or focus their assessment method (Sadler, 1985), Wyatt-Smith and Klenowski propose a strategy with three components to support teachers' use of subject standards. The first is elaborated guidelines about on-balance judgement processes. The second is a suggestion to use examples of pupils' work that are indicative of the standards, i.e. illustrative of a particular achievement. The third is about descriptive reports of pupils' achievements that accompany the exemplars to give insight into the factors that influence the overall judgement. These reports are supposed to be written during a period of time when particular subject content is treated and they are then to be used by the teacher to make an overall judgement.

An additional problem connected to assessing, which is illuminated in a study from England (Radnor, 1995), is that teachers in compulsory school believed that a tighter frame from the National Curriculum and in particular the assessment arrangements that were implemented at the time of this study, limited their course of action as teachers as well as their working practices of administration and pupil organization.

CHRISTINA OSBECK, OLOF FRANCK, ANNIKA LILJA
AND ANNIKA LINDSKOG

Challenges in assessing subject matter knowledge in RE

To assess pupils´ work is a core task for teachers. Through assessments, pupils get opportunities to learn more in relation to current goals but the other side of the coin is that assessment also values and grades pupils' work. This is a part of the work that can be particularly demanding and associated with various difficulties.

Assessment and religious education is an area that has yet to be researched and the same goes for assessment and ethics. In the few studies that exist it has been argued that one of the compelling challenges for RE professionals is to develop authentic, performance-based approaches to assessment, which clearly reject accountability. This way of assessing offers RE teachers a route to a more challenging curricular activity than when assessing achievements through measured accomplishment (Baylock, 2000).

In another study, Grant and Matemba (2013) argue that in many cases the assessment of RE is generally out of focus for teachers in Scotland, largely due to the fact that the assessment is seldom based on religion but rather on other issues, which are important, yet should not, according to Grant and Matemba, be allowed to consume so much of the limited time that RE is given. Examples of issues that are assessed instead of the subject of RE are generic skills such as listening, working in groups and enthusiasm. The same study reveals that it is also evident that school practices in RE are failing to match the ambitions of the curriculum. The teaching in RE has remained at a lower cognitive level where the pupils are taught to identify, describe, colour and so on instead of being taught to discover, critique and challenge. A reason for this might be that the curriculum in Scotland is in its early years of implementation, which is similar to the situation in Sweden today. In Sweden, a corresponding implementation started in 2011 and one could, regarding certain dimensions concerning, for example, assessment issues, say that this implementation is in one sense still being carried out.

Another aspect revealed is an ambiguity regarding how values are to be assessed (ibid.). They suggest that teachers need to plan their teaching in such a way that it is possible to assess controversial issues such as values. An additional circumstance that might cause difficulties when assessing pupils' work is that the teachers are taking their pupils' personal lives into account in different ways, for example when grading

them (Klapp Lekholm, 2010; Rinne, 2013). Similar findings have been found in other studies. For instance Allal (2012) shows in her study how the teachers, although they ensured a certain standardization of their assessment procedures, also introduced adaptations to take into account specific details about individual pupils. Also Brown et al. (1997) state in a study on teachers' views about the validity of national testing that the teachers major concerns about the validity of their assessment was about the unfairness of the tests to specific types of pupils, and the poor match with classroom practice. For example, the teachers sometimes decided not to take into account the result on one test if it was not considered valid because of problems in the personal life of the pupil.

Data and method

The empirical material of this article consists of transcripts from audio recordings of teachers who have been asked to reassess pupil-responses to tasks in Ethics, which were part of the first National Test (NT) in RE in Sweden (2013). This reassessment was organized by the team commissioned by the National Agency of Education to conduct the NT in RE, of which the research team conducting this study are also members. The present study was done with a view to identifying challenges for teachers in assessing pupils' knowledge in Ethics, as it is demonstrated in national tests. Two year-six teachers participated in the study and assessed the same 15 pupil-responses and four year-nine teachers assessed the same 12 pupil-responses. The teachers of year six are referred to in the findings section of this article as 6a and 6b, and the teachers of year nine are referred to as 9a-d. In the responses of the pupils, whose identities were not known to the re-assessing teachers, all grades that have knowledge requirements (E – lowest, C – intermediate, A – highest) were represented, according to the previous evaluation made by their ordinary teacher, and responses that had been evaluated as failed, here called F, were also included. The re-assessing teachers in the current study did their assessment in relation to the assessment instructions and were asked to try to put their thoughts into words while their worked. They could take breaks if they wanted to but were asked to work with the task as an ongoing process. The study can therefore be described as a "think-aloud study" in line with how

such studies are usually conducted (e.g. Charters, 2003), a method that has been used successfully for studying difficulties with assessment (e.g. Crisp, 2010). The transcription of the audio recordings were made with the intention of being close to the speech of the participants in the study, to indicate larger pauses but without the kinds of oral sounds that may be present when one "thinks aloud". The aim of the study was not to conduct analyses where attention is paid to those kinds of utterances and therefore a less detailed transcription was preferred.

Table 1. Analysed material, participating teachers and background material

Material analysed in this article	Participating reassessing teachers	Background material: Responses that the teachers worked with
2 transcriptions of year-six teachers "thinking aloud" while doing assessments of 15 responses	2 year-six teachers	15 responses of year-six pupils (the sweat dilemma)
• 6a about the 15 responses to the sweat dilemma	• 6a	Originally the responses were assessed with the following grades:
• 6b about the 15 responses to the sweat dilemma	• 6b	A: 4; C: 4; E: 4; F: 3
4 transcriptions of year-nine teachers "thinking aloud" while doing assessments of 12 responses	4 year-nine teachers	12 responses of year-nine pupils (the forgiveness task)
• 9a about the 12 responses to the forgiveness task	• 9a	Originally the responses were assessed with the following grades:
• 9b about the 12 responses to the forgiveness task	• 9b	A: 3; C: 3; E: 3; F: 3
• 9c about the 12 responses to the forgiveness task	• 9c	
• 9d about the 12 responses to the forgiveness task	• 9d	

The curriculum requirement, pupil-tasks and assessment instructions

The assessment instructions and the pupil-tasks of the NT are directly related to the syllabus of RE in the curriculum Lgr11. In the syllabus aims, core content and knowledge requirements draw on each other. These connections are expressed in more detail in Table 2, in the appendix. Some central phrases in the syllabus concerning Ethics express the aim that pupils should be given "…opportunities to develop their ability to […] reason and discuss moral issues and values based on ethical concepts and models." The core content for year 4-6 (age 10-12) states among other things that the teaching should include some ethical concepts, daily moral questions and what it may mean to do good. These issues are also central for year 7-9 (age 13-15) but here more specific areas are also expressed such as consequential, deontological and virtue ethics.

For this particular study one task in Ethics in year six and another in year nine were chosen. The task from the year-six test included a short story with an ethical dilemma in which a girl is asked by some friends to tell her best friend that she smells of sweat and that she needs to do something about it. The pupils were asked to write about what they think that Keyla should do, why they think so, and what consequences it may have for the different parties. The discussion should include at least two of the following or similar concepts: responsibility, wrong, right, duty and fair. The task was constructed in order to test the knowledge requirement of the syllabus, which states (the separate wordings for the grade levels E, C and A are written in italics):

> Pupils can apply *simple/developed/well-developed* reasoning about everyday moral issues, and what it might mean to do good. Pupils make reflections that *basically relate to the subject/carry the reasoning forward/ carry the reasoning forward and deepen or broaden it* and use some ethical concepts in a *basically/relatively well/well* functional/functioning way (Swedish National Agency for Education, 2011, p. 185 f.).

In the task from the year 9 test, pupils were supposed to discuss why forgiving can be important for the person that begs for forgiveness as well as for the person who forgives. This task draws on the following knowledge requirement:

> Pupils can reason and argue about moral issues and values by applying *simple and to some extent/developed and relatively well/well-developed and well-informed* reasoning, and use ethical concepts and models in a *basically/relatively well/well* functional/functioning way (Swedish National Agency for Education, 2011, p. 187 f.).

These knowledge requirements were operationalized in assessment instructions, presented in the Table 2.

Method reflection and data analyses

Since the aim of the study is to identify challenges in assessing pupils' knowledge in Ethics, it was considered important to design situations that were as close as possible to a real situation where a pupil-response is assessed and where a teacher may experience different kinds of challenges. Therefore the current think-aloud study was conducted. Even if think-aloud studies are sometimes said to explore thought processes (Charters, 2003), this seems to be an unrealistic expectation. However, it is reasonable to maintain that this kind of study reveals some of the reflections which occur to teachers while working with pupil-responses, the kind of reflections which they are aware of, which they want to share and are able to express verbally in their ongoing work (see, for example, Crisp, 2010). Since the processes are authentic in the sense that actual assessment is taking place, the method can be assumed to give better answers to the aim of the study – to identify challenges in assessing pupils' knowledge in Ethics – than for instance interviews would have done. In interviews, the time between actual work and utterances about the work can be expected to limit the scope of the challenges that the teachers are able to express. Even though the samples of teacher- and pupil-responses in this study are limited, the variation in the material (different kinds of teachers, tasks and levels of pupil-responses) makes it reasonable to assume that the main kinds of challenges that teachers experience while assessing pupils' knowledge in Ethics are captured in this study, especially in relation to the knowledge requirements of the Swedish curriculum and the national tests. The teachers participating in this study had had no experience of reassessing pupil-responses before they conducted the audio-recordings of the think-aloud study.

Since this kind of study is new in this particular knowledge area, it has been important to maximize the scope of knowledge present in the material and to conduct inductive and open analyses. The focus of the hermeneutical close-readings of the material has been the aim concerning the different kinds of challenges that the teachers experience in assessing Ethics. All the challenges that were identified have been given labels. Different kinds of challenges have been carefully related to each other and a system of categories has emanated from the material. The space or scope of findings is characterized in three main and overarching categories of challenges that the teachers can experience while assessing pupils' knowledge in Ethics.

It is possible to understand the utterances of the assessing teachers as utterances made up of different layers. First it is possible to identify what the teachers are doing, how they for instance assess a task, what grade they give, and compare this act with the acts of the other teachers and the assessment instructions. The analysis can in that sense reveal difficulties in the assessing processes that seem to have given rise to the differences. Second one can also identify what the teachers explicitly talk about as difficulties when they are in the process of evaluating the tasks. Third, one can focus on the utterances of teachers that are of a meta-nature, i.e. when the teachers take a step backwards and talk about the process of evaluating the tasks more generally. The systematic analysis of this article, described above, refers to the second level of utterances, i.e. what the teachers explicitly talk about as challenges. However a few illustrative examples from the other levels of analyses are also shown.

Findings

Three kinds of challenges in assessing pupils' knowledge in Ethics have emanated from the close-readings of the "think-aloud material". The first group of challenges includes those challenges that are inherent in interpreting the wording of the assessment instructions. The second group of challenges is related to the fact that teachers can bring other ideas into their assessment work about what should be valued when it comes to Ethics and the kinds of tasks that the pupils have been given. The third group of challenges regards the fact that the teacher

is a human being who has to navigate between his/her concern for the well-being of the dependent young pupils and his/her commission from the state to differentiate between the children.

Challenges regarding the assessment instructions

One conclusion that can be drawn from what the teachers say while assessing the pupils' tasks in Ethics is that *qualities of knowledge in Ethics*, the qualities of the responses which the assessor on the basis of the instructions should look for, are not obvious, are *not something that can be taken for granted*. Quite a few of the difficulties that the teachers experience are related to fact that the qualities that the assessment instructions prescribe are not taken for granted by the assessing teachers. The teachers who assess tasks for year six are struggling with the meanings of "ethical concepts" and which concepts can be accepted as "other similar concepts", with the meaning of "description of consequences" and how developed such a description must be in order to count, and also with the meaning of "change of perspective" and how developed such a change and its adherent descriptions have to be (Skolverket, 2013b, p. 10, i.e. Assessment Instructions for year 6). For the assessing year-nine teachers, the meanings of "problematize" and "consequences" get most attention (Skolverket, 2013a, p.12, i.e. Assessment Instructions for year 9). Furthermore, both teacher groups are concerned about how they should interpret the requirement for the pupils to "reason [resonera]" and how developed answers should be in this respect in order to count.

The challenges that the assessing teachers experience with *ethical concepts* are related to how specific the concepts must be and whether highly colloquial words can be accepted. Is for instance being "sulky" an ethical concept, is it possible to consider being an "asshole" equal to being mean and is "not being a bitch" of ethical importance (6b)? The assessment instructions do not give any clear answers to these questions and the teachers do not have any previous knowledge on this matter to guide them. The findings may show something important about Ethics as a school knowledge field. In which other subject areas of compulsory school are the concepts of the field so uncertain?

A *description of consequence* is, according to the assessment instructions year six, something that characterizes an answer that

deserves at least a grade C. One might therefore imagine that any utterance that describes a consequence should be interpreted as a C. This is not, however, a foregone conclusion since the assessment instructions exemplify an E-answer with an utterance that includes consequences (Skolverket, 2013b, p.10). At the same time descriptions of consequences are said to be something through which the pupil *develops* his/her way of reasoning [our italics], which can be understood as a remark about an expected quality. But what makes the required quality of the descriptions of consequences more complicated to grasp is that the assessment instructions for C-level first state "*possible* consequences" and later a "*relevant* description of consequence" [our italics]. The utterances of the assessing teachers show that they pay attention to different parts of the assessment instructions but sometimes they also express the contradictions in the instructions explicitly. Furthermore, the teachers' way of treating the instructions and the tasks shows that the question occurs as to whether arguments have to be added to a written consequence in order for it to constitute an approved description of consequence. "There is a description of consequences, the consequence is mentioned, that is true, but there is no description of *why* they would stop being friends" (6a). How developed or "exhaustive" (9d) an answer must be in order to count constitutes a challenge also among the year-nine-teachers. Here the difficulty is complicated by an unclearness in the assessment instructions as to whether a grade C-response should include consequences concerning both forgiving and being forgiven. The first part of the instructions stresses that:

the pupil should in a developed way reason regarding forgiveness with respect to both aspects… (Skolverket, 2013a, p. 12)

but a later part states that:

the reasoning concerning at least one of the aspects is relatively developed… (Skolverket, 2013a, p. 12).

Different ways of interpreting what constitutes a reasonable description of consequences seem to explain to some degree why the same answer is given different grades by different teachers.

The difficulties in interpreting what characterises a *change of perspective* or a developed *way of reasoning* are similar to the difficulties described in relation to descriptions of consequences, although one difference seems to be that it is easier to miss a change of perspective than to miss a description of consequence. An unclearness concerning change of perspective, not explicitly mentioned by the teachers but which can be seen in the material, is whether a change of perspective must mean that the pupils reason from the perspectives of different actors or whether it can also mean that the pupils reason about a particular act from different perspectives. *Problematizing* raises similar uncertainties concerning how developed the answers must be. However when it comes to problematizing, the think-aloud utterances of the teachers reveal more specific ideas about what "problematize" stands for. For instance, there is the opinion that problematizing includes examining a phenomenon from different angles. One should not only point out advantages of forgiving but should also emphasize the difficulties inherent in forgiving (9b). The teachers also appear to believe that it is advantageous to show what the opposite would mean, e.g. "if one does not forgive" (9c), or to specify certain conditions in relation to their perspective, for example, by showing that the good effects of forgiving are also related to the time aspect: "that one should be quick in forgiving" (9c).

Besides the fact that the assessment stresses competences that are not taken for granted or easily understood by the teachers, the main difficulty related to being commissioned by the assessment instructions concerns *the relation between part and whole in the responses*. The assessment instructions focus on individual aspects to be assessed and thereby ignore the fact that answers in their entirety also communicate different qualities, according to the teachers. One teacher in his/her meta-discussion regrets that "you partly get caught up in counting" (6a) that all aspects are present. Other teachers express that they are sorry when a good answer gets a lower grade because it is lacking in one part although it is excellent in other parts (9b; 9c). "I cannot give more than – with these assessment instructions – an E since the answer doesn't consider both aspects" (9b).

Challenges related to having additional or alternative ideas of qualities of knowledge in Ethics

Some of the assessing teachers have additional ideas or alternative ideas to those in the assessment instructions regarding competences in the knowledge field of ethics. Sometimes such ideas seem to have consequences for the outcome of the assessment – the grades given – but the ideas are also mentioned as merely frustrating to a greater or lesser degree. These ideas concern the *content* but also the *form* of the answers and the ideas can affect pupils' grades in both a positive and a negative way.

Content aspects of the pupil-responses that should be given credit, according to the teachers, are altruism, ethical insights, and certain precise and characterizing concepts. For instance one teacher states as a positive quality of a response that the responsibility that the pupil argues in favour of "not only [applies to you] as a friend but also as a fellow human being" (6b). The answer thereby expresses the importance of *altruism*. In a similar way, another teacher stresses as negative the fact that the pupil seems to argue for the importance of forgiveness "from an egotistic angle" (9b). Occasionally assessing teachers also emphasize a quality of a response that might tentatively be called *ethical insight*. For instance, one teacher states that the answer is good and has depth since it expresses that "when one asks for forgiveness one forgives oneself at the same time" (9c). Another teacher emphasizes as a "nice answer" how the pupil writes about humbleness: "you swallow your pride and show yourself to be humble towards the human being who actually manages to forgive you" (9d). An answer that is evaluated as having been written in an "ethically very nice way" expresses that the main character in the task "perhaps has noticed it [herself]" and that there "may be a reason and an explanation" (6a). The fact that some teachers associate good quality in the field of Ethics with the use of distinct *concepts* may be related to the RE syllabus, which explicitly stresses the importance of using ethical concepts. For the pupils in year six it was also a part of the task for which they were being assessed while *neither the task instructions for the year nine pupils required this nor the assessment instructions, with the exception of the concept of forgiveness*. Despite the fact that the assessment instructions do not advocate giving credit for additional

concepts, some of the teachers in year nine stress the use of concepts as a strength. The material does not reveal whether they consider the use of concepts to be an important factor in the knowledge field of Ethics in particular or whether they regard it as an important competence in every academic field. In any case, some of the teachers consider it to be particularly praiseworthy when the pupils characterize forgiveness as something "strong" or "mature" (9d; 9c) and link forgiveness to concepts such as "humbleness", "remorse" and "liberation" (9d; 9c). The use of these concepts is also said to contribute to an answer being characterized as well-written (9b).

Something that should be researched further is the fact that some of the year-nine teachers (9b, 9c, 9d) had difficulties with interpreting religious explanations of why forgiveness is important, such as "in order to avoid punishment from God". The teachers assessed the answers in different ways. One rejected the answer and considered it not to be sufficiently developed (9d), but others expressed the religious dimension as almost unfavourable "Yes, even if it is mostly about God, that God should forgive, I think that the pupil shows that [he/she] understand what it is about". (9c).

When *form aspects* of answers were mentioned they were mainly considered to be unfavourable factors even if exceptions occurred where a form aspect also could work as a favourable factor: "A longer answer. It does not always necessarily mean that it is a good one, but…" (9d). Examples of unfavourable form aspects that the assessing teachers stressed are shortness (e.g. 6b), lack of structure (e.g. 9d), unclearness (e.g. 9b), presence of highly colloquial expressions, of which it is said that it is "hard for a teacher to see beyond" (6a), bad language and composition (e.g. 9a). In relation to many of these remarks concerning unfavourable form aspects of answers, the conclusion is expressed that it is hard to understand what is stated in an answer. But there are also exceptions, such as when the assessing teacher concludes that "there are thoughts which the person in question has difficulties writing down but you can read the spirit…" (9b). The outcome of an assessment sometimes seems to be related to the attitude of the assessor.

Challenges related to being squeezed between the commissioner and a dependent pupil

Reading the "spirit" or the underlying intention of a student whose written answer is hard to understand is one example of how the assessing teachers indirectly express that they certainly want the very best outcomes for the pupils. Even if this is a specific and research-related situation, the teachers express an everyday knowledge that the pupils are dependent on the assessments. Some utterances of the assessing teachers can be interpreted as expressions of ambivalence in relation to their task. The teachers are in a sense squeezed between the pupils and the commission by the state to differentiate between these young people.

Empirically the ambivalence in relation to the commission is revealed through, on the one hand, expressions of feeling for the pupils and acts of loyalty to the pupils. On the other hand, there are expressions of how the assessing teacher hesitates to give high grades, which can be seen as an act of loyalty towards the commissioner. If every pupil gets a high grade, the test will not be discriminating.

Examples of comments where the teachers apparently feel for the pupils are "sometimes the work of the pupil can be *helped* by a more developed reasoning concerning the context…"(6a [our italics]), "Therefore, sadly, it will be after all an E for this pupil" (9c), "So even if there are a few linguistic misses and it is a bit difficult to interpret, I do think, probably, that it is... pretty developed anyway, rather developed through the work. And he mentions at least briefly one of the aspects that make it a weak C. But if I give him the benefit of the doubt I think... which is what one should do." (9c). The last example almost communicates a kind of suffering even if the teacher here does not express it as explicitly as the next one "I feel terrible but I do not think it is enough…" (9d).

The ambivalence in relation to the commission can be identified in the same person. The teacher who felt terrible when he/she could not accept the pupil's response does, for instance, express doubts when it comes to giving high grades as well. "I would give an A here *actually*" [our italics]. "Now I am back with X to whom I will give an A. Finally". (9d). In such a hesitation, the assessment instructions can

be presented as a kind of alibi. "This pupil response I consider, on the basis of the assessment instructions, as grade A." (9b).

The challenges that we have identified in the utterances of the assessing teachers are of different kinds; some are related to the task and the assessment instructions, and others are related to the specific challenges of the knowledge area of Ethics. But there are also difficulties that are related to the very nature of being a teacher, of being a teacher in general. The focus on the specific commission to assess answers in ethics has surprisingly pointed towards the broader difficulty of being a teacher. The profession includes both relational and instrumental tasks, to both help pupils to obtain the necessary qualifications and to differentiate among them.

Conclusions

The knowledge field Ethics in the Swedish school context has an unclear character, partly due to the fact that the object for analysis and reflection is what is considered to be right and praiseworthy. This means that Ethics, on the one hand, is a school knowledge field like most others that aims at developing young pupils' formal skills. But on the other hand, Ethics is to a larger degree than other subjects affected by the normative project of Swedish compulsory school. There are certain values that education shall impart and establish, values that the curriculum describes as "fundamental". In this way these values also come to constitute the object for analysis and reflection in the subject since these values are what all pupils are supposed to embrace and consider "right and praiseworthy". The Swedish curriculum seems on an overarching level to prescribe values education of both a "critical" and a "conservative" kind (see Jones, 2009).

Against the background of this ambiguous position of the subject, it is not surprising that quite a few of the challenges that the teachers of our empirical study express are particularly related to interpretations and understandings of the knowledge field, even if there are also challenges of a more general kind. Such general challenges have been described in previous research. Teachers may experience, for instance, a tension between their compassion for the children as fellow beings and their task as assessing civil servants, something which often seems

to mean that teachers take into account more and wider aspects than the ones that the assessment instructions stress (see e.g. Allal, 2012; Grant & Matemba, 2013; Klapp Lekholm, 2010; Rinne, 2013;). One variety of this overarching challenge is the tension between a wish to conduct a more holistic assessment of the pupil-responses, and the instructions to evaluate on the basis of specific aspects, to make an analytic assessment (Sadler, 1985; 2009). According to previous research, this is a tendency that may vary among different kinds of teachers (Wyatt-Smith & Klenowski, 2012).

Something that has not been described in previous research before, and therefore an important contribution to the research field, are the special challenges that are related to assessing pupils' responses in the knowledge field of Ethics. Some of these can be interpreted with regard to the ambiguous position concerning descriptivity and normativity within Ethics. The ambiguity is not easy to resolve since it is certainly difficult, or perhaps even impossible, to make a distinction between *reasoning* concerning an object, what constitutes a good action for instance, and the *character* that this object is given. When, for instance, Lawrence Kohlberg, with reference to a Kantian tradition, stresses the importance of the individual setting its own ethical principles to live by, the perspective is still normative due to the fact that these principles are supposed to be universal, altruistic, maintain the general social order and protect human rights. These are the perspectives that are considered to be most developed in this tradition (e.g. Bergling, 1982). The teachers in this empirical study who seem to have other ideas of creditable knowledge in the field of Ethics than those in the assessment instructions, the ones who stress altruistic values in the responses of the pupils and want to discourage egocentricity, can be understood as belonging to this tradition. They resist separating reasoning concerning an object from the character that the object is given.

With regard to the challenges in assessment that seem to be related to teachers having ideas of what constitutes knowledge of the field that differ from what is written in the instructions, another such idea is what here has tentatively been called ethical insight. Such a competence can, on the basis of the findings, be described as an ability to produce nuanced interpretations of a situation, i.e. the different wishes and demands that are in question in a situation, but also an ability to express this in a

way where the complexity and nuances of the situation are retained in the utterance. Such an ability can be related to Løgstrup's ethics of closeness or situated ethics. Here the central ability of the ethical subject is to grasp the silent ethical demand that is directed towards me from the need of the other. But the central ability also includes one's own response, taking responsibility for the demand, the situation and consequently for one's neighbour (Løgstrup, 1994). Løgstrup's perspective here points towards an action-related understanding of ethical competence with potential to widen the current discussion of Ethics as a knowledge field further (see e.g. Almers, 2009).

The challenges that the teachers appear to experience most frequently when assessing the responses of the pupils are unclearness concerning the assessment instructions and the ethical qualities that the assessment instructions commission the teachers to search for. Here descriptions of consequences and changes in perspectives are highlighted. Since the knowledge field is not very clearly staked out and familiar in a Swedish school context, the teachers are not acquainted with the qualities described in the instructions. They have not been taught these competences when they were pupils themselves; they have not been explicitly educated about them in teacher education. It would be interesting to compare this with the situation concerning, for example, the knowledge field of mathematics for instance. Vagueness in the school subject of Ethics is, not, however, a specifically Swedish phenomenon. The fuzzy character of the field has been a theme for research debate in itself. In moral education the object of the field, 'morality', is said to be unclear. It can be disputed what it means to be "good at" morality and thereby what a desirable progression in this area means (Wilson, 2000).

National tests and assessment criteria are often said to influence school practice since they define concretely what creditable knowledge and perspectives in an area are. This influence can be considered larger in a knowledge field with a fuzzy character. When the RE syllabus's rather general requirements concerning abilities to reason in a simple, developed or well-developed way, are interpreted into pupil-tasks and assessment instructions whose definitions are repeatedly used by teachers, a process of constituting the knowledge field of Ethics in school is clearly going on. In a situation where the knowledge field is

as vague as it is in Ethics, there is evidently a need for basic research concerning different conceptions of ethical competence. Such a variety of conceptions could be used for a critical discussion of competence in this field. It seems important to visualize perspectives from other countries and to show different perspectives within established ethical theory. But it also seems important to find ways to emphasize valuable qualities in pupils' existing perspectives and to highlight teachers' ideas regarding creditable knowledge in this field, as we have done in this study by focusing on challenges for teachers in assessing pupils' knowledge in Ethics.

Table 2. Core content and knowledge requirements of the Swedish curriculum in relation to parts of assessment instructions in the two National Tests in RE 2013.

Core content (ethics) year 6	Core content (ethics) year 9
Some ethical concepts such as right and wrong, equality and solidarity. Daily moral question concerning the identities, roles of girls and boys, and gender equality, sexuality, sexual orientation, and exclusion and violation of rights. Questions about what a good life can be, and what it may mean to do good.	Daily moral dilemmas. Analysis and argumentation based on ethical models, such as consequential and deontological ethics. Views of the good life and the good person are linked to different kinds of ethical reasoning, such as virtue ethics. Ethical questions and the view of people in some religions and other outlooks on life. Ethical concepts which can be linked to questions concerning sustainable development, human rights and democratic values, such as freedom and responsibility.

Knowledge requirements (ethics) year 6			Knowledge requirements (ethics) year 9		
Grade level E	Grade level C	Grade level A	Grade level E	Grade level C	Grade level A
Pupils can apply simple reasoning about everyday moral issues, and what it might mean to do good. Pupils make reflections which basically relate to the subject and use some ethical concepts in a basically functional way.	Pupils can apply developed reasoning about everyday moral issues, and what it might mean to do good. Pupils make reflections which carry the reasoning forward and use some ethical concepts in a relatively well functioning way.	Pupils can apply well developed reasoning about everyday moral issues, and what it might mean to do good. Pupils make reflections which carry the reasoning forward and deepen or broaden it and use some ethical concepts in a well functioning way.	Pupils can reason and argue about moral issues and values by applying simple and to some extent informed reasoning, and use ethical concepts and models in a basically functioning way.	Pupils can reason and argue about moral issues and values by applying developed and relatively well informed reasoning, and use ethical concepts and models in a relatively well functioning way.	Pupils can reason and argue about moral issues and values by applying well developed and well informed reasoning, and use ethical concepts and models in a well functioning way.

(Table 2 continues on next page)

(Table 2 cont.)

Instructions for assessment, item 3 (ethics), year 6			Instructions for assessment, item 12 (ethics), year 9		
The pupil discusses the situation by describing how K could act. The pupil uses at least two concepts in a basically functioning way by showing an acceptable understanding of their meaning. The pupil uses at least two concepts in a basically functioning way by showing an acceptable understanding of their meaning.	The pupil discusses the situation by observing possible consequences of K's actions. The discussion must contain a relevant description of consequences or one or more changes of perspectives. The pupil uses at least two concepts in a relatively well functioning way by partly showing that the concepts are relevant in the situation.	The pupil discusses the situation by observing both possible consequences of K's actions and how the situation can be perceived by different people. The discussion must contain a relevant description of consequences and one or more changes of perspectives. The pupil uses at least two concepts in a relatively well functioning way by consistently showing that the concepts are relevant in the situation.	The pupil holds a simple discussion of the concept of forgiveness in regard to at least one of the two aspects mentioned in the task: the active, "to forgive", and the passive, "to be forgiven", for example by clarifying the meaning of this aspect.	The pupil holds a developed discus-sion of the concept of forgiveness in regard to both as-pects mentioned in the task: the active, "to forgive", and the passive, "to be forgiven". The dis-cussion about one of the aspects is relati-vely elaborate con-cerning how the meaning of forgive-ness might be per-ceived. E.g., it may contain a developed comment about passive, "to be forgiven". The dis-cussion about one of the aspects is relati-vely elaborate con-cerning how the meaning of forgive-ness might be per-ceived. E.g., it may contain a developed comment about what the meaning of this aspect might be, or about what con-sequences an application of it may have concretely. The discussion about one of the as-pects may be more brief and expressed in a few words or consist of a naming of an example.	The pupil holds a developed discussion of the concept of forgiveness in regard to both aspects mentioned in the task: the active, "to forgive", and the passive, "to be forgiven". The discussion about both aspects is elaborate and points at how both aspects in a more exhaustive sense can be problematized and be related to one another concerning meaning and consequences.

References

Algotsson, Karl-Göran (1975). Från katekestvång till religionsfrihet: debatten om religionsundervisningen i skolan under 1900-talet. Diss. Uppsala. Uppsala Universitet.

Allal, Linda (2013). Teachers' professional judgement in assessment: A cognitive act and socially situated practice. *Assessment in Education: Principles, Policy & Practice, 20(1), 20-34.*

Almers, Ellen (2009). *Handlingskompetens för hållbar utveckling: tre berättelser om vägen dit.* Jönköping: Högskolan för lärande och kommunikation.

Bergling, Kurt (1982). Moralutveckling: bedömning av rätt och orätt från barndom till vuxen ålder. Stockholm: LiberFörlag.

Blaylock, Lat (2000). Issues in Achievement and Assessment in Religious Education in England: Which Way Should We Turn?, *British Journal of Religious Education, 23(1), 45-58*, DOI: 10.1080/0141620000230106

Brown, Margaret, McCallum, Bet, Taggart, Brenda & Gipps Caroline. (1997). The Validity of National Testing at Age 11: The teacher's view, *Assessment in Education: Principles, Policy & Practice, 4:2, 271-294*, DOI: 10.1080/0969594970040205

Charters, Elizabeth (2003). The Use of Think-aloud Methods in Qualitative Research. An introduction to Think-aloud Methods. *Brock Education, 12(2), 68-82.*

Connolly, Stephen, Klenowski, Val, Wyatt-Smith, Claire (2012). Moderation and consistency of teacher judgement: Teachers' views. *British Educational Research Journal, 38(4), 593-614.* Crisp, Victoria (2010). Judging the grade: Exploring the judgement processes involved in examination grading decisions. *Evaluation & Research in Education, 23(1), 19-35.*

Franck, Olof (2013a). Etisk litteracitet - om filosofiska förutsättningar för att utveckla ämnesdidaktiska perspektiv med hänsyn till kursplanen i religionskunskap (Lgr11). In Anders Marner & Hans Örtegren, (Red), *KLÄM. Konferenstexter om Lärande, Ämnesdidaktik och Mediebruk (*s. 178-197). *Umeå:* Umeå universitet.

Franck, Olof (2013b). Kunskapskrav och bedömning. In Olof Franck, Magnus Hermansson Adler & Inger Björneloo, *Samhällsämnenas didaktik F-åk 3* (s. 127-136). Lund: Studentlitteratur.

Gipps, Caroline, McCallum, Bet, Brown, Margaret (1996). Models of teacher assessment among primary school teachers in England. *The Curriculum Journal, 7(2), 167-183*. DOI: 10.1080/0958517960070204

Grant, Lynne, Matemba, Yonah H (2013). Problems of assessment in religious and moral education: The Scottish case. *Journal of Beliefs & Values: Studies in religion & Education, 34(1), 1-13.*

Jones, Tiffany Mary (2009). Framing the framework: discourses in Australia's national values education policy. *Educational Research For Policy & Practice*, 8(1), 35-57.

Klapp Lekholm, Alli (2010). Lärares betygsättningspraktik. *Forskning om undervisning och lärande, 3/2010, 20-29.*

Lindskog, Annika (2013). Nationellt bedömningsstöd i religionskunskap. I Falkevall, Björn. Att undervisa i religionskunskap - en ämnesdidaktisk introduktion. (s. 162- 169). Stockholm: Liber.

Lundgrén-Laine, Heljä, Salanterä, Sanna (2010). Think-Aloud Technique and Protocol Analysis in Clinical Decision-Making Research. *Qualitative Health Research, 20(4), 565-575.*

Løgstrup, Knud Ejler (1994). Det etiska kravet. Göteborg: Daidalos.

Osbeck, Christina (2004). Värdegrund i skolan. Uppdrag, elevperspektiv och didaktiska reflektioner. I Gunnel Colnerud & Solveig Hägglund (Red.), Etiska lärare – moraliska barn. Forskning kring värdefrågor i skolans praktik (s. 9-27). Linköping: Linköpings universitet, Institutionen för beteendevetenskap.

Osbeck, Christina (2013). Social Studies – A part of a Context Constructing Central RE Knowledge in Sweden. In Geir Skeie, Judith Everington, Ina ter Avest, Siebrem Miedema. *Exploring Context in Religious Education research,* (s. 147-169). Münster: Waxmann.

Olivestam, Carl-Eber (2012). Religionsdidaktik – om perspektiv, teori och praktik i religionsundervisning. Stockholm: Remus.

Radnor, Hilary A, Poulson, Louise, Turner-Bisset, Rosie.(1995). Assessment and teacher professionalism. *The Curriculum Journal, 6(3), 325-342.*

Rinne, Ilona. (2013). Att bedöma och bemöta. *Nordic Studies in Education, 33(3), 171-186.*

Sadler, Royce D (1985). The origins and functions of evaluative criteria. *Educational Theory, 35*(3), *285-297.*

Sadler, Royce D (1987) Specifying and Promulgating Achievement Standards, *Oxford Review of Education, 13:2, 191-209*, DOI: 10.1080/0305498870130207

Sadler, Royce D (2009) Indeterminacy in the use of preset criteria for assessment and grading, *Assessment & Evaluation in Higher Education, 34(2), 159-179*, DOI:10.1080/02602930801956059

Selander, Sven-Åke (1993). *Undervisa i religionskunskap.* Lund: Studentlitteratur.

Skolverket (2013a). Ämnesprov läsår 2012/2013, Bedömningsanvisningar, Religionskunskap, Årskurs 9. Stockholm: Skolverket. [WWW dokument]. URL http://www.sobedömning.se/npsoportal/laddaner.html (2014-12-18).

Skolverket (2013b). Ämnesprov läsår 2012/2013, Bedömningsanvisningar, Religionskunskap, Årskurs 6. Stockholm: Skolverket. [WWW dokument]. URL http://www.sobedömning.se/npsoportal/laddaner.html (2014-12-18).

Skolverket (2013c). Ämnesprov läsår 2012/2013, Delprov A, Årskurs 6. Stockholm: Skolverket. [WWW dokument]. URL http://www.sobedömning.se/npsoportal/laddaner.html (2014-12-18).

Skolverket (2013d). Ämnesprov läsår 2012/2013, Delprov A, Årskurs 9. Stockholm: Skolverket. [WWW dokument]. URL http://www.sobedömning.se/npsoportal/laddaner.html (2014-12-18).

Smaill, Esther (2013). Moderating New Zealand's National Standards: teacher learning and assessment outcomes, *Assessment in Education: Principles, Policy & Practice, 20(3), 250-265*, DOI: 10.1080/0969594X.2012.696241

Swedish National Agency for Education (2011). *Curriculum for the compulsory school, preschool class and the leisure-time centre.* Stockholm: Swedish National Agency for Education.

Taylor, Monica (1994). Overview of values education in 26 European countries. In Monica Taylor (Ed.), *Values education in Europe: A comparative overview of a survey of 26 countries in 1993* (pp. 1–66). Dundee: Scottish Consultative Council on the Curriculum.

Thornberg, Robert (2004). Grupprocesser och social påverkan: socialpsykologi med fokus på skolan. Linköpings universitet, Institutionen för beteendevetenskap, Vol. 237.

Tillson, John (2011). In Favour of Ethics Education, Against Religious Education. *Journal of Philosophy of Education*, *45*(4), 675-688.

Van der Schaaf, Marieke, Baartman, Liesbeth, Prins, Frans (2012). Exploring the role of assessment criteria during teachers' collaborative judgement processes of students' portfolios. *Assessment & Evaluation in Higher Education, 37*(7), 847-860.

Wilson, John (2000). Methodology and Moral Education. *Oxford Review of Education, 26*(2), 255-62.

Wyatt-Smith, Claire, Klenowski, Val (2013). Explicit, latent and meta-criteria: types of criteria at play in professional judgement practice. *Assessment in Education: Principles, Policy & Practice, 20*(1), 35-52.