Modelling Assessment of Mathematical Modelling – a Literature Review

Peter Frejd
Department of Mathematics, Linköping University

This paper presents a critical review of literature investigating assessment of mathematical modelling. Written tests as described in the reviewed papers draw on an atomistic view on modelling competencies while studies adopting more holistic approaches are rare. To assess the quality of students' work with mathematical models an elaborated view on the meaning of quality of mathematical models is needed.

Introduction

The role of mathematical modelling in the official Swedish upper secondary curriculum guidelines has increased during the last 50 years (Ärlebäck, 2009). In the present curriculum mathematical modelling is one of seven mathematical abilities to aim for, which is the ability to "interpret a realistic situation and design a mathematical model and to use and validate a model's properties and limitations" (Skolverket, 2010, p. 88, my translation). Modelling competence is assessed in the national course tests (NCT), which are developed in line with the curriculum. However, there are indications that modelling is difficult to assess. Frejd (2011) analysed test items in the NCT and concluded that there existed an uneven emphasis on different aspects of mathematical modelling. Frequently occurring aspects, such as to use an already existing model to calculate a result, were put in favour over other aspects that occurred sparsely or were left out, such as to critical asses conditions and validate results. Not a single item was found in this study that assessed all aspects of modelling (holistic view).

Looking at the result from Frejd (2011) and following the principle by Blomhøj and Hoff Kjeldsen (2006) that "the pedagogical idea behind identifying mathematical modelling competency as a specific competency is exactly to highlight the holistic aspect of modelling" (p. 166), raises some questions related to national assessment in Sweden regarding mathematical modelling. Is it possible to assess all aspects of modelling in the NCT? If yes how? If it is not possible to assess modelling in written tests in a holistic way, what other assessment modes are being used or suggested? The aim of this paper is to review a selection of literature focusing on mathematical modelling in mathematics education in order to analyse approaches used or suggested to assess students' modelling competence.

Mathematical modelling, modelling competence, and assessment

There is not one single unambiguous definition of the notion of mathematical modelling shared in mathematics education, but rather the definitions used and/or descriptions of modelling given depend on the theoretical perspective adopted (Frejd, 2011). In this paper I have chosen the approach taken by Blomhøj and Højgaard Jensen (2003) to describe mathematical modelling. Their description consists of the following six sub-process: *formulating a task in the domain of inquiry*; *selecting the relevant objects, relations and idealising*; *translation into a mathematical representation*; *using mathematics to solve the corresponding mathematical problem*; *making an interpretation of the results in the initial domain of inquiry*; and *evaluating the validity of the model* (p. 125). In addition, Blomhøj and Højgaard Jensen write that "[b]y mathematical modelling competence we mean being able to autonomously and insightfully carry through all aspects of a mathematical modelling process in a certain context" (p. 126), a definition I have adopted for this paper. The motivation for using this framework is that a similar view of modelling forms the basis of the construction of the guiding questions used in Frejd (2011) to examine NCT items.

While the term assessment is frequently used in educational research in mathematics education, Niss (1993) points out the complexity of assessment endeavours in mathematics education. Instead of declaring a specific definition of assessment the intention in this paper is to be open to investigate everything that is called assessment and relates to mathematical modelling.

Methodolodgy

According to Bryman (2004), central features of any qualitative literature review are first that relevant and adequate research literature related to the aim is identified, and secondly that the analysis of the literature serves the purpose(s) to relate, organise, and connect the literature (constructing intertextual coherence) as well as to point at research literature that seems to be incomplete (problematizing the situation). Therefore, the method used includes three phases: 1. Identifying adequate papers, 2. Categorising the papers, and 3. Extended analysis.

Phase 1, to identify relevant literature addressing modelling assessment from the vast amount of papers, articles, and books about mathematical modelling in mathematics education, is difficult. An overview or a state of the art in this research domain may be found in the *ICMI14 Study: Modelling and applications in mathematics education* (Blum et al., 2007) [1]. Other cited works relate to the proceedings from *the international conference on the teaching and learning of mathematical modelling and applications* (ICTMA), as well as the proceedings from the thematic working groups explicitly focusing on issues related to mathematical modelling in teh *Congresses of the European Society for Research in Mathematics Education (CERME)*. In this survey the ICMI14 Study, all 15

ICTMA proceedings, and all proceedings from the 'modelling working group' at CERME have been examined. In addition, the special issues of ZDM 2006 focusing on mathematical modelling (issues 38(2) and 38(3), respectively) and one other 'older' paper (Berry & O'Shea, 1982) have been analysed. The method used to identify the relevant articles in the chosen literature are based on key words in the titles (assessment, assessing, evaluating, etc), an examination of all abstracts in book sections relating to assessment, and a search for the word assessment in the index.

There are several alternatives to synthesise coherence between different works (phase 2) and to analyse the identified papers (phase 3). One alternative is to use grounded theory (Strauss & Corbin, 1998). A second alternative could be to design questions and argue for its relevance. A third alternative could be to use already designed questions used for a similar purpose.

Niss (1993 argues that some questions in relation to assessment of mathematical modelling are "highly relevant" (p. 50) to answer. Those concern issues about *why* assess modelling, *what* should be assessed using *which* kinds of *tasks*, *who* (individuals, groups, ...) should be assessed *when* and *how* (design, mode, outcomes, reports) and by *whom* (Niss, 1993, p. 49).

However, these questions, which seem to be derived from well articulated common sense arguments, are quite general and fundamental in all types of assessment. The choice for phase 2 (synthesis coherence) and 3 (further analysis) in this survey is therefore a combination of the alternatives discussed above. Phase 2 is based on a coding strategy inspired from grounded theory, in order to be open for a wide set of alternatives of categories. However, it is not possible to describe the actual coding process here, due to limited space, which is a limitation of this paper. Similarly, the presentation of phase 3 is also constrained by the limited space. The further analysis of the articles focuses on different *modes,* mainly in written tests, as a consequent of the initial concern as stated in the introduction. *Mode* in this paper refers to different types of assessment methods and relates to Niss' (1993) question *how*, but as *how* depends on *why* and *what,* these are also implicitly analysed.

Results and discussion

Underpinnings of assessment frameworks

To analyse approaches for assessing modelling, it is important to know about types of research studies to understand whether the criteria used in a framework or mode of assessment are derived from a theoretical analysis, based on literature, ad-hoc constructions, experience from assessment situations or empirical studies of students' work. 75 articles relating to modelling assessment were identified (out of total 707 papers from the ICTMA-, ICMI study14-, the CERME 'modelling group'-proceedings) and categorised as *empirical* (case studies), *philosophical* (based on argumentations), *theoretical* (building frameworks), *overviews* (literature reviews, introductions), *description of practice* (curriculum descriptions, course descriptions) and *competitions* (see figure 1).

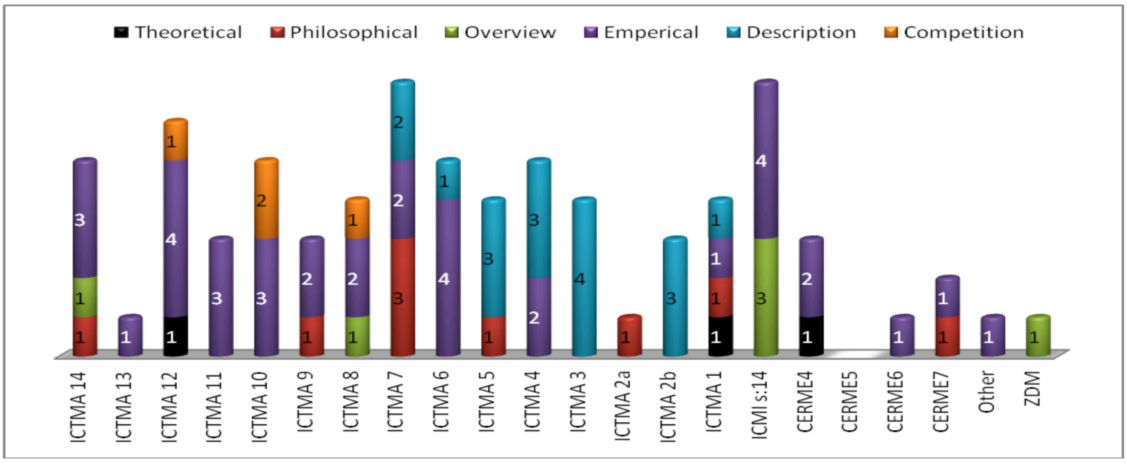

Figure 1. The frequency of the identified categories

The most frequent category in figure 1 is *empirical studies* (36 out of 75) and it includes a large variety of empirical studies, such as: developing and evaluating research tools assessing modelling (e.g. Zöttl, Ufer,& Reiss, 2011; Izard, 2007); analysis of test items in written tests (Frejd, 2011b; Turner, 2007; Stillman, 1998; Naylor, 1991); investigations of differences between teachers' marking of students work (Berry & O'Shea, 1982); poster assessments (Houston & Breedon, 1998; Houston, 1997; Wake, 2010); investigating how middle school students solve linear pictorial patterns (Amit & Neria, 2010); comparative study of modelling outcomes between two groups, where one group worked with the modelling process and the other group worked with just examining models (Legé, 2007). There are six literature *overviews* identified: an introduction to assessment chapters (Blomhøj, 2011; Galbraith, 2007); a review of assessment methods used (Houston, 2007); an overview of the expert and novice issue related to modelling (Haines & Crouch, 2007); overviews more in general about modelling, with parts related to assessment (e.g. Galbraith et al., 1998). Nine papers refer to more *philosophical* issues, such as why, whom, when, etc. (Burton, 1997; Niss 1993; Oke & Bajpai, 1986); descriptions and arguments on 'how to assess' based on examples and experience (e.g. Henn, 2011; Brown, 2001; Izard, 1997); and assessing mathematical models (Jablonka, 1997). There are three *theoretical* papers attempting to develop some type of framework for assessment and marking (Højgaard Jensen, 2007; Hall, 1984; Henning & Keune, 2006). Finally, there are four papers relating to modelling *competitions* like the China Undergraduate Mathematical Contest in Modelling (Jiang, Xie, & Ye, 2007) and 17 papers related to descriptions of practice, 'this is the way we work with assessment at our school' (e.g. Batteye & Challis, 1997; Swam 1991); or ‘description of the curriculum situation in for example northern Ireland/Australia (e.g. Coxhead, 2007; Money & Stephens, 1993).

To summarize, of the investigated proceedings every tenth paper relates to assessment, and out of these identified papers the most number of papers were categorised as *empirical* studies and the least number of papers as *theoretical* studies. This is m true for most research domains in mathematics education, where there is a need for a large number of empirical studies to explore, compare and evaluate complex issues as well as to underpin the development of theoretical frameworks. However, in this study, there seem to be ‘very’ few papers focusing on theoretical aspects on assessment (i.e. trying to create frameworks) and the development of these frameworks is not grounded in case studies. Notable is also the large proportion of identified papers (almost 30% of the investigated papers) relating to descriptions of practice and descriptions of competitions.

In the third phase of the analysis, there are several different *modes* identified in the sample. These are *written tests* (35 papers), *project reports* (21 papers), *project reports including oral presentations* (7 papers), *contests* (4 papers), *poster sessions* (3 papers), and *students’ portfolio* (2 papers). Twelve of the papers above involve more than one mode. The written tests analysed or used refer to multiple choice questions, shorter tasks, extended tasks, unseen tests (traditional tests) and seen tests, final exams, and shorter classroom tests, etc.

Written tests

Every third paper (11 out of 31 papers) relating to a written test is dealing with multiple choice questions which stem from Haines, Crouch and Davis (2000). This multiple choice test originally consisted of 12 questions, each with five alternatives, grouped in pairs to assess six aspects of the modelling process (i.e. to be used in pre- and post settings). The number of test items has been extended to a total of 22 items testing 8 aspects. The aspects that are tested relate to a modelling cycle and address phases such as making simplifying assumptions, formulating the problem, assigning variables, parameters, and constants, etc. The test, together with a partial credit assessment model (the scores 0, 1, 2), makes it “possible to obtain a snapshot of student’ [modelling] skills at key developmental stages without the student carrying out a complete modelling exercise” according to Haines et al. (2000, p. 10). This test instrument has been used in a variety of settings with different aims such as to investigate the levels of students’ modelling competencies (Frejd & Ärlebäck, 2011). However, even if this test instrument is widely used some critique has been raised. Frejd and Ärlebäck (2011) found in their study that only two of the pairs of items were comparable in the respective aspect of the modelling process. They also argue about the lack of ICT and collaborative work, which are other important aspects of modelling, though the main issue refers to the atomistic view. Haines and Crouch (2007) notice that "these items do not address the full range of modeling skills, for they do not, as yet, cover *solving mathematics*, *refining a model* and *reporting*" (p. 420).

Another written test to assess modelling competency is developed by Zöttl, Ufer and Reiss (2011). Their test consists of 12 (chosen from a total of 36) items divided in four categories, where three categories relate to different aspects of the modelling process and the fourth category to "short, but complete modelling tasks" (p. 432). They argue that "adequate modelling tasks should always require the performance of a complete modelling process" (p. 428). An example of a 'complete modelling task' is to "[e]stimate the total area of Spain by using the [given] map" (p. 432) where Portugal's total area is stated. To theoretically underpin the test Zöttl et al. (2011) use Højgaard Jensen's (2007) three dimensions of modelling competency; degree of coverage, radius of action and a technical level. Degree of coverage relates to which parts of the modelling process are to be used and to what extent the students perform autonomously and use reflections. The radius of action concerns the range of contexts in which a student may perform his/her modelling ability, and the technical level refers to how advanced the mathematics is which the student uses. Højgaard Jensen (2007) illustrates these levels by a geometrical box (i.e. as three independent vectors). One may question this illustration, because in case of modelling the three levels of modelling competency do not work independently, but are (as the author sees it) intimately interwoven (especially since one aspect of modelling is about 'pure' mathematics). Thus, if a student possesses a large tool box (good technical level) it will imply more mathematical options to solve a problem and create more opportunities for reflections (a better degree of coverage) than for a student that possesses a low technical level. Zöttl et al. (2011) only measure the degree of coverage and the technical level, since they let the radius of action be specified concerning geometry. To assess the items, they used a dichotomous scoring, because they find partial credits too complex. How they scored the 'complete modelling task' stated above is not described, but it appears to be related to PISA tasks, because it is similar to a PISA item discussed by Turner (2007). However, no references are given, and it can be questioned to what extent it is a complete modelling task: A clue is given in the text, "you can draw onto the map if it helps you" (p. 432), and a solution to the task seems straightforward (e.g to estimate the area of Spain by covering the map of Spain with copies of the map of Portugal).

The PISA framework (OECD, 2009) is a large scale activity for international students' assessment. The role of modelling in this framework is however problematic, according to Jablonka and Bergsten (2010), who call it "a circular construction" (p. 30): On the one hand mathematising, which is the primary building block of mathematical literacy, is described as a modelling process, while on the other hand, this primary building block consists of eight competencies, where one of these competencies is modelling competence and equal to mathematising. There are some papers about PISA items in the identified sample, and according to Turner (2007) the items do promote an interest to modelling but have a low level of complexity of modelling activity. Henning and Keune (2006) have reformulated PISA items to assess three levels of modelling competence. The three hierarchic levels, which Henning and Keune (2006) have developed by adopting “the competence levels of mathematical literacy” (p. 1667) are: 1. Recognize and understand modelling (to describe the modelling process); 2. Independent modelling (to solve a modelling problem and interpret the result); and 3. Meta-reflection on modelling (to critically analyze and reflect upon the modelling process). Henning & Keune seem to think that meta-reflection is an “extra” activity, which is not needed when someone makes a model. They state “[a]t this third level of competence, the overall concept of modelling is well understood” (p. 1669), but at the same time “[a]t this level, it is not absolutely necessary to have previously solved problems by means of modelling techniques” (p. 1669). This statement is contradictory to other research like Blomhøj and Hoff Kjeldsen (2006) who claim that “[m]odelling competency is developed through the practice of modelling” (p. 166)

Vos (2007) used the alternative practical assessment tasks (hands-on tasks) developed by TIMSS in a case study. The test was composed by several tasks where the students instead of solving realistic problems described in words were given concrete equipment such as rubber bands to work within a laboratory-like situation. It was found that both students and teachers had a positive attitude towards the test, but the coding of the open-ended questions was problematic.

Projects

It is often argued that using written tests is not the best way to test modelling. According to Berry and Le Masurier (1984, p. 59) “*project* is the ideal method”, and Niss (1993) claims that the use of traditional modes (i.e. written test) is difficult or even impossible, but that “[o]ne particular appropriate type [of assessment] is *projects*” (p. 47). There are also some evidence, in relation to the assumptions above, presented by Antonius (2007) who argues that ”the different competences seem to be more visible in project examination” (p. 414) than in traditional examination because it is more extensive (includes both written reports and oral examinations).

Descriptions of projects as part of assessment are quite frequently appearing among the identified papers. The projects are extended ‘realistic’ problems that students try to solve during a longer time period (like 40 hours in the case of Berry and O’Shea, 1982). The students explain their solutions in written reports and in some cases (in 7 papers) defend them in an oral presentation. The main question regarding projects is how they can be assessed. Approaches to assess projects found in this sample are impression marking (Gillespie, Binns &

Burkhardt, 1989), formal marking schemes (e.g. Berry & O'Shea 1982), grounded theory (Maaß, 2007), and observations (Dunne & Galbraith, 2003; Herring, 1991). The most frequent approach is the use of more or less formal marking schemes (i.e. a set of defined criteria to be followed). However, none of the marking schemes used in the papers are justified in the sense of describing why just these criteria are being used and no others. The fact that most criteria are quite generally written in order to cope with many different projects makes it difficult to apply them and accounts for considerable variation between markers (e.g. Berry & O'Shea 1982; Haines, 1991). Hall (1984) argues for double blind marking to increase reliability together with his framework to calculate the final outcome (i.e. geometric mean value), but according to Haines' (1991) case study Hall's 'geometric model' does not differ much from traditional marking.

A project may also include a poster session, which is discussed in three papers which also relate to peer-assessment. Wake (2010) does a case study focusing on formative assessment and poster presentation. He argues that peer-assessment provides feedback in a language that the students understand and that learning is most effective if students are aware of the objectives to be learned. A conclusion drawn is that the use of a modelling approach on teaching and learning with formative assessment in day to day practice changes the teachers' roles, so that "both [teacher and student] are now active with learners struggling to solve a task and make reflective judgments about their ability to do so using new rules of assessment that focus on process as opposed to outcomes" (p. 2093).

Concluding remarks

What is found in this study clearly relates to the complexity of any assessment endeavor (Niss, 1993a), illustrated by Izard's (1997) statement that "[n]o single assessment method is capable of providing evidence about the full range of achievement" (p. 109). The written tests as described in the reviewed papers draw on an atomistic view of assessment focusing more on the product than on the process. A confirmation of Frejd's (2011) findings of a lack of a holistic approach (in the NCT) was also found in other countries (Stillman, 1998; Naylor, 1991). The question still remains if it is possible to construct justifiable holistic approaches to assess modelling in national tests.

According to Jablonka (1997), the most crucial aspect to assess in students' work with modelling is to judge the quality of a mathematical model. The framework developed in Jablonka (1996; summarised on pp. 209-212) for analysing mathematical models is also intended to be used for assessing students' work with mathematical models. She lists a number of critical questions to ask to the modelling work, organised under the two main headings evaluation of efficiency ("To what extent does the model fulfill its main goal?") and assessment of usefulness ("What is the contribution to the solution of the main problem and how can the goals and consequences be evaluated?”). While the efficiency questions can be used to form the basis for developing a set of assessment criteria for students' models, assessment of students' work with mathematical modelling also needs to take into account the extent to which the students have considered the usefulness questions while studying or developing mathematical models.

When it comes to assess within a holistic approach the use of projects is suggested (e.g. Niss, 1993). Here assessment criteria could be guided by the critical questions suggested by Jablonka (1996). Recent developments draw on projects (Antonius, 2007), formative assessment (Wake, 2010), and alternative assessment (Vos, 2007). However, while in “pure” mathematics one usually has some common mathematical ground forming the base for classifying and assessing students’ work, without shared views on how to judge the quality of a mathematical model one cannot expect shared views in a debate about assessing students’ mathematical modelling.

Note

1. Due to the large number of references, papers published in ICTMA-proceedings and the ICMI-14 Study proceedings are listed in the appendix.

References

Blomhoj, M., & Jensen, T. H. (2003). Developing mathematical modelling competence: conceptual clarification and educational planning. *Teaching Mathematics and its Applications, 22*(3), 123-139.

Blomhøj, M., & Hoff Kjeldsen, T. (2006). Teaching mathematical modeling through project work. *ZDM, 38*(2), 163-177.

Blum, W., Galbraith, P. L., Henn, H., & Niss, M. (Eds.). (2007). *Modelling and applications in mathematics education. The 14th ICMI study*. New York: Springer.

Bryman, A. (2004). *Social research methods* (2. ed.). Oxford: Oxford University Press.

Frejd, P. (2011). *Mathematical modelling in upper secondary school in Sweden an exploratory study*. Licentiate thesis. Linköping: Linköpings universitet.

Haines, C., Crouch, R., & Davis, J. (2000). *Mathematical modelling skills: A research instrument*. Technical Report No. 55. Univ. of Hertfordshire: Dpt. of Mathematics.

Henning, H., & Keune, M. (2006). Levels of modelling competence. In M. Bosch (Ed.), *Proceedings of CERME4* (pp. 1666-1673). Barcelon: Universitat Ramon-Llull.

Jablonka, E. (1996). *Meta-Analyse von Zugängen zur mathematischen Modellbildung und Konsequenzen für den Unterricht*. Dissertation. Berlin: Transparent-Verl.

Jablonka, E., & Bergsten, C. (2010). Theorising in mathematics education research: differences in modes and quality. *Nordic Studies in Mathematics Education*, *15*(1), 25-52.

Niss, M. (1993). Assessment in mathematics education and its effects: An introduction. In M. Niss. (Ed.), *Investigations into assessment in mathematics education: An ICMI Study* (pp. 1-30). Dordrecht: Kluwer.

OECD. (2009). *Learning mathematics for life: a perspective from Pisa*. Paris: Author.

Skolverket. (2010). *Förordning om ämnesplaner för de gymnasiegemensamma ämnena*. http://www.skolverket.se/content/1/c6/02/39/50/Gymgemensamma.pdf

Strauss, A. L., & Corbin, J. M. (1998). *Basics of qualitative research: techniques and procedures for developing grounded theory (2. ed.).* Thousand Oaks, Calif.: SAGE.

Wake, G. (2010). Modelling and formative assessment pedagogies mediating change in actions of teachers and learners in mathematics classrooms. In V. Durand-Guerrier, S. Soury-Lavergne, & F. Arzarello (Eds.), *Proceedings of CERME 6* (pp. 2086-2095). Université Claude Bernard: France.

Ärlebäck, J. B. (2009). *Mathematical modelling in upper secondary mathematics education in Sweden. A curricula and design study.* Dissertation. Linköping: Linköpings universitet.

Appendix

Amit, M., & Neria, D.	**(2010)**	Assessing a modelling process of a linear pattern task.
Antonius, S.	**(2007)**	Modelling based project examination.
Battey, A., & Challis, M.	**(1997)**	Deriving learning outcomes for mathematical modeling units within an undergraduate program.
Berry, J., & O'Shea. T.	**(1982)**	Assessing mathematical modelling.
Berry, J., & Le Masurier, D.	**(1984)**	O.U. Students do it by themselves.
Blomhøj, M.	**(2011)**	Modelling competency: teaching, learning and assessing competencies-Overview.
Brown, R.	**(2001)**	Formulating and assessing a mathematical modelling problem in a technological environment
Burton, L.	**(1997)**	The assessment factor - by whom for whom, when and why.
Coxhead, C.	**(1997)**	Curriculum development and assessment in north Ireland.
Dunne, T., & Galbraith, P.	**(2003)**	Mathematical modelling as pedagogy-Impact of an Immersion program.
Frejd, P., & Ärlebäck, J. B.	**(2011)**	First results from a study investigating Swedish upper secondary students' mathematical modelling competencies.
Galbraith, P.	**(2007)**	Assessments and evaluation-Overview.
Galbraith, P., Haines, C., & Izard, J.	**(1998)**	How do students' attitudes to mathematics influence the modelling activity?
Gillespie, J., Binns, B., & Burkhardt, H.	**(1989)**	Assessment of mathematical modeling.
Haines, C.	**(1991)**	Project assessment for mathematicians.
Haines, C., & Crouch, R.	**(2007)**	Mathematical and applications: Ability and competence frameworks.
Hall, G.G.	**(1984)**	The assessment of modelling projects.
Henn, H.W.	**(2011)**	Why cats happen to fall from the sky or on good and bad models.
Herring, M.J.	**(1991)**	The use of mathematical modelling in a program of integrative assignments.
Højgaard Jensen, T.	**(2007).**	Assessing mathematical modelling competency.
Houston, K.	**(1997)**	Evaluating rating scales for the assessment of posters.
Houston, K.	**(2007)**	Assessing the "phases" of mathematical modeling.
Houston, K., & Breedon, B.	**(2003)**	Secondary pupils creating posters.
Izard, J.	**(2007)**	Assessing progress in mathematical modelling.
Jablonka, E.	**(1997)**	What makes a model effective and useful (or not)?
Jiang, Q., Xie, J., & Ye, Q.	**(2007)**	An introduction to CUMCM.
Lege', J.	**(2007)**	To model or let them model? That is the question?
Maaß, K.	**(2007)**	Modelling in class: What do we want the students to learn?
Money, R., & Stephens, M	**(1993)**	Linking applications, modelling and assessment.
Naylor, T.	**(1991)**	Assessment of a modelling and applications teaching module.
Niss, M.	**(1993)**	Assessment of mathematical applications and modeling in mathematics teaching.
Oke, K.H., & Bajpai, A.C.	**(1986)**	Assessment in Mathematical Modelling
Stillman, G	**(1998)**	The emperor's new clothes? Teaching and assessment of mathematical applications at the senior secondary level.
Swam, M.	**(1991)**	Mathematical modeling for all abilities,
Turner, R.	**(2007)**	Modelling and applications in PISA
Vos, P.	**(2007)**	Assessments of applied mathematics and modelling: using a laboratory like environment.
Zöttl, L. Ufer, S. & Reiss, K.	**(2011)**	Assessing modelling competencies using a multidimensional IRT approach.
The ref. to all ICTMA proceedings and the ICMI- study can be found at http://www.ictma.net/literature.html		