**A Rubric Development Study for the Assessment of Modelling Skills**

**Abstract**

The purpose of this study is to develop a rubric named “A Rubric for Assessment of the Modelling Skills (RAMS)” for assessing cognitive modelling skills. The dimensions of the RAMS were: understanding the problem, simplifying, mathematizing, working mathematically, interpreting and validating. Twenty three 6th grade and twenty 11th grade students took part in the study and their solutions on the modelling tasks were examined. During the ten weeks of data collection process, the participants' solutions were analysed and the dimensions were constructed. RAMS offers a detailed assessment and scoring for different solutions to be arose in any modelling task solution.

**Key words:** analytic scoring rubric, cognitive modelling skills, skill assessment, mathematical modelling.

**Introduction**

Nowadays, it is not enough for students to be successful only in lessons; it is important for students to be ready for the real world and to cope with the problems that they can encounter. Mathematical modelling, which provide solutions to the real life problems by using mathematical methods, are integrated into the lessons. It is stated that individuals have to pass through a modelling process for constructing model/s which enable them to elicit a relation between real life and mathematics (Kaiser, Blomhoj & Sriraman, 2006). Kaiser (2007) stated that the development of students’ modelling competencies will be achieved through the integration of modelling into lessons. Similarly, Authors (2013-a) pointed out that students have to use their modelling competencies to be successful in modelling. Depending on this, it is necessary to discuss the meaning of the modelling competency required for students to be able to model real life situations.

Modelling competencies are defined as the skills and abilities which are necessary for complementing the modelling process purposively and properly, and in this process the individual should be willing (Kaiser & Maaß, 2007; Kaiser & Schwarz, 2006; Maaß, 2006). In other definition, the modelling competencies are expressed as the ability to conduct the modelling process in an independent way (Maaβ & Gurlitt, 2011). In the literature, the different words such as competency/competence, skill and ability were used to explain the student approaches on modelling process. We preferred to use the modelling skill term to explain students' solution processes in this study.

The researchers handled different modelling processes in defining modelling skills (Blomhoj & Højgaard Jensen, 2003; Blomhoj & Kjeldsen, 2006; Borromeo Ferri, 2006; Cabassut, 2010; Henning & Keune, 2007; Houston, 2007; Maaß, 2006; Stillman, Brown & Galbraith, 2010). The differences were caused by the researchers’ perspective about understanding and interpreting of modelling and sometimes by the structure of modelling tasks (Borromeo Ferri, 2006). The content of the modelling process is directly related to the cognitive modelling skills (Maaβ, 2006). The cognitive modelling skills require the conscious direction of modelling process and the approaches performing in steps of the modelling process (Blomhoj & Kjeldsen, 2006). In our study, we choose the Modelling Cycle under a Cognitive Perspective (Borromeo Ferri, 2006) as theoretical framework. The cognitive modelling skills consisted of understanding the problem, simplifying, mathematizing, working mathematically, interpreting and validating. Blum (2011) reveals the importance of the modelling skills in teaching by stating that one of the main aims of mathematics teaching is to develop the students’ modelling skills. The development of the cognitive modelling skills was carried out with modelling applications (Blomhoj & Kjeldsen, 2006) and Blomhoj and Højgaard Jensen (2003) proposed two approaches such as holistic and atomistic to provide this development. Students work throughout all stages of modelling process in holistic approach and in atomistic approach they work at certain stages of the modelling process (Blomhoj & Højgaard Jensen, 2003). In order to reveal the development of students’ cognitive modelling skills, the question of how they can be assessed comes to the forefront.

**The Assessment of the Cognitive Modelling Skills**

While assessing the cognitive modelling skills, different tests and tasks have been used. One of these tests, Galbraith and Haines (1997) used the questions in parallel with the modelling process. The test questions were arranged in multiple-choice format atomistically. It was tried to assess the skills of making assumptions to construct mathematical model in the first questions, revealing the purpose of real model and constructing mathematical model in the second one, and so on. In scoring, the choice with the most appropriate response was assessed with 2 points; the choice with the appropriate response to some degree was assessed with 1 point and the choice with the inappropriate response was assessed with 0 point. In another study, Maaβ (2006) made use of the modelling test to reveal the cognitive modelling skills of 7th graders according to the atomistic and holistic approaches. The evaluation of the test was carried out as the correct answer of each problem would be 1 point. Similarly, Maaβ and Mischo (2011) developed a test to determine students’ modelling skills within the scope of STRATUM project. A total of six questions were asked in accordance with holistic and atomistic approach and five choices were arranged for each questions, and it was tried to make it possible for students to select more than one choices. The open-ended problems were assessed with 0, 1, 2 points according to the modelling process and each point respectively had the meanings of “the response is not understandable”, “the response is partially correct/understandable, “the response is understandable and is explained with its reasons”. In another study, Biccard and Wessels (2011) carried out model eliciting activities for 7th graders. They evaluated the progress of a particular skill between 0-3 points according to its status in the previous implementation. Grünewald (2012) handled the modelling skills in three groups as simplifying-structuring-mathematizing, working mathematically and interpreting-validating. She evaluated these groups with open-ended and multiple-choice questions.

In literature to assess students’ cognitive modelling skills by tasks, the usage of rubrics was noticed. For assessing whole modelling process, Berry and O’Shea (1982) developed a marking scheme which contains the main sections named “abstract, formulation, initial model, data, revisions to the model, conclusions, presentation”, and their contents and marks. Galbraith and Clatworthy (1990) described a rubric whose dimensions were “specify the problem clearly”, “formulate an appropriate model”, “solve the mathematical problem including interpretation, validation, and evaluation/refinement” and “communicate results in a written and oral form” at three different levels. For evaluating the modelling projects, Keck (1996) developed a rubric with seven dimensions named identification of problem, formulation of assumptions, construction of model, solution and interpretation, validation, communication and mathematics. Each dimension was defined with different number of levels. For instance, the formulation of assumptions was evaluated at four levels. Level 0 was coded when the assumptions were not explained. In this rubric the social aspects were also be evaluated besides the cognitive skills. Chan, Ng, Widjaja and Seto (2012) defined three levels about the dimensions named assumptions, interpretation of task using real-world knowledge, and mathematical reasoning and computation in their rubric. For example, in the assumptions dimension, Level 1 referred no/incorrect assumptions made. Level 2 referred at least two assumptions stated which are relevant to the model and explained based on the real-world interpretations of task. Level 3 meant comprehensive list of assumptions made which were relevant to the model and explained based on the real world interpretations of task.

To assess the modelling skills, Leong (2012) composed the dimensions and levels depending on the modelling cycle of Common Core State Standards for Mathematics, but he did not defined what the levels mean. A scoring between 0 and 4 was carried out in his Modelling Cycle Scoring Rubric. The rubric contained the scores and weights but there were no detail explanation about them. To score the modelling process, Anhalt and Cortez (2015) determined categories named the constructed model, the solution, and students’ reflections in constructing mathematical models. They used a rubric named “A General Rubric for Assessment of Modelling Tasks” to score the categories. The categories handled as explanations, connections, working, reasoning, concepts and calculations and the six levels ranging between 0-5 points were used to score these categories.

When the literature was examined, two approaches prominently emerged. The multiple choice modelling tests which is one of them may not give realistic information to us. Because modelling requires a process, dealing with the steps of the process in a separate way from one another may lead us to make incorrect decisions. Especially the open-ended structure of the modelling tasks may reveal more than one model depending on the assumptions. Therefore, a model constructed by the student may not be included in the items of the multiple-choice test. In this case, the student cannot make any choice and then misunderstanding could be occurred such as s/he could not have model construction skill, although s/he made proper approaches in model construction indeed. Similarly, when the student selects the correct choice by chance, we may think as if s/he has the model construction skill. Also, the assessment of students’ cognitive skills in a holistic way will give better insights about what is in their mind in fact. Likewise, Büchter and Leuders (2005) emphasized the possibility that the student’s cognitive process may be quite different from what is needed to reach the correct solution when a choice is made from the existing options. We have adopted that modelling skills should assess through modelling tasks. We have focused on the development of a rubric to assess students’ solutions while working on the modelling tasks. In this context, the purpose of this study is to develop an analytic scoring rubric for assessing students’ cognitive modelling skills. We realized that the level definitions did not include in detail in the rubrics included in literature. We set out from the necessity that the rubric to be developed should be comprehensive, should be understood in same way by everyone and its levels should be clearly defined. Besides, the dimensions of some rubrics include one or more skills instead of including each of them separately. In evaluating the modellers’ solutions, indecisiveness may appear about what dimension and level those that have been performed will be dealt with. In addition to this, Vos (2013) stated that there was not any scoring guidelines that will enable different people to make similar assessments in evaluating the modelling skills. So, an analytic scoring rubric is developed in this study so as primarily to fill in the gap in the literature and to propose a detailed assessment instrument that both researchers and teachers can use in the assessment of cognitive modelling skills. With the RAMS, it is possible to determine students' cognitive modelling skills and to what extent they have these skills in the solution of modelling tasks.

**Method**

While developing RAMS, twenty three 6th grade and twenty 11th grade students were included. By composing groups with four/five students, their solution approaches in solving the modelling tasks were examined. They presented their solutions to whole class. The solution and presentation processes were recorded by a video camera. The data were collected from the transcripts of the video records, the groups' solution papers and the researchers' observation notes. The process of data collection took 10 weeks. The implementations were performed first in a middle school then in a high school with one week intervals. These implementations were realized during two lesson hours in the selective course named Mathematics Applications in the middle school and in regular mathematics course in the high school. The groups solved the given task in the first lesson and in the second lesson they presented their solutions and in class-discussions were carried out. In discussions, the students were requested to assess the solutions of other groups and scaffoldings were provided by the researchers so as to enable the students to study in cognitive modelling skills. The modelling tasks implemented weekly were given in Table 1.

Table 1

*The Modelling Tasks*

|  |  |  |
| --- | --- | --- |
| Preparation Phase | 1st Week | Middle School: Step Problem (Author, 2011) |
| 2nd Week | High School: Step Problem (Author, 2011) |
| 3rd Week | Middle School: Bed Problem (Borromeo Ferri, 2014) |
| 4th Week | High School: Bed Problem (Borromeo Ferri, 2014) |
| 1st Phase | 5th Week | Middle School: Apple Pie Problem (adapted from Schukajlow, Leiss, Pekrun, Blum, Müller & Messner, 2012) |
| 6th Week | High School: Obesity Problem (Author, 2013-b) |
| 7th Week | Middle School: Invoice Problem (Author, 2015) |
| 2nd Phase | 8th Week | High School: Ancient Theatre Problem (Author, 2010) |
| 9th Week | Middle School: Time in School Problem (Maaß & Mischo, 2011) |
| 10th Week | High School: Fuel Problem (Author, 2012) |

**The Role of the Researchers**

The researchers conducted the study depending on their eight year experiences and implementations in mathematical modelling at elementary, secondary and undergraduate levels. The researchers also developed several rubrics peculiar to the modelling tasks applied within the modelling implementations to evaluate students’ solutions (Author, 2014; 2017). These studies formed a basis for what we could systematically do to develop a more comprehensive rubric. The teacher of class and we conducted together implementations throughout 10 weeks. Before the implementation started, we gave information about the content of the study by working together with the middle and high school teachers. Then we observed some lessons to obtain information about the students. Besides, we made video records in lessons and tried to enable the students to get used to the implementation and to feel themselves relaxed before the camera. While the groups were solving the modelling tasks during the implementations, we took observation notes and gave scaffoldings to reveal the students’ cognitive modelling skills. We asked probe questions to the students to support discussion about their solutions.

**RAMS Development Process**

**Preparation Stage**

We determined Borromeo Ferri’s (2006) framework as theoretical framework. The most important reason for deciding on this frame is the fact that it cognitively dealt with the student performances in modelling and it had a definition for each skill in solution. We carried out a comprehensive research about how the cognitive modelling skills were assessed. We observed that there were different rubrics for assessing cognitive modelling skills. We determined their advantageous and disadvantageous aspects and considered them in the development of the RAMS. Depending on them, we enabled each modelling skill to be assessed independent of one another and performed assessments to be as detailed as possible. Firstly, a rubric format was determined (Table 2). In this format, we selected one level for each skill and gave example from students’ solutions which was evident for our decision. Also we wrote our evaluations to compare with each other.

Table 2

*The First Rubric Format*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Skills | Level 1 | Level 2 | Level 3 | Samples from students’ work | Evaluation by the researcher |
| Understanding the problem |  |  |  |  |  |
| Simplifying |  |  |  |  |  |
| Mathematizing |  |  |  |  |  |
| Working mathematically |  |  |  |  |  |
| Interpreting |  |  |  |  |  |
| Validating |  |  |  |  |  |

There were no definitions for levels. The reason for this was the requirement for the definitions to be formed concerning each skill depending on the opinions obtained after students’ solutions were examined. After the first rubric format was arranged, two modelling tasks were applied. The Step Problem was solved with the 6th graders. After the implementation, the transcriptions were prepared. We separately examined the solution papers and the observation notes, and evaluated each group by using the rubric format. Then other modelling tasks were performed. After each implementation, the assessment process performed in the first implementation was repeated. As a result, we decided that three levels were insufficient. For instance, when the solutions of the middle school students were dealt with in simplifying, we determined that not making assumptions referred to the first level, making assumptions partially appropriate for real life referred to the second level and making assumptions completely appropriate for real life referred to the third level (see Table 3).

Table 3

*An Example for Simplifying from the Middle School Students’ Solution to the Step Problem*

|  |  |
| --- | --- |
| Level 2 | Evaluation by the researcher |
|  | Because the group only benefitted from the measurements and did not consider average human values, they made an assumption partially appropriate for the real life. |

Unlike the middle school students, in the implementation of Step Problem in the high school students, the students considered the variables that could be needed for constructing a model as well as formulating the assumption (see Table 3).

Table 3

*An Example for Simplifying from the High School Students’ Solution to the Step Problem*

|  |  |
| --- | --- |
| Level 2 | Evaluation by the researcher |
|  | They made a partially appropriate assumption as to the fact that human length could be twice as much as the leg length. Depending on the assumption at issue, they considered the angle between the legs. In order to find the distance between the steps, they determined variables depending on their assumptions and constructed a model depending on these variables. |

In the Bed Problem implementations, similar findings were obtained from the students’ solutions about simplifying (see Table 4). The groups who did not consider real life tackled own lengths instead of their parents’ lengths in formulating their assumptions. The groups who made assumptions completely appropriate for real life solved the problem by putting forward their opinions about the lengths of an adult man and woman. In this implementation similar to the Step Problem, it was observed that the high school students determined the variables that could be needed for constructing a model as well as expressing their assumptions verbally.

Table 4

*Examples for Simplifying from the Middle and High School Students’ Solution to the Bed Problem*

|  |  |
| --- | --- |
| Level 3 | Evaluation by the researcher |
|  | (Middle School) The students made assumptions appropriate for real life about the length of parents and their widths. Depending on these assumptions, the students, who made guess about the distances of persons to the sides of the bed, counted the distance between mother and father. |
|  | (High School) The students made assumptions appropriate for real life about the lengths of their parents and their widths. Depending on the so-called assumptions and the angle between the legs, they identified the variables. |

At the end of the four week, the levels of the rubric were decided to be increased and the format was revised. Even if the modelling tasks and student levels differed, the students' approaches were seen to be fixed and the definitions of levels were made by the researchers (see Table 5).

Table 5

*The Rubric Format Where There Are Four Levels about Each Dimension*

|  |  |  |
| --- | --- | --- |
|  | Levels | Definition |
| Understanding the Problem | Level 1 | Includes the expressions showing that s/he did not understand the problem, not determining the givens and goals, and not forming or mistakenly forming a relationship between them. |
| Level 2 | Includes the expressions showing that s/he understood the problem to some extent, determining the givens and goals to some extent but not forming or mistakenly forming a relationship between them. |
| Level 3 | Includes the expressions showing that s/he understood the problem completely, determining the givens and goals but not forming or mistakenly forming a relationship between them. |
| Level 4 | Includes the expressions showing that s/he understood the problem completely, determining the givens and goals, and forming a relationship between them. |
| Simplifying | Level 1 | Not simplifying the problem, not determining the necessary/unnecessary variables and making wrong assumptions. |
| Level 2 | Simplifying the problem to some extent, determining the necessary/unnecessary variables to some extent but making wrong assumptions. |
| Level 3 | Simplifying the problem, determining the necessary/unnecessary variables and making partly-acceptable assumptions |
| Level 4 | Simplifying the problem, determining the necessary/unnecessary variables and making realistic assumptions. |
| Mathematizing | Level 1 | Not constructing or mistakenly constructing mathematical model/s. |
| Level 2 | Constructing correct mathematical model/s based on partly-acceptable assumptions. |
| Level 3 | Constructing incomplete/wrong mathematical model/s based on realistic assumptions and relating them to one another. |
| Level 4 | Correctly constructing the needed mathematical model/s according to realistic assumptions, explaining model/s and relating them to one another. |
| Working Mathematically | Level 1 | Not presenting a mathematical solution, solving the constructed models wrongly or trying to solve the wrong mathematical model. |
| Level 2 | Solving correctly the mathematical models constructed incompletely/wrongly. |
| Level 3 | Including deficiencies/mistakes in the solution of the correctly constructed mathematical models. |
| Level 4 | Achieving correct mathematical solution by solving the correctly constructed mathematical models. |
| Interpreting | Level 1 | Misinterpreting or not interpreting the obtained mathematical solution in real life context. |
| Level 2 | Correctly interpreting the erroneous/incomplete mathematical solution in real life context. |
| Level 3 | Incompletely interpreting the obtained correct mathematical solution in real life context. |
| Level 4 | Correctly interpreting the obtained correct mathematical solution in real life context. |
| Validating | Level 1 | Not validating or making invalid validation. |
| Level 2 | Validating completely, not correcting the determined mistakes. |
| Level 3 | Validating completely, correcting the determined mistakes to some extent. |
| Level 4 | Validating completely, correcting the determined mistakes. |

The students both in the middle and high school had difficulty in interpreting and validating. Even if the students interpret the obtained solution, this interpretation was incomplete or they handled the validation only as controlling the calculation. In the next implementations, it was decided to use different modelling tasks in the middle and high school levels for observing different solutions' approaches and varying the levels of rubrics.

**The First Stage**

In the first stage of the RAMS development process, two different modelling tasks were implemented at the middle and high school levels. After them, the dimensions of mathematizing and working mathematically were revised included in five levels (see Table 9). The reason for this, when the students made partly-acceptable assumptions in simplifying, they constructed incomplete/wrong mathematical models as well as correctly constructing mathematical models. An excerpt of the Apple Pie Problem’s implementation was given in Table 6.

Table 6

*Example for Mathematizing from the Middle School Students’ Solution to the Apple Pie Problem*

|  |  |
| --- | --- |
| Sample from the solution paper | Evaluation by the researcher |
| \*ESHOT is the transportation company in Izmir. | The students made a partially-acceptable assumption about the fact that the ESHOT fee is 1 TL (ESHOT fee is 0.90 TL). It was determined that they constructed incomplete mathematical models because they had forgotten to include the so-called fee in constructing their models. So, they constructed an incomplete mathematical model based on a partially-acceptable assumption. |

After the level definitions in mathematizing were formulated, it was understood that a similar situation was also valid for the skill of working mathematically. The second level of solving mathematical models was revised included in two level: solving the models incompletely and solving them accurately. Thus, five levels were defined for the skill of working mathematically (see Table 9).

During the first stage, the four different modelling tasks were implemented. After implementations, the groups were enabled to present their solutions and their solutions approaches were discussed. We decided to examine whether the students should display deeper approaches in interpretation and validation because they gained experience about modelling and because of in class discussions. In discussions, the validation approaches were handled by emphsazing the validation was not only meant controlling the calculations and correcting mistakes but also validating the assumptions, models and the process. Thus, the additional four modelling tasks were decided to be implemented so as to determine whether the levels about the interpretation and validation skills should be changed.

**The Second Stage**

When the Ancient Theatre Problem was implemented, it was observed that one group interpreted the problem in real world despite of their incorrect results. The students found the height of the Aspendos Ancient Theatre of 15 m., which height is 36 m in reality. They discussed the results in terms of their real life experiences and thought that the height of one step could be more (see Table 7).

Table 7

*An Example for Interpreting from the High School Students’ Solution to the Ancient Theatre Problem*

|  |  |
| --- | --- |
| A section from the students’ expressions | Evaluation by the researcher |
| Oğuz: The height of a step is 40 cm.  Seval: 55. The height of a step is 55 cm.  Volkan: Just here of a step is 60cm (showing the height of one step).  Oğuz: Isn’t 60 cm so high?  Veli: Steps in an ancient theatre are a little bit higher (means that he had been to an ancient theatre before).  Volkan: A step is 60 cm height. | While the students were questioning if their solution was logical or not in the real life, they discussed over the assumptions as to the height of a step by referring back to the solution process. In this context, they made interpretations that the steps in ancient theatres should be more depending on their real life experiences. |

The interpreting dimension was revised in five levels (see Table 9). When the students got incorrect/incomplete solution, they could interpret the results in an incomplete way depending on the real life. Similarly, the different validating approaches also elicited. The students who controlled only the calculation in the first stage came to validate their assumptions, models and solutions. They tried to correct their mistakes based on their experiences gained by the evaluation of in class presentations and the researchers' emphases on validation. In the solution of Time in School Problem, one group firstly considered the number of school days to find the time spent in school. While validating their solution, they thought that solving the problem by depending on hours was more reasonable and handled the hours spent at school during a day (see Table 8).

Table 8

*An Example for Validating from the Middle School Students’ Solution to the Time in School Problem*

|  |  |
| --- | --- |
| A section from the students’ expressions | Evaluation by the researcher |
| Nuran: [First writes down on the paper “We think that we do not spend most of the year in school”] Now, we will prove this thought. We’ll make calculation. How many hours of a day do we spend in school?  Ulaş: 6.  Nuran: The lessons start at half past 7 a.m. and finishes at haft past 1 p.m.  Pırıl: 6 hours.  Ulaş: But, we will count the days, look it doesn’t say the hours. There are 52 weeks in a year, we’ll multiply 52 with 50 and subtract that result from 365 [The students make calculation on the number of days].  …  Nuran: First, we’ll find how many hours we spend in school and then subtract the result from 24 then multiply it with 5.  Erdem: Why do we multiply with 5?  Nuran: We go to school on five days. Then, to find our free time, we’ll multiply the result with 5. [The students decide to make the calculation depending on the hours spent in school, afterwards, they consider the holidays as hours.] | The students first constructed and solved a model depending on the number of the days spent in school. Then, in validating their solution, they decided to consider the hours because they did not spent whole day in school. Depending on this, they corrected all assumptions about the time spent in school and on holiday. Afterwards, they were observed to correct their model and solution based on the so-called assumptions. |

The validating dimension was extended upon and divided in two main categories such as validating completely and validating partially. The students’ approaches who had validated all their assumptions, models and solutions were evaluated as validating completely, while the students' approaches who had considered one or few of these were evaluated as validating partially. Some groups tried to correct their mistakes by validating but some others completed the solution process without correcting their mistakes. Thus, the validating skill was defined at seven levels (see Table 9). Finally, the RAMS constructed with five levels for understanding the problem, four levels for the simplifying, five levels for mathematizing, five levels for working mathematically, five levels for interpreting and seven levels for validating.

Table 9

*The Final form of the RAMS*

|  |  |  |
| --- | --- | --- |
|  | Levels | Definition |
| Understanding the problem | Level 1 | Not understanding the problem, not determining the givens and goals, and not forming or mistakenly forming a relationship between them. |
| Level 2 | Understanding the problem partly, determining the givens and goals to some extent but not forming or mistakenly forming a relationship between them. |
| Level 3 | Understanding the problem completely, determining the givens and goals, not forming or mistakenly forming a relationship between them. |
| Level 4 | Understanding the problem completely, making unimportant mistakes in determining the givens and goals, not forming a relationship between them. |
| Level 5 | Understanding the problem completely, determining the givens and goals, and forming a relationship between them. |
|  |  |  |
| Simplifying | Level 1 | Not simplifying the problem, not determining the necessary/unnecessary variables, and making wrong assumptions. |
| Level 2 | Simplifying the problem partly, determining the necessary/unnecessary variables to some extent, and making wrong assumptions. |
| Level 3 | Simplifying the problem, determining the necessary/unnecessary variables and making partly-acceptable assumptions. |
| Level 4 | Simplifying the problem, determining the necessary/unnecessary variables and making realistic assumptions. |
|  |  |
| Mathematizing | Level 1 | Not constructing or mistakenly constructing mathematical model/s. |
| Level 2 | Constructing incomplete/wrong mathematical model/s based on partly-acceptable assumptions. |
| Level 3 | Constructing correct mathematical model/s based on partly-acceptable assumptions. |
| Level 4 | Constructing incomplete/wrong mathematical model/s based on realistic assumptions and relating them to one another. |
| Level 5 | Correctly constructing the needed mathematical model/s according to realistic assumptions, explaining model/s and relating them to one another. |
|  |  |
| Working mathematically | Level 1 | Not presenting a mathematical solution, solving the constructed models wrongly or trying to solve the wrong mathematical model. |
| Level 2 | Including deficiencies/mistakes in the solution of the mathematical models constructed incompletely/wrongly. |
| Level 3 | Solving correctly the mathematical models constructed incompletely/wrongly. |
| Level 4 | Including deficiencies/mistakes in the solution of the correctly constructed mathematical models. |
| Level 5 | Achieving correct mathematical solution by solving the correctly constructed mathematical models. |
|  |  |
| Interpreting | Level 1 | Misinterpreting or not interpreting the obtained mathematical solution in real life context. |
| Level 2 | Incompletely interpreting the erroneous/incomplete mathematical solution in real life context. |
| Level 3 | Correctly interpreting the erroneous/incomplete mathematical solution in real life context. |
| Level 4 | Incompletely interpreting the obtained correct mathematical solution in real life context. |
| Level 5 | Correctly interpreting the obtained correct mathematical solution in real life context. |
|  |  |
| Validating | Level 1 | Not validating or making wrong validation. |
| Level 2 | Validating partially, not correcting the determined mistakes. |
| Level 3 | Validating partially, correcting the determined mistakes to some extent. |
| Level 4 | Validating partially, correcting the determined mistakes. |
| Level 5 | Validating completely, not correcting the determined mistakes. |
| Level 6 | Validating completely, correcting the determined mistakes to some extent. |
| Level 7 | Validating completely, correcting the determined mistakes. |

**The studies of Validity and Reliability**

Creswell (2013) stated that researchers should spend long time in the field to increase the validity of qualitative researches. We tried to assure this through ourr experiences regarding modelling, implementations with students, ten-week data collection, constant comparative analysis, and eight-month period in which the data were examined again. In data collection and analysis, triangulation (Creswell, 2013) was used depending on validity strategies. Based on this, the transcriptions and the researchers' observations were utilized. We defined the RAMS development process with all stages and tried to increase the reliability of the study. While the levels were explained, the reliability was provided by stating the excerpts on students' expressions and solutions.

The data analysis made depending on a theoretical framework makes the reliability of the study richer (Yıldırım & Şimşek, 2008). Defining the dimensions and levels based on the framework was one of the increasing the study’s reliability. Opinions from the experts of mathematics education studying on modelling were taken for the RAMS and we asked for them to make the needed corrections on the statements by reading the definitions of the levels. We aimed that this rubric should be understandable in a same way for every user. Thus, the definitions for all levels were structured and standardized and the statements were made to be more obvious.

The validity and reliability were increased by being realized several modelling tasks with more than one groups at different grades. The RAMS was more valid and reliable because the researchers gave scaffoldings to enable the cognitive modelling skills to appear in the course of the implementations.

Patton (2014) stated that validity and reliability could be enabled with multiple coders. We increased validity and reliability by handling the evaluations of two experts’ mathematics educator who used RAMS. The experts studying on modelling independently evaluated the students' solutions on modelling tasks. We gave them the transcripts on different groups’ solutions about each task. Then, their evaluations were compared, and the percentage of agreement (Miles and Huberman, 1994) was calculated by considering which skill was handled at which level. The agreed evaluation number was 22 for all tasks. Because the percentage of the agreement proved to be 73.3%, the latest stage of the RAMS was formulated by agreeing on it. Four tasks were coded as T1, T2, T3, T4 respectively in Table 10 and the levels selected by the two experts was given.

Table 10

*The Percentage of Agreement about the Evaluations of the Experts with RAMS*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | The Evaluation of the First Specialist | | | | The Evaluation of the Second Specialist | | | |
| Skills | T1 | T2 | T3 | T4 | T1 | T2 | T3 | T4 |
| Understanding the problem | L5 | L 5 | L5 | L5 | L5 | L5 | L5 | L5 |
| Simplifying | L4 | L4 | L4 | L4 | L4 | L4 | L4 | L4 |
| Mathematizing | L5 | L4 | L5 | *L5* | L5 | L4 | L5 | *L4* |
| Working mathematically | L5 | L2 | L5 | *L5* | L5 | L2 | L5 | *L3* |
| Interpreting | *L5* | L2 | L4 | L3 | *L4* | L2 | L4 | L3 |
| Validating | L7 | *L1* | L7 | L6 | L7 | *L3* | L7 | L6 |
|  |  |  |  |  |  |  |  |  |
| Agreement ratio between the evaluations |  | | | | 5/6 | 5/6 | 6/6 | 4/6 |
| Total agreement ratio | 22/30 = %73,3 | | | | | | | |

Note. The different evaluations were shown in italic. L5 means Level 5.

According to the stability method used to ensure the reliability of the data analysis, it is stated that a certain period after the first analysis of the data, a second analysis should be carried out by the same person (Krippendorff, 1980; Weber, 1985). To that end, the data obtained from four modelling tasks applied at the second stage were evaluated again with the RAMS by a researcher about six months later. So, it was also explained the reliability of the RAMS.

**Scoring Suggestions for the Assessments with the RAMS**

The only purpose of rubrics is not to assess the students’ performance but also informing the students about their own development based on the results of rubric (Black & William, 2009). It is necessary to give feedback to the students in the consequence of the assessments carried out with RAMS. After developing RAMS, we studied over how the students’ solutions would be scored. We thought that modelling applications could be tackled with two approaches as atomistic and holistic. We decided five different scoring related to the RAMS through the opinions of experts. Level 1 was scored with 0 in each skill dimension. According to the atomistic approach, when a modelling task is solved about a specific skill, a scoring is made with t the dimensions and levels of this skill. When a modelling task is solved according to the holistic approach, the researcher/teacher is primarily determined the maximum score that s/he can give to assess the solution. Each skill has equal score, then, s/he determines the scores for each dimension by dividing the total score with 6. After the score of each dimension is determined equally to one another, divisions are performed according to the number of level. In this second approach, the students’ solutions are assessed and an equiponderant scoring is enabled to be carried out about each skill. Because the levels of dimensions are different, a difficulty can come up in scoring since the scores obtained at the end of the divisions are decimal expansions. Thirdly, when a task is solved according to the holistic approach, the researcher/teacher may want to assess each level by starting from 0, in which case the maximum score to be taken with RAMS is 25. In calculation of scores with this assessment, a trouble may not be occured. Because the level numbers of each dimension are not equal, there can be a disadvantage in favor of greater number of levels such as the validation skill. When a modelling task is solved according to the holistic approach, the researcher/teacher may carry out an assessment by taking greater the score of whichever skill s/he wants to emphasize. Alternatively, when a task is solved in the holistic approach, the scoring of the skills apart from the understanding skill the task can be sufficient. In modelling applications, students start the solution after they have understood the problem. If the students do not understand the problem, they want help from the teacher and the teacher should enable them to understand correctly with scaffolding. The scoring of the understanding skill can be ignored in the assessment. An assessment could be made in which the rest of the five skills get equal scores (each 20 points). Thus, 20 points of each skill is equally shared according to the levels. For example, when considering the validating dimension, 20 points should be divided to six because this dimension includes six levels. The total score obtained can be accepted as the score of the modelling skill.

**Conclusion, Discussion and Suggestions**

In this study, to assess students’ cognitive modelling skills, the RAMS and its development process were introduced. RAMS can be used for assessing different modelling tasks because it was formulated from the implementations of several modelling task. Several researchers and teachers who want to determine their students’ cognitive modelling skills can take the advantage of RAMS as an assessment tool because the skill levels were defined clearly.

While the strengths and weaknesses of students in terms of modelling skills are disclosed with the RAMS, the opportunity to make both quantitative and qualitative assessment is provided. While determining the students’ modelling skills with quantitative assessment, the teachers may convert these values into points. Which approaches students display for modelling skills is disclosed in a clear and detailed way with the qualitative assessment. Thus, students' difficulties in skills may clearly observed and to overcome these obstacles effective modelling applications may be planned.

When literature is reviewed in terms of the rubrics developed for assessing cognitive modelling skills, it can be said that RAMS has many aspects different from them and constructed in a more detailed way. In the rubric of Berry and O’Shea’s (1982), even though the dimensions were handled as the cognitive modelling skills, the explanations about the dimensions' levels were not included. There were not explanations on how students’ solutions would be assessed. The researchers stated that assessment should be made with maximum five points considering the assumptions, simplifications and determined important features. However, the points to be taken from the five points regarding the factuality level of students’ assumptions or to what extent they make simplification were not determined. This may raise dilemmas for evaluator and could prevent the objectivity among different evaluators. Besides, the researchers pointed out that presentations of the solutions could be assessed in terms of clarity and layout. On the other hand, in RAMS the social skills of students were not assessed in the presentation; instead, the students’ approaches are cognitively dealt with in the presentations because it was aimed to evaluate the cognitive modelling skills.

In the rubric of Galbraith and Clatworthy (1990), although three dimensions about the cognitive modelling skills were, in these dimensions more than one skill were assessed at the same time. For example, the second dimension contained assumptions and eliciting models, and the third dimension contained interpretations and validation as well as performing the mathematical solution. Even though the levels' definitions of the dimensions were available, the evaluator may have difficulty deciding the students’ level. For instance, the first level of the third dimension includes solution of the task with no assistance, the second level includes solution with little or no assistance and no correction of model, and the third level includes the interpretation and validation of model by solving task independently. A trouble may create for determining student's level who have solved task with little assistance, and interpreted and validated constructed model. Different from the RAMS, the rubric of Galbraith and Clatworthy (1990) aims to assess cognitive and social skills of students based on their written and verbal solutions. However, RAMS may make a more detailed assessment possible since it focuses only on cognitive skills.

Keck’s (1996) rubric measures usage of mathematics and communication skills as well as cognitive modelling skills. When considering cognitive modelling skills, the dimensions of the rubric were observed to have similarities with RAMS. It differs from the RAMS because of its fourth dimension where mathematical solution and interpretation were assessed together. To the levels' definitions of dimensions, interpretation of solution may be assessed without working mathematically because it was only assessed the compatibility of mathematical solution. There was not any dimension or level definition about the assessment of solution of the model. Because of this, the rubric of Keck may be deficient in assessing the cognitive modelling skills. He explained that the dimension of identification of problem is assessed in three levels and all other dimensions in five levels. The RAMS is more comprehensive than Keck’s rubric, because in the validation skill, the seven levels were handled as correction of assumptions, models and solutions.

The rubric of Chan, Ng, Widjaja and Seto (2012) contained an assessment under three dimensions as assumptions, interpretation of task and solutions, and mathematical reasoning and computation. Each dimension was assessed under three levels however the levels were remained restricted while explaining dimensions. For instance, the level of mathematical reasoning and computation dimension evaluated the number of discussed variables, the compatibility of the used mathematics, the reasonability and explicitness of mathematical reasoning and computation at the same time. The fact that the levels deal with abilities different from one another suggests the question of how the assessment will be carried out when one or some abilities are not observed. Therefore, it was thought that the evaluators who use this rubric could have difficulty in classifying the students’ solutions under appropriate categories and making an assessment according to this classification.

Leong (2012) used a modelling process in his rubric was similar to that in RAMS. Yet, Leong’s rubric was thought not to provide an instruction clear enough for the researchers/teachers who wishes to assess the cognitive modelling skills because of the fact that dimensions in the rubric measures more than one abilities, definitions of levels about the dimensions were not made, and the reason for the difference in the weights of dimensions was not explained.

Anhalt and Cortez (2015), in their rubric, considered the students’ abilities in the solution of the modelling tasks but did not especially assess cognitive modelling skills. The rubric's levels assessed the student explanations in the solution, connections between concepts, student works, reasoning, used representations and mathematical concepts, and calculations. Although there were some criteria for the cognitive modelling skills in levels' definitions, it was observed to differ from RAMS because an assessment parallel to the skills was not especially carried out.

RAMS provides opportunity for the assessment of cognitive modelling skills, while it presents offers of scoring to the researchers/teachers who wish to combine their assessments with quantitative or qualitative techniques depending on their purposes. The RAMS may address to different audience and purposes. The strengths and weaknesses of RAMS may be examined from a different modelling perspective by making teachers use the RAMS in modelling implementations. Besides the assessment of cognitive modelling skills, the assessment of different social abilities in modelling process can be developed. By this means, it will be assured that different skills of students working in modelling tasks should be handled. The studies may be made about how student works on a modelling task will be assessed by exemplifying the student approaches in all dimensions of RAMS. Besides, the researches which were enriched with the sample student studies will make contribution to the field by focusing on one dimension of RAMS for the purpose of atomistic assessment of different modelling tasks.

**References**

Anhalt, C., & Cortez, R. (2015). Mathematical modeling: A structured process. *Mathematics Teacher*, *108*(6), 446–452.

Author (2010).

Author (2011).

Author (2012).

Author (2013-a).

Author (2013-b).

Author (2015).

Berry, J., & O’Shea. T. (1982). Assessing mathematical modelling. *International Journal of Mathematical Education in Science and Technology*, *13*(6), 715-724.

Biccard, P., & Wessels, D. (2011). Development of affective modelling competencies in primary school learners. *Pythagoras*, *32*(1), 1-9. doi: 10.4102/pythagoras.v32i1.20.

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment Evaluation and Accountability*, *21*, 5–31.

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

Blomhøj, M., & Kjeldsen, T. N. (2006). Teaching mathematical modelling through project work. *Zentralblatt für Didaktik der Mathematik-ZDM*, *38*(2), 163-177.

Blum, W., & Niss, M. (1991). Applied mathematical problem solving, modelling, applications, and links to other subjects - state, trends and issues in mathematics instruction. *Educational Studies in Mathematics*, *22*, 37-68.

Blum, W. (2011). Can modelling be taught and learnt? Some answers from empirical research. In G. Kaiser, W. Blum, R. Borromeo Ferri, & G. Stillman (Eds.), *Trends in Teaching and Learning of Mathematical Modelling.* (pp. 15-30). New York: Springer.

Borromeo Ferri, R. (2006). Theoretical and empirical differentiations of phases in the modelling process. *Zentralblatt für Didaktik der Mathematik-ZDM*, *38*(2), 86-95.

Borromeo Ferri, R. (2014). Mathematisches Modellieren Lernen und Lehren [Mathematical Modelling for Teachers and Learners]. *3rd Mathematics Education Applications Mathematical Modelling Activities Workshop*. 1-4 April 2014. İzmir, Turkey.

Büchter, A., & Leuders, T. (2005). From students’ achievement to the development of teaching: Requirements for feedback in comparative tests. *Zentralblatt für Didaktik der Mathematik-ZDM*, *37*(4), 324-334.

Cabassut, R. (2010). The double transposition in mathematisation at primary school. In V. Durand-Guerrier, S. Soury-Lavergne & F. Arzarello (Eds.), *Proceedings of the 6th Congress of the European Society for Research in Mathematics Education CERME 6* (pp. 2156 - 2165). Lyon, France: Service des publications, INRP.

Chan, C. M. E., Ng, K. E. D., Widjaja, W., & Seto, C. (2012). Assessment of primary 5 students' mathematical modelling competencies. *Journal of Science and Mathematics Education in Southeast Asia,* *35*(2), 146-178.

Creswell, J. W. (2013). *Nitel, Nicel ve Karma Yöntem Yaklaşımları Araştırma Deseni*. [Qualitative, Quantitative, and mixed methods approaches.] (S. B. Demir, Trans.) Ankara, Turkey: Eğiten Kitap.

Galbraith, P.L., & Clatworthy, N.J. (1990). Beyond standard models: meeting the challenge of modelling. *Educational Studies in Mathematics,* *21*(2) 137-163.

Galbraith, P. L., & Haines, C. R. (1997). Some mathematical characteristics of students entering applied mathematics courses. In S. K. Houston, W. Blum, I. Huntley & N. Neill (Eds*), Teaching and Learning Mathematical Modelling* (pp. 77-92). Chichester, UK: Albion Publishing.

Grünewald, S. (2012). Acquirement of modelling competencies – first results of an empirical comparison of the effectiveness of a holistic respectively an atomistic approach to the development of (metacognitive) modelling competencies of students. *12th International Congress on Mathematical Education*, 8 July-15 July 2012, COEX, Seoul, Korea.

Haines, C., Crouch, R., & Davis, J. (2001). Recognizing students’ modelling skills. In J. F. Matos, W. Blum, S. K. Houston, & S. P. Carreira (Eds.), *Modelling and mathematics education - ICTMA 9: Application in Science and Technology* (pp.366-380). Chichester, UK: Horwood.

Henning, H., & Keune, M. (2007). Levels of modelling competencies. In W. Blum, P. L. Galbraith, H. W. Henn, & M. Niss (Eds.), *Modelling and Applications in Mathematics Education The 14th ICMI Study* (pp. 225-232). New York: Springer.

Houston, K. (2007). Assessing the “phases” of mathematical modelling. In W. Blum, P. L. Galbraith, H. W. Henn, & M. Niss (Eds.), *Modelling and Applications in Mathematics Education The 14th ICMI Study* (pp. 249-256). New York: Springer Science + Business Media, LLC.

Kaiser, G., & Maaß, K. (2007). Modelling in lower secondary mathematics classroom – problems and opportunities. In W. Blum, P. L. Galbraith, H. W. Henn, & M. Niss (Eds.), *Modelling and Applications in Mathematics Education: The 14th ICMI Study* (pp 99-108). New York: Springer Science + Business Media, LLC.

Kaiser, G., & Schwarz, B. (2006). Mathematical modelling as bridge between school and university. *Zentralblatt für Didaktik der Mathematik-ZDM*, *38*, 196-208.

Kaiser, G. (2007). Modelling and modelling competencies in school. In C. Haines, P. Galbraith, W. Blum, & S. Khan (Eds.), *Mathematical modelling (ICTMA 12): Education, Engineering and Economics: Proceedings from the Twelfth International Conference on the Teaching of Mathematical Modelling and Applications* (pp. 110-119). Chichester, UK: Horwood.

Kaiser, G., Blomhøj M., & Sriraman B. (2006). Towards a didactical theory for mathematical modelling. *Zentralblatt für Didaktik der Mathematik-ZDM*, *38*(2), 82-85.

Keck, H. L. (1996). The development of an analytic scoring scale to assess mathematical modelling projects. (Doctoral dissertation). Retrieved from Scholar works University of Montana. (Accession or Order no.)

Krippendorff, K. (1980). *Content Analysis: An Introduction to its Methodology*. Beverly Hills, CA: Sage Publications.

Leong, K.E. (2012). Assessment of mathematical modeling. *Journal of Mathematics Education at Teachers College*, *3*(1), 61-65.

Maaß, K. (2006). What are modelling competencies?. *Zentralblatt für Didaktik der Mathematik-ZDM*, *38*(2), 113-142.

Maaß, K., & Gurlitt, J. (2011). Designing a teacher questionnaire to evaluate professional development in modelling. In V. Durand-Guerrier, S. Soury-Lavergne & F. Arzarello (Eds.), *Proceedings of the 6th Congress of the European Society for Research in Mathematics Education CERME 6* (pp. 2056-2065). France, Lyon: Service des publications, INRP.

Maaß, K., & Mischo, C. (2011). Implementing modelling into day-to-day teaching practice-the project stratum and its framework. *Journal Für Mathematik-Didaktik*, *32*(1), 103-131.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative Data Analysis.* Thousand Oaks, CA: Sage Publications.

Oprins, E., Burggraaf, E., & Van Weerdenburg, H. (2006). Design of a competence-based assessment system for air traffic control training. *The International Journal of Aviation Psychology, 16*(3), 297-320.

Patton, M. Q. (2014). *Nitel Araştırma ve Değerlendirme Yöntemleri.* [Qualitative Research & Evaluation Methods.] Ankara, Turkey: Pegem Akademi.

Schukajlow, S., Leiss, D., Pekrun, R., Blum, W., Müller, M., & Messner, R. (2012). teaching methods for modelling problems and students’ task-specific enjoyment, value, interest and self-efficacy expectations. *Educational Studies in Mathematics,* *79*, 215-237.

Stillman, G., Brown, J., & Galbraith, P. (2010). Identifying challenges within transition phases of mathematical modelling activities at year 9. In R. Lesh, P. L. Galbraith, C. R. Haines, & A. Hurford (Eds.), *Modeling Students’ Mathematical Modelling Competencies: ICTMA 13* (pp. 385-398). New York: Springer Science + Business Media, LLC.

Vos, P. (2013). Assessment of modelling in mathematics examination papers: Ready-made models and reproductive mathematising. G.A. Stillman, G. Kaiser, W. Blum & J. P. Brown (Eds.), *Teaching Mathematical Modelling: Connecting to Research and Practice, International Perspectives on the Teaching and Learning of Mathematical Modelling* (pp. 479-488). New York: Springer Science + Business Media, LLC.

Weber, R. P. (1985). *Basic Content Analysis, Quantitative Applications in the Social Sciences*. Beverly Hills, CA: Sage Publications.

Yıldırım, A., & Şimşek, H. (2008). *Sosyal Bilimlerde Nitel Araştırma Yöntemleri* [Qualitative Research Methods in Social Sciences]. Ankara, Turkey: Seçkin Yayıncılık.

Yin, R. K. (1987). *Case Study Research Design and Methods.* London, UK: Sage Publications Inc.