Studies in ***Learning, Evaluation Innovation*** and ***Development***

http://sleid.cqu.edu.au
6(3), pp. 60–79. December 2009

A model for alternative assessment of students' problem solving processes: Conceptual and experimental validation

Dr. Rama Klavir, Kaye College of Education, Israel, klavir@zahav.net.il
Prof. Malka Gorodetsky, Ben Gurion University, Israel, malka@bgu.ac.il

Abstract

The paper provides a qualitative assessment model for monitoring students' solution processes and their outcome while being engaged in problem solving and learning of analogical problems. The model is structured as a mapping sentence the facets of which are based on procedural components that have been found to be part of the solution processes of analogical problems by students of varying excellence. The model has the capacity to assess, independently, both the quality of the process and that of the solution outcome. The paper presents the validation of the model and possible connections between the process components and the outcome (solution). It was found that though performance on the process components and the outcome are correlated, the connection is not linear, i.e., the interrelationship between the solution processes and the outcome is not direct. The model was found to be sensitive to students' learning of analogical problems solving. The potential of the model as an assessment and learning tool for teachers' and students is suggested.

Keywords: Model for alternative assessment; Mapping sentence; Problem solving; Selectivity of solution components, Analogical learning.

This article has been peer-reviewed and accepted for publication in *SLEID*, an international journal of scholarship and research that supports emerging scholars and the development of evidence-based practice in education.

ISSN 1832-2050

Introduction

School evaluation and many other competitive tests are mainly aimed towards students' classification and ranking. These are mainly quantitative problem solving tests, the scoring of which is predetermined on the basis of externally expected achievements. While it is agreed that assessments that mainly address **outcomes** are an inefficient way to monitor students' learning, and although it is also agreed that insights into the learning process can provide a better understanding of students' learning difficulties, the latter are still very scant in the educational system (Chappell & Thompson, 1999). Nonetheless, the recognition of the educational potential of insight into learning processes for students' advancement, led to the development of "alternative assessment." Within this approach, alternative ways for grasping students' processes while coping with new, complex and often ill-defined problems for the exposure of learning difficulties as well as an insight into idiosyncratic knowledge have been developed. Gehrke and Dickinson (2002) mention a long list of possible alternative assessment processes, e.g., "*answer an open-ended question, work out a solution to a problem, perform a demonstration of a skill, or in some way produce work rather than select an answer from choices on a sheet of paper. Portfolios and instructor observation of students are also alternative forms of assessment.*" These have been implemented for the assessment of the learning of subject, even those that are perceived as closed and quantitative in nature, e.g., mathematics. In the assessment of learning mathematics portfolios, mathematical journals, teachers' observations, students' self-assessment, different forms of communications, surveys, interviews and performance assessment have been applied (Baker et al., 1993) as well as authentic assessment (Thomasenia, 1998).

As mentioned above, although processes of alternative assessment have been developed and their suitability for gaining better insight into students' understanding and learning have been confirmed, there seem to be deep rooted reasons that prevent the infusion of alternative assessment into the classroom. Sanders and Horn (1995) suggest theoretical reasons for these preventions such as (a) the shortcomings of alternative assessments for generalising information for comparison purposes over time and on a large-scale basis, as demanded in schools. Alternative evaluation is still mainly conceived as individualised and subjective assessment; (b) Alternative evaluation has not developed indicators that measure distinctive features that are in high correlation with other features to enable validated inferences as demanded by meaningful evaluation. Other possible reasons that hamper the infiltration process of alternative evaluation into the school are practical, and include teachers' difficulties in translating alternative assessment processes into daily classroom life. Seeley (1994) cites some of the difficulties teachers face:

- How to incorporate multiple sources of information into a single score?
- How to maintain tools that could adequately replace the conventional grades as to student's knowledge and degree of progress?
- How to maintain a grading continuum that expresses the entire process of learning and problem-solving strategies?
- How to alter the alternative assessment tools to changes in situations and time?

The present paper is yet another attempt to provide a qualitative model for monitoring students' solution processes while engaged in problem solving and learning. The uniqueness of the model lies in the inclusion of components of the solution process and the outcome (the solution) into a unified model and yet enabling insight into each of the components separately. The model makes use of solution components that were documented as being responsible for the efficient and elegant solution of analogical problems (problems that share a common solution structure), whereas their absence in the process was attributed to failure in achieving a solution. These components reflect students' thinking along the solution process. Thus, the model can be used as a measure to understand possible relationships between performance along the solution process and its outcome, from a theoretical point of view, and as a pragmatic tool for students and teachers in deducing possible difficulties and further needed learning.

http://sleid.cqu.edu.au
6(3), pp. 60–79. December 2009

The assessment model

The assessment model provided in this paper consists of a mapping sentence (see diagram no. 1) within which each of the facets stands for a different component involved in the solution process. The model can be used as a reflective instrument at the will of the teacher or student, or as a pre-post instrument for the follow up of learning. It can be applied for the assessment of **individuals** as well as **groups** at different points in time, e.g., pre-post learning.

The model includes six facets, each one standing for a different component that was found to be involved in the solution processes of excellent solvers, and a seventh facet that describes the quality of the solution outcome. The chosen components are:

- encoding
- retrieval
- relating
- combination
- goal directness
- awareness and
- correctness of the final solution.

Their choice is based on the research literature of excellent problem solvers, i.e., gifted and experts. Gifted students were found to be very good solvers (e.g., Rabinowitz & Glaser, 1986) of content-free problems, e.g., riddles, insight problems (Davidson, 1986) and problems in specific domains, e.g., mathematics (Sowell, 1993). Likewise, experts were also found to excel in solving general problems such as heuristic problems (Schoenfeld, 1985) and those involving scientific reasoning (Schunn & Anderson, 1999). Both populations were also found to be excellent learners, i.e., they gain more knowledge than their groups of comparison from different learning situations, e.g., learning from examples and analogical learning (for gifted, Davidson, 1986; for experts, Novick, 1988; for expert children, Gobbo & Chi, 1986). Learning in these cases pertains to better achievements on the final solution and/or in refinement of the solution process towards a more selective and elegant process (Novick & Holyoak, 1991). It has to be emphasised that these components were chosen by us for this study; however, their number or combination can be changed according to the needs of the assessment or research (e.g., Gorodetsky & Klavir, 2003; Klavir & Gorodetsky, 2001)..

In the research literature, the performance on these components was found to be executed on different levels of ***selectivity***, a feature that reflects the range of possible performances on a specific component, ranging from the most selective performance on a specific component to the least selective. Thus, the model is applicable for the analysis of (a) the quality of the performance, i.e., the nature of the solution outcome, and (b) analysis of the selectivity levels of the solution process components that were enacted. As a result, a solution profile can be ascribed to each solver that includes the selectivity levels employed in each component.

Model of a mapping sentence

	A. Encoding		B. Retrieval
In the solution process, the solver encodes mainly	1. Deep structure 2. Deep and surface structure 3. Surface structure	items, retrieves	1. Previous knowledge 2. Minimal pieces of previous knowledge 3. No pieces of previous knowledge

	C. Relating		D. Combination
that refers mainly for the interpretation of the	1. Deep structure 2. Deep and surface structure 3. Surface structure	of the problem, and executes, mainly a combination that can be described as	1. Integrative 2. Replicative 3. Distortive

	E. Goal Directness		F. Awareness
in a process that	1. Proceeds directly to the final goal 2. Proceeds by a systematic search 3. Proceeds by a random search 4. Proceeds directly in a wrong direction	and along this process she/he indicates that he/she	1. Is aware of 2. Is partially aware of 3. Is not aware of

	G. Correctness of Final Solution
the correctness of the final solution that can be objectively assessed as	1. Correct 2. Partially correct 3. Erroneous 4. No solution

The components of the solution process

The present model includes the following components:

A. Encoding: This component reflects the information extracted by the solver from a problem to be solved. The gifted and expert solvers were found to display more selective encoding than average children (for gifted, Davidson, 1986; for expert children, Gobbo & Chi, 1986). Both populations of excellence were found to encode relevant information (deep structure), whereas the average populations tend to encode surface structure features.

B. Retrieval: This refers to the activation of concepts and terms from the LTM that enable the interpretation, re-presentation, of a problem into terms with which the solver is familiar (Anderson, 1995). Fan, Mueller, and Marini (1994) distinguished between the retrieval of terms essential for the interpretation of the text and the retrieval of procedural knowledge needed for a solution. This distinction is applied in this paper. Thus retrieval refers only to semantic knowledge that is associated either with the surface or deep structure of the problem. It is assumed that retrieval of previous knowledge is an indication for a more intelligent (selective) process as the translation of the new problem into more familiar terms has a higher potential to lead to a correct solution (Gobbo & Chi, 1986; Mannes, 1994).

C. Relating: This component reflects the features of the problem for which previous knowledge was retrieved and led to the re-presentation of the problem. The assessment of the retrieved knowledge is based on its application, e.g., whether it is utilised to represent the deep structure of the problem or its surface structure (Gobbo & Chi, 1986, p. 225)

D. Combination: This reflects the solver's attempt to integrate the fragmented pieces of knowledge into a combined solution structure. Based on Anderson (1995), the solution structure is composed of a sequence of combinations ('productions') between conditions ('if') and relevant actions ('then'). Meaning, in the solution process fragmented pieces of encoded and retrieved knowledge are moulded into a sequence that is leading to the solution. In this process the experts exhibit a higher proficiency than novices, as their combinations are based on more selective and compiled productions. Likewise, gifted students were also found to be more selective in their combinations as compared to average ones (Davidson, 1986).

E. Goal Directness: Experts are faster than novices in reaching solutions (Simon & Simon, 1978) and their solution processes tend to be goal-directed (Novick & Holyoak, 1991); however, in the solution of difficult problems a switch to means-end analysis was observed (Rabinowitz & Glazer, 1986). Like the experts, the gifted are also capable of monitoring their solution pathway and moving from one pathway to another when necessary (Maniatis et al., 1998).

F. Awareness: Self-awareness is a reflective meta-cognitive component that relates to the solver's capability to monitor, regulate, plan, interpret and judge his solution process (Fernadez-Duque et al., 2000; Sperling et al., 2004). This process guides the solver towards changes in his attempts to cope with the solution. The reflections are executed on the content as well as on the process towards the solution (Dienes & Perner 2002; Rosenthal, 2000). These processes are supported by the solver's previous knowledge and the new knowledge he accumulates over the solution process and involve high level thinking skills (Nelson, 1999). Experts are more effective than novices in monitoring their processes, probably because of their rich and well-organized knowledge that guides their solution process (Alexander & Judy, 1988). Similarly, gifted students were found to be more aware of their solution process than regular ones (Robinson & Clinkenbeard, 1998).

G. Correctness of the Solution: As mentioned above, the gifted and the expert children were defined as excellent problem solvers who solve new and complicated problems more successfully than their comparative groups (regular and novice students, e.g., Rabinowitz and Glaser (1986).

Analogical learning

The chosen situation for assessing students' learning was analogical learning that refers to the transfer of learned knowledge in the past to new analogical problems. Effective analogical transfer was found to be a feature of the excellence of gifted (Davidson, 1986) and that of experts (Novick, 1988). We believe that analogical learning is an appropriate context for the present study as its implementation can be studied over a wide range of behaviours, from the mere transfer of a known solution schemata to creative – intelligent behaviour (Low & Over, 1992). Analogical learning was also found to be effective in promoting learning towards improvement in the solution processes (including in the selectivity of the components) and the solution outcomes (Schmid et al., 2003).

Validation of the assessment model

The very deduction and construction of the assessment model on past research findings, provides a theoretical validation of its potential for tracking performance (process and product) of analogical problem solving in a given situation. The experimental validation relates to the potential of the model to capture the nature and the relations between the components and the final solution in the context of analogical problem solving and learning. This validation was achieved through the following study.

The assessment tasks

The tasks for the validation of the model are naturally analogical problems and the learning context is analogical learning. It is essential that the problems be new to the solvers and sufficiently complex to ensure the generation of new solution processes rather than simple transfer of solution schemas (Schunn & Anderson, 1999). The solution process of the problems should be complex and involve high order thinking skills that can be acquired through analogical learning. The specific knowledge content for analogical learning and analogical problem solving was chosen in the field of mathematics. The rational for this choice is based on the findings that logical–mathematical excellence characterises both populations of excellence, that of gifted and that of experts the solution components of which were incorporated in the model.

The validation steps

The experimental validation of the assessment model as an instrument for the follow-up of the solution processes of analogical problems by different populations, in the context of analogical learning, involved the following steps:

1. Validation that the assessment model does indeed capture the quality of the process (levels of selectivity) and the quality of the outcome (correctness of the solution) of analogical problem solving by different students.
2. The validation that the quality of selectivity of the components is indeed a measure that is associated with the quality of the final solutions by different students.
3. The validation that the assessment model captures analogical learning that is disclosed in the improvement of the selectivity of the solution process and the correctness of the solutions of different students.

Method

General design

The design included three phases:

In the first phase students received a questionnaire in which they were asked to solve one problem (source problem) and write down additional information regarding their solution process. The questionnaire included questions that referred indirectly to their nature of coping with the solution. The questions were:

1. Write a letter to your best friend and describe in your own words the problem you just read.
2. What are the most relevant details for the solution?
3. Does the problem include details that in your opinion are not important, marginal or do not contribute to the understanding of the problem?
4. Please solve the problem.
5. Please explain the problem to a student that does not understand it, give him an explanation: how did you solve it and guide him how to solve it.
6. Do you think that you have solved the problem correctly?
7. From the moment you received the problem to finding the solution – how would you describe the solution process that you undertook (choose only one possibility)?
    a. I felt like I was in an unfamiliar space without knowing how to proceed. I kept returning to points I had already been at. I kept repeating the same mistakes. Finally I found the solution.
    b. I started to solve the problem in a certain way. When I realised that it did not lead me to the solution I changed direction till I reached the solution.
    c. I read the problem and immediately knew the solution.
    d. Other:

In the second phase the students received two solved problems for learning. One problem was analogous to the source one and the other was a distracting problem with a similar surface structure but with a different deep structure. Students were asked to learn the problems and their solution processes and then to recite the problems and their solutions in their own words. This active processing manipulation of the analogical and distracting problems follows previous researches as a means to ensure the participants' engagement in the learning process (Reiter-Palmon et al., 1997).

The third phase was undertaken two weeks later. At this phase the students were asked to solve again the source problem and also to answer the same questions as in the first questionnaire. Students that did not complete the second phase seriously, meaning there was no indication for a learning process of the analogous and the distracting problems were excluded from the analysis. This move was undertaken to ensure that all subjects have learned to solve the problems, i.e., have the appropriate knowledge.

The validity of the use of post-solution protocols is supported by Siegler (1989) and Geary and Brown (1991), who claim that students are capable of describing the strategies they employed, even in a post-solution protocol. Because the reports were provided after solving the problem, it is impossible to assign a chronological sequence in attending to the different components in the assessment model.

The students

The research students included two age groups that comprised four groups of ability: two groups of excellent students, i.e., gifted (7-9th graders) and experts in mathematics (11th graders), and two groups of regular students matching in age and background to these groups.

The gifted students were those defined by the Ministry of Education as being above IQ=135 (N=125). They were chosen from educational settings that specifically catered to this population. Their comparison group was regular students (N=148) from a similar urban, middle-class background that attended the same comprehensive school. The latter were defined as neither gifted nor as special education students and studied in regular classes.

The expert students were defined in the field of mathematics and included students that took an advanced course in mathematics (5 credits in the Israeli matriculation system) and achieved a grade of over 70 (N=63). The novice students (N=101) were students that studied mathematics only at the ordinary level (3 credits) and achieved a grade of over 70. The choice of advanced students as experts rests on criteria used in previous studies (Berger & Wild, 1988) that have established the legitimacy of addressing populations of relative expertise as experts.

These four groups establish a continuum of populations of two age groups varying on their problem solving capabilities and learning.

The assessment tasks

The set of problems for the gifted/regulars and that for the experts/novices were based on similar reasoning skills, but their content knowledge was adjusted to the different age populations. Thus, although the problems had similar underlying features – a common denominator – they differed for both populations with respect to the required knowledge for the solution and the surface structure.

Each of the students solved only one problem but each group solved three problems. The problems included an insight-mathematical problem), a complex mathematical problem and an insight-verbal (not based on specific mathematical knowledge) problem.

Analysis

Students' answers in the questionnaires served as the basis for the analysis of the components involved in the solution process and for evaluating the level of the correctness of the solution. The criteria for deciding whether a component and its level of selectivity are reflected in students' answers were defined a priori and served as guidelines for the analysis. If more than one category seemed to be applicable, the most dominant one was chosen. For each student, only one category for each component was assigned. This enabled the construction of a solution profile for each student in each problem that consisted of a string of the accumulated selectivity levels for each component.

Lack of information for the analysis of a certain component was assigned as a missing value. When information was available the analysis followed the following guidelines:

Encoding – Facet A: The number of words and concepts that were used by the solver to describe the problem were counted. The different levels of encoding were assigned on the basis of dominancy of deep or surface structure items. Three levels of selectivity were assigned. The encoding of only deep-structure features was considered the most selective level and was assigned as level 1. The second level of selectivity was the provision of a mixture of features from the deep and surface structure and was assigned as level 2. Finally, the encoding of only surface structure features was considered as the least selective and was assigned as level 3.

Retrieval - Facet B: New words or concepts (not included in the problem) that appeared in students' replication of the problem story in addition to those of the original problem, were collected. The different selectivity levels of retrieval were assigned on the basis of the number of the re-presented items. Three levels of selectivity were defined: The retrieval of previous knowledge sufficient for the re-presentation of the problem in familiar terms was assigned – level 1. Retrieval of only minimal pieces of knowledge was assigned – level 2, and the absence of any retrieved previous knowledge was assigned – level 3.

Relating - Facet C: The retrieved new terms in students' stories were categorised on the basis of their relevancy to the deep or surface structure. The retrieval of information that relates to the data and terms that are associated only/mainly with the deep structure of the problem were assigned – level 1. The retrieval of information that relates to the deep and surface structures was assigned – level 2, and the retrieval of information that relates to only/mainly the surface structure was assigned – level 3.

Combination - Facet D: This component was defined as creating a logical/mathematical connection between at least two items in the problem. For analysis students' answers to questions 4 and 5 in the questionnaire were considered. Our analysis was based on search for production systems (‘if – then’) suggested by Anderson (1995). For the labelling of selectivity levels we used the categorisation provided by Peled and Wittrock (1990). The higher the number of selective productions in a solution process the combination was considered as more selective. The category that was assigned for a solver reflects the most abundant combinations he/she had performed. The most selective combination was considered as integration, a less efficient combination was considered as a replication, and the least selective combination was assigned as distortive, i.e., a combination that involved erroneous logical/mathematical combinations between the ’if-then’ parts and/or the addition of a non-relevant item that changed the solution structure.

Goal Directness – **Facet E:** This facet relates to the sequential solution steps of the solver. The assignment of the selectivity level was based on student’s answers to questions 6 and to the multiple-choice question 7. Four selectivity levels were distinguished: The most selective process was the one directed to the correct final solution, i.e., the solver chose to describe his solution process as (7c) and in his answer to question 6 in the questionnaire he stated that he believed his solution was correct (category 1). A less selective process was that which entailed a systematic search, i.e., when the solver chose (7b) (category 2). Even a less selective process was one that involved a random search, i.e., when the solver chose (7a) (category 3), and the least selective category was that which was directed to a final but erroneous goal, i.e., the solver chose (7c) but in his answer to question 6 he seemed to believe that his solution was erroneous (category 4). The open-ended possibility (7d) was analysed as to its correspondence to one of the categories: 1 – 4.

Awareness – **Facet F**: The categories for the levels of selectivity on this component were assigned on the basis of a match between the solver’s judgment regarding the nature of her/his solution – subjective correctness (as expressed in the answer to question 6) and the actual nature of the solution (the analysed answer to question 4) as judged by the researchers – objective correctness. The nature of the match between both measures of correctness was the criteria for assigning the level of selectivity of awareness. The most selective level was assigned when there was a complete match between the subjective and the objective judgments, i.e., both judgments recorded a correct solution, a partially correct or erroneous one. The next level of selectivity was assigned when there was a partial match, i.e., there was a small disparity between the subjective and the objective judgments, e.g., the solver stated that his final solution was correct but actually it was partially correct, and similar. The third level was assigned in situations that reflected a large disparity between the subjective and the objective judgments, e.g., the solver wrote that in his opinion "the solution is correct" but according to the researchers his solution was erroneous. An answer like "I don't know" was also assigned to this level.

Correctness of the solution – **Facet G**: The categories in this facet were assigned according to the level of the correctness of the solution. When the final solution was correct it was assigned as level 1, when a solution included half or more correct steps towards the solution it was considered a partially correct one (level 2) and when it included more than half incorrect steps it was considered an erroneous one (level 3). Level 4 was assigned to cases that did not provide any solution though there was an indication that the student made an attempt to solve the problem.

Analysis of the questionnaires

Analysis of the questionnaires was performed by one of the researchers (RK) and a group of research assistants (10) that were chosen following an interview. The assistants underwent two days of intensive training that included a common experience in the analysis of each component (facet), discussions and a follow-up of comparative analyses. Cases of disagreement were discussed with the researcher until an agreement was reached. After the training each of the assistants analysed about 10% of the questionnaires. A random sample from the analysed questionnaires by each research assistant was re-analysed. The agreement between the analyses of the facets ranged between 84%–93%.

Definition of variables

•Level of Correctness (of the final solution) was scored on the basis of Facet G. A fully correct solution was scored as 3, a partially correct solution was scored as 2, and an erroneous solution was scored as 1. A score of 0 was assigned to cases that indicated an attempt by the solver to cope with the problem but a solution was not provided.

•Selectivity levels of each of the components: The most selective component in facets A-F in the mapping sentence (encoding, retrieving, relating, combination, awareness, respectively) was scored as 3. The second level of selectivity in these facets was scored as 2. The least selective sub-process in these facets was scored as 1. A score of 0 was given in cases in which the student did not provide information for the specific component. In facet E, Goal Directness that included 4 levels of selectivity, the most selective attempt was scored as 4 and the others in decreasing value.

•Level of selectivity of the entire process: The number of selective components in the solution process of each student was counted. Thus the selectivity profile ranged between the highest level of 6 (selective components in all 6 facets, A-F) and of the lowest of 0 (no selective components). The selectivity level for each solution refers to the sum of the most selective facets in the solution.

Results

The results are presented under research questions that together provide an answer as to the validity of the proposed assessment model and to possible new understandings it provides.

1. Is there a correlation between the quality of the process – the selectivity levels of the components and the levels of correctness of the solution?

For this analysis an X2 test for the level of correctness of the final solution and the selectivity level of the solution process was performed. It was found that there is a significant correspondence between the 6 levels of selectivity of the solution process and the 4 levels of correctness of the solution product. Tables 2A and 2B provide the results for pre and post analogical learning correspondingly (See Appendix.). Before learning, X2=214.58 d.f.=18 $p<0.001$; and after learning, X2=281.15 d.f.=18 $p<0.001$, indicating that there is a connection between the number of the selective sub-processes and the correctness of the solution, or vice versa.

This result connotes a relationship between the solution process and the final solution establishing the existence of a general connection between selectivity and correctness of the solution. The results also indicate that correct solutions can be achieved without exhibiting the highest selectivity level in all six components.

2. What is the relationship between the selectivity level in each of the six components and the level of the correctness of the solution?

A Pearson correlation test was done between the correctness of the solution and the selectivity levels of each of the component. The results (see Table no. 3 in Appendix.) show a significant correlation ($p<0.001$) between the level of correctness and the selectivity level (the quality) of each of the components before and after analogical learning.

3. Does learning affect both the quality of the final solution (levels of correctness) and the selectivity of the solution process?

A Wilcoxon Signed Ranks Test was performed for the comparison between the ordinal levels of (a) the levels of correctness and (b) levels of selectivity, before and after analogical learning.

It was found (Table no. 4 in Appendix) that there is a significant difference between the quality of the solution and that of the process, before and after analogical learning. Learning had an effect on both, i.e., analogical learning led to improvement in the correctness of the solutions (from a mean of 1.61 to 1.87) and also to more selective levels of performance on the components (from a mean of 1.83 to 3.04).

4. Is the assessment model sensitive in detecting excellence in problem solving of excellent populations (gifted, experts)?

A Wilcoxon Signed Ranks Test was performed to compare between the ordinal levels of (a) levels of correctness, (b) levels of selectivity of gifted vs. regular, experts vs. novices and excellent solvers as a combined group vs. the comparative combined group of regular solvers. It was performed separately for the situation "before analogical learning" and "after analogical learning" (see Table no. 5 in Appendix). It was found that there was a significant difference between the quality of the solutions and the selectivity of the solution processes for both groups of excellence with respect to their comparative groups. It is of interest that although after learning there was an increase in the correctness of the solutions and in the quality of the process (the selectivity levels of the components), the change was more salient in the quality of the solution process than that in the correctness of the solution.

Discussion

The paper provides a validation of an alternative assessment model for the analysis of the process and product in problem solving and learning in the context of analogical thinking. The model is structured as a mapping sentence that provides the teacher (assessor) and student (learner) with a tool that can be used for multiple purposes. The application of the model to the analysis of the solution process supports a relationship between the selectivity levels of the components, i.e., the quality of the process and the level of the correctness of the solution, i.e., the quality of the product. The relationship was found to be positive, i.e., with the increase in the level of the selectivity of the process there was also an increase in the quality of the final solution. Though correct solutions can probably be achieved through non-elegant and non-selective processes as indicated by the results, the positive correlation between selectivity and the solution reinforces the importance of selective performance on the components and justifies attention to these as a major objective in teaching and concurrent follow-up through assessment. It should be stressed that the assessment model is flexible and provides the user with the choice of the components according to her/his needs. Thus, the problems to be solved, the components, the populations for follow-up, the learning situation, can be changed according to the assessor's or teachers' knowledge and needs. It should be stressed that the provided facets of the follow up are based on the research literature of analogical problem solving and thus are suitable for application in such a context.

The assessment model was also found to be sensitive to the gains from analogical learning. As a result of analogical learning there was an increase in the selectivity levels of the solution components as well as an improvement of the final solutions.

Engross, the model can be useful for two major orientations in the school life. We believe that good pedagogical assessment tools should also be good educational prescriptive instruments. Thus the model has the potential to serve both evaluative and prescriptive needs of those involved in the educational system. It can be applied for the evaluation of problem solving (process and product); for the follow up of learning (process and product) and the follow up of teaching. These processes can be addressed on two levels, that of the individual solver or learner or teacher and that of the group. Table no. 1 summarises these possibilities on the individual and collective (classroom or a larger group) levels.

Table 1. *The multiple educational uses of the model*

		Individual level	Collective level
Alternative Evaluation	(a) as compared to the learner	*Student*	*Classroom, cohorts etc.*
(Pre-post influence of learning)	(b) vs. an optimal solution process	*Student*	*Classroom, cohorts etc.*
Prescriptive	(a) learning	*Student*	*Classroom, cohorts etc.*
	(b) teaching	*Teacher*	*Classroom, cohorts etc.*

The possibility of the individual student to self-follow-up, step-by-step, of components in his learning/problem solving advancement has the potential to lower his/her feelings of frustration and incompetence via three possible mechanisms: 1. The division of the solution process into sub-tasks, components, increases students' chances to succeed on some of the components and thus lowering their frustration level. 2. Students' self monitoring of the impact of learning (that influences more the selectivity than the final solution (see results for the research question no. 4) can provide instant self-encouragement along the solution process, even if they fail on the final solution. Students' higher involvement in their cognitive processes can also affect their dispositions towards learning and problem solving in general. This is important in preventing students' possible disengagement from school activities (Pajares & Schunk, 2001).

The assessment model can, of course, be used by teachers to call students' attention to their successes and motivate them to learn. This kind of motivation was found to be very effective in students' willingness to cope with complex learning situations (Martin et al., 1999; Zimmerman, 2002).

The assessment model can also be used for explicit formal instruction of solution processes, a strategy that was found to be profitable in the learning of disabled students. McCleery and Tindal (1999) found that LD students that received detailed instructions on the processes to be undertaken when coping with complex situations coped better than regular students with no instruction. Similarly Carnine and Carnine (2004) found that instruction of junior high school students, that explicated ways of coping with complex scientific subjects, led to a better understanding of the issues and the development of high order thinking skills.

We have attempted to provide and validate an assessment model for the follow-up of students' performance in analogical problem solving. We hope that this model will be found by teachers not only as a source of descriptive information for problem solving, on the individual and class levels, but also as a prescriptive tool that can be helpful in their attempts to promote students' learning and self-efficacy. The suggested assessment model that was structured as a mapping sentence has a potential to provide teachers with a generic model for dividing a complex solution process into sub-processes. As such it can serve teachers in constructing their own models to different kinds of problems, different fields, different students and different kinds of learning situations.

References

Alexander, P., & Judy, J. (1988). The interaction of domain-specific and strategies performance. *Review of Educational Research*, *58*(4), 375–404.

Anderson, J.R. (1995). *Cognitive psychology and its implications*. New York: Freeman.

Baker, E.L., O'Neil, H.F., & Linn, R.L. (1993). Policy and validity prospects for performance-based assessment. *American Psychologist*, *48*(12), 1210–1218.

Berger, D.E., & Wilde, J.M. (1988). A task analysis of algebra word problem. In D.E. Berger, K. Pezdek, & W.P. Banks (Eds.), *Applications of cognitive psychology: Problem solving education and computing* (pp. 123–137). Hillsdale, NJ: Erlbaum.

Carnine, L., & Carnine, D. (2004). The interaction of reading skills and science content knowledge when teaching struggling secondary students. *Reading and Writing Quarterly, 20*(2), 203–218.

Chappell, M.F., & Thompson, D.R. (1999). Reflections on practice: Modifying our questions to assess students' thinking. *Mathematics Teaching in the Middle School*, *4*(7), 470–74.

Davidson, J.E. (1986). Insight and giftedness. In R.J. Sternberg, & J.E. Davidson (Eds.), *Conceptions of giftedness* (pp. 201–222). Cambridge, England: Cambridge University Press.

Dienes, Z., & Perner, J. (2002). The metacognitive implications of the implicit-explicit distinction. In P. Chambres, P.J. Marescaux, & M. Izaute (Eds.), *Metacognition: Process, function, and use* (pp. 241–268). Dordrecht, Netherlands: Kluwer Academic Publishers.

Fernadez-Duque, D., Baird, J., & Posner, M. (2000). Awareness and metacognition, *Consciousness* and *Cognition*, *9*(2), 324–326.

Fan, N., Mueller, J.H., & Marini, A.E. (1994). Solving difference problems: Wording primes coordination. *Cognition and Instruction, 12*(4), 355–369.

Geary, D.C., & Brown, S.C. (1991). Cognitive addition: Strategy choice and speed - of processing differences in gifted, normal and mathematically disabled children. *Developmental psychology, 27*(3), 398–406.

Gehrke, N., & Dickinson, D. (2002). *Assessment terminology: A glossary of useful terms*. Retrieved from www.newhorizons.org/strategies/assess/terminology.htm

Gobbo, C., & Chi, M. T. H. (1986). How knowledge is structured and used by expert and novice children. *Cognitive Development*, 1, 221-237

Gorodetsky, M., & Klavir, R. (2003). What do gifted/non-gifted problem solvers do? *Learning and Instruction, 13*(3), 305–325.

Klavir, R., & Gorodetsky, M. (2001). The processing of analogous problems in the verbal and visual-humorous (cartoons) modalities by gifted/average children. *Gifted Child Quarterly*, *45*(3), 205–214.

Low, R., & Over, R. (1992). Hirarchical ordering of schematic knowledge relating to area-of-rectangle problems. *Journal of Educational Psychology, 84*(1), 62–69.

Mannes, S. (1994). Strategic processing of text. *Journal of Educational Psychology, 86*(4), 577–588.

Maniatis, E., Cartwright, G.F., & Shore, B.M. (1998). Giftedness and complexity in a self-directed computer-based task. *Gifted and Talented International*, *13*(2), 83–89.

Martin , W.E. Jnr, Swartz-Kulstad, J.L., & Madison, M. (1999). Psychosocial factors that predict the college adjustment of the first-year undergraduate student. *Journal of College Counseling, 2*(2), 121–33.

McCleery, J.A., & Tindal, G.A. (1999). Teaching the scientific method to at-risk students and LD students through concept anchoring and explicit instruction. *Remedial & Special Education, 20*(1), 7–18.

Nelson, T.O. (1999). Cognition versus metacognition. In R.J. Sternberg (Ed.), *The nature of cognition* (pp. 625–641). Massachusetts Institute of Technology: The MIT Press.

Novick, L.R. (1988). Analogical transfer, problem similarity and expertise. *Journal of Experimental Psychology: Learning, Memory and Cognition, 14*(3), 510–520.

Novick, L.R., & Holyoak, K.J. (1991). Mathematical problem solving by analogy. *Journal of Experimental Psychology: Learning, Memory and Cognition, 17*(3), 398–415.

Pajares, F., & Schunk, D.H. (2001). Self-beliefs and school success: Self-efficacy, self-concept, and school achievement. In: R. Riding., & S. Rayner (Eds.), *Perception* (pp. 239–266). London: Ablex Publishing.

Peled, Z., & Wittrock, M.C. (1990). Generated meaning in the comparison of word problems in mathematics. *Instructional Science*, *19*(3), 171–205.

Rabinowitz, M., & Glaser, R. (1986). Cognitive structure and process in highly competent performance. *Learning Research and Development Center: University of Pittsburgh*, 15, 75–98.

Reiter-Palmon, R., Mumford, M.D., O'Connor Boes, J., & Runco, M.A., (1997), Problem construction and creativity: The role of ability cue consistency, and active processing. *Creativity Research Journal*, *10*(1), 9–23.

Robinson, A., & Clinkenbeard, P.R. (1998). Giftedness: An exceptionality examined. *Annual Review of Psychology*, *49*(23), 117–119.

Rosenthal, D.M. (2000). Consciousness and metacognition. In: D. Sperber (Ed.), *Metarepresentation: Proceedings of the Tenth Vancouver Cognitive Science Conference*. New York: Oxford University Press.

Sanders, W.L., & Horn, S. (1995). Educational assessment reassessed: The usefulness of standardized and alternative measures of student achievement as indicators for the assessment of educational outcomes. Retrieved from http://epaa.asu.edu/epaa/v3n6.html

Schmid, U., Wirth, J., & Polkehn, K. (2003). A closer look at structural similarity in analogical transfer. *Cognitive Science Quarterly*, *3*(1), 57–89.

Schoenfeld, A.H. (1985). *Mathematical problem solving*. New York: Academic Press.

Schunn, C.D., & Anderson, J.R. (1999). The generality/specificity of expertise in scientific reasoning. *Cognitive Science*, *23*(3), 337-370.

Seeley, M.M. (1994). The mismatch between assessment and grading. *Educational Leadership*, *52*(2), 4–6.

Siegler, R.S. (1989). Hazards of mental chronometry: An averaging data over strategies: An example from subtraction. *Journal of Educational Psychology*, *81*(4), 497–506.

Simon, D.P., & Simon, H.A. (1978). Individual differences in solving physics problem. In R.S. Seigler (Ed.), *Children's thinking: What develops?* (pp. 325–361). Hillsdale, NJ: Erlbaum.

Sowell, E. J. (1993). Programs for mathematically gifted students: A review of empirical research. *Gifted Child Quarterly*, *37*(3), 124–132.

http://sleid.cqu.edu.au
6(3), pp. 60–79. December 2009

Sperling, R., Howard, B., & Staley, R. (2004). Metacognition and self-regulated learning constructs, *Educational Research and Evaluation*, *10*(2), 117–139.

Thomasenia, A.L. (1998). Alternative assessment in elementary school mathematics. *Childhood Education*, *74*(4), 220–224.

Zimmerman, B.J. (2002). Becoming a self-regulated learner: An overview. *Theory into Practice*, *41*(2), 64–67.

Appendix

Table 2 ***The relationship between the levels of correctness of the solution and the levels of selectivity, (A) before and (B) after analogical learning, beyond populations***

	Selectivity levels (no. of selective facets) / Levels of correctness	0 (No selectivity)	1	2	3	4	5	6 (High selectivity)	Total
A. Before learning	**Correct solution**	**1**(1.0[a]) (1.1[b])	**4**(3.9) (3.7)	**14**(13.7) (23.3)	**34**(33.3) (55.7)	**15**(14.7) (50)	**22**(21.6) (95.7)	**12**(11.8) (92.3)	**102**(100) (26.6)
	Partially correct solution	**3**(9.1) (3.4)	**7**(21.2) (6.4)	**9**(27.3) (15.0)	**4**(12.1) (6.6)	**8**(24.2) (26.7)	**1**(3.0) (4.3)	**1**(3.0) (4.3)	**33**(100) (7.7)
	Erroneous solution	**81**(33.1) (93.1)	**98**(40) (89.9)	**36**(14.7) (60)	**23**(9.4) (37.7)	**7**(2.9) (23.3)	**0**(0) (0)	**0**(0) (0)	**245**(100) (64.0)
	No final solution	**2**(66.7) (2.3)	**0**(0) (0)	**1**(33.3) (1.7)	**0**(0) (0)	**0**(0) (0)	**0**(0) (0)	**0**(0) (0)	**3**(100) (0.8)
	Total	**87**(22.7) (100)	**109**(28.5) (100)	**60**(15.7) (100)	**61**(15.9) (100)	**30**(7.8) (100)	**23**(6.0) (100)	**13**(3.4) (100)	**383**(100) (100)
B. After learning	**Correct solution**	**0**(0) (0)	**2**(1.3) (2.9)	**6**(3.9) (13.0)	**12**(7.8) (28.6)	**21**(13.7) (60)	**29**(19.0) (93.5)	**83**(54.2) (100)	**153**(100) (43.0)
	Partially correct solution	**3**(7.5) (6.1)	**10**(25.0) (21.7)	**10**(25.0) (21.7)	**8**(20) (19.0)	**7**(17.5) (20)	**2**(5.0) (6.5)	**0**(0) (0)	**40**(100) (11.2)
	Erroneous solution	**39**(31.0) (79.6)	**48**(38.1) (68.6)	**19**(15.1) (41.3)	**16**(12.7) (38.1)	**4**(3.2) (11.4)	**0**(0) (0)	**0**(0) (0)	**126**(100) (35.4)
	No final solution	**7**(18.9) (14.3)	**10**(27.0) (14.3)	**11**(29.7) (23.9)	**6**(16.2) (14.3)	**3**(8.1) (8.6)	**0**(0) (0)	**0**(0) (0)	**37**(100) (10.4)
	Total	**49**(13.8) (100)	**70**(19.7) (100)	**46**(12.9) (100)	**42**(11.8) (100)	**35**(9.8) (100)	**31**(8.7) (100)	**83**(23.3) (100)	**356**(100) (100)

Table 3: *The correlations between the level of correctness of the solution and the level of selectivity for each of the six facets, before and after learning, beyond populations*

The facets / Levels of correctness	Encoding	Retrieval	Relating	Combin-ation	Goal Directn-ess (G.D.)	Awaren-ess
Final solution *before learning*	**0.313**[a] (0.000 [b]) *383* [c]	**0.241** (0.000) *383*	**0.343** (0.000) *383*	**0.781** (0.000) *381*	**0.522** (0.000) *381*	**0.459** (0.000) *370*
Final solution *after analogical learning*	**0.520** (0.000) *356*	**0.480** (0.000) *356*	**0.543** (0.000) *356*	**0.844** (0.000) *354*	**0.737** (0.000) *354*	**0.424** (0.000) *347*

[a]Pearson Correlation [b]Significance(2 tailed) [c]Number of Cases

Table 4: *The effect of analogical learning on the quality of the final solution and the selectivity of the solution process, beyond populations*

	The quality of final solution *[Level of correctness (Facet G)]*	**The quality of solution process** *[Level of selectivity of the 6 facets (A-F) within the profiles]*
Before Learning	1.61[a] (0.89 [b]) 383[c]	1.83 (1.63) 437
After Learning	1.87 (1.09) 356	3.04 (2.17) 356
Z[d](P[e])	6.857 (0.000)	10.090 (0.000)

[a]Mean Score [b]s.d. [c]Number of cases [d]Wilcoxon Signed Ranks Test [e]Significance

Table 5: ***The capability of the assessment model in reflecting excellence in the solution and the process***

	The quality of the solution		The quality of the solution	
	Before Learning	**After Learning**	**Before Learning**	**After Learning**
Gifted vs. Regulars	M[a]=1.9(117[b]) vs. M=1.3(121) U[c]=4914.500 W[d]=12295.500 Z=4.740 P=0.000	M=2.4(114) vs. M=1.4(105) U=2852.000 W=8417.000 Z=7.164 P=0.000	M=2.7(125) vs. M=1.2(148) U=4660.000 W=15686.000 Z=7.206 P=0.000	M=4.2(114) vs. M=1.7(105) U=2229.500 W=7794.500 Z=8.137 P=0.000
Experts vs. Novices	M=1.8(58) vs. M=1.4(87) U=1991.500 W=5819.500 Z=2.597 P=0.009	M=2.2(56) vs. M=1.6(81) U=1592.000 W= 4913.000 Z=3.161 P=0.002	M=2.3(63) vs. M=1.3(101) U=1882.500 W=7033.500 Z=4.543 P=0.000	M=4.1(56) vs. M=2.3(81) U=1177.0500 W=4498.000 Z=4.853 P=0.000
Excellent solvers vs. Regular solvers	M=1.9(175) vs. M=1.4(208) U=13262.500 W=34998.500 Z=5.396 P=0.000	M=2.3(170) vs. M=1.5(186) U=8923.500 W=26314.500 Z=7.596 P=0.000	M=2.6(188) vs. M=1.3(249) U=12519.000 W=43644.000 Z=8.522 P=0.000	M=4.2(170) vs. M=2.0(186) U=6852.000 W=24243.500 Z=9.370 P=0.000

[a]M=Mean score. [b]Number of cases [c]U = Mann-Whitney U. [d]W = Wilcoxon W