What teachers assess and its consequences

Sabiha ODABASI CIMER[1] and Atilla CIMER

Karadeniz Technical University, Fatih Faculity of Education
Department of Elementary Science Education
61335-Sögütlü-Trabzon, TURKEY

E-mail: sabihaodabasi@gmail.com and cimeratilla@yahoo.com

[1]Correspondence author

Received 1 Mar., 2010
Revised 21 Sept., 2010

Contents

Abstract

This study aimed to investigate teachers' perceptions and practices related to student outcomes and the criteria of student assessments in secondary school biology classes. Teacher interviews and documents provided the data for this study. According to the results of this study, teachers mainly assessed cognitive achievement by evaluating students' knowledge of basic facts, involving recall and recognition. Teachers also assessed non-cognitive characteristics to assign end-of-term grades. In addition, there were discrepancies between what teachers

claimed to assess and what they actually assessed. Furthermore, teachers' expectations of students changed with the perceived level of student ability. Consequences of these findings on student learning are discussed throughout the paper, and suggestions for teacher education and practice are also provided.

Keywords: Assessment, science learning

Introduction

In education, "assessment" is a broad term that includes all activities that teachers use to help students learn and to gauge student progress. Recently, assessment has been a major discussion topic in the reform of education to increase student learning and achievement in schools. The main theme of these discussions is that assessment can be an effective tool to monitor and shape student learning toward intended learning outcomes (Black and Wiliam, 1998).

There is ample evidence suggesting that students adopt an approach to learning depending upon what is expected of them (Marton and Saljo, 1976; Thomas and Bain, 1984; McKeachie, 1986; Biggs, 1987; Entwistle, 1988; 2000; Crooks 1988; Biggs, 1992; 1995; Ramsden, 1992; Biggs and Moore, 1993; Scouller, 1996; Cizek, 1997; Gerstman and Rex, 2001). In other words, students use study strategies according to the knowledge and skills assessed. Furthermore, "*if a concept, skill or knowledge chunk is deemed to be examinable, then a high priority is given in the learning strategy of the student*" (Bennett, 2002).

Scouller (1996) investigated students' perceptions of and preparation for different assessments and indicated that there were distinct patterns of learning for various assessment methods. If students perceived an assessment that targeted lower levels of knowledge, they were more likely to employ surface learning approaches. In contrast, they were likely to employ deep learning approaches when preparing for assessments that they expected would assess higher levels of knowledge.

These results were also supported by Bol and Strage (1996). Bol and Strage (1996) studied the assessment practices of teachers in the United States and found that students did not engage in more advanced study skills because the course exams and other assignments did not demand it. They concluded that the lack of correlation between achievement goals and assessment practices could explain why "*students do not develop the study skills necessary to tackle more complex and higher order kinds of instructional tasks that requires problem solving and critical*

thinking " (p. 159). In addition, Cotton (1988) synthesized 37 research reports on the relationship between the classroom questioning behavior of teachers and the quality of student outcomes. They found that when a teacher asked higher cognitive questions and increased wait time, the cognitive sophistication of student responses increased.

Therefore, asking higher-order questions helps students develop higher-order thinking skills (Hernstein, et. al., 1986; Robinson, 1987; Cotton, 1988, Mathews, 1989; Baum, 1990; Bol et al., 2000; Schouler, 2000). As research has shown, assessment that encourages students to think for themselves, apply their knowledge to new contexts, and solve problems may shift students toward a deep learning approach (Schouler, 2000). In contrast, asking questions requiring no more than the accurate reproduction of information can lead to a predominantly surface learning approach (Entwistle, 1988). Changing the method of assessment in a course can offer an effective mechanism for changing student approaches to learning (Brown, Bull and Pendlebury, 1997).

Thus, what teachers assess has a significant influence on what and how students learn. In other words, what is learned is determined by decisions students themselves make and, more importantly, the decisions the students make are informed by classroom assessments. Therefore, it is important that assessments demand and reflect the kind of learning that teachers want their students to develop. To promote deep learning in students, it is crucial that assessments reflect the requirement for deep learning (Gerstman and Rex, 2001). Furthermore, assessments that require recollection, in which teachers and the textbooks provide answers without students needing to understand the underlying concept, should be reduced to a minimum.

Based upon this research evidence, the present study aimed to reveal teachers' perceptions and practices related to student outcomes and criteria against which students are assessed in secondary school biology classes. For this purpose, teacher interviews were conducted, and the questions that teachers asked to their students were also analyzed. The overarching research question in this study was what are the valued student outcomes and criteria against which students are assessed in secondary school biology classes?

Method

This study involved a small-scale survey that employed both qualitative and quantitative approaches.

1. Data collection instruments

Data for the research were collected through semi-structured interviews and document (i.e., examination papers) analysis.

1.1. Interviews

To answer the research question and to make the analysis more manageable, a semi-structured interview was used for the study (Robson, 1993).

Twenty-four interviews with teachers were conducted, and all of the interviews were tape-recorded. The interview schedule was mainly guided by the research question. The interviews included two sections. The first section of each interview included questions about the teacher's background information and also assessed the teacher's broad perceptions. This section sought to obtain biographical information, each teacher's perception of their role as a teacher, and a list of professional development activities. The information from this section of the interview was intended to provide background information to generally explain teachers' reported actions and perceptions of what they assessed.

The second section of the interview included questions related to the kind of learning and skills that teachers assessed. Previous studies have stressed the importance of assessment in targeting higher-order thinking skills (Brown et al., 1997). This is also an official requirement of assessment policy in Turkey. Thus, in this section of the interview, teachers were asked about the tasks that they assigned, the kind of questions they asked in the classroom, and the questions that they asked in their examinations. The goal of this section of the interview was to further to elicit their criteria for the judgment of learning and understanding.

1.2. Documents

All of the teachers interviewed provided examination papers that they had graded. Specifically, six teachers provided two copies of two examinations they had conducted at different times. Thirty examination papers were collected, and there

were a total of 370 questions in these papers (230 short answer questions and 140 multiple-choice test questions).

2. Participants

Twenty-four biology teachers from 16 state secondary schools in Ankara, Turkey, participated in the study.

The schools were selected from the Ministry of Education's directory of state schools. Schools were selected from three districts that have different socio-economic levels. Sixteen schools in total were randomly selected from the directory, including five schools from District 1, six schools from District 2 (the largest district of the three), and five schools from District 3. The schools sampled were all state high schools.

Each school had two to eight biology teachers, for a total of 62 biology teachers across the studied schools. Twenty-four of the teachers volunteered to be interviewed. Each teacher was assured of confidentiality.

The table below summarizes the biographical information of the teachers that participated in the study.

Table I: *Teacher characteristics*

		Interview sample	
		Frequency (n=24)	*Percentage (%)*
Gender	*Male*	*11*	*46*
	Female	*13*	*54*
Educational qualification	*B.Ed.*	*21*	*88*
	Master's Degree	*3*	*12*
	Ph.D.	*0*	*0*
Years of teaching experience	*6-10*	*9*	*37.5*
	11-19	*9*	*37.5*
	20-25	*6*	*25*

The sample included 11 male teachers and 13 female teachers. The teaching experience of the teachers varied from 6 to 25 years. Thus, the majority of the

teachers were experienced teachers. Three teachers held a Master's Degree in biology.

3. Data analyses

The interview questions were analyzed using Bloom's taxonomy of knowledge levels. The author and another expert in this field first analyzed each question separately. Next, the results were compared, and when there was a discrepancy between the results, the item was discussed until an agreement was reached.

The interview data were analyzed following step-by-step guidelines advocated by Krueger and Casey (2000), Miles and Huberman (1994) and Bogdan and Biklen (1992).

After completing the data collection, the raw data were transformed into a readable form for analysis. Tape-recorded interviews were transcribed verbatim as soon as possible and then were entered into a word processing document.

The transcripts were read carefully several times to identify emerging themes, topics or concepts that had the potential to answer the study's questions (*i.e.*, Bogdan and Biklen, 1992; Miles and Huberman, 1994). These themes, topics, and concepts (hereafter referred to as "codes") were written in the right margin of the paper. This search was continued until no new codes emerged. The two researchers first conducted this analysis separately. Next, the results were compared and discussed. Consensus was reached on all items.

The codes were then grouped together to form broader categories with strong commonalities (*i.e.* Miles and Huberman, 1994). Two main categories of "cognitive" and "non-cognitive" criteria emerged from the qualitative data analysis. There were also sub-categories. For example, the category of non-cognitive criteria contained codes that categorized students as "respectful," "completes homework," "obeys rules," "polite," and so on.

Throughout this study, each teacher was identified as teacher (T), followed by a number to protect the teacher's identity.

Results and discussion

To understand what each teacher's assessment valued, teachers were invited to talk during the interviews about the nature of the questions they asked on their exams and the qualities they wanted to see in their students. The examination papers they provided also supplied tangible evidence of what they assessed.

Analysis of the data suggested that the teachers based their assessments on two types of criteria: cognitive and non-cognitive.

1. Cognitive criteria

Cognitive achievement is related to the acquisition of knowledge and learning skills. The kind of knowledge and skills assessed in Turkish schools is defined in the official assessment policy, which states that:

> …in addition to assessing the acquisition of knowledge, qualities of comprehension, application, analysis, synthesis and evaluation should also be assessed (MONE, 2001b: Article 5k).

The official policy emphasizes the importance of assessing students' higher-order thinking skills. However, the application of this approach in the classroom is dependent upon whether teachers' assessments consider these different levels of thinking.

In this study, the cognitive achievement that teachers assessed was apparent both in their definitions of their assessments during the interviews and in the examination papers they provided. Two of the teachers (T11 and T17) defined biology as a subject "*based on text, so memorization is inevitable.*" All of the other teachers emphasized that they wanted students to go beyond memorization. One of the teachers (T7) said:

> Sometimes they say, "let's not do the lesson. You just give us a list of questions, and on the exam ask from this list." They will memorize the answers. What I try to do is to prevent memorization. How can I do this? How can I teach them think critically? How can I see their abilities and understanding differently? I do not think I do this but I try. While I try to do this students tend to do exactly the opposite. I do not believe that memorization is any good for students. They

> just memorize but do not know what it means. They need to know how to use their brains.

Although most teachers indicated that they did not want to encourage memorization, analysis of the interview data and the examination papers together suggested otherwise. The data showed that students were expected to absorb knowledge transmitted by the teacher or found in the textbook in a ready-to-use format and then reproduce that information on the exams. The main concern of the teachers, thus, was not to ascertain what the students knew but to determine whether specific facts were known. One teacher's (T10) comments exemplified this:

> I present them the important parts of the topic and I ask them on the exam. During the lessons, I tell them that "I ask this…I ask that on the exam, be careful.

This explanation defines assessment as a process of cueing rehearsed answers in the lessons (Wiggins, 1990).

The ability to remember important information was frequently mentioned by the teachers:

> I ask short-answer, easily-remembered important details. I already dictate notes during the lessons and on the exams I try to ask them… (T23).

This highlights the overemphasis of the assessment of low-level skills (*i.e.*, recall versus recognition).

The practice of providing students with an answer that they can simply memorize in preparation for the exam reduces the cognitive demands associated with the exam (Strage, Tyler, Thomas and Rohver, 1987; Bol and Strage, 1993). Through this approach, the teacher sends a message to students about what is important. Therefore, it is not surprising that the students only learn the answers to these questions (Wiliam, 2003). This means that students are engaging in less productive types of studying. Furthermore, the goals of science learning that call for critical thinking and more sophisticated study strategies may not be realized (Bol and Strage, 1996).

It is interesting to note that two of the teachers admitted that they did not prefer asking questions requiring interpretation or critical thinking in their exams,

especially for the low-achieving groups of students. These teachers expressed concern about the inclusion of critical thinking questions in their exams, pointing out that their students were not able to cope with such questions. One of the teachers said:

> I do not ask questions requiring explanation or interpretation… They cannot do that… It should be asked though…. Yet students cannot even write what is written in their textbook or a sentence they wrote in their notebook during the lessons, which I had dictated to them or they had copied from the blackboard, let alone adding something on to that…. in some high achieving classes sometimes I do that… but especially Year 9 students… for example, they are very bad at interpreting graphs… (T16).

This comment highlights the teacher's stance regarding the weaknesses of the students. Although s/he thinks that questions requiring explanation and interpretation by the students should be asked, s/he avoids such questions, and students are not challenged because s/he see students as unable to answer such questions.

Hence, students are sometimes not asked interpretation or thinking questions because of the concern that they cannot cope. Furthermore, assignments requiring research are avoided because the students do not know how to do research. Additionally, because students are not seen as knowledgeable enough or able to make valuable contribution to the lessons, their input is not asked for in the lessons. One of the teachers said, "*when I call on students they only repeat what is written in the textbook... they do not say anything different at all*" (T16). One teacher indicated that previous years of students were not taught in an environment where they were invited to contribute or discuss their ideas, which suggests that the students in general are not used to these teaching and learning processes and that they do not have thinking abilities. In addition, the social and family background of the students and the students' lack of ability were also indicated as the reasons for their low expectations of teachers. In any case, underachievement was not thought of as the responsibility of the teacher, and the teachers repeatedly indicated that they could not do much to remedy accumulated weaknesses in the students.

The analysis of the questions in the exam papers also supported the teachers' emphasis on assessing basic facts and asking what is given in the textbook and by the teacher in the lessons. In other words, the exams supported the practice of

cueing rehearsed answers in the lessons. Table II provides the distribution of the questions in the exam papers according to Bloom's taxonomy of cognitive levels (Bloom, 1956) and provides the criteria on which the questions were grouped.

As can be seen in Table II, 249 out of 370 questions (67%) were knowledge questions. These questions mainly asked for the knowledge of basic facts, details, terms and definitions, which essentially required students to write pre-defined answers. The words identified in these questions were: "what," "where," "name," "state," "define" and "list." According to the analysis results, the questions in this group fall into three groups: definition questions, list questions and questions asking general facts. Examples of the questions include the following: "What is muscle tonus?" (definition); "Name the tropism movements" (list); and "In which canal is the corti organ found?" (general facts). These types of questions require students to recall the definition given by the textbook or the teacher.

Table II. *Distribution of the examination questions according to Bloom's taxonomy (n=370)*

		Questions		Total %
		n	%	
Knowledge	Definitions/descriptions	29	7.8	67.3
	List	37	10.0	
	General facts (Where is…?; How many…?)	183	49.5	
Comprehension	Comparison (What is the difference between…?)	12	3.2	16.7
	Explanation (why is this so?)	50	13.5	
Application	Application of standard algorithms and making simple calculations	42	11.4	11.4
Analysis	Read graphs, make inferences/deduce/conclude	10	2.7	4.6
	Relationships (How are these concepts related?)	7	1.9	

(Source: Examination papers)

Comprehension questions test students' understanding of scientific knowledge, and their knowledge of facts, theories and procedures are assumed. The questions in this group require students to make comparisons, order steps in a process,

recognize meaning or explain why. An example of a comprehension question is: “Why does glycolysis occur in the cytoplasm in aerobic respiration?” **(**explanation**)**. The answer to this question is: “Because its enzymes are in the cytoplasm.”

It is noteworthy that the answers to 35 of the 62 comprehension questions on the exam papers were drawn from the examples given in the lessons. Examples below illustrate this:

> Question: What is the role of the air sacs in the birds’ skeleton? (Explanation)
>
> Student’s answer: To be able to fly easier and faster using the air.

This question seems to be a good example of a question that requires students to think. However, at the same time it is the exact example that is written in the students’ notebooks: "Birds’ skeletons are adjusted to enable them to fly. In their skeletons, there are air sacs to reduce the body weight of the bird." Hence, the question requires nothing more than what is written in the student’s notebook.

Another example is:

> Question: What is the difference between regeneration in human epidermis and that in a sea star? Explain.
>
> (Comparison)
>
> Student’s answer: Different. Sea star can regenerate a new sea star from an arm that had broken off but, since humans have various tissues (muscle, skeleton), the wound just heals.

One might argue that the level of this question is "analysis" if the student has not previously been told what the difference is (or read it from the textbook). However, this was also an example directly from the students’ notebooks.Hence, the previously given answer is sought.

Application of knowledge questions were considered to be the questions that asked students to use known algorithms to conduct simple calculations, as illustrated below:

Question: There are 1500 nucleotides in a DNA molecule. If the ratio of A+T/G+C equals to 2/3, what are the numbers of A and G?

(apply known equations and conduct calculations)

The students are expected to calculate the values of A and G. To answer this question, students should know that A=T and C=G. The rest requires application of this simple mathematical calculation.

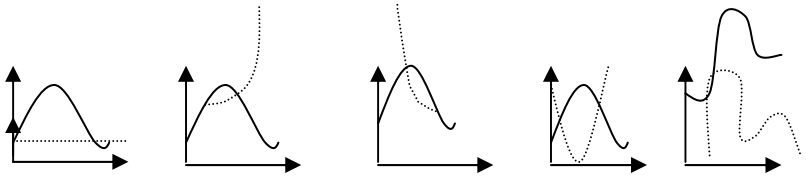
Analysis questions required reasoning and the identification of relationships between concepts; in other words, students were tested on their ability to explain how one thing is related to another thing. Only 17 questions fell into this category. Three main sub-groups could be identified from the analysis: considering the relationships between different concepts and with daily life, reading graphs and making inferences, and identifying the relationships between concepts.

The question below tests students' ability to identify the relationship between two concepts:

Question: Which of the following illustrates the relationship between contract-relax stretching of a muscle and the heat production? (reading graphs and identifying relationships)

(---------- muscle stretching, ---------- heat production)

amount



stimulus time

A **B** **C** **D** **E**

The question below tests students' ability to identify the relationship between aging and mitochondrial DNA:

The question below tests students' ability to identify the relationship between aging and mitochondrial DNA:

> Question: What is the relationship between aging and mitochondrial DNA? (relationships)

Overall, the analysis of the exam papers supported the interview data results that the knowledge of basic or "important" facts is seen as an important cognitive criterion and that the teachers mainly checked the extent of students' ability to recall (short answer) and recognize (multiple choice or true/false) these facts.

This emphasis on low-level knowledge or thinking skills is also suggested by other studies (Crooks, 1988; Linn, 1990; Gallagher, 1991; White and Gunstone, 1992; Bol and Strage, 1996; Dindar, 1996; Turan, 1998). Bol and Strage (1996) found that teachers excluded items that required higher-order thinking skills from their exams based on a perception that their students would be unsuccessful. Additionally, White and Gunstone (1992) reported that the majority of assessments used by many teachers measured only the recall of closed, low-level science information and algorithmic skills. They indicated that oral questions and written tests were often restricted to short answer questions, which do not measure understanding.

As indicated previously, the nature of learning is determined by the type of assessment that is anticipated (Crooks 1988; Ramsden, 1992; Biggs and Moore, 1993; Bol and Strage, 1996; Scouller, 1996; Cizek, 1997) because students inevitably use the strategy that they think will best achieve the highest grade. Hence, assessment has a crucial effect on what and how students learn. Given the evidence from previous studies, a likely consequence of teachers' assessments targeting low-level skills could be that they may encourage lower-level learning in students. They may also hinder students' ability to think and confine them to a very limited range of knowledge and activities embodied in the textbooks (Bol and Strage, 1996). The present study found that teachers' assessment predominantly targeted low-level skills. Therefore, this is important to consider in an attempt to improve student learning in schools.

2. Non-cognitive criteria

The term "non-cognitive criteria" is used to indicate the qualities that are not directly associated with students' academic progress. Official assessment policy states that:

> ….in addition to formal exam results while assigning end-of-term marks, teachers should consider students' qualities, such as participation in class activities, science attitudes, ability to observe, conduct research, take responsibility, work in groups and share ideas (MONE, 2001b; Article 5i).

Although judgment of cognitive achievement dominated their assessment practices, teachers in this study considered other additional factors. However, the factors that teachers considered in their assessments did not comprise all of the items in the policy. For example, qualities like the ability to observe, conduct research and work in groups were not an option for most of the teachers in the study because they did not conduct laboratory work, assign individual or group projects or use group work in their teaching. The qualities that the study's teachers considered were effort, general in-class behavior and personal characteristics.

2.1. Effort

The teachers in this study shared the view that effort was an important component of end-of-term marks, and 23 out of 24 teachers mentioned taking effort into account while assigning end-of-term marks. The teachers' main reason for considering effort in end-of-term marks was a concern for making the assessment fair by assigning extra marks to those who studied during the term. One of the teachers defined assessment as a means to show students that their efforts were appreciated by the teacher:

> In the end, we need to distinguish between students who study and who do not so that students will see that their effort is appreciated and rewarded. Assessment is the only way to do that (T17).

One of the teachers discussed how she distinguished between the students who studied and those who did not:

> ….. in the exams, I ask one or two questions they can answer easily and questions that can be counted as of medium difficulty, and then there are one or two difficult questions. I want to see how many of them can answer these questions. These last ones show who really studied (T8).

By asking questions of different difficulty levels on the exams, the teacher (T8) made judgments about his/her students regarding the effort they exerted. This teacher went on to say that "*these students are rewarded at the end-of-the term*" with extra marks.

It is noteworthy that six of the teachers indicated that students who tried hard or put effort in during the term would be given a passing grade, even if the formal exam results were failure. One of the teachers said:

> I have some students, I know that they study, they really put an effort but they cannot be successful on the written exams. I cannot fail them (T24).

Taking effort into account in this sense was a means to compensate for the effort expended. The interviewees mentioned that written exams might not always reflect the students correctly. They argued that students might not show their real performance in the written exams for different reasons, such as illness or lack of preparation, or that students might know the answer but could not remember it during the exam. More importantly, the teachers also indicated that students might not be the kind of students who were good at memorizing and that is why they received low marks on the written exams. Therefore, they believed that it was necessary to consider effort when assigning end-of-term marks.

Such a practice also seems to operate as a reward system based on the relationship between marks and student motivation. The underlying assumption is that if students know that a particular behavior will earn them a mark, they will behave accordingly. Indeed, students may be motivated by the prospect of getting high marks and expend extra effort. However, Stiggins and Conklin (1992) warned that such a relationship may hold true for good students, but for students whose experienced such a system and failed in the past, marks will be nothing more than a reminder of failure, and they will stop trying. Such practices may also lead to the students only trying for marks (Rowntree, 1987; Kohn, 1993) or, as one of the interviewees in this study indicated, result in students with a "materialist view of school education."

2.2. General conduct and personal characteristics

General conduct in the classroom (*i.e.* paying attention to the teacher, keeping quiet and obeying classroom rules) and personal characteristics (*i.e.* politeness and

showing respect to others) were also mentioned by all of the teachers as factors that they considered when assigning end-of term marks. One teacher said:

> At the end of the term while assigning marks, I consider their participation, effort and certainly their personal characteristics, too (T23).

Another teacher indicated:

> If a student's general behavior has been good during the term, we raise his/her end-of-term mark (T10).

The teachers' emphasis on students' general conduct and personal characteristics was mainly a reflection of societal values. All of the teachers interviewed indicated that they saw themselves mainly responsible for educating students and that the mission of teaching came later. In other words, "making good citizens" was the primary responsibility of teachers.

Furthermore, the teachers' practice of considering general conduct and behavior during the term when assigning end-of-term marks also served managerial purposes. Again, linking a reward in the form of extra marks to the end-of-term marks for good behavior would make it easier to manage student behavior in the classroom.

Overall, it seemed that marks were treated as currency in the teachers' hands and that good marks were paid to students for their good traits. However, marks were generally used in the positive way, as one teacher said: "The oral exam marks I assign are not lower than the average of the written exam marks, even higher than that" (T3). Thus, in general, the teachers explicitly admitted being generous with marks to students who showed some evidence of having worked hard and/or behaved well. The teachers frequently indicated increasing students' end-of-term marks based on the consideration of students' effort and behaviors. Similar findings were also reported by previous studies (Brookhart, 1991; Cross and Frary, 1996; Cimer, 2004).

Rowntree (1987) reported that students with personalities favored by their teachers were assigned better grades than their objective performances seemed to justify:

> The teachers' knowledge may be used with prejudice - positively or negatively. If he receives generally "good vibrations" from a student he may over rate the student on many constructs that contribute in no way to the "vibrations"

> conversely, the student from whom he has conceived a distaste or lack of affection may be unfairly "marked down" on constructs that were not influential in shaping the early poor impression (p. 115).

Brookhart (1991) showed that grading often consisted of a "hodgepodge" of attitude, effort and achievement. Similar findings concerning the mixed nature of grades were also reported by Cross and Frary (1996). They reported that teachers raised the grades of low-ability students and those students who showed high effort. Thus, they took student conduct and attitude into consideration when assigning grades. More importantly, the students largely confirmed and supported the use of conduct and attitudes for determining grades and indicated that the students' ability level should be considered to be fair to all students.

Conclusions and implications

Students' final marks consisted of cognitive achievement as well as non-cognitive criteria

Teachers' cognitive criteria were mainly comprised of basic facts, recall and recognition. Certainly, achievement requires knowledge of important facts and concepts. However, for this knowledge to be usable, pieces of information need to be interrelated conceptually (Ausubel, 1968).

Learning science is a process of building a knowledge structure. Assessing such structural aspects of knowledge through tests and short answer exams seems difficult because these assessments cannot show how key concepts are mentally organized by a student. A picture of the "cognitive structure" of the student is needed, and to see this picture, teachers need more than tests. Student explanations, for example, can be used. Asking students to explain a situation in their own words may reveal what students know as well as how they connect concepts in their knowledge structure and use the information to find solutions to problems.

In addition, students develop skills that are assessed, and we cannot expect students to engage in higher-order thinking skills unless this is demanded by the assessments. The assessments that teachers reported in the present study may produce students who have factual knowledge but may not foster individuals who reason, think and solve problems. Hence, assessment in schools should target the valued outcomes of science learning and teaching in today's world, which place

greater emphasis on the students' ability to inquire, reason scientifically, apply scientific concepts to real-world situations, and communicate. Students need to think, reason, and draw conclusions rather than trying to understand the conclusions that teachers draw for them.

The official assessment policy in Turkey also emphasizes the development and assessment of students' higher-order thinking skills. However, data from this study shows that this is an unrealistic expectation, at least from the teachers' point of view, because their pedagogies are not supportive to the development of these skills.

There need to be different types of assessments for different styles of teaching and learning. The teachers' views about learning seemed to reflect their approaches to assessment. From the data, it was clear is that the behaviorist approach underlined the teachers' perceptions of teaching and learning. Specifically, they viewed teaching as telling and learning as remembering. However, this approach is not compatible with current theories of learning and assessment or with the policies that Turkish policy makers are trying to establish.

Students' effort and interest in the lessons and their approved and disapproved behaviors were also assessed to control students' study behaviors or make them study during the term and to be fair at the end of the term. Thus, students' end-of-term marks may contain criteria that are not related to biology knowledge. In other words, students who are well-behaved, attentive and diligent are given extra marks. In the present study, the teachers' emphasis on favoring students who listened to the teacher during the lessons and seemed interested resulted in a tendency to assign passing grades to students even though their written exam results were not good enough.

Consequently, in view of these findings and data from previous studies, I recommend training teachers in assessment through in-service and pre-service courses. Furthermore, using assessments effectively for formative purposes should be a priority.

There were discrepancies between what teachers said they assessed and what they actually assessed.

The results of this study showed that there was misalignment between what teachers said they assessed and what they actually assessed. During the interviews, all but two of the teachers indicated that they expected students to learn meaningfully and did not accept assessing memorization, yet memorization questions appeared on their tests, and their comments during the interviews pointed out that they mainly assessed recall and memorization abilities. Thus teachers could not accurately judge the cognitive complexity of the questions that they asked in assessments.

The analysis of the examination papers provided tangible evidence of this result. Over 67% of the questions asked were knowledge-level questions, requiring recall and recognition, and only approximately 15% were application- and analysis-level questions. There were no synthesis- or evaluation-level questions. Because assessments demanded mostly recall and recognition, it is not realistic to expect students to develop more sophisticated study skills or engage in higher-order thinking.

During the interviews, the teachers also indicated that while teaching a particular concept, they warned students that that particular concept might appear on the exam. The practice of providing students with exact replicas of items that appear on the exam or giving students an answer that they can simply memorize for the exam sends a message to the students that this is what is important to learn and, as a result, learning becomes memorizing or recall, making the schooling all about grades and passing exams.

Thus, there is a need to plan and conduct effective professional development initiatives, including both pre- and in-service training, to transform teachers' epistemologies in line with the current theories of teaching, learning and assessment. Unless teachers' underlying assumptions are assessed and refocused, their instructional and assessment practices cannot be transformed.

Teachers' expectations of students changed based on the perceived level of student ability.

During the interviews, conversations about students' learning frequently resulted in a statement of frustration about students, referring their lack of ability or lack of interest. This frustration had important implications on what teachers assessed.

An important finding of the study is that teachers avoided exam questions that required higher-order thinking skills, based on a perception that their students would be unsuccessful. Thus, if students were perceived as having low abilities or being unable to cope with the demands of a particular task, teachers avoided such tasks, and the students were left unchallenged. Hence, a perception of a lack of skill or ability resulted in avoiding challenging tasks rather than taking responsibility to improve that skill.

The teachers' perceptions of the students' abilities not only limited their practice to focusing on low-level skills, such as recall versus recognition items in their written exams, but also might result in students developing these skills and adopting a learning strategy based on rote learning. As indicated earlier, there is empirical evidence suggesting that students study in the way that they think they will be tested (McKeachie, 1986; Crooks 1988). If students expect an exam focused on facts, they will memorize details. However, if they expect a test that will require problem solving or integrating knowledge, they will work toward understanding and applying information.

Hence, teachers' expectations of students may affect student outcomes (Good and Brophy, 1987). Positive teacher expectations are recognized as a key variable that separate teachers who produce good achievement gains from those who do not (Rowntree, 1987; Imants and DeBrabander, 1996). It is important that teachers believe that every student can learn and that they appreciate their different learning needs and consider this in their teaching and assessment (*i.e.* the alternative assessment movement).

References

Ausubel, D. P. (1968). *Educational Psychology, A Cognitive View*. New York: Holt, Rinehart and Winston, Inc.

Baum, R. (1990). *Finishing Touches-10 Top Programs. Learning*, *18*(6), 51-55. http://ncrtl.msu.edu/issue.htm

Bennett, S. W., (2002) Chemistry teaching and learning: Is there a future?, *Chemistry in Action*, 60, 14-19.

Biggs, J.B. (1987). *Student Approaches to Learning and Studying. Australian Council for Educational Research*. Melbourne, Australia.

Biggs, J. (1992). The psychology of educational assessment and the Hong Kong scene. *Bulletin of the Hong Kong Psychological Society*, *28/29*, 5-26.

Biggs, J. (1995). Student approaches to learning, constructivism and student-centred learning. *Paper presented at the Twentieth International Conference on Improving University Teaching*, Hong Kong.

Biggs, J., and Moore, P. (1993). *The Process of Learning*. Sydney, Australia: Prentice Hall.

Black, P. and Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy and Practice*, 5 (1):7- 74.

Bloom, B. S. (1956). *Taxonomy of Educational Objectives. Handbook I: Cognitive Domain*. New York: David McKay.

Bogdan, R. and Biklen. S. (1992). *Qualitative Research for Education: An Introduction to Theory and Methods.* Boston: Allyn and Bacon.

Bol, L., Nunnery, J., Stephenson, P. and Mogge, K. (2000). Changes in teachers' assessment practices in the new American schools restructuring models. *Teaching and Change*, *7*(2), 127-146.

Bol, L., and Strage, A. (1993). Relationship among teachers' assessment practices and their student outcome and study skill development goals. *EDRS. ED367 639*.

Brookhart, S. M. (1991). Grading practices and validity. *Educational Measurement: Issues and Practice*, 10(1): 35-36.

Brown, G., Bull, J. and Pendlebury, M. (1997). *Assessing Student Learning in Higher Education*, London: Routledge.

Cimer, A. (2004). *Effective Biology Teaching: Turkish Biology Teachers' and Students' Views.* Unpublished EdD thesis, University of Nottingham, U.K.

Cizek, G. J. (1997). Learning, achievement, and assessment: Constructs at a crossroads. In G.D. Phye (Ed.) *Handbook of Classroom Assessment: Learning, Achievement, and Adjustment.* New York: Academic Press.

Cotton, K. (1988). *Classroom Questioning: Close-Up, Northwest Regional Educational Laboratory*, May 1988.

Crooks, T.J. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research*, *58*(4), 438-481.

Cross, L. H., and Frary, R. B. (1996). Hodgepodge Grading: Endorsed by Students and Teachers Alike. Paper presented at the annual meeting of the *National Council on Measurement in Education*, New York.
Dindar, H. (1996). *Ortaogretim Kurumlarinda Biyoloji Egitiminin Yapi ve Sorunlari*. Doktora tezi. Gazi Universitesi Fen Bilimleri Enstitusu. Ankara.

Entwistle, N. (1988). Motivational factors in students' approaches to learning, In R. R. Schmeck (Ed.) *Learning Strategies and Learning Styles*. New York: Plenum Press. pp. 21-49.

Gallagher, J.J. (1991). Prospective and practising secondary school science teachers' establishing sound knowledge and beliefs about the philosophy of science. *Science Education*, *75*(1), 121-133.

Gerstman, J., and Rex, J. (2001). Improving learning outcomes in a multicultural classroom. Paper presented *AARE 2001 International Education Research Conference*- Freemantle, Australia. December 2-6 2001

Good, T. L. and Brophy, J.E. (1987). Teacher expectations. Chapter 4 in T. L. Good and J E. Brophy, *Looking in Classrooms*, New York : Harper and Row. pp:93-121.

Herrnstein, R. J. Nickerson, R. S. de Sanchez, M. and Swets, J. A. (1986). Teaching thinking skills. *American Psychologist*, *41*(11), 1279-1289.

Imants, J. G. M., and DeBrabander, C. J. (1996). Teachers' and principals' sense of efficacy in elementary schools, *Teaching and Teacher Education*, 12(2), 179-95

Kohn, A. (1993). *Punished by Rewards: The Trouble with Gold Stars, Incentive Plans, A's, Praise, and Other Bribes*. Boston: Houghton Mifflin.
Krueger, R.A. and Casey. M.A. (2000). *Focus Groups. A Practical Guide for Applied Research*, Thousand Oaks: Sage.

Linn, R. (1990). Essentials of student assessment: from accountability to instructional aid. *Teachers College Record*. 91(3): 422-436.

Marton, F. and Saljo, R. (1976). On qualitative differences in learning - I: Outcome and process. *British Journal of Educational Psychology*, *46*, 4-11.

Matthews, D. B. (1989). The effect of a thinking-skills program on the cognitive abilities of middle school students. *Clearing House, 62/5*, 202-204.

Maykut, P. and Morehouse, R. (1994). *Beginning Qualitative Research: A Philosophic and Practical Guide*. London: The Falmer Press

McKeachie, W. J., (1986). *Teaching Tips.* Lexington, MA, D.C. Heath.

McMillan J.H. (2001). Secondary teachers' classroom assessment and grading practices. *Educational Assessment: Issues and Prcatice*, *20*(1), 20-32.

Miles, M.B., and Huberman, M. (1994) *Qualitative Data Analysis: An Expanded Sourcebook*, Thousand Oaks: Sage.

MONE (The Ministry of National Education) (2001a). *2002 Yili Basinda Milli Egitim.* T.C. Milli Egitim Bakanligi Arastirma, Planlama ve Koordinasyon Kurulu Baskanligi. Aralik 2001.

Ramsden, P. (1992). *Learning to Teach in Higher Education*. London: Routledge.

Robinson, I. S. (1987). *A Program to Incorporate High-Order Thinking Skills into Teaching and Learning for Grades K-3*. Fort Lauderdale, FL: Nova University. (ED 284 689).

Rowntree, D. (1987). *Assessing Students: How Shall we Know Them?* London: Kogan Page.

Scouller, K. (1996). Influence of assessment method on students' learning approaches, perceptions and preferences: The assignment essay versus the short answer examination.. Different Approaches: Theory and Practice in Higher Education. Proceedings *HERDSA Conference*. Perth, Western Australia, 8-12 July 1996. http://www.herdsa.org.au/confs/1996/scouller.html

Semb, G., & Spencer, R. (1976). Beyond the level of recall: An analysis of complex educational tasks in college and univer-sity instruction. In L. E. Fraley & E. A. Vargas (Eds.), *Behavior Research and Technology in College and Univer-sity Instruction* (pp. 115-126). Gainesville, FL: Department of Psychology, University of Florida.

Stiggins, R.J. and Conklin, N. (1992). *In Teachers' Hands: Investigating the Practices of Classroom Assessment*. State University of New York press.

Strage, A., Tyler, A.B., Thomas, J.W. and Rohwer, W.D. (1987). An analytic framework for assessing distinctive course features within and across grade levels. *Contemporary Educational Psychology*, *12*, 280-302.

Thomas, P.R., and Bain, J.D. (1984). Contextual dependence of learning approaches: the effects of assessments. *Human Learning*, *3*, 227-240.

Turan, E. (1998). *The Problems of Teaching in Biology in High Schools*. Masters Thesis. Graduate School of Applied Science of Dokuz Eylul University. Izmir, Turkey.

Watson, A. (2003). *Potential Sources of Inequity in Teachers' Informal Judgements about Pupils' Mathematics.* http://www.nottingham.ac.uk/csme/meas/papers/watson.html.

White, R. and Gunstone, R. (1992). *Probing Understanding*. Falmer Press.

Wiggins, G. (1990). The case for authentic assessment. *ERIC Digest*. http://ericae.net/edo/ED328611.htm.

Wiliam, D. (2003). National curriculum assessment: How to make it better. *Research Papers in Education*, *18*(2), 129-136.