A DIAGNOSTIC STUDY OF COMPUTER APPLICATION OF STRUCTURAL COMMUNICATION GRID

Associate Professor Dr. Mehmet BAHAR
Science Education Department, Abant Izzet Baysal University
e-mail: mehmet.bahar@gmail.com

Research Assistant Fatih AYDIN
Science Education Department, Gazi University
e-mail: fatihaydin14@gmail.com

Assistant Professor Dr. Erol KARAKIRIK
Primary School Education Department, Abant Izzet Baysal University
e-mail: karakirik@gmail.com

ABSTRACT
In this article, Structural communication grid (SCG), an alternative measurement and evaluation technique, has been firstly summarised and the design, development and implementation of a computer based SCG system have been introduced. The system is then tested on a sample of 154 participants consisting of candidate students, science teachers and faculty staff members. A comparative study between paper-and-pencil and computerized versions of SCG has been conducted. 154 subjects, consisting of candidate students, science teachers and faculty staff members, have participated in this study. Participants' views are probed by using Software Evaluation Questionnaire and Open-Ended Interview Form. Findings of the study show that i) participants find SCG technique easier to use in computer medium, ii) they have positive attitudes towards this computer software, and iii) they prefer computerized version of SCG technique to the paper-based SCG technique. The findings have been discussed in terms of its contribution to computer aided measurement and evaluation methods by taking similar studies in the literature into account.

Keywords: Structured Communication Grid, Software Evaluation Questionnaire, Measurement and Evaluation

INTRODUCTION

The views of educators towards learning, teaching and evaluation have experienced significant changes in the last thirty years. There has been a shift from behaviorist learning approach to constructivist learning approach, from teacher-centered education to student-centered education, from memorization to meaningful learning. This shift has also been reflected to the measurement and evaluation as being a part of learning and teaching process. Constructivist approach, which argues that individuals actively build their new knowledge upon prior ones in a way distinct to themselves, has questioned the effectiveness of traditional measurement and evaluation techniques that only focus on the learning product, and highlighted the utilization of alternative measurement and evaluation techniques which focus on the learning process.

The Science and Technology, Mathematics, Social Sciences and Turkish curriculums have been redesigned in Turkey in line with the constructivist paradigms. Accordingly, constructivism has been embedded into courses which are to be spirally distributed thorughout the school years and an inter-disiciplinary approach has been adopted. The most important change, perhaps, has been in the perceptions regarding measurement and evaluation. This change is more apparent in Science and Technology programs (Bahar, 2006).

Multiple choice tests are still most widely used, and at the same time criticized, assesment and evaluation techniques all around the world. The main drawback of a multiple choice test is to inform participants in advance that each question in the test, no matter how many choices it have, has only one correct answer. Participants are required to find the desired answer by eliminating the incorrect choices. Knowing that the question has just one correct answer may lead participants to focus on guessing the answer rather than thinking over the question which contradicts the basic aims of the assessment in the fisrt place. Hence, the usage of new assessment and evaluation techniques have been advocated in the new Turkish curriculum. Structured Communication Grid in this regard has been mentioned under the category of alternative measurement and evaluation in the 2004 Turkish Science and Technology Program (MEB, 2004).

Structured Communication Grid (SCG) Technique

Structured Communication Grid (SCG) technique was proposed and developed as an alternative assesment technique to multiple choice tests by trying to eliminate their drawbacks. SCG differs from traditional multiple choice tests in this regard that participants are not in advance aware of how many correct answers a question

have. Hence, it is very difficult, if not impossible, to reach a correct answer just by guessing in a SCG test. Furthermore, to make guessing much more difficult, SCG tests have more choices than traditional multiple choice tests and the choices are displayed in a grid as the name implies (Bahar et al., 2006). For instance, while a 3x3 SCG has 9 choice a 4x4 SCG has 16 choices (See Figure 1).

First studies about SCG was commenced by Egan (1972) and then has been used by many researchers successfully (Johnstone et al., 2000; Bahar and Hansell, 2000; Bahar, 2001; Özatlı, 2006). In almost all of these studies, it has been emphasized that SCG is an assessment and evaluation technique that measures meaningful learning and helps in diagnosing misconceptions and erronues information of students.

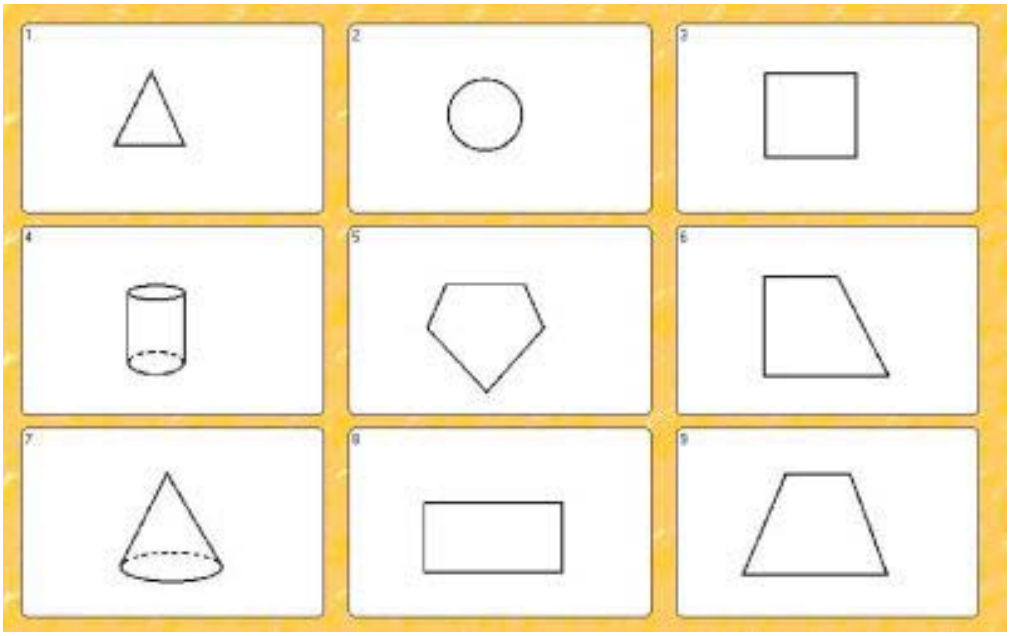

Figure 1: A 3x3 Structured Communication Grid

There is not an optimum grid size for a SCG test. 3x3, 3x4 or 4x4 SCG grids are prepared with respect to the age and school level of the audience and each cell of the grid is numbered in turn (Johnstone et al., 2000). In order to prepare a SCG, the teacher asks a fair question and places its answer randomly into one or a few of the cells. Then s/he asks another question and places its answer into the cells respectively. However, the cells belonging to the correct answer of a question may be amongst the cells for the correct answer of another question. In other words, some part of correct answers of a question may also be part of correct answers to another question. One might ask as many questions for a grid as s/he likes unless filling all the cells of the grid. What is required to complete a SCG successfully may ,indeed, resemble to composing a short essay or an article. While composing an essay, the ideas are first put forward and then organized and presented in a logical order and coherent way. While comstructing a SCG ,likewise, correct answers for each question first put forward and then distributed among the cells of the grid.

There are two types of questions that could be asked in a SCG, namely standart and ordering questions. When it is administired, depending on the type of the question the participants are required either

a) to select the cells appropriate to be an answer for the question or
b) to order the cells by enumerating each cell according to their logical and functional relations for a particular criteria (this option may not be suitable for every question of a grid).

As a diagnostic test, scores taken in a SCG test may not be used for a summative evaluation. Separate scoring mechanisms are used for each type of the question. In standart questions, total score of a question is divided into two parts and while 50% of the score comes from selecting correct cells for a question, the remaining 50% of the score comes from not selecting an incorrect cell for the question. In other words, participants are rewarded for each selected correct cell and punished for each selected incorrect cell. In ordering questions, the relative order of the correct cells are also taken into account. An example of SCG scoring mechanism can be seen in the Appendix A.

It will be appropriate here to mention about computer aided measurement and evaluation techniques since the development of a computer based SCG test is aimed in this study. Parallel to technological developments, the utilization of computers during instruction continues to increase; however, similar increase has not been observed in computer aided measurement and evaluation (Schwid & O'Donnell, 1992; Dillon & Clyman, 1992). Underwood and Brown (1997) also argue that while the number of computer aided measurement and evaluation instruments assisting class learning increased, its utilization in educational tests is still quite low.

Thelwall (2000) states that computer aided measurement and evaluation techniques, which might increase the effectiveness of teaching and learning, has a variety of usages: Employmet of the diagnostic tests before any teaching activity and employing formative assessment tests during the teaching activities are among them. Bennett (1998), furthermore, states that computer aided measurement and evaluation instruments have brought opportunuties for innovation in testing and measurement and evaluation. Computers could be utilised more effectively in evaluating any problem solving activities. Moreover, they are also effective in identifying students' process dependent abilities (Chung and Baker, 1997; Schacter et al., 1999; Stevens et al., 1999). Computer based measurement an evaluation instruments not only are a rich resource for helping participants to asses their process and knowledge level but also provide immediate and personal feedback (Mooney et al., 1998)

It may be argued that the comparative studies addressing the utilization of computerized and paper based tests are inadequate and inconclusive (Burke & Normand, 1987; Mazeo & Harvey, 1998 as cited in Dillon & Clyman, 1992). While some studies (Hicken, 1993; Vansickle & Kapes, 1993; Finegan & Allen, 1994) report that there is not a significant difference in student performances for these two tests, other studies (Dimock & Cormier, 1991; Mazzeo, 1991) report the contrary. Kapes and Vansickle (1992) report that computer based tests are more meaningful and reliable than the others.

Levine and Donitsa-Schmidt (1997) and Ogunkola (2008) state in their study that widely usage of computers facilitates positive attitudes towards computers. It is thought that several issues such as participants' demographic factors (gender, cultural background, age etc.) (Leutner & Weinsier, 1994) and their previous computer experiences (Maurer, 1994) may affect their performances during computer based measurement and evaluation. Lee (1986), in his study on college students, found that the students who have less experience with computers are less successful than others at computer based mathematics tests. However, it is argued that this inadequacy may be removed by the provision of a minimal computer experience opportunity.

Mazzeo et al (1991), on the other hand, found conflicting results in their studies on English and Mathematics tests. In one of their studies, for instance, the authors observed that there was a negative effect of previous experience with computers, i.e., the less experienced students got higher marks. In another study focusing on the previous computer experiences (Johnson and White, 1980), it has been observed that a group of participants who took a computer course got higher scores than a group of participants who didn't take any computer course before.

THE AIM OF THE STUDY

The paper form of SCG technique has been widely used in many studies. However, there is not any study, authors aware of so far, which utilizes this technique in the computer environment. The aim of this study is, hence,

i) to develop a computer based environment for the administering and evaluating SCG technique, and

ii) to collect participants' views about this environment and to determine positive and negative aspects of both SCG technique itself and its computerized version.

METHOD

Participants

One hundred and fifty participants, consisting of 118 candidate elementary school teachers, 24 candidate science teachers at the master program and 12 faculty staff members from the elementary education department of Faculty of Education at Abant Izzet Baysal University, Bolu Turkey, have taken part in this study. All of the participants filled a standard software evaluation questionnaire after using the computerized version of SCG technique. Face-to face interviews were conducted with 26 candidate elementary school teachers, 19 candidate science teachers and 4 faculty staff members in order to probe their views about SCG technique and its adaptation to the computer environment.

None of the participants in the study had reported any difficulty in using computers since they had already taken computer courses at graduate level. The teachers and faculty staff reported that they had their own personal computers and they had been using them as an integral part of their teaching.

Design of SCG Software

A computer system that facilitates the preparing, administering and evaluating a SCG technique has been developed. The system has two different modes, namely design and application modes. While design mode provides tools for the preparation of a SCG test, application mode helps administering a SCG test, collecting user responses and analyzing them instantly.

The design mode has very flexible facilities that ease the creation of a SCG test. There are several buttons on the upper part of the screen that provide easy access to the major functionalities of the system such as adding a new grid or deleting an existing grid, adding a new question for a grid and deleting a question from the grid, navigating between existing grids and questions, changing the question types etc. There is also a menu on top of the system that enables easy access to all the functionalities of the system. There are also several text labels that inform users about the current status of the system such as how many grids, questions it has, and the number of the current grid and the current question, the current question type , how many questions are needed to be answered or needed their correct cells to be assigned etc.

One is firstly required to construct a grid by stating how many cells it will consists of and what type of responses (text or image) it will hold. How many cells a grid contains can not be changed after its construction. If one wants to construct a grid with different number of cells, s/he has to add a new grid with the desired number of cells. However, one can easily change what type of a response each cell of a grid will hold. Some cells of the grid might hold images while some other cells of the same grid might hold text fields. By default, each cell of a newly constructed grid has a default placeholder. It has a specific picture for image cells and a default placeholder text indicating the cell number such as “Item 4” for text cells. One could easily change the content of a cell from a text to an image or from an image to text by pressing a specific key. One could place graphs, charts, pictures etc as an image. One could browse for an image file to load after selecting a cell and converting it to an image by pressing a key if it is a text cell. One could also change the font size, styles, text colours and background colours of cells containing texts. One could also create fixed cells that could be addressed in questions. Fix cells can not be clicked or highlighted during the design or application mode and left part of the cell is shown in red to denote it is fixed. This flexibility of the system makes the preparation of complex grids possible.

After a grid is constructed, one could ask as many questions as s/he wishes for a grid and assign their correct cells. The design mode of the system enables constructing both standard and ordering questions which are called A and B type questions respectively. One could ask just A type, just B type or both A and B type questions. One could change between question types by pressing the buttons labeled as “A” and “B” respectively on the upper left side of the screen and may enter his/her question into the question text field. If A is selected, the question becomes a standard question and if B is selected the question becomes an ordering question. A text field of a question by default contains placeholders for actual questions indicating the question number and the type of the question. For instance for a standard question it contains “This is standard question 1” while it holds “This is ordering question 1” for an ordering question. One could write his/her own questions by deleting these placeholders. If the text field of A or B type of a question is left empty, it is regarded as not having that type of the question and that type of the question is not taken into consideration during the application mode. The cells selected during the preparation of a question are regarded as its correct cells. One could easily change the correct cells of a question anytime during the design mode.

Pressing a mouse button over a cell is automatically selects or de-selects a cell. The cells are automatically sequenced in the order that they are selected. The selection of the cells is done in type A question and the order of the selected cells is changed in type B question. Selection of a cell is shown by highlighting either the actual number in standard questions on the left upper side of cell or the sequence in which it was selected in ordering questions on the left bottom side of the cell. If a change is desired to be done on automatic ordering, part of the cell that shows the number is either to be left-clicked in order to increase order number or to be right-clicked in order to decrease order number.

Grids prepared in design mode could be saved to a text file and could then be re-loaded again in design mode and modifications could be made.

Application mode of the system enables the administering of the prepared files by the system. Application mode starts with the selection of SCG file and entering some identification information about the user such as name, age, gender etc (See Figure 2).

Figure 2: A screenshot of identification information window in the application mode.

However, in application mode one is only allowed to make selections and change ordering of selections and can not modify any other part of the grids or questions. One can not exit the application mode without giving adequate responses for each question in every grid. Hence, there are also additional buttons providing specific functionality for the application mode such as to bypass a question, labeled as " I do not want to answer this question", or to go to the next unanswered question or information labels that indicate how many questions are not answered.

Application mode of the system also provides instant analysis of the users responses by both displaying it to the users and saving it to a file. Detailed analysis information about the test results give both participants and researchers instant feedback and provide opportunity for their self evaluation. Analysis of the results both include what responses users gave including correctly and incorrectly selected cells for each question and how long they deal with each question and grid as well as their raw and scales scores for each question. It also provides an average score for standard and ordering questions and produces a test score scaled to 100.

Facilities provided in application mode of the system simplifies the administration of a SCG test which is very difficult to administer and analyze manually by paper-pencil. Hence, the system might make life easier for teachers and enable them to obtain in-depth knowledge about their students.

Some screenshots of the system during the administration of a biology and mathematics tests could be seen in the following figures (Figure 3, Figure 4, Figure 5 and Figure 6).

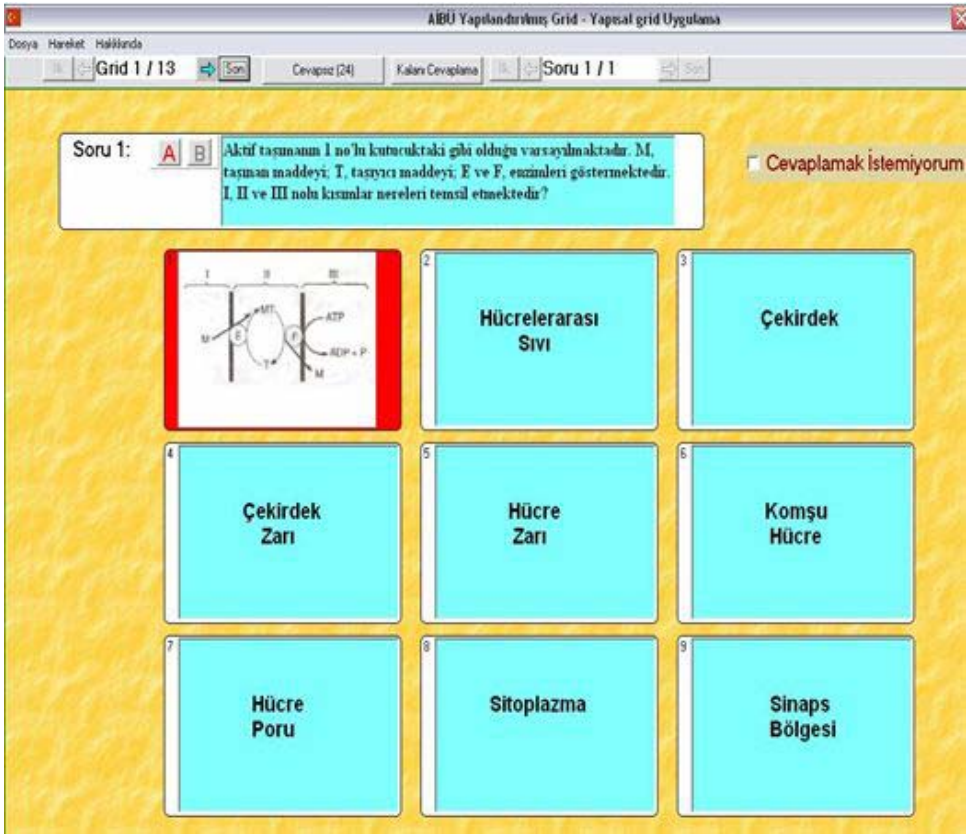

Figure 3: A SCG consisting of 8 text choices and a fixed image cell

Figure 4: A SCG consisting of 9 image choices

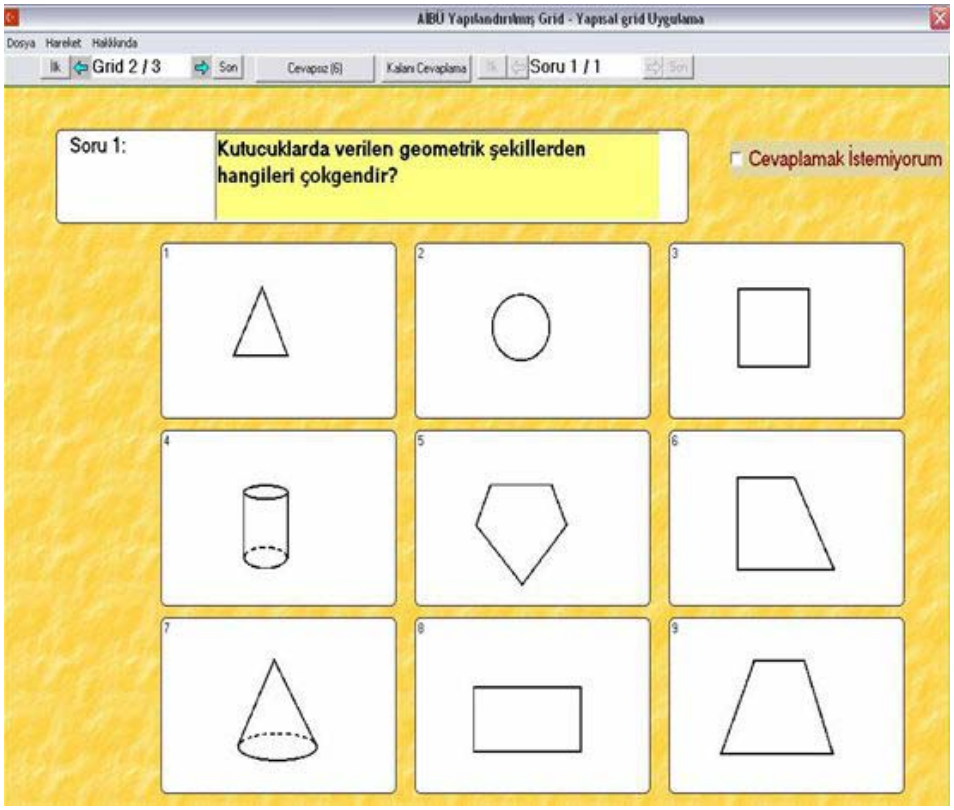

Figure 5: A SCG consisting of 9 geometric figures

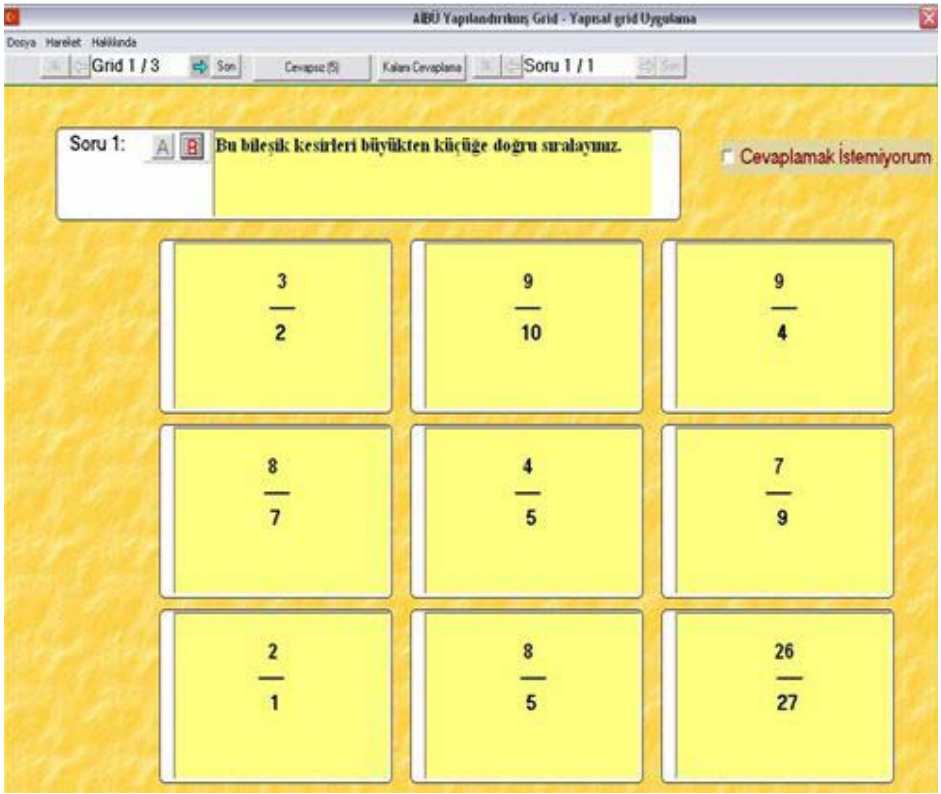

Figure 6: A SCG ordering question with 9 choices

Data Gathering Instruments and Analysis
Two research instruments are utilized in this study to collect data:

1) Software Evaluation Questionnaire and
2) Open-ended interview form.

*Software Evaluation Questionnaire: "*Software Evaluation Questionnaire (SEQ)", addressing to participants' views about SCG-software, has been adapted by us by using a "prototype software evaluation form" and a "final release evaluation form" (Akpınar, 1999).The SEQ consists of 20 items and 5-Likert Type Scaling was preferred. Its validity was provided by consulting to several experts. An open-ended question was added at the end of the questionnaire in order to provide the participants an opportunity to express their views about the system. Test-retest method was applied in order to determine the instruments' reliability (Cohen et al., 2005, p. 131; Can, 2000): A group of participants (N=77) were asked to fill the questionnaire after they use the system. The same questionnaire was given to the same participants 3 weeks later. Correlation coefficient was found to be 0.82. Moreover, internal consistency (Cronbach alpha) of the instrument has been found to be 0.90.

The responses given to SEQ were converted to a score for each item. In order to do this, the options were first scaled from 1 to 5 (1- I Strongly Disagree; 5- Strongly Agree). Then participant responses were converted to SPSS format. Average scores for each participant and each item were calculated in SPSS. These scores were re-scaled by assigning equal quantitative ranges to Likert-Type Scaling.

1.00-1.79; I strongly disagree	1.80-2.59; I disagree
2.60-3.39; neither agree nor disagree	3.40-4.19; I Agree
4.20-5.00; I strongly agree	

Participants' responses are intensifed in the ranges of 3.40 – 5.00 (I agree – I strongly agree).

Open-ended Interview Form:
In order to obtain participant views, two open-ended question were addressed to the participants:

1- *What do you think about positive and negative aspects of structured communication grid technique?*
2- *Would you compare positive and negative aspects of paper-based structured communication grids with computer-based structured communication grids?*

The data obtained from the interview forms have been analyzed using descriptive analysis method (Yıldırım & Şimşek, 2000, p. 158). The data have been grouped under different categories and participants' related views have been gathered together and were analyzed. By means of this, it has been aimed that participant views may be presented in a more coherent way and their positive and negative views about the program may be identified more clearly.

Implementation
SCG technique by paper and pencil, first, has been introduced to the students, teachers and faculty staff at separate sessions (Since the number of students were high, their presentation has been divided into two sessions). Then, they were required to prepare an example SCG test in a week in order to probe their level of understanding and to ensure that they fully understand the technique.

The SCG software has then been introduced to participants with a presentation in a computer laboratory. After the presentation, the students and teachers were required to complete a computer based SCG test consisting of questions from their own subject fields. Faculty staff has not attended this part of study. Researchers prepared the questions employed in the tests themselves and experts' views about them have been collected. Software evaluation questionnaires were administered to the participants after they complete the test.

FINDINGS AND DISCUSSION
The data gathered from participant responses to SEQ are presented in Table 1. It can be seen from this table that average scores of each group and consequently general average scores for each particular item ranges at the interval of 3.40 – 5.00 (I agree- I strongly agree). Moreover, the participants have, in average, strongly agreed (4.20-5.00) with twelve of twenty items. These findings suggest that the computer software prepared for SCG technique is sufficient in terms of the items addressed to participants.

It can also be seen from Table 1 that the average scores of the students and science teachers are slightly lower than the average scores of faculty staff for the item numbers 7 (The quantity of knowledge presented on the screen is enough), 8 (Transitions between screens are completely probable), and 10 (System errors, whose source is unknown, are not encountered). It may be argued that the reason for this difference may be due to the nature of these items which may necessitate more expert knowledge than the other items in the SEQ.

However, students' and teachers' views about these items may be taken into account and therefore the quantity of knowledge presented on the screen may be increased and the number of screen transitions and problems perceived as system errors may be minimized. Every software development processes is circular and definitely need feedback from end-users to make further revisions.

Another noticeable statement which has high average score in Table 1 is item number 20 (It can be used for program measurement and evaluation purposes). This is an expectable result for SCG technique due to its characteristics such as the provision of an opportunity for evaluating partial knowledge, decreasing chance factor by questioning meaningful learning as a result of its selection and ordering features, and maintaining immediate feedback (such as scoring, number of incorrect and correct answers, saving information etc.).

Table 1: Average score for participant responses to questionnaire items

	Students	**Science Teachers**	**Faculty Staff**	**Average**
	$\overline{X}$	$\overline{X}$	$\overline{X}$	
01- Screen layout of the program is clear and understandable..	4.16	4.20	4.66	**4,34**
02- Screen design helps understanding the program fully.	4.04	3.91	4.08	**4,01**
03- It is easy to learn how to use the program.	4.28	4.16	4.16	**4,20**
04- It is esy to remember how to use the functions and operations in the program.	4.11	3.83	4.33	**4,09**
05- It is easy to use the functions in the program.	4.24	4.12	4.50	**4,28**
06- Function names and their functions are consistent with each other.	4.11	3.91	4.41	**4,14**
07- The quantity of knowledge presented on the screen is enough.	3.81	3.45	4.33	**3,86**
08- Transitions between screens are completely probable.	3.63	3.91	4.33	**3,95**
09- The program is fast.	4.19	4.41	4.66	**4,42**
10- System errors, whose source is unknown, are not encountered	3.58	3.87	4.33	**3,92**
11- The program is appealing.	4.17	4.45	4.75	**4,45**
12- The program is motivating.	4.10	4.16	4.66	**4,30**
13- The program has characteristic to help in construction of knowledge.	4.02	3.95	4.16	**4,04**
14- Communication characteristic of the program is strong.	4.03	3.95	4.25	**4,07**
15- It is beneficial that the program saves the results.	4.45	4.58	4.75	**4,59**
16- The program is beneficial in terms of providing teachers opportunity to develop themselves.	4.18	4.29	4.50	**4,32**
17- It is convenient to study this program.	4.26	4.20	4.33	**4,26**
18- The program is in a characteristic that it can be controlled at any time..	4.33	4.33	4.33	**4,33**
19- Feedback characteristic of the program is beneficial.	4.26	4.29	4.66	**4,40**
20- It can be used for program measurement and evaluation purposes.	4.43	4.41	4.41	**4,42**

Especially high average score of faculty staff shows that this software meets experts' expectations of a computerized version of SCG. This finding shows that it is consistent with the usability criteria of a measurement instrument as Lien suggested (1976). Mooney et al. (1998) claim that computer aided measurement and evaluation techniques provide instant and individualized feedback as well as assisting evaluation of knowledge and process levels of students. Sluijsmans et al. (1998) and Bond (1995) claim that while self-evaluation may not be a new evaluation strategy, it is an effective way to increase a student's role as an active participant during self-learning. Both of these views are parallel to the questionnaire results about evaluation of a

computer based SCG technique and its feedback provision mechanism. That is, this mechanism provides feedback in a multidimensional format and saves it to a file which is very helpful in students' self-evaluation and reflection about the process.

In order to support the consistency of the findings in the SEQ qualitatively, participants' responses to open-ended interview forms have been provided below.

It has been seen that the participants mostly expressed positive views about 1st question in the open-ended interview form (What do you think about positive and negative aspects of structured communication grid technique?). Positive views are more apparent in the statements of faculty staff as well. The positive views may be categorized as "effectiveness of this technique as a measurement and evaluation tool", "assessment of partial knowledge", replacement of assessment of rote learning with assessment of meaningful learning", and "knowledge endurance". Some examples selected from participant views may be seen below:

"...it is difficult from the teachers' perspective since it requires much higher quality questions; however, it is a good way of evaluating the learning of students. As a candidate teacher, I am satisfied with it. It aims to facilitate meaningful learning of students. It will provide knowledge endurance since in case the students give correct answer, it makes them question why it is correct; that means it leads students to the discovery, contemplating, and analyzing. I like to use this method at the school that I will teach" (A student's view)

"...this is a nice method to assess the students who have prior knowledge and correctly evaluate their knowledge level. However, it may be difficult to implement at first. It is also very good in terms of providing alternative method to multiple-choice tests..." (A teacher's view)

"...SCG technique is excellent for elementary school and secondary school evaluations and especially for more precise examinations like UDS and KPDS. I think it is a multidimensional technique since it increases teachers' and students' motivation, gives students an opportunity to correct their mistakes, to complete their missing knowledge, and strengthen their correct understandings."

Some negative statements, besides to positive ones, have also been expressed by some students as can be seen from open-ended interview forms. These negative statements can be grouped into two categories: Lack of time issue during the preparation of the technique and restrictions due to the nature of the specific subjects. Some excerpts about students' negative views are below:

"...it requires too much time to prepare questions in SCG technique. Teachers' thinking time gets longer. In order to give it an exam status, the numbers of questions are to be increased; however, this also increases its cost. Moreover, there may be some operational mistakes during the evaluation, which in turn affects the assessment..."

"..it may be restricted to specific subjects. It may be impossible to prepare both standard and ordering questions and just ordering questions for each subject. I mean it may be difficult to apply in every subject."

According to 2nd question of open-ended interview form (Would you compare positive and negative aspects of paper-based structured communication grids with computer-based structured communication grids?), all participants expressed that computer-based SCG had more advantages than the paper and pencil version. These include the factors such as its user-usability, provision of fast and effective feedback, quick and easy scoring, attractiveness of design with different colours and shapes. Some statements about them are presented below:

"...you don't have to know the software in detail. It is nice to answer the questions with a few clicks or using keyboard, without dealing with paper and pencil. It is also nice that the answers are supported visually" (A Student's view)

"...its utilization by means of computer makes the process faster. It facilitates quick thinking. While using paper-based method, we may sometimes forget what we thought. This computer software, I think, is good from this point." (A teacher's view)

"The utilization of computer based type of SCG technique is quite appropriate for secondary and higher education...The most important aspect of these grids is its transfer of the principal "the more senses are addressed during the education, the more learning occurs" into the measurement and evaluation dimension of learning." (A Faculty staff's view)

While these positive thoughts have dominated participant views, they have also expressed concerns about the importance of computer laboratories and computer literacy of students and teachers for a successful application of computer-based SCG. Most of the participants agree with the following view:

"This software may be quite attractive for students; however, the preparation of an adequate SCG test requires a teacher both to be good at using computers and having enough time to prepare it. Therefore, I think, the utilization of it in Turkey may only be possible after years of endeavours since in our schools there is not enough computer equipment. However, if we somehow achieve to use this technique, I think it will be enjoyable for students and the concepts will be more durable since it addresses to visual senses..." (A teacher's view)

"Most of our teachers even don't know how to use computers. Due to this, it is too early to use this software. However, we may think is as a big step to increase education level..." (A student's view)

"The students are to be able to use computer to the some degree in order to use the software efficiently. I mean, objective results may not be obtained if it is used with the individuals who are reluctant to use or afraid of using computers." (A faculty staff's view)

The views of participants about computer literacy of students/teachers and technological infrastructure of schools, of course, are not to be overlooked. However, inadequacy in using computers is not only restricted to SCG technique. It is also applicable for all software programs utilized in computer-aided teaching. The studies about the effects of previous computer experience, however, give conflicting results. Most of the studies (for ex. Johnson and White, 1980; Mazzeo et al., 1991) show that computer experience influence achievements in computer aided programs or tests. However, some studies (e.g., Lee, 1986) suggest that computer experience does not have direct influence on achievement in computer-aided studies. The SCG technique used in our study does not require high level of computer literacy and a minimum level of experience may be enough in order to use it. It is observed that the participants had already that experience. The points raised on this issue may apply to those individuals who have no computer experience at all. However, if one takes probable widely usage of computers at our schools and homes in near future into account, it may be concluded that this concern is temporary.

As it has been mentioned in the introduction part, while some studies suggest that there are no differences in student performances for computer-based exams and paper-based exams (Hicken, 1993; Vansickle & Kapes, 1993; Finegan & Allen, 1994), others suggest that there are significant differences between them (Dimock & Cormier, 1991; Mazzeo, 1991). In our study, we have not reported participant performance for computer-based tests and paper-based tests; however, participants' views about major issues such as usability, provision of fast and effective feedback, easy and quick scoring mechanism, and attractive design with different colours and figures suggest that computer-based SCG technique may bring more success than its paper-based counterpart addressing the same subject.

Although there are some negative participant views about the feasibility of computer-based SCG, such as lack of computer literacy, positive participant views outnumbers the negative ones and the issues raised could be resolved over the time. This makes computer-based SCG technique a positive view dominated technique. Moreover, if one takes the studies indicating that computer-based tests are significantly more reliable than paper-based tests (Kapes and Vansickle, 1992) into account, it may easily be appreciated why computerized SCG technique is more favourable to paper-based one. If one compares the responses obtained by means of open-ended interview questionnaire with the ones obtained from SEQ, it may be concluded that they support each other. This also shows that they are consistent with each other.

Another point that is worthy of discussion concerning this subject is about the place of SCG technique during assessment of process among other alternative measurement and evaluation techniques mentioned in 2004 Turkish Science and Technology Program. There are, of course, many advantages of both paper-based and computerized SCG technique such as assisting assessment of partial knowledge, decreasing chance factor in reaching to correct answer, diagnosis of misconceptions or erroneous conceptions of students by means of incorrectly selected cells, identifying defected aspects of cognitive schemata of students, giving opportunity to fill and replace the cells with words, figures, numbers, equations, formulas by means of what visual and verbal thinking abilities of students enhanced. While these advantages are quite important, can SCG be thought to be in the same category with the process evaluation techniques such as portfolio and projects assessment, which address all characteristics of formative evaluation? Logical and operational sequencing of the selected cells in SCG technique may demonstrate that the process is an important aspect of it; however, we think that it may be more appropriate to regard SCG technique in the middle of traditional and alternative measurement and evaluation techniques.

As Bennett (1998) stated, computer-aided measurement and evaluation provides innovations for tests, and measurement and evaluations. The SCG technique, which highlights both diagnostic and formative aspects of assessment, and evaluates both the product and partially the process, is quite feasible area to be applied and developed by researchers at various areas.

REFERENCES

Akpınar, Y. (1999). *Bilgisayar destekli öğretim ve uygulamalar*. Anı Yayıncılık, Ankara.

Bahar, M. (2001). Çoktan seçmeli testlere eleştirel bir yaklaşım ve alternatif metotlar. *Kuram ve Uygulamada Eğitim Bilimleri Dergisi*, 1(1), 23–38.

Bahar, M. & Hansell, M. H. (2000). The relationship between some psychological factors and their effect on the performance of grid questions and word association test. *Educational Psychology,* 20(3), 349–364.

Bahar, M. (2006*). 4-8. sınıflar fen ve teknoloji öğretim programına genel bir bakış (Ed.) M. Bahar Fen ve teknoloji öğretimi içinde* (s. 433-450). Ankara, PegemA Yayıncılık.

Bahar, M., Nartgün, Z., Bıçak, B. & Durmuş, S. (2006). *Geleneksel ve alternatif ölçme ve değerlendirme öğretmen el kitabı,* 1. Baskı. Ankara: Pegem A yayıncılık.

Bennett, R. E. (1998). *Reinventing asessment: speculations on the future of large-scale educational testing.* Princeton, NJ: Educational Testing Service, Policy Information Center.

Bond, D. (1995). *Enhancing Learning Through Self-Assessment*. London: Kogan Page.

Burke, M.J. & Normand, A.L. (1987) Computerized psychological testing: overview and critique. *Professional Psychology Research and Practice*, 18, 42–51.

Can G. (2000). *Ünite 13. Eğitimde ölçme ve değerlendirme (Ed.) M. Gültekin. öğretimde planlama ve değerlendirme içinde* (s.207-222). Eskişehir: Anadolu Üniversitesi Yayınları.

Cohen, L., Manion, L., & Morrison, K. (2005*). Research methods in education*. Routledge Falmer.

Chung, G. K. W. K. & Baker, E. L. (1997). *Technology in action/Year 1 technology studies: implications for technology in assessment (CSE Technical Report No. 459*). Los Angeles, CA: University of California, National Center for Research on Evaluation, Standards, and Student Testing.

Dillon, G.F. & Clyman, S.G. (1992). The computerisation of clinical science examinations and its effect on the performance of third year medical students, *Academic Medicine,* 67, 66–68.

Dimock, P.H. & Cormier, P. (1991). The effects of format differences and computer differences on performance. *Measurement and Evaluation in Counselling and Development,* 24, 119–126.

Egan, K. (1972). Structural communication – A new contribution to pedagogy. *Programmed Learning and Educational Technology*, 1, 63–78.

Finegan, J.E. & Allen, N.J. (1994). Computerised and written questionnaires: are they equivalent?, *Computers in Human Behaviour*, 10, 483–496.

Hicken, S. (1993). Administering comprehensive examinations using computers, *Collegiate Microcomputer,* 11, 194–198.

Johnson, D., & White, C. (1980). Effects of training on computerized test performance in the elderly. *Journal of Applied Psychology*, 65, 357–358.

Johnstone, A. H., Bahar, M., Hansell, M. H. (2000). Structural communication grids: A valuable assessment and diagnostic tool for science teachers. *Journal of Biological Education*, 34(2), 87-89.

Kapes, J.T. & Vansickle, T.R. (1992). Comparing paper and pencil and computer-based versions of the Harrington O'Shea career decision making model. *Measurement and Evaluation in Counselling and Development,* 25, 5–13.

Lee, J. (1986). The effects of past computer experience on computerized aptitude test performance. *Educational and Psychological Measurement*, 46, 727–733.

Levine, T., & Donitsa-Schmidt, S. (1997). Commitment to learning: effects of computer experience, confidence and attitudes. *Journal of Educational Computing Research*, 16(1), 83–105.

Leutner, D. & Weinsier, P.D. (1994). Attitudes towards computers and information technology at three universities in Germany, Belgium and US. *Computers in Human Behaviour*, 10, 569–591.

Lien, A. J. (1976). *Measurement and evaluation of learning, Third Edition*, Wm. C. Brown. Company Publisher, USA.

Maurer, M.M. (1994). Computer anxiety correlates and what they tell us: a literature review. *Computers in Human Behaviour*, 10, 369–376.

Mazzeo, J. (1991). *Comparability of computer and paper and pencil scores for two CLEP general examinations*. College Board report no. 91–95 (New York, College Board Publications).

Mazzeo, J. & Harvey, A. (1988). The equivalence of scores from automated and conventional educational and psychological tests: *A Review of the Literature*. College Board Report 88–8 (New York, College Entrance Examination Board).

Mazzeo, J., Druesne, B., Raffeld, P., Checketts, K., & Muhlstein, A. (1991). *Comparability of computer and paper-and pencil scores for two CLEP general examinations (College Board report No. 91–5).* Princeton, NJ: Educational Testing Service.

Milli Eğitim Bakanlığı (2004). *İlköğretim fen ve teknoloji dersi öğretim programı ve kılavuzu (4–5. Sınıflar).* Ankara. MEB yayınları.

Mooney, G. A., Bligh, J. G., Leinster, S. J. (1998). Some techniques for computer based assessment in medical education. *Medical Teacher*, 20(6), 560-566.

Ogunkola, B. J. (2008). Computer attitude, ownership and use as predictors of computer literacy of science teachers in Nigeria. *International Journal of Environmental & Science Education,* (inpress), 3(2), 53-57.

Özatlı, N. S. (2006). *Öğrencilerin biyoloji derslerinde zor olarak algıladıkları konuların tespiti ve boşaltım sistemi konusundaki bilişsel yapılarının yeni teknikler ile ortaya konması*. Doktora Tezi, Balıkesir Üniversitesi Fen Bilimleri Enstitüsü.

Schacter, J., Herl, H. E., Chung, G. K. W. K., & O'Neil, H. F. (1999). Next generation performance assessments: Reporting performance and process data together using computationally designed assessments [Special issue]. *Computers in Human Behaviour*, 15(3–4), 403–418.

Schwid, H.A. & O'Donnell, D. (1992). Anaestheologists' management of simulated critical incidents. *Anaesthiology*, 76, 495–501.

Stevens, R., Ikeda, J., Casillas, A., Palacio-Cayetano, J., & Clyman, S. (1999). Artificial neural network-based performance assessments. *Computers in Human Behaviour*, 15(3–4), 295-313.

Sluijsmans, D., Dochy, F. and Moerkerke, G. (1998). Creating a learning environment by using self-peer and co-assessment. *Learning Environment Research*, 1(3), 293–319.

Thewall, M. (2000). Computer-based assessment: a versatile educational tool. *Computers & Education,* 34, 37-49.

Underwood, J. D. M., & Brown, J. (1997). *Integrated learning systems: potential into practice*. London: Heineman.

Vansickle, T.R. & Kapes, J.T. (1993) Comparing paper-pencil and computer based versions of the Strong–Campbell interest inventory. *Computers in Human Behaviour*, 9, 441–449.

Yıldırım, A., & Şimşek, H. (2000). *Sosyal bilimlerde nitel araştırma yöntemleri*. Seçkin Yayıncılık, Ankara.

APPENDIX A

A Sample of Structured Communication Grid Test and Its Scoring Mechanism

A structured communication grid test focusing on the subject of vertebrate and invertebrate animals is presented below. Please use the numbers assigned to each cell of a grid in order to answer the questions prepared for that grid. Each cell may be used as an answer of one or more questions.

1 Gnat	2 Dolphin	3 Butterfly
4 Crab	5 Trout	6 Lizard
7 Goshawk	8 Worm	9 Salamander

Q1- Which of the cells above include vertebrate animal names?

a) Select appropriate cells. (Correct Answer: 2, 5, 6, 7, 9)

b) Sequence the cells you selected in part a in the order of Fish, Amphibians, Reptiles, Birds, Mammals. (Correct Answer: 5, 9, 6, 7, 2)

Q2- Which of the cells above include invertebrate animal names?

Q3- Which of the cells above include the name of animal whose appearance and actions seems like a fish however it is categorized under mammals?

Q4- Which of the cells above show the names of the vertebrate/invertebrate animals that live on land?

Q5- Which of the cells above show the names of the vertebrate/invertebrate animals that live in water?

Scoring

Scoring mechanism of SCG technique will be shown here for an example question, namely Q1. Q1 has two parts: Part A is a standard question while part B is an ordering question.

Let us firstly illustrate the scoring of a standard question, Part A. In standart questions, total score of a question is divided into two parts and while 50% of the score comes from slecting correct cells for a question, the remaiming 50% of the score comes from not selecting a incorrect cell for the question. In other words, participants are rewarded for each selected correct cell and punished for each selected incorrect cell. Table 1 shows the formula for getting the raw score of a standard question.

Table 2: Scoring mechanism of a standard question

$$\frac{C1}{C2} - \frac{C3}{C4}$$

C1= The number of cells selected correctly
C2= Total number of correct cells
C3= The number of cells selected incorrectly
C4= Total number of incorrect cells

The formula above gives a raw score between -1.0 and +1.0. In order to convert it to a scale of 10, we first add it with 1 and then multiply it with 5.
As it has been stated above, the correct cells for Part A of Q1 are 2, 5, 6, 7, and 9. Suppose that a participant selected the cells 2,4,6,7 and 8 for the Part A. Then, one could compute his/her scores as follow:

Correct cells are 2, 5, 6, 7 and 9 and the total number of correct cells are 4 (**C2=5**).
Cells selected correctly are 2, 6 and 7, and the number of cells selected correctly are 3(**C1=3**).
Incorrect cells are 1, 3, 4 and 8 and the total number of incorrect cells are 4 (**C4=4**).
Cells selected incorrectly are 4 and 8 and the number of cells selected incorrectly are 2 (**C3=3**).

$$\text{Raw Score for Part A} = \frac{3}{5} - \frac{3}{4} = \frac{12-10}{12} = \frac{2}{12} = 0{,}10$$

If one would like to convert it to a scale of 10, then one should add 1 to the raw score and multiply it with 5.

$$\text{Score for Part A at a sclae of } 10 = (0{,}10+1).5 = (1{,}1).5 = 5{,}5$$

Let us then illustrate the scoring of an ordering question, Part B of Q1. In ordering questions, the relative order of the correct cells are also taken into account. It is difficult to score an ordering question. One could select all the cells correctly but put them in a wrong order. One could ask two Yes or No ordering question for every consecutive pairs of the correct cells and determine how many consecutive cells are in the right position. If one of the consecutive cells is not selected or if two consecutive cells are selected but the order is not right, two No responses is returned for two ordering questions. If two consecutive cells are selected and their order is correct but they are not consecutively selected, one Yes and one No responses is returned for two ordering questions. Hence, the raw score of an ordering question is computed by proportioning the number of Yes questions over the total number of Yes questions for thr right positions of consecutive pairs in the correct order.

$$\text{Row Score for an Ordering Question} = \frac{\text{Total Number of Ordering Questions}}{\text{Total Number of Yes responses}}$$

Please pay attention to the fact the number of Yes or No questions doubles as the number of correct cells in an ordering question increase. However, it is enough to multiply by 10 to scale this raw score to 10.

As it has been stated above, the correct sequence of order for Part B of Q1 is 5, 9, 6, 7 and 2. Suppose that a participant selected the cells 9, 6, 5, 7 and 4 in sequence for the Part B. Then, one could compute his/her scores as follow:

Following 4 consecutive pairs are found for Part B of Q1. (5,9), (9,6) , (6,7) and (7,2).
One could ask the following 8 ordering Yes or No questions whether or not the right positions for each consecutive pair exist in the response.

Q1: Is number 5 selected before number 9? **Q2:** If it is, are they ordered consecutively?
Q3: Is number 9 selected before number 6? **Q4:** If it is, are they ordered consecutively?
Q5: Is number 6 selected before number 7? **Q6:** If it is, are they ordered consecutively?
Q7: Is number 7 selected before number 2? **Q8:** If it is, are they ordered consecutively?

Our participant gets the following results for the above ordering questions. Note that if first part of the question is incorrect, then the second part is not taken into account.

Q1:	**No**	(5 after 9)	**Q2:**	**No**	(-)
Q3:	**Yes**	(9 before 6)	**Q4:**	**Yes**	(9 and 6 are consecutively selected)
Q5:	**Yes**	(6 before 7)	**Q6:**	**No**	(6 and 7 are not consecutively selected)
Q7:	**No**	(2 is not selected)	**Q8:**	**No**	(-)

Total Number of Ordering Questions is 8.
Total Number of Yes responses are 2.

$$\text{Row Score for Part B} = \frac{\text{Total Number of Ordering Questions}}{\text{Total Number of Yes responses}} = \frac{2}{8} = 0{,}25$$

If one would like to convert it to a scale of 10, one should multiply this score by 10.

$$\text{Score for Part B at a sclae of } 10 = (0{,}25).10 = 2{,}5$$

Total score of the question could be computed by taking the average of the standard and ordering questions. Hence,

$$\text{Total Score for Q1} = \frac{(\text{Score for Part A} + \text{Score for Part B})}{2} = \frac{(5{,}5 + 2{,}5)}{2} = 4$$

ACKNOWLEDGMENT

Correspondent Author:

Fatih AYDIN
Gazi University
Faculty of Education
Science Education Department

e-mail: fatihaydin14@gmail.com