

**Feasibility of Computer Assisted English Proficiency Tests in Turkey:
A Field Study¹**

Belgin AYDIN

Anadolu University, Turkey

baydin@anadolu.edu.tr

Murat POLAT

Anadolu University, Turkey

mpolat@anadolu.edu.tr

Serkan GERİDÖNMEZ

Anadolu University, Turkey

sgeridon@anadolu.edu.tr

Emel AKAY

Anadolu University, Turkey

eakay@anadolu.edu.tr

Abstract

The content and administration procedures of proficiency tests, the difficulties faced and the perception of “proficiency” in schools of foreign languages in Turkish context are issues worth focusing on. One of the comprehensive studies on this topic was conducted by Aydın, Akay, Polat and Geridönmez (2016) in which they conducted a series of interviews with the administrators of 12 schools of foreign languages in Turkey as a part of a Scientific Research Project of Anadolu University. One of the striking findings of this study was the willingness and enthusiasm of administrators for computer assisted language proficiency tests. They stated that this practice may bring certain advantages regarding work load and quality. Therefore, two international test providers which design and administer such tests were visited within the scope of this research project to gather information about the technical and procedural information regarding the following issues: the reasons of the transition from paper-based tests to computer-based ones; the decisions taken prior to transition and preliminary works; the advantages and difficulties of computer based tests; and automatic scoring practices. This paper aims to report the results of these interviews by suggesting a possible path to develop our nationwide computer based proficiency test. development of the student.

Keywords: Language testing, Language proficiency, Paper based and computer based testing, Proficiency exam

¹ Research paper, Received: 10.01.2017 Accepted: 13.02.2017

Introduction

It is a clearly established fact that testing is an inevitable component of language teaching process. Not surprisingly, it is also the mostly discussed phenomenon in Turkish higher education system (Altınışdört, 2010; Aydın, 2000; and Bayraktaroğlu, 2010). The variety of practices considering the foreign language proficiency are the main sources of the concerns. Language proficiency tests prepared by the institutions, lack of agreed content and the administration framework, and disparities of the concept of “proficiency” bring about several variations and drawbacks in practical and theoretical terms. Therefore, focusing on proficiency tests is an issue worth studying because of the above mentioned problems and concerns.

One of the recent studies on this topic was carried out by Aydın, Akay, Polat and Geridönmez (2016) in which they conducted a series of interviews with the administrators of 12 schools of foreign languages in Turkey as a part of a scientific research project of Anadolu University. The aim of this study was twofold. It aimed to explore the practices used or offered by the universities to prepare reliable and valid language proficiency tests, and to discuss the feasibility of these practices in their contexts. The study also aimed to gather opinions from these state universities in Turkey about the use of computer assisted assessment techniques in the assessment of language proficiency, as well as to identify the existing practices, if there are any. The universities involved in that study were chosen among those listed in Turkey Special Listing mentioned in University Ranking by Academic Performance – Turkey - URAP-TR (URAP, 2013). The criterion for the choice of these state universities was the highest number of students enrolled each year. The video recorded interviews were conducted with the directors and the testing unit members–if they were available– of the participant universities. Accordingly, a total of 12 directors and 26 testing unit members of School of Foreign Languages in the following state universities participated in the study: Anadolu,

Ankara, Boğaziçi, Çukurova, 9 Eylül, Erzurum Atatürk, Gazi, Hacettepe, İstanbul, İstanbul Technical University, Middle East Technical University and Selçuk. The data were collected through focus group interviews conducted by using a semi-structured interview form prepared by the researchers. These video recorded interviews were later transcribed and analyzed to determine the themes mentioned.

The findings revealed a detailed picture of the current practices of universities regarding language proficiency tests, the problems they encounter in the process of preparation, administration, and evaluation and their attitudes towards computer assisted language assessment. The most prominent findings of the study can be summarized as follows:

- The perceptions of the institutions about language proficiency differ to a great extent; there is a lack of commonly agreed standards. While most of the institutions state taking CEFR as the basis for their proficiency concept, the way they perceive proficiency is highly affected by their context-specific conditions and beliefs. Therefore, even if a student receives language education in another institution, the language level they receive is not accepted by most of the schools delivering similar education. Even nationally recognized YDS (Foreign Language Test administered by OSYM- Measurement, Selection and Placement Center) is believed to be inadequate in determining language proficiency.
- They all believe in the importance of involving all skills in a proficiency test; namely reading, listening, writing and speaking. Yet, most of them cannot assess speaking skill due to practical reasons.
- Most of the institutions stated not having sufficient human resources and technical equipment for the preparation, administration and assessment procedures of proficiency tests. These tests are mostly prepared and

administered by the instructors assigned for this job or volunteers to do it. The number of staff in testing units who received education in assessment and evaluation is quite low.

- They also stated experiencing certain problems in administration of proficiency tests. It is not possible to pilot the tests due to time limitations both for administration and assessment procedures. Due to high number of students, tests are in multiple-choice format and the statistical analysis of test results are not done by experts in most institutions because of the reasons mentioned above.
- All the universities that took part in the study emphasized the necessity of establishing certain standards in foreign language education. Also, all of them except one stated that they support the idea of developing a nation-wide proficiency test by using technology.

The data obtained from the leading universities of Turkey clearly reveals that there are problems and remarkable differences among universities regarding the preparation, administration and assessment of proficiency tests. Besides, it seems inevitable to benefit from technology to catch up with developed countries in terms of test preparation and assessment processes and to prepare a test that is close to internationally recognized language tests in terms of validity, reliability and usability. One of the striking findings of this study is the willingness and enthusiasm of the administrators for computer assisted proficiency tests. They stated that such a practice will bring certain advantages in terms of work load and quality of the tests.

The second step of the project was visiting two of the internationally recognized testing organizations abroad - namely, Educational Testing Service (ETS) and Pearson. The aim of this visit was to collect information on technical and procedural aspects of the administration

of international computer based tests. The first visit was on August 24th 2015 to ETS in Princeton, New Jersey and the second one was on 27th and 28th August 2015 to Pearson in Menlo Park, California. The participants were the directors and experts responsible for various stages of the preparation, application and various stages of the development, operation, and validation of their tests. Due to the security and confidentiality policies of the organizations, these interviews were not recorded or taped. The researchers took notes during the interviews.

The aim of this paper is to report the findings from the interviews conducted with the professionals in the testing organizations visited. The ultimate aim is to suggest a possible path to develop our nationwide computer based proficiency test. The content analysis of the results combining the applications of both institutions will be reported in the following categories suggesting some implications for Turkish context: The reasons of the transition from paper based tests to computer based ones; the decisions to be taken prior to transition and preliminary works; the advantages and difficulties of computer based tests; and automatic scoring practices.

Findings

The reasons of the Transition from Paper Based Tests to Computer Based Ones

Deciding to digitalize language testing is an important milestone for the institutions operating in the field of testing and assessment and for educational organizations. Since the ultimate aim of these visits is to collect information about the feasibility of computer assisted English proficiency tests in Turkey, one of the most important issues to consider is the motives of these institutions for the transition from paper based tests to computer based ones. A considerable number of reasons stated during the meetings will be presented in detail in this section.

One of the main reasons of such a transition was the changes in the policies of the institution regarding test contents and administration procedures. Due to the recent trends and developments in educational assessment, both institutions visited in this study conducted research and held meetings to discuss necessary changes in their policies to keep up with these current trends and developments. Both institutions stated this as the most important motive for such a transition. They believed that this was also a necessity to keep or increase market share in the sector.

The second reason mentioned by these institutions was the potential of increased security in test administrations. It is certain that test security is one of the most important issue that test designers and administrators should take into consideration. Advanced computer technology now provides high levels of security in terms of storing and transferring data; therefore, it is possible to store test items in the form of a question bank and transferring the test data to and from test centers in a more secure and faster way compared to paper based test practices thanks to developed crypto technology. The institutions also stated that this practice reduced paper work and shipping procedures to a great extent.

Thirdly, digitalized test design is assumed to enable test designers to include certain functions and constructs in the content such as interaction and voice recording, which is not practical and feasible in paper based test administrations. Many more functions can also be added as technology develops.

Another motive mentioned regarding the transition to computer based testing was innovation. Both institutions stated that integrating innovative trends and practices into test designs would bring a new dimension to testing. They also mentioned that innovation brings a new vision and motive for both test designers and test takers since individuals are enthusiastic about trying new and modern practices.

Similar to that motive, increasing positive public opinion towards the integration of technology into life practices was another factor triggering digitalization in the field of testing. Today, individuals strongly support the idea of computer use in every field since such practices save time and reduce workload in many applications. Therefore, it is inevitable that the use of computer in language testing will bring similar advantages to the field and practices. In short, the institutions felt the need to meet such a demand from the public by digitalizing their test designs.

Still another important reason stated was the wish to prepare and administer more efficient tests in terms of time, workload and content. It is believed that computer-assisted testing enables test designers and administrators to save time, reduce workload and to prepare more interesting contents.

Finally, faster scoring was mentioned as a very important motive for transition to computer-assisted testing. Since paper based tests requires the transfer of booklets and writing papers to the center via shipping and later the assessment procedures by human raters take a long time, these institutions searched for the ways to make assessment of writing and speaking parts of the tests faster by benefitting from advanced technology. These attempts resulted in “automated scoring”, which made it possible to make assessment of writing papers and speaking interviews in split seconds. This was actually the most important practical outcome of computer-assisted testing practices, as stated by the interviewees.

In summary, keeping up with recent trends, increasing security, having more flexibility in the content of the tests, addressing public demand, preparing more efficient tests in a shorter time with reduced workload, more interesting content and finally more efficient and faster scoring were the main reasons for the transition from paper based tests to computerized tests.

The Decisions To Be Taken Prior to Transition and Other Preliminary Works

It is clear that transition to digitalized language testing is an important step for institutions, so it is essential to make certain decisions and do some preliminary analyses prior to transition. The interviews in these institutions revealed some of these decisions and steps to be taken regarding prior to such a transition.

One of these was whether to link the scores taken from computer-assisted testing applications to traditional paper based test or design them as separate tests testing different constructs. This decision certainly affects test design, its content and administration procedures. The second decision to consider was whether continuous testing or timed-date testing would be a better option for computer-assisted testing. Both have advantages and disadvantages depending on administration procedures and facilities. In continuous testing, test takers have the chance to take the tests whenever they wish to, and it is easier to solve problems in continuous testing since it is an ongoing process and there is time to make necessary changes when problems occur. However, it has also certain disadvantages such as high levels of item exposure, higher security risks, the need for a large pool of items and cost. As for timed-date testing, there are dates and times scheduled for the tests and test takers apply for each session and take the tests. Moreover, it is easier to control item exposure and to carry out post equation. However; it is necessary to have more tests seats, and scoring is slower compared to continuous testing. Therefore; these conditions should be taken into consideration while planning to initiate a transition to computer-assisted testing.

In addition, financial issues are one of the important factors to consider during transition to computer-assisted testing. Since this type of testing depends on advanced technology to a great extent, expenses are high for most applications. Automated scoring requires

sophisticated and costly software and programming. As a result, a careful and detailed financial plan must be made before digitalizing tests.

The next decision to make is whether the test will be adaptive or linear. In adaptive tests, the software releases the test items according to the level of the test taker. It is obvious that this application requires more advanced and expensive technology. As for linear testing, test items are the same for each test taker and do not change according to test takers' level. The factors to be considered for this decision are the number of students, test items, technical specifications and configuration of computers.

In addition, educational institutions should carry out a detailed analysis regarding their facilities and abilities for such a transition to determine the advantages and disadvantages of the transition. Besides, the institutions should make sure that they have the necessary infrastructure and expertise to carry out equation of the tests since it is important to have equal tests for each test takers. Finally, the procedures for the administration of the test should be planned according to the diverse conditions of test centers. Such diversities may require different procedures.

As a result, how the scores will be evaluated and reported to the test takers and the institutions, when the test will be delivered, what kind of financial plan will be followed and what facilities test centers should have are all important decisions to be taken before transmitting to computer based tests.

Advantages of Computer-assisted Testing

The most important motive for these two institutions to plan a transition from paper based tests to computer based ones is the advantages of digitalizing test design. One of these advantages of computer-assisted testing is that it allows continuous testing and more flexible time for scoring since test takers take the test in different sessions at flexible times. It is clear

that this flexibility decreases the amount of the tests or papers to be graded at one time and the workload on graders to a great extent. This situation also gives test takers flexibility to take the test according to their own schedule. Secondly, it was mentioned that a question bank can be created more effectively and detailed statistical analyses are done more easily and quickly since test data are stored digitally and can be accessed anytime easily. Thirdly, tests take shorter time if they have adaptive formats since test items are released according to the level of test takers, and low level learners do not have to answer all the items in the test. In addition, it is possible to prepare different forms of the test to reduce the risk of item exposure and have higher security. The other advantages are related to assessment procedures. Computer-assisted testing makes central assessment easier and more effective. Finally, it is technically possible to apply automated scoring in computer-assisted testing. To summarize, a flexible schedule in both taking and evaluating the tests, a less workload for graders, a more effective test preparation and analysis, a less time period necessary to take the test and finally easier and more effective assessment procedures can be listed as the most important advantages of computer based tests.

Disadvantages of Computer-assisted Testing

Although computer-assisted testing has considerable advantages, there are some disadvantages that should be considered prior to the transition to computer-assisted testing. One of these drawbacks is the necessity of a considerably large pool of test items in order to have parallel forms of the test, especially when continuous assessment is preferred. At this point another difficulty arises, which is the challenge of preparing parallel forms of the test. This is very important for the reliability and validity of the test. Tests designers must be sure that each test taker is exposed to parallel items in each test. This process requires detailed analyses before and after the administration of the tests and hard work. It is recommended to have experts to carry out this process.

Another concern is the security of question bank and test administration process. Although this is also a concern for paper-based tests, computer-stored test items and test data are vulnerable to cyber-attacks. Therefore; it is recommended to install special software to ensure security of the data, which also increases the cost to some extent.

Still another challenge was mentioned as potential technical problems such as power cut, software and hardware problems and computer configuration and compatibility issues. Such technical problems might cause serious problems during the administration of the test, so it is essential to make sure that they will not occur anytime. Such problems might also affect the reputation of the institution negatively.

In addition, it was stated that test cost might be a concern if the number of test takers is low. Since it is a costly setup, a low number of participants could not adequately compensate for the expenses that would need to be outlaid.

As for the increased validity and reliability, it is necessary to scientifically compare the scores with those taken from traditional paper based tests. Such an analysis and equation is a good indicator for the validity and reliability of the test.

Another problem may occur for those test takers who are not good at using computers and do not feel comfortable while writing and speaking to a computer screen since this situation may affect their real performance. However, it is predicted that technology-competent new generation will solve this problem easily. Besides, there might be some pre-training sessions to increase the familiarity of test takers for this new testing environment.

Finally, the cost and complexity of automated scoring might cause burden for both test designers and institutions although it decreases the work load and grading time at the end. It is the institutions' decision to make whether it is worth doing automated scoring or not. It was

also stated that, in the near future, due to advancing technology and the availability of experts in the field, this concern might not be a problem for institutions and test designers anymore.

Automated Scoring

One of the striking innovations brought by computer-assisted testing is automated scoring. With the advances in new technology and artificial intelligence, automated scoring is becoming more popular and more reliable in testing environments. Although this practice is still developing, it is predicted that it will be the most common way to assess writing and speaking performance of test takers to make grading process far faster than human rating.

Conclusion and Implications

The results of the study by Aydın et.al (2016) and this study explicitly imply the possibility of a transition from paper based test system to a computer based one. It is clear that due to the problems experienced in assessing the proficiency level of language learners at the preparatory programs in Turkey, a more efficient, practical and scientific solution is essential. The solution could be establishing a computer based testing system from which all institutions can benefit. In order to do this, there are some implications we need to get as the experts in the field point out.

Computer based assessment has certain advantages which give us very good reasons for a transition from paper based to computer based testing such as increased security, more flexibility and efficiency in terms of content, delivery and scoring. However; jumping into a system without taking necessary precautions and making necessary planning would bring more problems rather than solving the existing ones. There are decisions to be made and serious questions to be answered considering all the stages of such a transition process starting from pre-planning to assessment:

Pre-planning

- Do we really want to have a transition from paper based testing to computer based one? Is it necessary for our context? What would be our reasons to do so? Will it bring a solution?
- Do we have sufficient financial resources to realize such a transition?
- Do we have experts in the area?
- What facilities and abilities do we have?

Planning

- What will be included in the tests? Who will review them? When?
- Who will prepare the questions?
- How will the questions be stored?
- How will the tests be delivered?
- When will the tests be delivered?
- How often will the tests be delivered?
- How will the reliability and validity be constructed?
- How will the tests be evaluated (human, machine or blended)?
- What kind of a program will be used in automated scoring? Who will prepare it?
- Who will be the raters? How will they be selected, trained and monitored?
- Who will give feedback to the test items? When and how?

- What kind of rubric(s) will be used for assessment? Who will develop it or them? How and when?

Delivering

- Who will be responsible for the delivery?
- How will the center be controlled?
- How will the arising problems be solved?

Assessment

- Who will evaluate the results?
- When will the results be evaluated?
- How will the results be delivered?

All of the questions mentioned above should be answered before making the decision to prepare computer based tests. However, the most important issue even before taking such a decision is having national standards for English language teaching in Turkey. As mentioned very frequently nowadays, the Turkish higher education system has problems in language teaching. Although a great deal of time, energy and money is spent on this issue, the results of the meetings and the studies all pinpoint the problems experienced in language teaching and the unsatisfactory success gained in Turkey (YÖK, 2015, Yabancı Diller Yüksekokulları Yöneticileri Toplantısı, 2015, British Council, 2015, Aydın, et al, 2016). All the references imply the necessity of having national standards as the first initial step of solving this problem.

With this study and the previous one conducted in this research project, we are suggesting that a team including professionals from English language teaching work together to come up with

our own solution. The group needs to include practitioners working in the area as well as the theoreticians. We strongly believe that by getting expert opinion and support from internationally accredited institution, Turkey has the potential of finding out its own context specific solution to address the needs of its own students.

We asked many questions in this paper and answering all of them is out of the scope of our aims and also out of our limits and expertise. We hope at these questions and many others not addressed in this study can act as preliminary step in finding answers to them.

References

- Altıncı, G. (2010). Yabancı Dil Öğretiminde Nasıl Bir Ölçme-Değerlendirmeye İhtiyaç Vardır? *Kuramsal Eğitimbilim*, 3(2), 175-200.
- Aydın B., Akay E., Polat M., Geridönmez S. (2016). Türkiye'deki Hazırlık Okullarının Yeterlik Sınavı Uygulamaları ve Bilgisayarlı Dil Ölçme Fikrine Yaklaşımları. *Anadolu Üniversitesi Sosyal Bilimler Dergisi*, 16 (2).
- Aydın, S. (2000). İngilizce Öğretiminde Uygulanan Testlerle İlgili Sorunlar ve Çözümler. *Atatürk Üniversitesi K.K. Eğitim Fakültesi dergisi, Dil Bilimi ve Dil Öğretimi Özel Sayısı*, 4(1), 182-192.
- Bayraktaroğlu, S. (2010). Yabancı Dil Eğitimi Gerçeği, Yabancı Dilde Eğitim Yanılgısı. *Cumhuriyet Bilim ve Teknik* (CBT1221/18).
- British Council. (2015). *The State of English in Higher Education in Turkey: A Baseline Study*. Yorum Basın Yayın Sanayi Ltd. Şti, Ankara.
- URAP. (2013). *Ekim 2013 Basın Açıklaması*. Ankara, Urap Araştırma Laboratuvarı, ODTÜ Enformatik Enstitüsü. [Available online at <http://tr.urapcenter.org>], Retrived on November 25th 2015.
- Yabancı Diller Yüksekokulları Yöneticileri Toplantısı (2015). [Available online at <https://ydyotr.wordpress.com/toplantilar>] Retrieved on December 15th 2015.
- YÖK. (2015). *Bologna Süreci Nedir?* Yükseköğretim Kurulu Başkanlığı, Uluslararası İlişkiler Birimi, Ankara. [Available online at <http://www.yok.gov.tr/web/uluslararasıiliskiler/bologna-sureci-nedir>] Retrived on November 27th 2015.