

Kocaeli Üniversitesi

Eğitim Dergisi

E-ISSN: 2636-8846

2019 | Cilt 2 | Sayı 2

Sayfa: 134-142



**Kocaeli University
Journal of Education**

E-ISSN: 2636-8846

2019 | Volume 2 | Issue 2

Page: 134-142

İngilizce öğretmeni seçiminde Praxis diğer dilleri konuşanlara yönelik İngilizce (ESOL) sınavı ile Türkiye Kamu Personeli Seçme Sınavı'nın (KPSS) karşılaştırmalı geçerlik analizi

The comparative validity analysis of Praxis English to Speakers of Other Languages (ESOL) test and Turkish Public Personnel Selection Examination (KPSS) for selecting English language teachers

Pınar KIR, <https://orcid.org/0000-0002-8612-667X>

Fatih Sultan Mehmet Vakıf Üniversitesi, Yabancı Diller Bölümü, pkartal@fsm.edu.tr

ARAŞTIRMA MAKALESİ

Gönderim Tarihi
18 Haziran 2019

Düzeltilme Tarihi
12 Aralık 2019

Kabul Tarihi
24 Aralık 2019

Önerilen Atıf

Recommended Citation

Kır, P. (2019). The comparative validity analysis of Praxis English to Speakers of other languages (ESOL) test and Turkish public personnel selection examination (KPSS) for selecting English language teachers. *Kocaeli Üniversitesi Eğitim Dergisi*, 2(2), 134-142. <http://dx.doi.org/10.33400/kuje.579561>

ÖZ

Yabancı dil öğretmenlerine olan talebin artmasıyla birlikte, kurumlar işe alım süreçlerinde öğretmenlere farklı testler uygulamaya başladı. Bazı ulusal değerlendirme merkezleri, Türk Kamu Personeli Seçme Sınavı (KPSS) gibi kendi sınavlarını geliştirirken, diğer kurumlar Praxis Diğer Dilleri Konuşanlara Yönelik İngilizce (ESOL) Sınavı gibi uluslararası testleri kullanmaktadır. Alandaki testlerin güvenilirliğini araştıran birçok çalışma olmasına rağmen, testlerin geçerliğini belli kriterlere göre inceleyen sadece birkaç çalışma olduğu görülmektedir. Bu nedenle, bu çalışma ESOL ve KPSS sınavlarını sistematik olarak gözden geçirerek farklı sınavların geçerliğini karşılaştırmayı ve sınavların geçerlik derecelerini analiz etmeyi amaçlamaktadır. Bu amaçla, her iki sınav da dört geçerlik türüne göre analiz edilip, karşılaştırmalı nitel vaka çalışması yöntemi takip edilmiştir. Veriler internet sitelerinden toplanmış ve doküman analiz çalışması yaklaşımı ile analiz edilmiştir. Bu araştırmanın sonuçları, her iki sınavın da geçerlik kriterleri açısından bazı benzerlikleri ve farklılıkları olduğunu göstermiştir. ESOL Sınavı tüm geçerlik türlerinde daha tutarlı maddeler sunarken, KPSS Sınavı birçok test geçerliği gereksinimini karşılamamaktadır. Bu sonuçlar, test merkezlerinin zaman ve pratiklik kısıtlamaları ve ayrıca sınav katılımcılarının sayısı ile açıklanabilir.

Anahtar Sözcükler: Geçerlik, Praxis diğer dilleri konuşanlara yönelik İngilizce (ESOL) Sınavı, İngilizce öğretmeni, KPSS, Uluslararası sınav

ABSTRACT

Along with the increasing demand for language teachers, institutions started to apply different tests to recruit teachers. While some of the national assessment centers develop their own tests such as Turkish Public Personnel Selection Examination (KPSS) for Selecting English Language Teachers, other institutions use international tests such as Praxis English to Speakers of Other Languages (ESOL) Test. Although there are many studies researching reliability of tests in the field, literature shows that there are only a few studies to determine a test's validity based on the validity specifications. Therefore, this study aims to compare and contrast validity of different tests, systemically reviewing ESOL Test and KPSS Test in order to find out whether they are valid or not. For this purpose, both tests were analyzed based on four types of validity. Comparative qualitative case study method was followed. The data were gathered from the websites and analyzed through document analysis study approach. The results of this investigation showed that both tests have some similarities and differences in terms of validity criteria. Whereas ESOL Test offers more valid test items in all validity types, KPSS Test does not meet many requirements of test validity. These results can be explained through the time and practicality constraints of test centers and also the number of test takers.

Keywords: English language teacher, International test, Praxis English to Speakers of Other Languages (ESOL) Test, Turkish Public Personnel Selection Examination (KPSS), Validity

INTRODUCTION

With the demands of globalization, English has become the lingua franca of the 21st century. Apart from the inner circle where English is spoken as the first official language (e.g., the UK, the USA) and the outer circle where English is spoken as the second language (e.g., Philippines, South Africa), English has increased its dominance in the outer circle where English is taught as a foreign language (e.g., Turkey, Germany). It has been acknowledged as the language of education, science, aviation, medicine, and trade all over the world and it created a common communication ground for all countries. In that sense, English language education has gained great importance from kindergarten to higher education institutions all over the world. Following this situation, the demand for English language teachers has increased. Both national and international institutions have developed some tests to use in the recruitment process of English language teachers. Praxis English to Speakers of Other Languages (ESOL) Test and Turkish Public Personnel Selection Examination (KPSS) for Selecting English Language Teachers are two examples of these tests. ESOL Test is an international test that institutions benefit during their recruitment process in all over the world. Likewise, Turkish Public Personnel Selection Examination (KPSS) for Selecting English Language Teachers is the Turkish national test that is used in the selection of English language teachers. To compare and contrast these tests, this study systemically reviews the data for those tests aiming to provide their validity degrees.

LITERATURE REVIEW

Praxis English to Speakers of Other Languages (ESOL) Tests

Praxis English to Speakers of Other Languages (ESOL) Test is a computer-delivered exam that aims to evaluate the pedagogical and linguistic knowledge of prospective English teachers who are going to work in primary and middle schools. ESOL Tests are applied in the international test centers such as universities and some other institutions. Candidates can apply to international centers to take the exam. The test takes 2 hours, including two sections. The first section contains 20 selected response listening questions. Questions cover oral grammar, vocabulary, and pronunciation. The second section includes 100 selected response questions. That part tries to measure the knowledge and skills of teachers under four main parts: Foundations of linguistics and language learning; planning, implementing and managing instruction; assessment and cultural and professional aspects of the job. Specifications of these main titles are explained in detail in the exam guide. Candidates can study for the exam following the instructions in the guide and solving the mock test.

Turkish Public Personnel Selection Examination (KPSS) for Selecting English Language Teachers

Turkish Public Personnel Selection Examination (KPSS) for Selecting English Language Teachers is a high-stake exam that aims to evaluate the pedagogical and linguistic knowledge of prospective English teachers. The test is used to decide to test taker's future workplace. The prospective English teachers are assigned to state schools in direct proportion to their test results. It is administered by the Turkish Ministry of National Education to select the English teachers who are going to work in public schools. The test is held in July every year. The test is written by Assessment Selection and Placement Center (OSYM). It was first initiated and used in 2013. The exam with 50 multiple choice questions including 16 linguistics, 14 literature, and 20 subject matter knowledge questions takes 75 minutes. To be able to take the test candidates should have completed four years of language teaching education in the Faculty of Education or they should be the seniors in the faculty.

Validity

Validity which is one of the most important features of a test determines whether the test measures the specific characteristics that it intended to measure (Hughes, 1989, p. 26). Cureton

Pınar Kır

The comparative validity analysis of Praxis English to speakers of other languages (ESOL) test and Turkish public personnel selection examination (KPSS) for selecting English language teachers

(1951) suggests that validity is the fundamental feature of the test to show whether the test functions in the same way that it needs to function. It is acknowledged as the most critical component of the exam quality. Messick explains validity as “an integrated evaluative judgment of the degree to which empirical evidence and theoretical rationales support the adequacy and appropriateness of inferences and actions based on scores or other modes of assessment” (Messick, 1989, p. 1).

High test validity indicates that the test is appropriate for its intended focus. A considerable number of high-stakes tests examine whether the test items are interrelated with the intended job or profession. If a test does not require the needs of validity, it does not demonstrate the results of the skills and knowledge required for the profession. American Psychological Association identified validity framework and the types of validity in 1954 for the first time (American Psychological Association, 1954). Then, Lado (1961) and Clark and Watson (1995) concerned with the importance of validity while developing each language test although they had shown an interest to few types of validity such as content validity. In 1980, Guion (1980) suggested “Holy Trinity” to define three forms of validity: criterion-related, content, and construct. His influential idea became more popular with Messick’s unitary validity model. Nonetheless, the main weakness of Messick’s model was its failure to create concrete examples for language testers. To solve this problem, Bahman and Palmer applied five qualities: construct validity, reliability, authenticity, interactiveness, and impact (Bachman & Palmer, 1996). In 1957, Loevinger discussed the importance of construct validity indicating its superiority over the other validity types. That was the alteration from the multiple types of validity to one type of validity. Today, validity is regarded as the essential notion as it shows how well prepared a test is. Exclusively, with the increasing number of language tests all over the world, validity is fast becoming a key instrument in language test development. It is accepted as the fundamental feature of each test.

Construct validity

Cronbach and Meehl first defined construct validity in 1955. After this date, construct validity attributed to the determination of main constructs of a test. It suggested how much a test measure the constructs that it claimed to measure. Shaw and Crisps (2011) defined construct validity as a bridge between the pre-learned knowledge and test performance. According to them, the test performance of a student and their curriculum-based theories should be steady and specifications of a test construct should be determined in the initial stages of testing. In 1986, Anastasi claimed the view that construct validity is the top category of both content and criterion-related validity and Smith (2005) approved her idea defining construct validity as an umbrella term that includes test validation processes.

Content validity

Content validity refers to the representativeness of test items following the defined curriculum and syllabus. To be able to have content validity, a test should cover all specific topics and structures intended to cover as it is the critical factor of participant performance. According to Hughes (1989), content validity should be developed at the early stages of test preparation. Also, he emphasizes that there is a significant positive correlation between content validity and washback effect.

Face validity

Face validity concerns with the medium of a test. Hughes states that a test has a face validity only if it reflects the skill that it aims to measure (Hughes, 1989, pg. 33) For instance, in a writing test, participants ought to compose an essay. In the case of high-stake tests, face validity is about how the people perceive the test. Hughes concludes that face validity is getting critical attention from the students, teachers, and academics as it affects many people. On the other hand, McDowell (1995) claims that face validity and washback effect are very similar to each other.

Criterion-related validity

Criterion-related validity refers to which degree a measure can be related to its potential outcomes. It tries to assess the relationship between criterion and candidates' ability. It is one of the most crucial validity types as it is used for calibration of a test for standards. It can be best treated under two headings: concurrent validity and predictive validity. Concurrent validity tends to correlate the score of a test to another score of an established test. Concurrent validity is determined when criterion and possible outcomes are determined at the same time. Predictive ability presents the potential future performance of a candidate based on their test results. Whether the candidates are going to be successful in their departments after taking the tests is about the predictive validity of the exam. If high achievers perform better in their jobs, this situation approves the high predictive validity of the test.

Qualitative validity

Lynch (2003, p.157) claims that common standards for quality does not meet the expectation of people in qualitative studies. The same situation can be transferred to the assessment process of tests. Although many tests offer the basic validity criteria, some points such as the real quality of test items do not satisfy the test takers and administrators in large scale exams. Especially, while comparing two tests qualitative validity is crucial. For this reason, the qualitative validity of the questions ought to be high.

METHOD

A qualitative comparative case study analysis was used to investigate test validation of two high stake tests to summarize phenomenology with grounded theory. To analyze the data, document analysis was adapted. Document analysis which is a qualitative research method offers an area for researchers to be able to interpret and assess the data (Bowen, 2009). As the document, mock exams, retired questions and webpages were checked and interpreted by the researcher on the basis of comparative analysis criteria.

FINDINGS and DISCUSSION

Construct Validity

Regarding the notions of construct validity, a test should measure the construct that it is intended to measure. Having receptive (listening and reading) and productive (writing and speaking) skills and being able to teach these skills to the students are the essential features of language teachers. Also, a competent language teacher should have in-depth planning, management, and assessment abilities. To be able to measure all these fundamental constructs, a test should have high construct validity, especially in high stake testing. In that sense, KPSS is different from ESOL Test in terms of construct validity. Regarding the notions of construct validity, ESOL Test measures listening, reading, planning, management and assessment constructs of English teachers. Nonetheless, an English teacher should be competent in each of the four skills (writing, reading, listening and speaking) and the test should measure all of them. ESOL Test assesses only receptive skills and does not examine any real-life teaching tasks. Professional aspects of the job are evaluated with multiple choice questions, which does not represent the full range of educational objectives. All these drawbacks decrease the construct validity of the exam. In contrast to ESOL Test, KPSS Test only includes 50 multiple questions. It does not measure any of the productive skills (listening and speaking) or professional aspects of the job like planning, measurement, and assessment in multiple choice questions. Due to these constraints, construct validity of the KPSS Test can be attributed as low. The best indication of teaching competence is to present classroom performance in the teacher recruitment process; still, none of the tests apply this process to their constructs.

Content Validity

Based on the definition of content validity, the content of tests should be relevant to English language teaching curriculum taught at universities. Also, the tests should require to be representative of all tasks and texts that students encountered during their undergraduate education. To generate a piece of evidence for the proposition of content validity in KPSS and ESOL Tests, the curriculum of English Language Teaching Departments both in Turkey and abroad were checked. Linguistics, English and American Literature, Language Acquisition, Language Teaching and Language Testing courses are the must undergraduate level courses in all national (e.g., Bogazici University, Foreign Language Teaching Curriculum) or international language teaching programs (e.g., University of Stuttgart, English Teaching Qualification Curriculum). By the investigation of curriculums and test contents, both of the tests were similar to each other regarding content validity as both tests include these topics in their content. As an example, both tests have questions measuring the knowledge of language acquisition as stated below:

Which of the following statements is not a characteristic of Broca's aphasia?

- Patients' utterances convey the meaning they wish to communicate.
- Patients have great difficulty in comprehending speech.
- Patients reveal speech production breakdown and phonological deficits.
- Patients' poor articulation results from the injuries of the front (anterior) part of the left side of the brain.
- Patients' speech is telegraphic because they omit function words and grammatical morphemes. (e.g., ÖSYM, 2015)

Mr. Levitt overheard a conversation in the teacher's lounge regarding bilingual education. One teacher remarked to another that a student's learning of a second language is much different than learning the first language. Which of the following would be an appropriate response by Mr. Levitt?

- Learning a second language is not different because it is essentially like learning new vocabulary words, which is done in any language
- Learning a second language is not much different than learning the first language; the language acquisition for L1 and L2 are very similar
- The acquisition of a second language is much different from learning the first language due to the advanced cognitive development of students
- The acquisition of the second language is much different for students because of the cultural assimilation required to become proficient in a new language (e.g., Praxis ESOL, mock exam)

Also, following the curriculum objectives, tests assess the knowledge of language testing with different question types as in the examples below:

A teacher is monitoring her class while the students are involved in a group activity exploring the size of angles in a set of triangles. She moves from group to group, pausing and watching the group dynamics. What the teacher is doing can best be described as:

- formal summative assessment.
- informal summative assessment.
- informal formative assessment.
- formal formative assessment. (e.g., Praxis ESOL, mock exam)

When we interpret scores from language tests as indicators of test-takers' language ability, a crucial question is, 'To what extent can we justify these interpretations?' The ---- refers to the extent to which performance on tests is consistent with the interpretations we make on the basis of a theory of abilities like proficiency, fluency and communicative competence. The ----, on the other hand, concerns whether a test actually samples the subject matter about which conclusions

are to be drawn. If a course has ten objectives but only two are covered in a test, then this type of validity suffers.

Choose the option that completes the sentence.

- a) construct validity / criterion-related validity
- b) predictive validity / face validity
- c) concurrent validity / content validity
- d) predictive validity / criterion-related validity
- e) construct validity / content validity (e.g., ÖSYM, 2015)

Despite their similarity, ESOL Test has a more definite content and topic match with the university curriculums. ESOL Test has "planning, implementing and managing instruction; assessment and cultural and professional aspects of the job" parts in its content, which is more related and including with the curriculum, in addition to having linguistics and subject matter knowledge questions. However, the content of KPSS test including only linguistics, literature and subject matter knowledge cannot provide a comprehensive review of "planning, implementing and managing instruction; assessment and cultural and professional aspects of the educational job processes" due to the restricted number of questions." Concerning lack of some content covered in the university curriculum, content validity rate of KPSS Test is less than ESOL Test.

Face Validity

Regarding face validity, both tests have a significant rate of face validity. ESOL Test assesses listening skill with a listening recording and questions, and the other skills and teaching knowledge with a multiple-choice exam. All of these parts are perceived as a normal test design by candidates, which increases the face validity of the exam. However, the KPSS test only has multiple questions and looks like a multiple-choice test. It does not include any listening or writing part.

Criterion-Related Validity

Following criterion-related validity notions, KPSS test does not support the idea of criterion-related validity as it is a norm-referenced test. The points of teachers are determined by considering the standard deviation of participants in norm-referenced tests. Therefore, it is impossible to demonstrate the real criterion validity of the test related to the test results of participants. On the other hand, there is not an equivalent recruitment test in Turkey to compare concurrent validity. For predictive validity, the test does not guarantee the future success of teachers in their job, as it does not assess their real-life teaching performance. Nonetheless, ESOL Test is not a large-assessment test, so it does not have norm-referenced criteria. Each candidate's performance can be assessed regarding their test result relevant to the criterion. That increases the criterion-related validity of the test. For predictive and concurrent validity, ESOL Test lacks evidence like KPSS.

CONCLUSION

The comparison of two high stake exams reveals that KPSS Test is not a well-constructed high-stake assessment comparison to ESOL Test. Regarding four types of validity, the test does not offer any significant validity rate. Also, qualitative validity of the questions is very low. The test does not require the demands of a valid test both for prospective teachers and purpose. On the other hand, although ESOL Test supplies some criteria of validity types, it is not an ideal valid test as it does not assess all the elements of language teaching. It seems possible that these results are due to time and practicality constraints of test centers and also the number of test takers.

Finally, this study has a number of limitations. Firstly, the study has only investigated validity based on the curriculum and exam comparison. However, there was no participant interviewees to get feedback from. Also, the research has only qualitative data. Further studies can be

triangulated with quantitative data and participants. Another limitation was about ESOL Test examples. As it is an online test, there were no retired questions available, so mock exams were taken into consideration.

REFERENCES

- American Psychological Association. (1954). Technical recommendations for psychological tests and diagnostic techniques. *Psychological Bulletin*, 51, 201-238.
- Anastasi, A. (1988). *Psychological testing* (6th ed.). New York: Macmillan Publishing Company.
- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests*. Oxford: Oxford University Press.
- Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative Research Journal*, 9(2), 27-40. doi: 10.3316/QRJ0902027
- Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, 7(3), 309-319. doi: 10.1037/1040-3590.7.3.309
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52(4), 281-302. doi: 10.1037/h0040957
- Cureton, E. E. (1951). Validity. In: E. F. Lindquist (Ed.), *Educational measurement* (621-694). Washington, DC: American Council on Education.
- Guion, R. M. (1980). On Trinitarian doctrines of validity. *Professional Psychology*, 11(3), 385-398. doi: 10.1037/0735-7028.11.3.385
- Hughes, A. (1989). *Testing for language teachers*. Cambridge and New York: Cambridge University Press.
- McDowell, C. (1995). Assessing the language proficiency of overseas-qualified teachers: the English language skill assessment (ELSA). In G. Brindley (Ed.), *Language Assessment in Action* (11-27). Sydney: National Center for English Language Teaching and Research.
- Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13-104). New York, NY: American Council on Education and Macmillan.
- Lado, R. (1961). *Language testing: The construction and use of foreign language tests: A teacher's book*. Bristol, Inglaterra: Longmans, Green and Company.
- Lynch, B. K. (2003). *Language assessment and programme evaluation*. Edinburgh: Edinburgh University Press.
- Loevinger, J. (1957). Objective tests as instruments of psychological theory. *Psychological Reports*, 3, 635-694. doi: 10.2466/pr0.1957.3.3.635
- ÖSYM. (2015). 2015 Kamu personel seçme sınavı (KPSS) *Public Personel Selection Examination (PPSE)]. Ankara: Öğrenci Seçme ve Yerleştirme Merkezi. Retrieved from <http://dokuman.osym.gov.tr/pdfdokuman/2015/KPSS/KILAVUZ14062017.pdf>
- Shaw, S. and Crisp, V. (2011). Tracing the evolution of validity in educational measurement: Past issues and contemporary challenges. *Research Matters: A Cambridge Assessment Publication*, 11, 14-17.
- Smith, G. T. (2005a). On construct validity: issues of method and measurement. *Psychological Assessment*, 17(4), 396-408. doi: 10.1037/1040-3590.17.4.396

GENİŞLETİLMİŞ ÖZET

Giriş

Praxis Diğer Dilleri Konuşanlara Yönelik İngilizce (ESOL) Sınavı ve Türk Kamu Personeli Seçme Sınavı (KPSS) öğretmen işe alım sınavlarına iki örnektir. Praxis ESOL Sınavı, kurumların işe alım sürecinde dünyanın her yerinde yararlanılan uluslararası bir sınavdır. Aynı şekilde, KPSS, Türkiye’de İngilizce öğretmenlerinin seçiminde kullanılan ulusal sınavıdır. Bu çalışma, mevcut sınavları ve sınavların geçerlik derecelerini karşılaştırmak için sınav verilerini sistematik olarak gözden geçirir.

Yöntem

Bu çalışmada fenomenolojiyi, temelli teoriyle özetlemek için iki yüksek riskli sınavın test geçerliği araştırılmış ve nitel bir karşılaştırmalı vaka çalışması analizi kullanılmıştır. Verileri analiz etmek için doküman analizi uyarlanmıştır. Doküman olarak, örnek sınavlar, çıkmış sorular ve web sayfaları karşılaştırmalı analiz kriterleri temelinde araştırmacı tarafından kontrol edilip yorumlanmıştır.

Bulgular

Bu çalışmanın sonuçları, Praxis ESOL ve KPSS’nin geçerlik türleri açısından bazı benzerlikleri ve farklılıkları olduğunu göstermiştir. Yapı geçerliği bakımından, KPSS, Praxis ESOL Sınavı’ndan farklıdır. Yapı geçerliği kavramlarına göre, ESOL Sınavı İngilizce öğretmenlerinin dinleme, okuma, planlama, yönetim ve değerlendirme yapılarını ölçer ancak beceri öğretimi değerlendirmesi yapmaz. ESOL Sınavı’nın aksine, KPSS Sınavı, üretken becerileri (dinleme ve konuşma) veya işin planlama, ölçme ve değerlendirme gibi profesyonel yönlerini çoktan seçmeli sorularda ölçmez. İçerik geçerliği ile ilgili olarak, müfredatların ve sınavın içeriğinin araştırılmasıyla, her iki sınavın da birbirleriyle aynı olduğu görülmüştür, çünkü her iki sınav içeriğinde öğretim programı konuları dahil edilmiştir. Benzerliklerine rağmen, ESOL Sınavı üniversite müfredatları ile daha kesin bir içeriğe ve konu eşleşmesine sahiptir. ESOL Sınavı, dilbilim ve konu bilgisi sorularına ek olarak, içeriğinde, dersle ilgili daha fazla olan ve müfredatla birlikte “öğretimi planlama, uygulama ve yönetme; işin değerlendirilmesi ve kültürel ve profesyonel yönleri” bölümlerine sahiptir. Yüz geçerliği ile ilgili olarak, KPSS Sınavı sadece çoklu sorulara sahiptir ve çoktan seçmeli bir test gibi gözükmemektedir. Herhangi bir dinleme veya yazma bölümü içermez. Bununla birlikte, ESOL Sınavı bir dinleme kaydı ve sorusu ile dinleme becerisini ve çoktan seçmeli bir sınavla diğer becerileri ve öğretme bilgisini değerlendirir. Kriter geçerliği kavramlarını takiben, KPSS Sınavı, norm referanslı bir sınav olduğu için kriter geçerliği fikrini desteklememektedir.

Tartışma ve Sonuç

İki yüksek riskli sınavın karşılaştırması, KPSS Sınavı’nın ESOL Sınavı ile iyi yapılandırılmış bir yüksek riskli değerlendirme olmadığını ortaya koymaktadır. Dört geçerlik türü ile ilgili olarak, sınav önemli bir geçerlik oranı sunmamaktadır. Ayrıca, soruların niteliksel geçerliği çok düşüktür. Sınav hem öğretmen adayları hem de amaç için geçerli bir sınav beklentilerini karşılamaz. Diğer yandan, ESOL Sınavı bazı geçerlik türleri kriterlerini sağlasa da dil öğretiminin tüm öğelerini değerlendirmedikten ideal bir geçerli sınav değildir. Bu sonuçların, test merkezlerinin zaman ve pratiklik kısıtlamaları ve ayrıca sınava giren katılımcıların sayıları gibi nedenlerden ötürü olması muhtemeldir.