Journal of Educational Technology & Online Learning
Volume 7 | Issue 1 | 2024
http://dergipark.org.tr/jetol

Measuring cognitive loads and attitudes of pre-service teachers in computer based testing environment

Mehmet Mevlüt Odacı [a] *, Erman Uzun [b]

[a] Niğde Ömer Halisdemir University, Türkiye.
[b] Mersin University, Türkiye.

Suggested citation: Odacı, M. M. & Uzun, E. (2024). Measuring cognitive loads and attitudes of pre-service teachers in computer based testing environment. *Journal of Educational Technology & Online Learning*, 7(1), 84-101.

Highlights

- The participants stated that they spent less cognitive effort in the CBT environment; it is thought that the use of multimedia elements contributed to generation of less cognitive load on them.
- The students also stated that the CBT environment offered them a more flexible and individualized study environment compared to the classroom environment.
- There is a low and negative relationship between the exam duration and the cognitive load score. Text-based exam question in computer based testing environment, which participants found distracting and took long to answer the related questions.
- There is a negative correlation between cognitive load and exam scores. It is possible to recommend that using multimedia integrated exam questions would be effective in terms of achievement in the computer based testing.

Abstract

This study aimed to reveal the attitudes of the students who were subjected to measurement and evaluation in an online testing environment towards the computer based testing (CBT) platform and the factors that affect their attitudes. It also examined the students' cognitive loads in the exam designed with multimedia elements from message design principles by using the facilities of this platform. The study was conducted with a case study design. The participants consisted of 32 students. Participants took a previously used and redesigned English achievement test based on the multimedia design principle. Then, cognitive load scale and attitude towards computer scale were implemented. In addition, seven volunteer participants were interviewed. The results of the study showed that factors such as students' equipment advantage, time, success, flexibility, individuality and less mental effort affected the participants' attitudes towards the CBT environment. A significant difference was found between the scores of the participants in the cognitive load scale and the duration of the exam. There was a low and positive correlation between the cognitive load score of the participants for the listening test and the time to complete these questions, and a low and negative correlation between the correct answer scores of the participants in the listening test.

Article Info: Research Article

Keywords: *Cognitive load theory, principles of message design, online testing, computer based testing (CBT), case study.*

1. Introduction

Nowadays, proliferation of universities and the increase in the number of students in universities have brought difficulties in the management, monitoring and evaluation of learning. Technology can provide some solutions to overcome these challenges. Another area where technology is being used extensively in

* Corresponding author. Mehmet Mevlüt Odacı. Niğde Ömer Halisdemir University, Türkiye.
e-mail address: mevlutodaci@ohu.edu.tr
This research is based on a M.A. thesis submitted to the Institute of Educational Sciences, Mersin University (Odacı, 2019) which was supported by Mersin University, Scientific Research Project Office, under the project number 2019-2-TP2-3524.

Doi: http://doi.org/10.31681/jetol.1366359
Received 26 Sep 2023; Revised 21 Oct 2023; Accepted 3 Nov 2023
ISSN: 2618-6586.

higher education is measurement and evaluation processes. Tests can be created using technology in different environments and by using different tools in the field of measurement and evaluation. Computer aided tests, web-based tests, and individualized tests are technology based exam platforms that are widely used today.

Although these different exam platforms are used at different levels and for different purposes, their important common feature is the use multimedia elements. The effective and efficient use of multimedia requires to pay attention to design principles. The use of multimedia in the learning environment or materials positively affects learning processes. Mayer (2003) states that multimedia elements where text and pictures are used together encourage learning. It is necessary to consider some design principles to achieve this positive effect. The relevant literature reports these principles as cognitive load theory message design principles (Yıldırım, 2016). These design principles are also needed to be taken into consideration in exam environments, because the interaction with multimedia elements and the way how they are presented can create cognitive load on students (Paas, Renkl, & Sweller, 2003). In order to prevent this increase in cognitive load level, it is necessary to pay attention to the message design principles under the cognitive load theory.

Cognitive load theory is an approach that guides the design of teaching materials and aims to enable learners to use their cognitive resources efficiently (Kılıç, 2009). Cognitive load theory assumes a working memory that is related to long-term memory and is limited in capacity (Baddeley, 1986, as cited in Kirschner, 2003, p.3). The capacity of the working memory is limited; therefore, the material or message to be prepared should not exceed the load limits of the working memory of the learners (Kirschner, 2003, p.3). Cognitive load occurs when the material or message exceeds the load limits of working memory. Message and design principles within the cognitive load theory guide the designer while preparing the material so that the material does not exceed the load limits of working memory, and it is assumed that the materials prepared according to these principles reduce the cognitive load.

The multimedia principle is also one of the message design principles. This principle recommends the use of text and graphics or videos rather than written text (Yıldırım, 2016). In this study, the questions in testing were supported by a multimedia element and presented in a design language suitable for each question. The multimedia element was presented using video and audio. Cognitive load emerges as an important variable since multimedia is used in testing. Therefore, the use of multimedia elements in learning and teaching processes requires that the cognitive load factor be considered because the design and presentation of multimedia have a positive and negative effects on the cognitive load.

A multimedia item is a presentation consisting of words and pictures that are designed to encourage meaningful learning. According to the definition of the multimedia principle, (a) the multimedia element includes words and pictures, and (b) is designed to promote meaningful learning. The multimedia element used in this study includes dynamic graphics and voice. In addition, since listening texts should be optionally presented to the participants in long listening sessions, these texts were presented during the exam process.

Cognitive load occurs while information is processed in the working memory, which is assumed to be in the human brain. Difficulty in processing information from an audio source to working memory creates a cognitive load in working memory. The relevant literature reports various definitions of cognitive load. The working memory has a limitation which should be taken into account when designing instruction (Kirscher, 2003).

Cognitive load refers to the load on the cognitive systems of learners due to the limitation in working memory (Kılıç & Karadeniz, 2004). Information is stored in long term memory after first being attended to and processed by working memory. The cognitive load should be low in order to facilitate this process or to transfer information to long-term memory (Kılıç & Karadeniz, 2004, p. 564).

Limited working memory is one of the defining aspects of human cognitive architecture and accordingly all instructional designs should be analyzed from the perspective of cognitive load. Otherwise, instructional

designs made without considering the limitations of memory are insufficient (Sweller, van Merrienboer, & Paas, 1998).

There are three types of cognitive load: Extraneous cognitive is the state where elements in a material are not related to learning objectives. Intrinsic cognitive load is associated with the number of elements required to learn the instructional topic (Yıldırım, 2016). Germane cognitive is generated by the activities performed to construct and automate mental schemas. The existing working memory capacity can be used in construction of schemas by reducing the extrinsic cognitive load (Sweller, van Merrienboer, & Paas, 1998).

The literature review shows that there are many studies on cognitive load theory. These studies make a great contribution to the literature as well as being a guide for future studies. Mayer (2003) conducted a study to examine the effect of the principles of multimedia, integrity, spatial proximity, and individualization. This study revealed that a multimedia is more effective than a printed text containing only text. In another study, Mayer (1989) presented a learning object describing how brakes work according to the multimedia principle, and then observed a 79% increase in students' creative problem solving (as cited in Mayer, 2003). It was proved that the multimedia principle, one of the principles of message design, increases success.

Dindar, Kabakçı Yurdakul and Dönmez (2015) examined the use of fixed animation and graphic animation in the computer based English achievement test in terms of cognitive load theory. The researchers measured students' response times, correct answer rates, self-expression rates in terms of cognitive load, and secondary task approaches. The study conducted with 7th grade students (n = 303) used an achievement test and the cognitive load scale developed by Paas and van Merrienboer (1994) and adapted into Turkish by Kılıç and Karadeniz (2004) as measurement tools. Also a monitoring test and the web environment were used collect information about visual search, scanning, processing speed, mental flexibility, and executive functions. The study found that the fixed graphics group completed the exam in a shorter time than the animated graphic group and there was no significant difference between the groups in terms of the rates of correct answers. The scores of the group with animation graphics were higher than the fixed graphics group in the secondary tasks and there was no significant difference between the groups in terms of their self-expression levels.

Takir (2011) redesigned a course according to the cognitive load theory. New algebra books were prepared for teachers and students within the scope of this design study. The study investigated the effect of the redesigned course on students' success in algebra and their cognitive load. It was carried out using the experimental research method and assigning 40 students to the experimental and control groups each. The lesson was taught to the experimental group using the newly designed form. The control group was taught the lesson using the current program of the Ministry of National Education. Also, a questionnaire—a rating scale—was administered to the experimental group and the participants were interviewed. The study indicated that the instruction that was redesigned according to the principles of cognitive load theory had a positive effect on the teaching algebra.

The use of computers in the field of education, as in many other fields, has become very popular in recent years. Computers are used in many areas, from presenting course contents to evaluating students (Akdemir & Oğuz, 2008). Individuals take exams CBT systems. In recent years, there has been an increase in the transition from paper and pencil based tests (PPT) to CBT. Accordingly, more and more institutions are now passing from PPT to CBT (Bugbee, 1996). Large scale institutions such as OECD-PISA, OECD-PIAAC, NAEP now conduct CBT (Lemmo & Mariotti, 2017).

CBT also allows the use of multimedia, which is unlikely in PPT. In addition, It makes it possible students to be to stimulated not only through the visual but also the auditory channel. For instance, Özturan (2016) proved the students' positive attitude towards CBT. It was investigated whether there was a change in the scores of the participants who took CBT and PPT. The results revealed that the experimental group showed a positive attitude towards CBT; however, the control group did not show interest in such testing. Most of the students in the experimental group thought that CBT increased their performance. In addition, the experimental group thought that if they took the exams on the computer, they could use the system easily.

Similarly, Prisacari and Danielson (2017) examined the relationship between test types (CBT and PPT) and cognitive load. The results support that online tests can be implemented in educational settings without imposing additional cognitive load on students. Based on the results of the study, CBT can be claimed to create less cognitive load. Nonetheless, further research is needed in different context and with the use of different online testing tools. Because each of them will have new features and capabilities to support assessment processes.

1.1 Research Problem

The study was carried out in the Measurement and Evaluation Research and Application Laboratory of a higher education institution. This laboratory is also the center where CBT activities, which are open to the use of the whole university. However, exams for English preparatory students at the School of Foreign Languages within this university are done in the classrooms using cassette players. The use of cassette players in testing can actually be considered as integration of technology into education. However, it is not a very effective and efficient educational technology today, where computers are quite common and accessed very easily.

One research problem is to investigate the attitudes of the participants towards the effective use of the CBT environment and the effects of the multimedia elements used in the exam on the exam duration, exam score and cognitive load of the participants. Another problem is the examination of the cognitive load factor since multimedia elements are also used in the testing. A multimedia element can have a positive effect on cognitive load; however, a poorly designed multimedia element can have negative effects. Therefore, it was also examined whether the multimedia used in the testing had a positive effect on cognitive load.

1.2. Aim

This study aims to measure the participants' cognitive loads in a CBT environment and determine their attitudes towards CBT. Thus, it reveals the opinions of the participants in order to expand the use of CBT environment in higher education institutions. The study provides feedback for greater integration of technology into education.

1.3. Research Questions

- What are the participants' attitude towards CBT environment and the factors that affect their attitudes?
- What is the relationship between the participants' exam score, exam duration and cognitive load levels in CBT environments?
- Is there a significant difference between the cognitive load levels of the participants, the test scores and the exam duration?

2. Methodology

2.1. Research Design

This study examined the attitudes of the participants towards the CBT environment and the factors affecting their attitudes as well as their cognitive load for the exam, which was redesigned using the multimedia element, one of the message design principles. In this context, the case study, a qualitative research method, was used as the study design. The reason why this method is preferred is that the case study is based on 'how' and 'why' questions and provides an opportunity to examine a phenomenon or event in depth (Yıldırım & Şimşek, 2018, p. 289). Woodside (2010) states that the main purpose of the case study is to provide in-depth information about the participants' perceptions and the reasons of their activities, and that data should be collected from more than one source in the same situation. According to Büyüköztürk, Kılıç Çakmak, Akgün, Karadeniz, and Demirel (2019) a case study is regarded as a distinctive approach used in seeking answers to scientific questions.

2.2. Research Process

The research problem was determined, then the participants were selected using the convenience sampling method. The questions to be included in the English knowledge test, which was going to be done in the CBT environment, were taken from Mersin University School of Foreign Languages by taking experts' the opinions in the field. It was paid attention to the fact that the questions had been previously used and their answers were accessible in order to conduct distinctiveness and difficulty analysis. First of all, distinctiveness of the questions and item difficulty analyses were made with the expert in the field of English. Based on the analysis results, six questions whose biserial item total test correlation values were lower than 0.30 were excluded from the test. Then, the questions that were decided to be used in the exam were rearranged by adding multimedia elements together with an expert in the field of material design.

While transforming the questions into multimedia elements, the multimedia principle, a cognitive load message design principle, was taken into consideration. Yıldırım (2015) defines the multimedia principle as presenting the text and the figure together rather than the written text alone. According to the multimedia principle, individuals learn better when the text and related figures are used together. However, the multimedia used must be associated with the topic. Multimedia items that are used only for motivation can negatively affect learning. In addition, the stylistic principle should be considered in the application of the multimedia principle (Yıldırım, 2016). According to the stylistic principle, written information should be given verbally when there is both a written and a formal source to lighten the load on the visual channel. Since both the text and the figure are directed to the visual channel, all the information will be tried to be processed in a single channel, which will create a load on the visual channel. When the text is given verbally, it will be processed in the auditory channel and the load of the visual channel will be lightened. In addition, when verbal narration is used, it should be ensured that the listening item can be listened again, and long listening texts should be supported with the option of written text (Yıldırım, 2016).

2.3. Data Collecting Tools

This section introduces the data collection tools used in the study. In the study, in which both qualitative and quantitative data collection tools were used, the data were collected using the cognitive load scale, an interview form, English knowledge test and attitude scale towards computer scale.

2.3.1. Cognitive Load Scale

There are different methods to measure cognitive load. A cognitive load rating scale was used in this study. There are physiological measurement and task-performance-based assessment methods (Sweller, van Merrienboer, & Paas, 1998, as cited in Kılıç, 2009). Subjective rating scale (SRS) was developed by Paas and van Merrienboer (1994) and adapted into Turkish by Kılıç and Karadeniz (2004). The adapted scale's Cronbach's alpha internal consistency coefficient was calculated as .90 and the cut-off point was 5. Values below 5 were classified as low cognitive load, and those above 5 were classified as high cognitive load (Kılıç & Karadeniz, 2004). The rating scale consists of a single item and is presented after each question, and at the end of the exam to evaluate overall cognitive load over students in the testing environment of this study.

2.3.2. Interview Form

This semi-structured interview form was prepared by the researcher. Expert opinion was taken while the interview questions were prepared. Initially, 12 questions were prepared; however, two more questions were included in the form. The interview questions were rearranged based on the expert opinions. The questions were about the CBT experiences of the participants, the usability of the CBT platform, the reasons for the participants to prefer the CBT, the participants' use of multimedia in the CBT environment and their mental efforts, and their suggestions for the CBT. The attitudes of the participants towards the multimedia elements were also included in the questions because the exam questions had been rearranged according to the multimedia principle.

2.3.3. English Knowledge Test

The English knowledge test was selected from the question pool of the School of Foreign Languages of the university where the study was conducted. Questions had been previously used were preferred. They were nine single-paragraph reading passages, and each passage included four to six questions. Of these reading passages, two were omitted because they did not fit the scope of the exam. Then, distinctiveness of the questions and item difficulty values of the remaining questions were calculated; accordingly, the items whose point biserial values were below .30 were excluded. The reading passages were arranged in a way that two of them would include videos and texts, three listening items and texts, and two only texts. Opinions of the experts working in the department of foreign language education were taken to assess the suitability of the questions for the students.

After the questions were selected, in the stage of adding the multimedia elements, firstly, an expert in vocalization of English texts voiced the questions. According to the multimedia principle, if the vocalization is likely to take a long time, the text and the multimedia should be presented together; therefore, they were given together in the TAO platform. Accordingly, two reading passages including video items and texts, three including listening items and texts, and two paragraphs including just texts were presented to students.

In this context, a group of paragraph questions received and found appropriate to be used were added to the CBT platform in the forms of

- text only
- text and listening item
- text and video voiceovers.

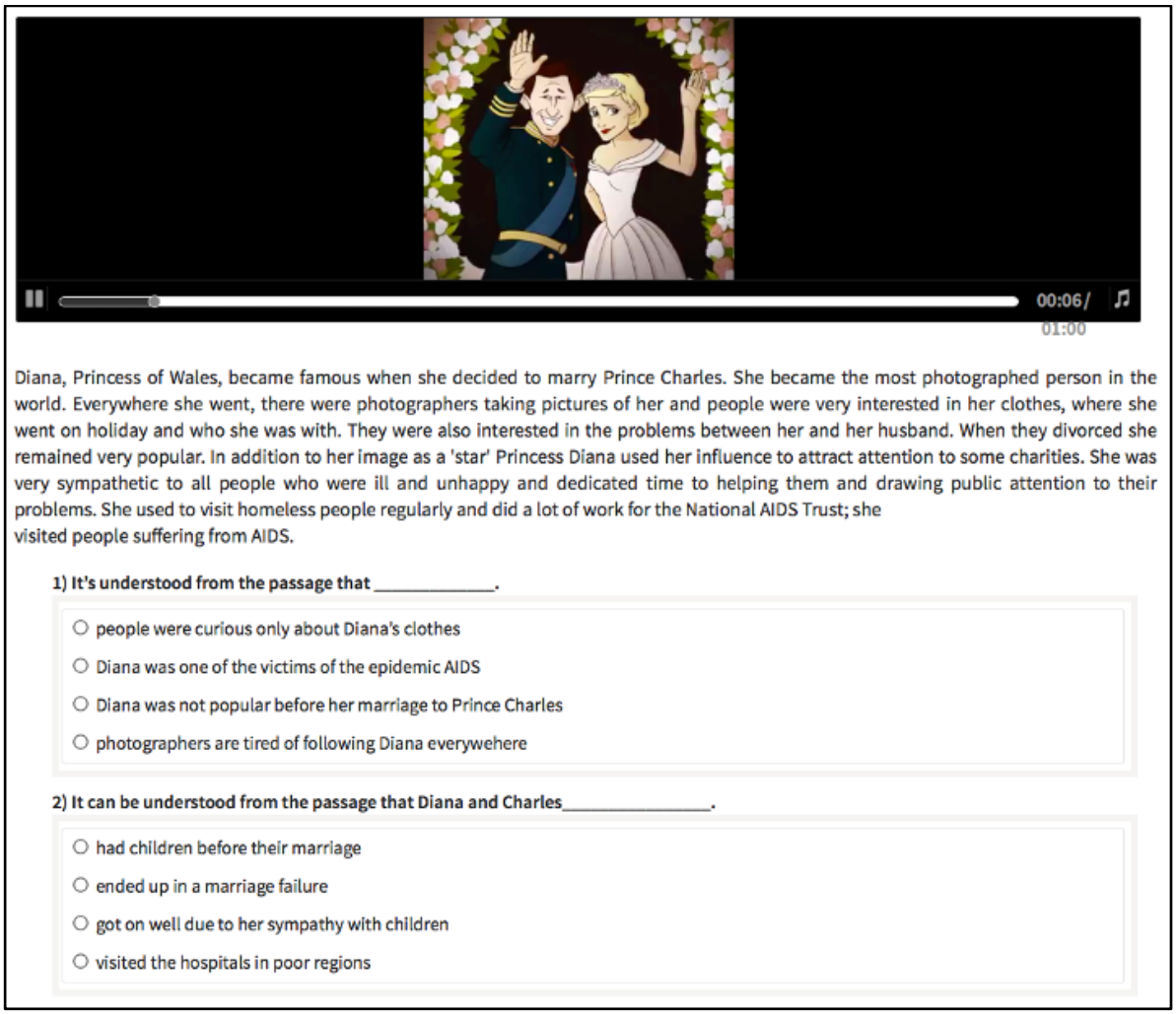

Fig. 1. Screen shot of an exam question

2.3.4. Attitude Towards Computer Scale

The attitude scale towards computer applied to the participants was developed by Aybek (2012). The scale items represent cognitive, affective, and behavioral elements and half of the items report positive and half report negative attitudes. The scale consists of 23 items. Cronbach Alpha value of the scale was calculated as .90. While selecting the scale items, the factor value loading criterion was taken as .30 and the items below .30 were not used in the original version of the scale. In the pilot application of the scale, one item below .30 was removed. This scale was used to reveal the participants' attitudes towards computer and to reveal whether there is a relationship between attitudes towards computer and other variables.

2.4. Study Group

The participants consisted of 32 preparatory English language teaching students from the school of foreign languages in a higher education institution where the study was conducted. Of the students, 26 were female and 6 were male. The participants were in the same class and department.

Of the participants, seven—four females and three males— who were selected using the convenience sampling method were interviewed after the CBT. Fraenkel, Wallen, and Hyun (2012) state that the advantage of convenience sampling method is that it is easy, but the people who accept to be interviewed should have strong views on the subject. For this reason, only volunteer participants were selected for the interview.

2.5. Data Collection Process

Study data were collected from the students at the School of Foreign Languages of a higher education institution. The students took the exam in the CBT hall affiliated with the Assessment and Evaluation

Application and Research Center of this higher education institution. There are numbered lockers in the exam hall for students to put all their belongings. The participants left their belongings in the lockers, then were taken to the exam hall. Computer desks in the exam hall are designed in cubic form to study individually. The exam hall is equipped with surveillance cameras, and they are monitored in the control room. In addition, the mainframe computer and server that control the whole system are also located in this room. TAO platform is installed in this server.

In the exam, a total of 36 questions assigned to seven reading passages were asked to the students on the TAO CBT platform. Students were told that they could only see each question only once during the exam, that they could not return to the previous question, and that they could not access the exam again when the exam was over. After passing each reading passage in the exam, the participants completed the cognitive load scale which aimed to measure their cognitive load levels. Then, at the end of the exam, the scale was repeated for the students to indicate their cognitive levels for the whole exam. After the exam, the Attitude towards Computer Scale was administered and the participants were informed that interviews would be conducted with those who volunteered. The contact information of the volunteer students was obtained and interviews were held with the students on the specified days and hours. The interviews were recorded after the participants gave their consent. Participating students were informed about the study and the process at the beginning of the exam, before they completed the questionnaires and scales, and at the beginning of the interviews.

2.5.1. TAO Testing Platform

TAO is an open-source online testing environment that combines assessment and learning data and turns them into a new action. The platform includes about 20 types of questions (multiple choice, matching, multimedia, sequencing, open-ended, file uploading, short-answered questions, hotspot questions, gap-filling, gap-filling from the drop-down menu, long-answered questions). In addition, question banks can be created, online exams can be published and their results can be obtained in reports using this platform.

2.6. Data Analysis

2.6.1. Analysis of the Qualitative Data

An interview form developed by the researcher was used to collect qualitative data. The participants were informed that data would be collected through interviews at the end of the testing. The contact information of the volunteers was obtained from the participants. Afterwards, the participants were contacted and appointments were arranged for the interviews. Interviews were conducted in a quiet and appropriate environment where the participants would feel comfortable during a time when they were available.

The interviews were recorded. Then, these records were transcribed in computer environment and analyzed. The transcribed text was coded by the researcher using the inductive method. Çetin (2016) states that the coding scheme in content analysis can be created in 3 ways: inductive, deductive and inductive-deductive methods. In the inductive method, codes are derived from the transcribed texts and categories and subcategories can be created from these codes (Çetin, 2016).

After the analysis of the data in this study, three main categories emerged: the attitude towards physical environment, the availability of the exam platform, and the student. The exam environment included two sub-categories: opportunities of the CBT environment and the limitations of the PPT environment. There were five sub-categories under the availability of the CBT platform: efficiency, effectiveness, satisfaction, the use of multimedia, and the novelty of platform. Two sub-categories emerged under the category of attitude towards the examination process: anxiety and stress, and technology proficiency.

2.6.2. Analysis of the Quantitative Data

Quantitative data were collected from cognitive load scale, demographic information questionnaire and attitude scale towards computer. The results of the cognitive load scale were obtained from the TAO exam platform. The statement of overall rating scale indicates the score obtained from the scale for the entire testing process.

3. Findings

3.1. Attitudes of Participants towards CBT Environment and Factors Affecting These Attitudes

The categories and subcategories that emerged as a result of the interviews with the participants within the scope of the study are shown in Table 1.

Table 1.

Categories and Subcategories

Attitude towards Physical Environment	CBT Center Environment
	PPB Testing Environment
Attitude towards CBT Platform	Platform Efficiency
	Platform Effectiveness
	Platform Satisfaction
	Use of Multimedia on the platform
	Novelty of the Platform
Attitude Towards the Examination Process	Anxiety and Stress
	Technology Proficiency

3.1.1. Attitude towards Physical Environment

Factors that affected students' attitudes towards the physical environment were examined under two headings: opportunities of the CBT center and limitations of the PPT environment.

3.1.1.1. CBT Center Environment

P2 said that the listening activities were better than the classical exam and added: "…The only difference here was that we were isolated from the outside sound by means of headphones and we listened more comfortably…". Considering the problems related to hearing in the PPT, P2 also said that having headphones was an advantage in terms of equipment.

P4 said that using headphones was better in terms of listening to the voices, voice quality and focusing on the voice compared to the classroom environment.

The factor that affected the attitudes of the participants was that the CBT environment was more suitable for exams including listening activities. Participants did their listening activities individually using headphones. In addition, they had the opportunity to pause, resume and rewind listening tracks. This also affected the attitudes of the participants. In fact, the importance of the integration of technology into education emerged.

3.1.1.2. Limitations of the PPT Environment

Another factor that affected students' attitudes towards the physical environment was the limitations of the PPT environment. The participants talked about some of the problems they experienced while doing the exam in PPT environment.

Some participants regarded the lack of equipment as a limitation. They said that the classrooms did not have sufficient technical equipment for listening exams. Regarding the lack of equipment, P1 said, "One thing, you know, while we listen to the listening material from the cassette player, too many things happen. There is a humming sound" and "If we sat in the back, we would hear no sound, we always have problems."

P1 noted that the use of cassette players in listening exams was actually insufficient for the transmission of sound.

P2 said, "No sound comes to the back seats in the classroom and the sound is too high for the students in the front desks. It is crowded and the sound cannot be heard clearly." P4 said that having headphones was an advantage compared to the classroom environment in terms of hearing the listening material. P5 said that they experienced problems hearing the sounds in the listening activities conducted in the classroom due to the cassette players. "When we do this in the classroom, we do it with an old cassette player. Students in the front rows can hear it, but those in the back row cannot. There is buzzing, humming and so on..." . P6 noted that the classroom environment was not suitable for listening activities using cassette players. "If we don't sit in front rows, we can't understand anything because the sound cannot be heard clearly in the back. We ask the teacher to turn up the volume, then those in the front become disturbed."

3.1.2. Attitude towards CBT Platform

Five main factors that affected students' attitudes towards the BDS platform emerged: Efficiency of the platform, effectiveness of the platform, platform satisfaction, use of multimedia on the platform, and the novelty of the platform. Details on these factors alongside the participants' comments are reported below.

Efficiency, effectiveness and satisfaction that influence attitude towards the CBT platform are based on the ISO 9241-210:2010 usability standard in the literature. According to this standard, effectiveness is users' correct and complete achievement of certain goals. Efficiency is users' correct and complete achievement of certain goals and the efficacy of the relationship between the resources used to achieve these goals. Satisfaction is users' comfort and their positive attitude towards the use of the system (ISO 9241-210, 2010). Details on these will be addressed below.

3.1.2.1. Platform Efficiency

Efficiency refers to users' correct and complete achievement of certain goals and the efficacy of the relationship between the resources used to achieve these goals. Here, the participants talked about the advantages of the CBT environment and that more activities could be done in less time.

In this regard, P2 noted that it reduced the mistakes, saying "…I mean, it minimizes the mistakes that we can make…". Mentioning that CBT facilitated the process, P2 added, "…But it could have been much easier if we had taken a writing test on the computer…."

Noting that exams taken in the CBT platform were easier than those done in other platforms, P3 said "I could understand more easily, I could see the questions more easily." P3 also drew attention to individuality and convenience aspects of the CBT saying, "I understood better because I saw and listened to the questions here individually; I think I could do better."

3.1.2.2. Platform Effectiveness

P2 said "… I listened to the text and read it at the same time. It helped me to solve the questions at once" to indicate that CBT enabled them to read and listen to the text at the same time during the exam and this helped them understand the questions at once. P2 added, "…Both watching the video and reading the paragraph while listening to the audio recording shortened the time. I mean, instead of reading it five times, I read it twice and listened…" and pointed out that the multimedia and listening items used in reading test shortened the time to complete the exam quite significantly.

3.1.2.3. Platform Satisfaction

One participant, P4, noted that the CBT environment is better for improving listening skills in PPT saying "I think it is more effective to use these methods to improve this skill" and indicated that the CBT environment is especially effective in improving listening skill.

3.1.2.4. Use of Multimedia on the Platform

P2 said, "Presentation of text and audio together at the same time or emphasizing the emojis in the video made them more memorable." and drew attention to visuality and memorability along with visuality.

Regarding the single channel code, P1 said, "We couldn't see the text when we were listening to the items in the listening section in the PPT, we only had to hear it." Here, the participant stated that they made more effort in the PPT environment. Video questions cause less effort because they address the visual channel as well as the auditory channel. This also refers to the modality principle.

3.1.2.5. Novelty of the Platform

The fact that the platform is new for students also positively affected the attitudes of the participants. For example, P4 said, "For example, people always have paper psychology. This is something different. You know, we partly overcame it. We were nervous before taking the exam, but this anxiety went away through the end of the exam", pointing out the positive effect of novelty of the platform.

3.1.3. Attitude Towards the Examination Process

Two main factors that affected the attitude towards the exam process emerged: anxiety and stress and technology proficiency. Anxiety and stress and technology proficiency sub-categories emerge as factors affecting the attitudes of the participants towards the CBT process.

3.1.3.1. Anxiety and Stress

Regarding the attitude towards the CBT, the participants mentioned the anxiety and stress situations they experienced during the CBT process.

P1 stated that he/she was nervous at the beginning of the exam because it was his/her first experience. Expressing the fear of the CBT, P1 said "I've never had such an experience before, so I was a little afraid." P4 said, "At first, it made us nervous because we've never had this experience before. But then we could see that it was effective." P4 was worried because of being inexperienced; however, they noticed the effectiveness of CBT in the first experience.

Participants also stated that they felt anxiety in terms of grades in the CBT environment. P1 expressed their concern about the grades, saying "I was thinking that I would fail to do it because it would affect our grades."

Time anxiety refers to the anxiety associated with time experienced by the participant during the CBT. P4 noted that he/she was worried because of the duration since the exam platform was new, saying "I felt anxious about controlling the time, I had difficulty because it was new for me."

3.1.3.2. Technology Proficiency

Technology proficiency has emerged as an important factor affecting participants' attitudes towards the testing. In this regard, the participants stated that they were not sufficient in terms of using computers and CBT experience.

Answering the question about studying with the CBT system, P5 said “It is easy for me because I am quite used to it” and expressed their familiarity with computers. P7 expressed that they did not have feelings such as fear, uneasiness, anxiety etc. and said, “It’s because, perhaps, this comfort is caused by the fact that we are always in front of computers.”

P7 who added “My experience may be due to the fact that I spend most of my day in front of the computer." indicated that the reason of not having a difficulty in TAO exam platform stemmed from being experienced.

Of the participants, four had CBT experience and it was observed that these participants did not have any difficulties in using computers during the exam process.

3.2. Participants Question based Cognitive Load Scores

Statistical information about the participants' answers to the question-based cognitive load scale is shown in Table 2. The table shows the minimum and maximum values and averages of the scale responses for the question-based scale and general evaluation.

Table 2.
Descriptive Statistics Table for Participants Question based Cognitive Load Scores

	N	Minimum	Maximum	Total	Average	Variance
Cognitive Load-Q1	32	2,00	7,00	164,00	5,1250	1,339
Cognitive Load-Q2	32	1,00	7,00	120,00	4,7500	2,710
Cognitive Load-Q3	32	1,00	6,00	92,00	2,8750	2,242
Cognitive Load-Q4	32	1,00	9,00	128,00	4,0000	3,355
Cognitive Load-Q5	32	1,00	9,00	138,00	4,3125	4,222
Cognitive Load-Q6	32	1,00	8,00	124,00	3,8750	3,403
Cognitive Load-Q7	32	1,00	8,00	122,00	3,8125	2,415
Cognitive Load-Overall	32	2,00	7,00	151,00	4,7188	1,757
Evaluation						
Valid N (listwise)	32					

According to this table, except for Question 1, all other questions created low cognitive load for the participants. Question 1 created a high cognitive load. Kılıç and Karadeniz (2004) state that the highest score from the rating scale is 9, the lowest score is 1 and the cut-off score of the scale is 5. A score above 5 indicates a high cognitive load and a score below 5 indicates a low cognitive load.

Table 3 shows the results of the Mann Whitney U test comparisons of the participants from the general cognitive load scale to total exam time, total exam score and attitude towards computer scale. Values below 5 were classified as 1 to represent low cognitive load, and those above 5 were classified as 2 to represent high cognitive load.

Tablo 3.
Mann Whitney U test comparisons of the participants from the overall cognitive load scale to total exam time, total exam score and attitude towards computer scale

	Overall Cognitive Load Level	**N**	**Rank Means**	**Rank Totals**	**U**	**p**
Total Exam Time	1,00	24	14,25	342,00	42	,018
	2,00	8	23,25	186,00		
Total Exam Score	1,00	24	18,17	436,00	56	,086
	2,00	8	11,50	92,00		
Attitude Towards Computer Scale	1,00	24	16,33	392,00	92	,881
	2,00	8	17,00	136,00		

When Table 3 is analyzed; it is seen that there is a significant difference between the participants' time to complete the exam according to the total rating scale score (U=42, p<0,05). According to the total rating scale score, there is no significant difference between the rank averages of the participants' exam score (U=56, p>0.05). According to the total rating scale score, it is seen that there is no significant difference between the mean ranks of the participants' attitude towards computer scores (U=92, p>0,05).

3.3. Results on the Correlation of the Exam Time, Exam Scores and Cognitive Load Scores of the Participants

The findings on whether there was a relationship between the exam time, exam score and cognitive load scores of the participants, which is the second research question of the study, are shown in Table 4.

Table 4.

Correlation between Exam Time, Exam Score and Cognitive Load Scores

		Total Scale Score	Total Exam Time	Total Exam Score
Total Scale Score	Pearson Correlation	1	.308	.200
	Sig. (2-tailed)		.086	.273
	N	32	32	32
Total Exam Time	Pearson Correlation	.308	1	-.432*
	Sig. (2-tailed)	.086		.014
	N	32	32	32
Total Exam Score	Pearson Correlation	-.200*	-.432*	1
	Sig. (2-tailed)	.273	.014	
	N	32	32	32

*Correlation is significant at the 0.05 level (2-tailed)

The relationship between the participants' time to complete the exam, their exam scores, and responses to the rating scale was examined using the Pearson product-moment correlation coefficient. According to the analyses, the correlation between the exam completion durations of the participants and their exam scores was low and negative ($r = -.432$, $n = 32$ and $p < .05$).

The findings on the relationship between participants' rating scale score of in the listening questions, the completion time of the exam and the variable of exam score obtained from the listening questions are shown in Table 5.

Table 5.

Correlation between the Listening Items, Scale, Time and Scores

		Subjective Rating Scale for Listening Items	Completion Time for Listening Items	Scores for Listening Items
Subjective Rating Scale for Listening Item	Pearson Correlation	1	.387*	-.330*
	Sig. (2-tailed)		.029	.065
	N	32	32	32
Completion Time for the Listening Item	Pearson Correlation	.387*	1	-.357*
	Sig. (2-tailed)	.029		.045
	N	32	32	32
Scores for Listening questions	Pearson Correlation	-.330*	-.357*	1
	Sig. (2-tailed)	.065	.045	
	N	32	32	32

*. Correlation is significant at the 0.05 level (2-tailed).

The relationship between the participants' time to complete the exam, their exam scores, and responses to the rating scale was examined using the Pearson product-moment correlation coefficient. According to the analysis, there was a low and positive relationship between the participants' completion time of listening questions and the rating scale score they received from the listening questions ($r = 387$, $n = 32$ and $p <, 05$); on the other hand, there was a low and negative ($r = -.387$, $n = 32$ and $r < .05$) relationship between the completion time of the listening questions and the test scores.

4. Discussion and Suggestions

In this study, an English knowledge test was redesigned according to the design principles of the cognitive load theory and administered to the students studying at the School of Foreign Languages in the higher education institution where the study was conducted. In this context, the study aimed to reveal whether the design principles reduced the cognitive load of the students and to determine the correlation between the cognitive load of the participants, their exam completion time and exam scores. The study determined the attitudes of the participants towards the CBT and the reasons why the CBT is preferred. Using an attitude scale towards computers, the study also investigated whether participants' cognitive load scores, exam completion times, scores on the variables in the questionnaire, and attitudes towards computers varied according to some variables. Accordingly, the following results were obtained.

The expressions of the participants about the advantage of equipment used show that the tools used in the CBT environment provide comfort compared to the classroom environment. In addition, the advantage of using equipment provides flexibility and individuality to the participants during the exam. Especially the presence of individual headphones has made the CBT environment more preferable than the classroom environment. This is due to the fact that headphones enables the participants to hear more easily and that they do not have to lean over the exam papers on the desk for a long time as in the way they do in the classroom.

Moreover, students have had to hear and listen to the recording coming from a single sound source in a single area in a large classroom in the PPT. Sometimes they could not hear or understand clearly because of external factors such as sounds coming from desks, human voices or noises coming from outside of the building. Every participant has their own headphones in a CBT environment, so they can focus and hear better. In addition, earphones enable noises caused by external factors to be isolated.

Regarding the time spent in testing, the participants thought that the fact that the exam was conducted in the CBT environment and that multimedia was included in it actually shortened the exam completion time. Although the difficulty level of the questions asked in the exam conducted within the scope of this research was almost the same, the students completed listening questions in a shorter time than the questions that included listening items. The relevant literature reports that shortening of the total exam time is an indicator of cognitive load; for example, Brünken, Seufert, and Paas (2010) stated that time is an indirect indicator of cognitive load in activities in multimedia environments. Therefore, it can be argued that the participants spend less time in the CBT environment than in the classroom environment due to the fact that the CBT environment decreases the cognitive load in multimedia supported questions.

Participants also stated that they preferred the CBT because of the use of multimedia in CBT environment. Some of the participants stated that they have benefited from their visual memory. Participants benefited from their visual memory and understood the questions easily, which shows that this ability had a positive effect on their success.

The CBT environment also provides flexibility to the student since the accuracy rates of the questions can be increased by saving time in the listening sections in the CBT. Students can listen to the listening items on the cassette player twice at most in the classroom environment. In addition, they do not have a chance to forward, reverse or pause and continue the recording by themselves. Therefore, if students miss a few sentences during listening, this may cause them to answer a few questions incorrectly. This flexibility in the CBT gives students the possibility to answer the questions more accurately because they can listen to the part they missed again. According to Hamad (2023), it was reported that the use of computer based tools (like web 2.0 tools) to facilitate online/blended teaching and learning would be beneficial. That means computer based tools to support learning environment provides flexible tool with multimedia support. In this case, it was a computer based testing platform that creates flexible environment for students.

In the adaptation study of the scale, the cut-off point was determined as 5. Scores above 5 indicates high cognitive load, those below 5 indicate low cognitive load (Kılıç & Karadeniz, 2004). When the opinions of the participants obtained through the interviews are examined alongside the results, it is seen that the multimedia integrated questions in the exam created a low cognitive load.

Difference between the levels of cognitive load of the participants related to different types of exam questions. According to the results of the study, adding multimedia elements created less cognitive load in the exam. When all the materials were evaluated in the context of cognitive load, it was found that the material that created the most cognitive load according to the vast majority of the participants was the text only exam questions, while less cognitive load was created by the multimedia integrated exam questions. There is a negative correlation between cognitive load and exam scores. It is possible to recommend that using multimedia integrated exam questions would be effective in terms of achievement in the computer based testing platform. Learning with multimedia elements is more effective for achievement than learning with textual material only, according to another study on the effect of learning with multimedia on achievement and metacognitive skills (Lindner, Eitel, Barenthien & Köller, 2018). Similarly, according to Çaka & Dursun (2022), visual intensive infographics significantly increases the achievement scores of the students. Both in the exam or course material using visualization would be effective for students' achievement.

According to the other correlation result, there is a low and negative relationship between the exam duration and the cognitive load score supports this finding. As it was stated above, text based exam questions had a high level of cognitive load. In other words, as the density of visual element increased, the cognitive load level and the time spent on the exam questions decreased. There are studies in the literature reporting similar in the literature. Çaka & Dursun (2022) reported that the visual intensive infographics created less cognitive load.

To summarize, in reference to the findings obtained from the data of the current research, which was carried out to determine the text-based exam question in computer based testing environment, which participants found distracting and took long to answer the related questions. Moreover, CBT provided a more effective learning experience and supported the meaningful learning process with their multimedia features. Nonetheless, in another study Şenel (2019) reported that most of the Information Technology teachers mainly preferred to use paper and pencil tests and multiple-choice items. Even that teachers gave programming course, they prefer traditional assessment tools. Although, teachers felt themselves inadequate about alternative measurement methods (Şenel, Pekdağ ve Günaydın, 2018), our participants had positive attitude towards physical environment of the testing center and CBT platform. Moreover, they were also able to cope with the anxiety and stress of the exam, despite their limited knowledge of the technology.

In future research, this exam platform should be tested with respect to different points of view. In fact, TAO testing platform support many different question types with many other multimedia support. Different question types can be tested in different courses to clarify the multimedia effect on the cognitive load in an CBT platform.

Acknowledgements

This research is based on a M.A. thesis submitted to the Institute of Educational Sciences, Mersin University (Odacı, 2019) which was supported by Mersin University, Scientific Research Project Office, under the project number 2019-2-TP2-3524.

References

Akdemir, Ö. & Oğuz, A. (2008). Computer-based testing: an alternative for the assessment of Turkish undergraduate students, *Computers and Education, 1*, 1198-1204. https://doi.org/10.1016/j.compedu.2007.11.007

Aybek, E. C. (2012). *Kâğıt-kalem formu ve bilgisayar ortamında uygulanan genel yetenek testinin psikometrik özelliklerinin karşılaştırılması* [Master's thesis, Ankara University]. YÖK Açıkbilim. https://acikbilim.yok.gov.tr/handle/20.500.12812/85211

Brünken, R., Seufert, T., & Paas, F. (2010). Measuring Cognitive Load. In J. L. Plass, R. Moreno, & R. Brünken (Eds.), *Cognitive Load Theory* (pp. 181–202). Cambridge University Press.

Bugbee, A.C. (1996). The Equivalence of Paper-and-Pencil and Computer-Based Testing. *Journal of Research on Computing in Education, 28*(3), 282. Retrieved September 19, 2023 from https://www.learntechlib.org/p/80961/.

Büyüköztürk, Ş., Kılıç Çakmak, E., Akgün, Ö. E., Karadeniz Oran, Ş., & Demirel, F., (2019). *Bilimsel Araştırma Yöntemleri*. Pegem Akademi.

Çaka, C. & Dursun, Ö.Ö. (2022). Evaluation of the effectiveness of different infographic designs. *Journal of Educational Technology & Online Learning*, 5(3), 519-534. https://doi.org/10.31681/jetol.1075352

Çetin, İ. (2016). Nitel içerik analizi. In Özden, M. Y. & Durdu, L. (Eds.), *Eğitimde Üretim Tabanlı Çalışmalar İçin Nitel Araştırma Yöntemleri* (pp. 125-148). Anı Yayıncılık.

Dinar, M., Kabakçı Yurdakul, I. & Dönmez, F. (2015). Measuring Cognitive Load in Test Items: Static Graphics versus Animated Graphics. *Journal of Computer Assisted Learning*, 31, 148-161. https://doi.org/10.1111/jcal.12086

Fraenkel, J. R., Wallen, N. E., ve Hyun, H. H. (2012). *How to design and evaluate research in education* (8th ed.). McGraw-Hill.

Hamad, W. (2023). Students' experiences of learning in virtual classrooms facilitated by Google Classroom. *Journal of Educational Technology and Online Learning*, 6 (2), 362-383. http://doi.org/10.31681/jetol.1250095

Kirschner, P.A. (2003). Cognitive Load Theory: Implications of Cognitive Load Theory on the Design of Learning, *Learning and Instruction*, 12, 1-10. https://doi.org/10.1016/S0959-4752(01)00014-7

Kılıç, E. (2009). *The effects of cognitive load in learning from goal based scenario designed multimedia learning environment for learners having different working memory capacities* (Publication No. 255259) [Doctoral Dissertation, METU]. YÖK Tez, https://tez.yok.gov.tr/UlusalTezMerkezi/TezGoster?key=NtBAevXNhYaNqJFoAcdBdoczOCd_zbRfD8kVAKPsDubMuN-FCZVksYHFE-wH0v9e

Kılıç, E. ve Karadeniz, Ş. (2004). Hiper ortamlarda öğrencilerin bilişsel yüklenme ve kaybolma düzeylerinin belirlenmesi. *Kuram ve Uygulamada Eğitim Yönetimi*, 40, 56—579. Retrieved from https://dergipark.org.tr/tr/pub/kuey/issue/10357/126808

Lemmo, A., Mariotti, M-A. (2017, 1-5 February). From paper and pencil to computer-based assessment: some issues raised in the comparison, In. Dooley, T. & Gueudet, G. (Eds.), *Proceedings of the Tenth Congress of the European Society for Research in Mathematics Education*, DCU Institute of Education & ERME, 3540 – 3547.

Lindner, M. A., Eitel, A., Barenthien, J., & Köller, O. (2018). An integrative study on learning and testing with multimedia: Effects on students' performance and metacognition. *Learning and Instruction*, 1(2), 1-9. https://doi.org/10.1016/j.learninstruc.2018.01.002

Mayer, R.E. (2003). The Promise of Multimedia Learning: Using The Same İnstructional Design Methods Across Different Media, *Learning and Instruction*, 13, 125-139. https://doi.org/10.1016/S0959-4752(02)00016-6

Özturan, T. (2016). T*he impact of computer assisted assessment on the exam success and attitudes of prospective English teachers.* (Publication No. 435226) [Master's thesis, Hacettepe University]. YÖK Tez, https://tez.yok.gov.tr/UlusalTezMerkezi/TezGoster?key=cbOXH84ZayrLjc0tI-QXKtTWLLFqln0MiUoLfO-MxUyK83GNRlTkrzR5Bs7ix3I1

Paas, F. G., & Van Merriënboer, J. J. (1994). Instructional control of cognitive load in the training of complex cognitive tasks. *Educational Psychology Review*, *6*, 351-371. https://doi.org/10.1007/BF02213420

Paas, F., Renkl, A. & Sweller, J. (2003). Cognitive load theory and instructional design: recent developments, *Educational Psychologist, 38*(1), 1-4. https://doi.org/10.1207/S15326985EP3801_1

Prisacari, A.A. & Danielson, J. (2017). Computer-based versus paper-based testing: Investigating testing mode with cognitive load and scratch paper use, *Computers in Human Behavior*, 77, 1-10. https://doi.org/10.1016/j.chb.2017.07.044

Şenel, S. (2018). Programlama Konusu "Kazanımları" ile "Sınıf İçi Ölçme Süreçlerinin" Bilişsel Düzeylerinin Karşılaştırılması. *International Journal of Computers in Education*, 1 (2), 1-20. Retrieved from https://dergipark.org.tr/en/pub/ijce/issue/42323/486158

Şenel, S., Pekdağ B., & Günaydın, S. (2018). Kimya öğretmenlerinin ölçme ve değerlendirme sürecinde yaşadıkları problemler ve yetersizlikler. *Necatibey Eğitim Fakültesi Elektronik Fen ve Matematik Eğitimi Dergisi (EFMED).* 12(1). 419-441. https://doi.org/10.17522/balikesirnef.437824

Sweller, J., van Merrienboer, J.J.G. & Paas FGWC. (1998). Cognitive Architecture and Instructional Design, *Educational Psychology Review*, 10(3), 251-296. https://doi.org/10.1023/A:1022193728205

Takır, A. (2011). *The effect of an instruction designed by cognitive load theory principles on 7th grade students' achievement in algebra topics and cognitive load*. (Publication No. 255259) [Doctoral Dissertation, METU]. YÖK Tez, https://tez.yok.gov.tr/UlusalTezMerkezi/TezGoster?key=EEdeQgIdFRxX5NbvVau-AhjV-QzHr495E11DeqZt2L6gI88s6pwJxJAMh4_IhC9C

Woodside, A. G. (2010). *Case Study Research: Theory, Methods, Practice*. Emerald Group Pub. Ltd., ebrary Inc.

Yıldırım, A. & Şimşek, H. (2018). *Sosyal Bilimlerde Nitel Araştırma Yöntemleri.* Seçkin Yayıncılık.

Yıldırım, Z. (2016). Öğretim teknolojileri ve ileti tasarımı. In Çağıltay, K. & Göktaş, Y. (Eds), *Öğretim Teknolojilerinin Temelleri: Teoriler, Araştırmalar, Eğilimler* (279-296). Pegem Akademi.