

Anxiety and Test Form: The Differences of Test Anxiety Levels in Terms of Test Form

Lussy Dwiutami Wahyuni, Rezza Citraini, Bijak Aditia Hutomo,
Gungum Gumelar Fajar Rakhman

Department of Psychology, State University of Jakarta, Indonesia

lussysf@unj.ac.id

Abstract

Education is one of the most important aspects of human life. In the education file, examination is an important part to evaluate the progress of the students. The students need to fulfill the test parameters and indicators in order to move to the higher level of class or higher degree. The common feeling while students feel anxious in facing the test is called test anxiety. There are several factors that affect test anxiety, and one of the most common sources is the test form. This study aimed to determine the differences in test anxiety levels in students in terms of multiple-choice and matchmaking tests. The population of this study was 142 students of the Faculty of Psychology Education that registered in 2018. The design of this study was a quasi-experimental research with two-group pretest-posttest design. The instrument used in this study was the Reaction to Test scale by Irwin G. Sarason. The results showed that the multiple-choice test anxiety level was higher than the matching test, although the difference was small. The multiple-choice test anxiety level was ($M = 2.17$) and matchmaking ($M = 2.07$).

Keywords: Test anxiety, test form, multiple choice test, matchmaking test.

Abstrak

Pendidikan adalah salah satu aspek terpenting dalam kehidupan manusia. Dalam ranah pendidikan, ujian merupakan bagian penting untuk mengevaluasi kemajuan peserta didik. Para peserta didik perlu memenuhi parameter dan indikator tes untuk berpindah ke kelas yang lebih tinggi atau tingkat yang lebih tinggi. Ada perasaan umum dimana peserta didik merasa cemas dalam menghadapi tes yang disebut Kecemasan Tes (Test Anxiety). Ada beberapa faktor yang mempengaruhi Kecemasan Tes, dan salah satu sumber yang paling umum adalah bentuk dari suatu tes. Penelitian ini bertujuan untuk menentukan perbedaan tingkat Kecemasan Tes pada siswa dalam bentuk tes pilihan ganda dan tes menjodohkan. Populasi penelitian ini adalah 142 mahasiswa Fakultas Pendidikan Psikologi yang terdaftar pada tahun 2018. Desain penelitian ini adalah penelitian kuasi eksperimental dengan desain pretest-posttest dua kelompok. Instrumen yang digunakan dalam penelitian ini adalah Reaction to Test Scale oleh Irwin G. Sarason. Hasil penelitian menunjukkan bahwa tingkat kecemasan tes pilihan ganda lebih tinggi daripada tes menjodohkan meskipun perbedaannya kecil. Tingkat kecemasan tes pilihan ganda adalah ($M = 2,17$) dan perjodohan ($M = 2,07$).

Kata Kunci: Kecemasan tes, bentuk tes, tes pilihan ganda, tes menjodohkan.

Introduction

Test is an activity that is held for individual placement, consideration for policy-making, and evaluation process in teaching and learning activities (Febrayosi, 2012). The application of the test is often connected with some anxious feelings about the test that is called test anxiety. Test anxiety is a primary human emotion that indicates the uncertainty or threat in an environment influenced by stress, cognitive impairment, and interacting evaluative situations (Sarason, 1977). Meanwhile, Shechter and Zeidner (1990) had conceptualized test anxiety as a result of a stimulus of a possible hazardous future outcome and a response to a stressful condition. The criteria of test anxiety are found in various ways. Sarason (1990) explained the requirements of an individual who assesses a threatening and challenging situation, considers himself inadequate in completing tasks, concentrates on undesirable consequences, self-deprecating, and from the start, the individual has anticipated the failure that will be experienced. Test anxiety also consists of two main components, worry and emotionality. Worry refers to the tendency for a test taker's negative thoughts about their performance while taking the test. Negative thinking may include thinking about the consequences of the test results, comparing himself to others, and criticizing himself for not being able to take the test. Meanwhile, emotionality refers to the autonomic nervous reaction when stressed on a test. These reactions can include palpitations, dizziness, and sweating (Deffenbacher, 1986; Hembree, 1988; Morris 1981 in Zeidner 1998).

Test Anxiety

Generally, anxiety is an emotional blurry and unpleasant state. It is shown by the characteristics of fear of something, feeling daunted, and uncomfortable (Reber & Reber, 2002). A similar opinion is expressed by Nelson-Jones (cited in Sobur, 2016) that anxiety is a general feeling of fear and apprehension that is associated with a particular person or situation, one of them is test anxiety. The facts from Las Cruces Sun-News (2017) that reveal the data sourced from the American Test Anxieties Association, 16 out of 20 American students have test anxiety with 18% of students trapped with severe anxiety test problems. As well as revealing data on test anxiety, the article from Las Cruces Sun-News also interviewed Brandon Dixon, a physics student at New Mexico State University who is experiencing test anxiety. Brandon Dixon had difficulty sleeping at night before the test, even though he had studied the previous two weeks. According to Brandon Dixon, this often happens when he is about to face the tests at his campus.

In the implementation of tests in Indonesia, some sources underline the implementation of the national exam. Tempo.co on April 16, 2016, revealed the data on the selling of answer keys for national exams at the high school level in East Java, particularly in Mojokerto and Lamongan. There is a report on the sale and purchase of National Exam (UN) answer keys that reach a price of IDR 14,000,000, and an average student is charged IDR 50,000 as a joint venture. Whereas in the Capital Region of Jakarta, this practice pegged at a price between IDR 14,000,000 to IDR 21,000,000. The students also pay a fee of IDR 50,000 to IDR 100,000 per person as a joint venture. This case shows that students are indirectly looking for ways to get grades that meet standards and their desire regardless of the consequences of not passing the exam. Even though the students have studied for months, even some of them have undergone a national try out, but still rely on the answer keys to take exams. This phenomenon indicates the symptom of test anxiety, they do not believe in their abilities and worry about the consequences that they will have if they fail. The effects that will be accepted include social embarrassment and having to repeat the test.

Miller (2018), Assistant Professor at the University of British Columbia in his article "Test Anxiety" published at anxietycanada.com on August 19, 2018, presented the results of his research regarding the implementation of achievement tests. According to his study, almost all students experience anxiety when facing the test. This feeling exists because many students worry about the consequences if they fail.

In this study, the symptoms of test anxiety also presented, (1) Physical: headaches, dizziness, diarrhea, extreme changes in body temperature, excessive sweating, shortness of breath, fainting, rapid heartbeat, dry mouth. (2) Emotional: feelings of fear, disappointment, anger, depression, crying and feelings of helplessness. (3) Behavior: restlessness, pacing, substance abuse, and avoidance of evaluative situations. (4) Cognitive: difficulty concentrating, blank thoughts, negative self-talk, feelings of fear, and comparing oneself with others.

Based on Irwin G. Sarason's research in 1976 entitled "The Test Anxiety Scale: Concept and Research", the objective of this study was to measure the level of test anxiety using the TAS (Test Anxiety Scale). The number of research subjects is 520 students at the University of Washington, consisting of 283 males and 237 females. The results of this study indicated the mean average anxiety level for men was 16.72, while women it was 19.74. The results of this study showed that women's anxiety was higher than men's, even though they were still in the range of moderate anxiety levels. This research is also supported by meta-analytic research on 562 studies on test anxiety. The research conducted by Hembree in 1988 found that (1) female students experienced higher anxiety than male students, (2) test anxiety directly related to feelings of dislike of tests, ineffective learning, and fear of taking the test.

According to Revilla (2009), some main factors triggered the test anxiety in its implementation among others: (1) Time constraints, limited time can cause stressful situations for students because this requires the students to answer under time pressure; (2) The level of material difficulty, this is a test anxiety factor that affects cognition and causes an orientation of future failures and can lead to a tendency to think of oneself as incapable of doing the task; and (3) Test instructions and test question forms. Confusing test instructions and test forms such as subjective or objective tests can have an impact on test anxiety in its implementation. According to Anastasi and Urbina (1997), an achievement test is a test that measures the effect of an instructional or training program. One of the causes of test anxiety is the repeated failure when the individual is given various tests (Slameto, cited in Elfiana 2005). Some individuals tend to think of the tests as a measure of success because they see the fact that tests and grades are used as the basis for obtaining various social rewards, such as environmental acceptance, promotion, good work, more money, and so on. Based on the previous research, test anxiety is one of the parts in human beings. It shows that anxiety comes when the people are facing something that may affect their future as one of the examples. It is not a surprise that the test can be a force to motivate the person to achieve success. However, if someone has repeatedly failed during the test, then the new failure will reduce the motivation of the person. Therefore, this study will focus on test anxiety and triggered by the test form.

Test Form

According to Arikunto (1992), there were two tests, namely subjective and objective tests. A subjective test or description test is a test with a long discussion to get the answer. The subjective questions or descriptions require students to remember in detail, recall, and possess high creativity in answering the questions that need analysis. This requirement indicates that the use of subjective tests mobilizes the ability of test-takers to describe the things clearly. According to Reynolds, Livingston, and Willson (2009), the essay test is divided into two, namely the restricted-response essay and the extended-response essay. Limited response descriptions are structured descriptions that require students to list, determine, describe or provide reasons with limited answers. At the same time, the extender-response description is a description with a more flexible answer, and there are no limitations in responding to the questions. This test also requires test takers to write, summarize, formulate, compare and interpret. Based on the characteristics of subjective tests, it can conclude that subjective tests require good language skills and high analytical thinking.

The objective test is a test conducted by choosing one of the available alternatives or by filling in the correct answer with several words or symbols. An objective test requires the test takers to select one of

the most appropriate answers from the choices. The objective test questions also tend to require the ability to remember and recognize with many answer choices as a distraction. The test-takers need accuracy in choosing the right answer. There are four classifications of the objective tests: (1) true-false test; (2) matchmaking test; (3) filling or completion; and (4) multiple choice.

In this study, the form of the test was multiple choice and matchmaking. It is considered because both have similarities in characteristics, in which both are objective tests. Another reason for choosing the type of objective test as the focus of this problem is because the research between subjective and objective tests had been conducted. Research on test anxiety in terms of subjective and objective test types, one of which was conducted by Stephanie Elfiana from Atmajaya Catholic University, Jakarta (2005) with the research title "Differences in Test Anxiety between Multiple Choice Test and Essays". The results showed that the test anxiety of an essay test was higher than the multiple-choice test. The results of this study can be explained in Zeidner's research in 1987 "Essay Versus Multiple-choice Type Classroom Exams: The Students Perspective". This study aimed to see the students' perspective of the form of essay test and multiple choice. In this study, it found that 80% of respondents chose multiple-choice formats as a form of test that was easier to do than descriptive form and 83% of students felt more anxious about having to make an essay form rather than multiple choice. This finding illustrates the comparison of the level of test anxiety through multiple-choice forms, and the actual description is not balanced to be used as a comparison because the structure and way of answering them are too different. Another supporting factor for the choice of multiple-choice test forms and matchmaking as the basis of this study is that research was still rare in Indonesia compared with the level of Test Anxiety on these two test forms.

The previous research conducted by Oktaviani and Sastrawijaya (2017) from the State University of Jakarta compared the form of multiple-choice true or false forms. The results of this study proved that the group given the multiple-choice test had a higher level of anxiety than the true-false form because the multiple-choice format requires test takers to fill in gaps in sentences and has more distractors than true and false forms. Meanwhile, in the true-false format, the test takers face questions in the form of whole sentences without having to match an overlapping sentence with the answer.

The finding provides an overview of the structure of the test and the level of difficulty triggered the test anxiety for the participant. The multiple-choice and true-false tests are classified in the objective test form. However, multiple-choice aims to see how someone mastered a material through gaps while true-false forms measure the right and wrong choices. A large number of comparisons of test anxiety levels that focus on comparing the form of the test as the basis of research so far have rarely reached the form of the filled-in test with other forms of the test. Arikunto (1992) defines the form as a form of test that includes direct statements or free answers from imperfect sentences. Based on this, even though the form of the test is an objective test, it has a similar tendency to the subjective test. From the description of the comparison of the test forms, it can conclude that the structure of the test looks simple, but it has an impact on test anxiety. Therefore, the tests used in this study will arrange the rules for achievement tests and the level of difficulty that has been generalized. Through this explanation, a multiple-choice and matchmaking test is determined with the similarities and differences in structure.

The multiple-choice form is often used due to short correction time. A multiple-choice test generally consists of a one-sentence statement or question sentence called a stem, and several answer choices have alternatives or the options. One of these alternatives is the right or best answer and is called the answer key, while the other options are the answers called distractors (Azwar, 2016). According to Kuechler & Simkin (2003), the advantages of multiple-choice tests are: (1) It test large amounts of knowledge practically; (2) it takes short time to evaluate; (3) it is easy to process the results of the assessment; and (4) it is easy to analyze the level of difficulty of each item. Multiple-choice test also has some disadvantages, namely: (1) it takes a lot of time to make it; (2) the material tested does not test the ability to analyze and does not test creativity or unique ideas; and (3) the form of the test only encourages the

students to learn the surface in learning the material so that there is a tendency for students not to understand the concept as a whole.

The matchmaking test is used for relationship recognition and association making. The statement or question in the matchmaking test is called the premise while the alternative answer is called a response. The advantages of matchmaking tests are: (1) easy to make; (2) efficient in scoring; (3) effective for assessing definitions and relationships; and (4) saving the space. On the other hand, the disadvantages of the matchmaking test are: (1) it takes a lot of time to make item questions; (2) it is not suitable for measuring higher-order thinking analysis; (3) there is a tendency for someone to guess answers rather than understand the material; and (4) if the premise is with a response number is the same, so the last answer can be called a gift or bonus because that test takers no longer need to read questions to match the premise and the response (Penn State University, 2017); therefore the number of answers must be more than the premise.

Through the explanation, it can be understood that the similarities between multiple-choice and matchmaking forms are: (1) requires the test taker to choose an answer with many distractors; (2) test takers have a tendency to guess the answers; and (3) the level of difficulty between multiple-choice and matchmaking test considered as balance. According to the research conducted by Jumaeda (2016), there was no difference in difficulty level between the multiple-choice form test and the matching test in Arabic for class XI students at Madrasah Aliyah Negeri Waimital, West Seram Regency. This study showed that the difficulty level of the test is in the form of multiple-choice and matchmaking were balanced with the difficulty level of the question form.

At first glance, the multiple-choice and matchmaking forms are almost the same, but there are some differences: (1) The number of alternative answers provided. In the multiple-choice form, three to five alternative answers are provided in a question, while matching the number of responses is exaggerated from the number of premises. As an example, the matchmaking test consists of ten questions. The available choices must be more than ten questions to avoid a bonus score on the last premise; (2) How to face the test item. The multiple-choice form requires precision so as not to be confused with the number of options available. In contrast, in the matchmaking form, there are more alternative answers which makes the test-taker have to ascertain whether the premise has been attached to the right response because there are still many remaining responses; (3) The consequences of a wrong answer. The matchmaking form has more answer alternatives, so the error of matching one partner will result in another error in matchmaking other partners. In making the question, matching the placement of the response with the premise must be considered because placing the premise with the response too close will provide more opportunities for the test-taker to make it easier to match. Whereas in multiple-choice, the wrong answer to one question does not affect other questions. Based on this explanation, there is an assumption that there are differences in anxiety test level in the multiple-choice and matching test. Hence, the title raised in this study is "Differences in Test Anxiety Levels in terms of Test Form."

Methods

This study used a quantitative approach with experimental research. The data were presented in numbers form, then analyzed by inferential statistics. According to Surunjukri (2015), inferential statistics is a statistical section that discusses how to analyze data and draw conclusions based on population parameter estimates and hypothesis testing. Based on the problem, this research was a comparative study that aimed to test the comparison of the measurement results of certain variables in one group with another.

The research design used in this study was a quasi-experimental research with two-group pretest-posttest design. Two-group pretest-posttest design or before-after design is a form of research that

measures the dependent variable that has been manipulated by the subject and measures the dependent variable with the same measuring instrument. In this study, the population of the research subjects were students of the Faculty of Psychology Education that registered in 2018. The subjects were limited to the class of 2018 in order to find the same course tested during this study. Seniati, Yulianto and Setiadi (2011) suggest that the requirements for experimental research are manipulation, control, and randomization. This research fulfilled the requirements of experimental research.

The variable measured in this study was test anxiety, the same measuring instrument given before the test and after the test. In this study, randomization was carried out as a control for the proactive history of genetic factors, namely age and gender and to divide the form of the next test. Another control was the level of question difficulties with two forms, namely multiple-choice and matchmaking forms. It is the symbols and process (Table 1) of two-groups pretest-posttest design:

$$\text{Measurement } (O^1) \rightarrow \text{Manipulation } (x) \rightarrow \text{Measurement } (O^2)$$

Table 1. Experimental Research Design

Students (Randomization of test Form)	Multiple Choice Group	→	Measurement (O ¹) (Pretest)	→	Manipulation (x) (Multiple Choice Test)	→	Measurement (O ²) (Posttest)*
	Matchmaker Group	→	Measurement (O ¹) (Pretest)	→	Manipulation (x) (Matchmaker Test)	→	Measurement (O ²) (Posttest)

The pretest was given after the students were informed about the final semester test. The information was about two forms of test among others multiple-choice and matchmaking test and it had been randomized previously. After being informed, students were given the Reaction to Tests (RTT) instrument to measure the first reaction before obtaining the test form treatment. After the pretest, students were given the treatment in the form of a test form according to randomization. Students were given the Reaction to Tests (RTT) instrument for the second time after taking the test (Posttest).

The control in the implementation of this experimental research assisted by a lecturer of the tested subjects. The things controlled to generate behaviors are: (1) Equivalent form of the test, the type of objective test with the same difficulty level. Both test forms were analyzed using the Rasch Model, the difficulty level of the questions ranged from -2.20 logit to 2.60 logit; (2) The test was conducted during the final semester test, thus it had a higher percentage weight and heavier material than the quiz; (3) The final semester test schedule was informed a month before, the reason of readiness was not considered due to the readiness of the student during the test; and (4) In the final semester test, the supporting lecturer helps to supervise.

Expert judgment was carried out to test the process given to five students. The trials were carried out using the content validity approach of Lawshe's CVR and using the reliability of previous research that had been translated and back-translated. To find the validity of the RTT scale, trials were conducted through a small group approach with the content validity coefficient of Lawshe's CVR. The calculation of the validity coefficient of this content, using a calculator made by Prasasto on his website <http://prasasto.xyz> based on Saifuddin Azwar's 2015 book "Reality and Validity". The number of calculators provided was seven raters with one hundred points on a scale. In the trial, there were only five ratings used. Rater assessed the suitability of the scale with the blueprint by giving a score of 1-5 (Prasasto, 2018).

Lawshe (1975, Cited in Azwar 2018) is a statistic that reflects the level of content validity of items based on empirical data. The assessment was carried out by giving a number 1 to 5 (very not essential - very essential). CVR figures move between -1.00 to +1.00 when $CVR > 0.00$ means that 50% more of the panel stated the important items. The greater the CVR than the number 0, the more essential and higher

the content validity. The overall average of 0.75 means essential items and each item was above 0, it means each item is essential and can be used.

Reliability, known also as consistency, reliability, trustworthiness, stability and consistency, or called reliable measurement process (Azwar, 2018). Zeidner & Matthews (2003) reported the consistency of the internal reliability of the subscales ranging from $\alpha = 0.68$ to $\alpha = 0.81$ while for total reliability scale $\alpha = 0.78$ for 40 items. In 1986 Detlef H. Rost and Franz J. Schermer reported the Worry subscale $\alpha = 0.86$, Irrelevant Thinking $\alpha = 0.91$, Tension $\alpha = 0.93$, Bodily Reaction $\alpha = 0.76$ with a total average reliability of this scale is $\alpha = 0.94$. Based on Guilford's reliability rule as the basis for determining good reliability criteria, the reliability of the RTT scale was very reliable.

Results and Discussion

At the very first time, the subjects of the research were 172 students and divided into two groups, 87 students were in the multiple-choice group and 85 students were in the matchmaker group. In this study, each of the groups was only tested by one model of the test form, the multiple-choice form or the matchmaker form. The number of subjects was decreased to 142 students because of the processing data with the Rasch model and it found that 30 students were invalid. Thus, to maintain data stability, 30 invalid students were eliminated. From 142 subjects, it was divided to 74 students in the multiple-choice group and 68 students in the matchmaker group. The research data used in the Independent sample T-test was the gain-core data or the difference from the reduction in the posttest and pretest scores of the multiple-choice and matchmaker groups calculated using SPSS 24.0.

Table 2. Statistical Analysis of Multiple Choice and Matchmaker Test

No	Test Form	N	Mean*	Std. Deviation*	Std. Error Mean*
1	Multiple Choice	74	2.17	0.79	0.91
2	Matchmaker	68	2.07	0.36	0.44

*In Logit

According to the analysis (Table 2), it found that the mean of the multiple-choice group was 2.17 while the matchmaking group was 2.07 with a small difference that showed multiple-choice anxiety was higher than matchmaking. This finding is similar to the previous research held by Steven S. Shaha (1984). Previous research by Shaha found that multiple choice results were higher than the matchmaking form. Another research found that multiple choice tests had the possibility of higher anxiety level if the test had negative scoring or incorrect answers (Pamphlett & Farmill, 1995).

Based on the other analysis (Table 3), this research also showed that the data were homogeneous ($F = 2.30 > 0.05$), meaning that there was no variance between the multiple-choice and matchmaking groups, the variation of the data in both groups was the same. From the results, it can be seen that the difference was 0.05 ($t = 0.92 > 0.05$) the Null Hypothesis was accepted. There was no significant difference in Test Anxiety in terms of the form of multiple choice and matchmaking tests. The level of significance in hypothesis testing only describes the magnitude of the risk of being wrong in deciding to reject the null hypothesis and does not represent the effect of the parameters tested.

Table 3. Independent Sample T-Test Analysis

		Test Form	
		Equal Variances Assumed*	Equal Variances Not Assumed*
Levene Test for Equality of Variances	F	2.30	
	Sig	0.13	
T-test for equality of means	T	0.92	0.94

*In Logit

Conclusion

Based on the results, it can conclude that there was no significant difference in anxiety level between the groups working on multiple-choice and matchmaking. Even though there were no significant differences, there were several concerns in constructing the test. There are difficulties for educators in deciding the quality of the construct of the test that is influential in students' decision. The study limitation was the amount of the subject in this study. The subject of this study is only taken from one specific grade of the student. According to the limitation, future research tends to understand more about test anxiety from broadening grades, such as test anxiety in primary school, junior high school, senior high school, even graduate school. The future researcher also suggested finding the test anxiety in different developmental age or different cultures. Moreover, this research also suggests the teacher or lecturer to apply active methods to manage the anxiety before the test. Faculty members and heads of department should let the students learn study preparation and support their engagement in class activities and accomplishment.

References

- Arikunto, S. (1992). *Dasar-dasar evaluasi pendidikan*. Jakarta: Bumi Askara
- Ali, Linan (2012). *Test anxiety may leave some students unable to achieve on exams*. <https://www.lcsun-news.com/story/life/wellness/2017/11/21/test-anxiety-performance-anxiety-students-finals/800902001/>
- Anastasi, A & Urbina S. (1997). *Psychological testing*. USA:Prentice-Hall,Inc
- Azwar, S. (2016). *Tes prestasi , fungsi dan pengembangan pengukuran prestasi belajar*. Yogyakarta : Pustaka Pelajar.
- Azwar, S. (2018). *Validitas & realibilitas*. Yogyakarta : Pustaka Pelajar
- Elfiana, Stephanie. (2005). *Perbedaan kecemasan tes antara tes berbentuk pilihan ganda dan tes berbentuk esai*. Jakarta: Fakultas Psikologi, Universitas Katolik Atma Jaya Jakarta
- Febrayosi, Puti. (2012). Estimasi true score pada second order unidimensional data: sebuah studi simulasi monte carlo tentang dampak panjang tes, tingkat kesukaran, dan daya pembeda item. *Jurnal Pengukuran Psikologi dan Pendidikan Indonesia*. Vol 1(4) 243-262
- Jumaeda, ST. (2016). Komparasi bentuk tes pilihan ganda dengan tes menjodohkan (matching test) ditinjau dari tingkat kesukaran pada mata pelajaran bahasa Arab di MAN Waimital Kabupaten Seram bagian barat. *Keguruan Institut Agama Islam Negeri Ambon*. Vol.1, No.1, 55- 71. *STAIN Samarinda*.
- Jihad, Asep & Abdul Haris. (2008). *Evaluasi pembelajaran*. Yogyakarta: Multi Pressindo
- Kubiszyn, Tom & Gary Borich. (2003). *Educational testing and measurement: classroom application and practice*. India: Replika Press

- Kuechler, W. L. & Simkin, M. G. (2003). How well do multiple choice tests evaluate student understanding in computer programming classes? *Journal of Information Systems Education*. West Lafayette: Winter 2003. Vol. 14, Iss. 4; p.389
- Lynn Miller. *Test Anxiety*. Cited on October 1, 2018, https://www.anxietycanada.com/sites/default/files/Test_Anxiety_Booklet.pdf
- Naga, Dali S. (2010). *Ukuran efek dalam laporan hasil penelitian*
- Pamphlett, R., & Farmill, D. (1995) Effect of anxiety on performance in multiple choice examination. *Medical Education*. 29 (297-302)
- Penn State University. (2017, March 15). *personal.psu.edu*. Cited on October 1 2018, <http://personal.psu.edu/bxb11/QuizQuestions/QuizQuestions5.html>
- Reber, Arthur S., & Reber, Emily S. (2001). *Dictionary of psychology: Third Edition*. London: Penguin Books
- Reynolds, Cecil R, Ronald. B Livingston, Vivor Willson. (2009). *Measurement and assessment in education*. United States of America: Pearson
- Revilla, Lina. (2009). *Kecemasan menghadapi tes (test anxiety) dan dampaknya terhadap aktivitas belajar*. Samarinda: STAIN Samarinda
- Rest, Dehef H, & Franz J.Schermer. (1986). Reaction to Tests (RTT) Manifestations of Test Anxiety “(Dai- MAN): Same or Different Concepts? P.114-129
- Sarason, G. I. (1977). *The test anxiety scale: concept and research*. Washington: Departement of Psychology
- Sarason, Irwin G. (1982). *Stress, anxiety, and cognitive interference: reaction to test (RTT)*. Washington: Departement Of Psychology, University Of Washington
- Sarason, I. G., & Sarason, B. R. (1990). Test anxiety. In H. Leitenberg (Ed.), *Handbook of social and evaluative anxiety*. 475–496. New York: Plenum Press
- Sakti, Kartika Oktaviani & Yuliatr Sastrawijaya. (2017). *Bentuk tes objektif dan kecemasan pada pembelajaran membaca huruf hiragana bahasa jepang*. Jakarta: Universitas Negeri Jakarta
- Sangaji, Etta Mamang, & Sopiiah. (2010). *Metode penelitian*. Andi: Yogyakarta
- Sobur, Alex. (2016). *Kamus psikologi*. Bandung: Pustaka Setia
- Suryaratri, Ratna Dyah & Anna Armwini Rangkuti. (2015). *Statistika deskriptif, “untuk psikologi dan pendidikan”*. Jakarta: Prodi FIP-UNJ
- Seniati, Liche, Aries Yulianto, & Bernadette N. Setiadi. (2011). *Psikologi eksperimen*. Jakarta: Indeks
- Shechter, M., & Zeidner, M. (1990). Anxiety: Towards a decision-theoretic perspective. *Journal of Mathematical and Statistical Psychology*, 43, 15-28
- Raju, P. Mohan, Misganaw, Mesfin., & Esa Alia. (2010). Test anxiety scale: reliability among ethiopian students. *Psychological Reports*, 107, 3, 939-948
- Tarigan, Mitra. *Ini kecurangan ujian nasional 2015 versi sekretaris jenderal federasi serikat guru indonesia (FSGI)*. Cited on 4th November 2018. <https://nasional.tempo.co/read/658121/ini-kecurangan-ujian-nasional-2015-versi-fsgi/full&view=ok>
- Zeidner, M (1998). *Test anxiety: the state of art*. New York: Kluwer Academic /Plenum Publishers
- Zeidner, M. & Gerald, Matthews. (2003). *Test anxiety: encyclopedia of psychological assessment*. London: Sage Publication, 965-969. <http://dx.doi.org/10.1135/9780857025753.n2>