JP3I (Jurnal Pengukuran Psikologi dan Pendidikan Indonesia)
DOI: http://dx.doi.org/10.15408/jp3i.v11i1.22352
http://journal.uinjkt.ac.id/index.php/jp3i

Qutest Construction and Psychometric Evaluation as Test of Attention and Willpower for Employee Selection Screening

Iqbal Maesa Febriawan[1], Achmad Afrizal Fauzan[1], Chelsea Adelina[1], Hana Naufalinda Afifah[1], Razaf Pari[1], Aswin Januarsjaf[1]

Global Talentlytica Indonesia, Indonesia[1]

iqbal.maesa@talentlytica.net / iqbalmaesa@gmail.com

Abstract

Attention and willpower are some of the primary modalities for work, so it is necessary to evaluate those potentials. However, standard tests for assessing the two aspects, such as the Pauli or Kraepelin test, are bias-prone to numerical ability. Due to that concern, Qutest was designed as an attention and willpower (screening) test with symbolic stimulus adapted from the d2 test (Brickencamp, 1998). Two studies were conducted for Qutest psychometric properties evaluation. In the first study, 261 participants finished Qutest on a web-based platform with a maximum of 96 pages, each containing 24 symbols of key stimuli (a symbol Q with two dots) and distractors that had to complete in a maximum of 30 minutes. Analyses of internal consistency and correlation with Pauli test results were performed to establish test psychometric properties. The results supported adequate internal consistency and convergent validity claim for Qutest. In addition to internal consistency, a second study was conducted to develop further and evaluate parallel-form consistency between mobile and computer versions. In the second study, 371 participants completed a maximum of 192-page Qutest with a similar amount of time to the previous study. Participants took the mobile version first and then the computer versions of the test with a minimum of one lag day. Statistically significant correlation index of results across devices supported parallel-form reliability of both mobile and laptop/desktop Qutest version. A follow-up analysis of answering aids (mouse, trackpad, touch screen) for the computer version recommended using a mouse or touch screen instead of a trackpad due to statistically significant difference in performance among the three tools. Thus, Qutest can be used as a psychometrically adequate test and be flexible in various devices to assess attention and willpower for (prospective) employee screening.

Keywords: attention, qutest, screening, willpower

Abstrak

Atensi dan daya juang adalah beberapa modalitas utama untuk bekerja, sehingga potensi tersebut perlu dievaluasi. Namun demikian, tes standar untuk menilai dua aspek tersebut, seperti tes Pauli atau Kraepelin, cenderung bias terhadap kemampuan numerik. Oleh karena itu, Qutest dirancang sebagai tes skrining atensi dan daya juang dengan stimulus simbolik sebagai adaptasi dari tes d2 (Brickencamp, 1998). Dua penelitian dilakukan untuk evaluasi atribut psikometrik Qutest. Dalam studi pertama, 261 peserta menyelesaikan Qutest pada platform *berbasis* web *dengan maksimal 96 halaman, masing-masing berisi 24 stimulus kunci (simbol Q dengan dua titik) dan distraktor yang harus diselesaikan dalam waktu maksimal 30 menit. Analisis konsistensi internal Qutest dan korelasi dengan hasil tes Pauli dilakukan untuk menetapkan properti psikometrik tes. Hasilnya mendukung konsistensi internal yang memadai dan klaim validitas konvergen untuk Qutest. Selain konsistensi internal, studi kedua dilakukan sebagai tindak lanjut evaluasi konsistensi bentuk paralel antara versi ponsel dan komputer (*PC/*laptop). Dalam studi kedua, 371 peserta menyelesaikan Qutest maksimum 192 halaman dengan jumlah waktu yang sama dengan studi sebelumnya. Peserta mengikuti tes versi ponsel terlebih dahulu kemudian versi komputer dengan jarak waktu minimal satu hari. Indeks korelasi yang signifikan secara statistik di kedua perangkat mendukung keandalan bentuk paralel dari versi Qutest ponsel dan komputer. Analisis lanjutan dari alat bantu (mouse, trackpad, layar sentuh) untuk versi komputer merekomendasikan penggunaan mouse atau layar sentuh karena perbedaan hasil yang signifikan secara statistik dari penggunaan ketiga alat tersebut. Dengan demikian, Qutest dapat digunakan sebagai tes yang memadai secara psikometri dan fleksibel di berbagai perangkat untuk penilaian atensi dan daya juang untuk penyaringan (calon) karyawan.*

***Kata kunci**: atensi, daya juang, qutest, skrining*

Introduction

Regarding attention and willpower, work attitude assessment is necessary to identify the primary individual characteristic before evaluating the suitability of intelligence and personality profile with a particular job profile. There are two derivatives of attention, namely sustained attention and selective attention. Sustained attention refers to the ability to persist in doing repetitive, long, and monotonous tasks, while selective attention is the one to persist in doing relevant tasks by ignoring distractions (Gomez-Ramirez et al., 2016; Grahn & Manly, 2012). Reyes-Parra et al. (2017) found that employees naturally relied on sustained attention in work, while selective attention could help employees focus on work details, especially when their work starts.

Attention tests are primarily designed to digitally assess psychopathological conditions for detecting attention deficit hyperactivity disorder (ADHD) among adolescents (Luf & Fichman, 2012). Attention measurement for the adult (normal) population has been constructed but still on a limited scale and has not been aimed for occupational purposes (Ha et al., 2002). Working with much information requires (prospective) employees to select and maintain optimal attention to finish tasks well. Cohen et al. (2017) agreed that optimal attention impacted productivity and compliance to work safety. In other words, optimal attention impacts work quality and quantity. Organizations ideally use these two aspects to support employees in completing the job.

The Kraepelin and Pauli test are two of the most utilized attention tests in recruitment and selection. The two tests measure attention, concentration, willpower, and work stability. Although commonly used in recruitment and selection, the two tests have been used for practical purposes without further theoretical advancement for the last decade (e.g. Pradipta & Hidayat, 2019; Hasanah & Herliansyah, 2016). In addition to the initial purpose for detecting clinical populations, a critique of the tests as attention measurement has been raised against numerical stimulus so that the tests are prone to bias against numeracy skills. It implies that the test result may be higher for assessees with superior numerical ability than those with a low one. Due to that constraint, some developments of attention-and-performance tests utilized stimulus in a symbolic form (Brickenkamp & Zillmer, 1998; Lozano et al., 2015).

An example of an attention test with symbolic stimuli was created by Brickenkamp & Zillmer (1998), namely the d2 test. The test appears in symbols resembling the letters 'd' and 'p' with one to four lines either above or below the symbol. Participants are instructed to choose the symbol 'd' with two lines (further referred to as key stimulus) among the other available distractors. The d2 test produces several indices, including total chosen stimulus, correctly chosen key stimuli, errors (amount and percentage), and concentration performance (total correct answers minus errors). The test has been evaluated for internal (α = .61-.97) and test-retest reliability (intraclass correlation = .78 – .94) as well as structural and criterion-related validity in the general and clinical population (Brickenkamp & Zillmer, 1998; Bates & Lemay Jr., 2004; Lee et al., 2018). In addition to psychometric evaluation, Raghavendra and Telles (2012) found that the d2 test scores increased after the meditation intervention, which was previously claimed to increase individual attention. In other words, the d2 test has been empirically proven as a measure of attention. The use of the d2 test is mostly in clinical contexts and has not been widely explored in industrial-organizational selection settings (Bates & Lemay Jr., 2004; Lee et al., 2018). For the non-clinical purpose, Lord and Clarke (2005) found that the d2 test could predict driving proficiency to obtain a driver's license.

Due to the need for an attention-and-performance screening test for recruitment and selection, the d2 test was adapted and named Qutest. The test development was intended to solve bias proneness against Pauli and Kraepelin test's numerical abilities. In addition, Qutest was designed to measure (visual) attention and willpower. There are two approaches in the test construction for measuring attention, namely perceptual speed and visual search (Santacreu & Quiroga, 2015). Qutest is classified as a perceptual speed test because the stimuli are displayed all at once. Test participants are instructed to

respond as quickly and precisely as possible, and the test is time-limited. The d2 benchmark for developing Qutest because the stimulus form is more straightforward than the visual search test that uses complex images as the stimulus (cf. Lozano et al., 2015).

Based on the number of processed stimuli and (in)accuracies in selecting the key stimulus, Qutest produces two categories of index related to the quantity and quality of work. The variables' energy-of-will' and 'willpower' represent the quantity of work. Those variables are calculated from the total number of successfully processed stimuli, in which the greater of them represents a greater willingness to complete the task. The difference between energy-of-will and willpower is that the former is implied only from the number of successfully processed stimuli. In contrast, the latter uses time (in seconds) as a divider on each page so that a person's willpower may be concluded low when each page is completed in a relatively long time.

The variables of attention, concentration performance, and accuracy represent aspects of work quality. Like work quantity, attention-and-concentration performance variables both utilize the total number of correctly answered key stimuli as the basis for calculation. Concentration performance is distinguished from attention by utilizing time (in seconds) as a divider of the number of correctly chosen key stimuli. It means that the more participants answer correctly in a short time, the concentration performance may be considered as high. Accuracy is operationalized by the number of unchosen key stimuli plus the number of non-key stimuli chosen by participants.

Due to the increasing need for digital assessment, mainly due to the pandemic condition that hinders face-to-face assessment activities (Ghosh et al., 2012), Qutest was designed as an assessment method that can be carried out online through mobile web browsers or desktop/laptop computers. The option is more beneficial from the participant's perspective than using a dedicated application due to access on various devices without installation. The stimulus clicking activity during the test is also considered easy on both a cellphone and a computer/laptop. Regarding distractions that may occur, it can be analogous to environmental distractions that may arise and could not be controlled by the test administrator. Qutest's construction was intended to create an Indonesian version of a symbolic test of willpower and attention with adequate psychometric properties

Methods

Research Design

Qutest psychometric properties were evaluated using convergent validity with Pauli test results as the benchmark and testing the internal and parallel-form (mobile vs laptop/desktop) consistency of Qutest result. Similar to the d2 test scoring (Bates & Lemay Jr., 2004), 'item' was defined as scores per page (total chosen stimulus, number of correct answers, and number of incorrect answers). Thus, the first-study version of Qutest had 96 items, each of which has three indices (total chosen stimulus, number of correct answers, number of incorrect answers). Qutest in the second study extended participants' maximum performance to 192 items with the same number of indices on each page as in the first study. Each page item was used to assess the internal consistency of the test. The total items of chosen stimulus, correct answers, and incorrect answers became indices that would be correlated with analog indices from Pauli as a benchmark test. Generally, the two studies utilized a non-experimental design due to no experimental manipulation.

Item Construction

Qutest's creation began by designing the key items: a circle with a vertical line below (similar to the letter Q) with two dots above, below, or one point above and one point below the symbol. Then, distractor items were also created in addition to the key stimuli (e.g. a circle with a vertical line above and a circle with a vertical line at the bottom with one or three dots). Visually, the variety of stimuli that participants

can choose from is described in **Figure 3**. Key answers are in **Figure 1**. while the distractors are in **Figure 2**.

Figure 1. Key Stimuli

Figure 2. Distractor Stimuli

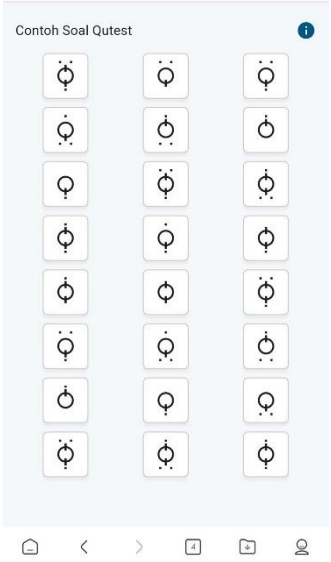

Figure 3. Qutest Appearance on Assessee's View

Study 1 – Validation and Internal Reliability

Data Collection Method

Data collection was carried out online through Talentlytica's web-based application from 7 September to 23 September 2020. Participants took the test via mobile phone connected to the internet. For initial development, Qutest was designed as many as 96 pages, with each page containing 24 stimuli containing key items and distractors, making the maximum total chosen stimulus was 2,304 symbols. Each page contained 11-12 key items and randomly placed distractors. Test time was limited to a maximum of 30 minutes. The index generated by Qutest was the total stimulus processed (number of pages completed multiplied by 24), the number of the correctly chosen key items, the number of total errors (selected distractors), and processing time (in seconds). In addition to Qutest performance, education level was collected as demographic data. After finishing Qutest, participants were given the classical Pauli test, used as a validation benchmark for the Qutest index.

Participants

Three hundred people completely participated in the data collection phase. Thirty-nine participants resulted in outlier Qutest indices (less than -2z and more than 2z), therefore their data were eliminated. Hence, the total number of participants included in the Qutest validation was 261 participants. The majority of the participants were high school or vocational school graduates (n = 126),

http://journal.uinjkt.ac.id/index.php/jp3i

followed by university graduates (n = 66), diploma graduates (n = 16), and finally junior high school graduates (n = 3). Fifty participants did not identify their educational background.

Data Analytics

The reliability evaluation of this test used an inter-item consistency approach (Cronbach Alpha), while the validity claim of this test utilized a convergent validity approach by correlating the Qutest score and the Pauli test as benchmark measures. Due to non-normal data distribution (multivariate $W = 0.426$, $p < .001$, bivariate $W = 0.105$-0.982, $p < .01$), correlational analysis was conducted through Spearman rank-correlation test rather than Pearson product-moment correlation test. Criteria for internal reliability is a Cronbach Alpha index above .7, and validity criteria are upheld when the correlation between the two test scores is statistically significant ($p < .05$) (Chadha, 2009). All data analyses were conducted with JASP 0.13.0 software.

Study 2 – Parallel Form Reliability

Data Collection Method

Data collection was conducted online through a web-based application from 21 June to 6 July 2021. Participants took the test via mobile phones and computers (PCs or laptops) connected to the internet. The items in the two types of digital devices met the requirements for testing consistency between formats, namely similar impression validity, based on the same theoretical basis, and similar answering format (Chadha, 2009). Qutest on the two devices were designed as many as 192 pages with each page still containing 24 stimuli containing both key items and distractors. The number of key items and distractors, their randomization, maximum time limit, and the index generated in this study were the same as the first study. Participants received an invitation to take the mobile version first and when they finished, they would receive an invitation the day after to do the desktop/laptop version. After finishing the computer version of the test, participants were asked questions about the assisting device with multiple choices (mouse, touch screen, or trackpad) to determine the effect of the aid device on the test takers' performance.

Participants

Four hundred and six people completely participated in the second study. Thirty-five participants produced outlier Qutest indices (below -2z and above 2z) to eliminate their data. A total of 371 participants' data were included in the parallel format reliability study of Qutest. The majority of the participants were university graduates (n = 316), followed by high school or vocational school graduates (n = 31), diploma graduates (n = 22), and finally junior high school graduates (n = 2).

Data Analytics

The evaluation of parallel form reliability was conducted by correlational analysis of Qutest scores on mobile phones and desktop/laptops. Similar to previous study, correlational analysis was conducted through Spearman rank-correlation test due to non-normal data distribution (multivariate $W = 0.612$, $p < .001$, bivariate $W = 0.469$-0.983, $p < .001$). Consistency across formats would be upheld when the correlation between the two measurements is significantly correlated ($p < .05$). In addition, an analysis of differences between the use of assistive devices on a computer (mouse, trackpad, and touch screen) was performed using the Kruskal-Walis test as the nonparametric alternative to one-way analysis of variance (ANOVA). A statistically significant difference was then followed by post-hoc analysis using the Bonferroni index. The index was chosen because of consideration for the unequal number of participants between groups (Langsrud, 2003). In general, using a mouse generated a higher number of Qutest performance than using a trackpad, while using a touch screen produced a non-significant difference compared to the two devices. Data analysis was conducted with JASP 0.14.1.

Results and Discussion

Study 1 - Validation and Internal Reliability

Participants took the Qutest for an average of 26.19 minutes (*SD* = 4.42). The total processed stimulus on average was 832.85 (*SD* = 217.19, Cronbach α = .990), the number of correctly chosen stimulus on average was 811.32 (SD = 212.99, α = .979), while the average error was 21.53 (SD = 18.97, α = .938). Based on the Cronbach alpha coefficient value, Qutest can be claimed as a highly reliable measurement in terms of inter-item consistency. The number of test-takers who completed Qutest in under 30 minutes was 96.55% of total participants (*n* = 252), so that the number of Qutest pages can be increased for the next development so that the speed test criteria, which cannot be completed in the allotted time, would be met.

The high internal reliability index for Qutest followed Bates and Lemay Jr.'s (2004) finding about the internal consistency of the d2 test was above 0.9. The finding can be understood by the characteristics of Qutest as a performance speed test with a uniform level of item difficulty. The high-reliability coefficient on the three aspects measured by Qutest was also analogous to the study of Steinborn et. al. (2018) on the d2 test scoring system where the total number of chosen stimuli, the number of correctly answered stimuli, and a number of errors were relatively consistent. Therefore, they could function as the main indices of the attention test.

The correlational analyses of Qutest and Pauli test scores were presented in **Table 1**. The result indicated that all indices (total chosen stimuli, total correct answers of Qutest, and total errors) were positively correlated with the respective variable on the Pauli test. Qutest total chosen stimulus and correct answer index were significantly correlated with the total worked-out stimulus and the total correct answer of the Pauli test can be explained based on the data that the average number of Pauli test performance (*M* = 2,539.98). This was not much different from the number of correct stimuli (*M* = 2,530.51), which was similar in pattern to Qutest's score, thereby increasing the possibility of the three indices being correlated with each other. In addition, the effect size of the correlation index was small-to-medium, partly because Pauli's score range is much larger, even up to four times the Qutest score range. (SD_{total} = 1,057.24, $SD_{correct}$ = 1,065.69, SD_{error} = 50.62, $SD_{correction}$ = 14.21).

Table 1. Pearson Correlation Index between Qutest & Pauli Test Score

Qutest Score – Pauli Score	Spearman ρ
Total chosen stimulus Qutest – Total worked out stimulus Pauli	.313***
Total chosen stimulus Qutest – Total correct answer Pauli	.314***
Total chosen stimulus Qutest – Total error Pauli	-.099 (ns)
Total chosen stimulus Qutest – Total correction Pauli	.061 (ns)
Total correct Qutest – Total correct answer Pauli	.328***
Total correct Qutest – Total worked out stimulus Pauli	.328***
Total correct Qutest – Total error Pauli	-.121 (ns)
Total correct Qutest – Total correction Pauli	.054 (ns)
Total error Qutest – Total error Pauli	.200**
Total error Qutest – Total worked out stimulus Pauli	-.054 (ns)
Total error Qutest – Total correct answer Pauli	-.055 (ns)
Total error Qutest – Total correction Pauli	.092 (ns)

Source: Primary data, $^{**}p < .01$; $^{***}p < .001$; ns = not statistically significant

Study 2 – Parallel Form Reliability

Due to most participants completed 96-page Qutest in the first study under 30 minutes, the number of pages was doubled to 192-page with the same amount of maximum processing time. Regarding

http://journal.uinjkt.ac.id/index.php/jp3i

answering aids used by participants with computers, the majority of participants used a mouse (58%), the second most used a trackpad (37%), and the rest used a touch screen (5%). **Table 2.** displays the correlation between Qutest scores on mobile and computer versions. Due to the statistically significant result correlation between formats, the claim of parallel-form reliability can be upheld.

Table 2. Pearson Correlation Index between Mobile & Computer Version of Qutest Score

Mobile & Computer Versions of Qutest scores	Spearman ρ
Total selected stimuli	.631***
Total correct stimuli	.650***
Total error	.569***
Total completed pages	.532***

Source: Primary data, ***$p < .001$

Regarding assisting device to answer aid of Qutest computer version, there were statistically significant differences in the number of selected stimuli ($H(2) = 20.364$, $p < .001$), the number of the correct answer ($H(2) = 19.846$, $p < .001$), and the number of completed pages ($H(2) = 11.677$, $p = .003$ among participants using a mouse, trackpad, and touch screen (based on **Table 2.**). The number of errors was not statistically significant ($H(2) = 3.352$, $p = .187$). Post-hoc analysis of the difference between assisting devices showed that Qutest performance with mouse exerted higher results than trackpad use. In contrast, the result with the touch screen did not differ significantly from the other two answering devices. Based on these results, it is recommended that Qutest participants on the computer version to use a mouse or touch screen as an answering device and not recommended to use a trackpad.

Table 3. Post-hoc Analysis of Computer version of Qutest Score with Mouse, Trackpad, and Touch Screen

Qutest Score	Tool 1	Tool 2	ΔM	T	Cohen's d
Total selected stimuli	mouse	trackpad	130.818	4,697***	0.521
	touch screen	trackpad	118,612	1,948	0.460
	mouse	touch screen	12.206	0.205	0.048
The total correctly-answered stimuli	mouse	trackpad	125.759	4,668***	0.518
	touch screen	trackpad	112,657	1,913	0.447
	mouse	touch screen	13,101	0.228	0.053
Total completed pages	mouse	trackpad	7,819	2,563*	0.279
	touch screen	trackpad	-0.806	-0.121	-0.027
	mouse	touch screen	8,624	1.325	0.325

Source: Primary data, *$p_{Bonferroni} < .05$, ***$p_{Bonferroni} < .001$

Conclusion

Qutest was designed as an alternative to the Pauli and Kraepelin test for evaluating work quantity and quality that describes the work attitude of (prospective) workers. The current study also extended the d2 test development to non-clinical populations or use. The psychometric properties of this test had met the quality for convergent validity, internal consistency, and parallel-form reliability. With a more accessible technology medium to work anywhere and anytime, Qutest can be an alternative for mass screening. Maximum allocated test-taking time is also an advantage for Qutest relative to the one-hour maximum time of the Pauli test. With a shorter test-taking time, inference on attention and willpower can still be upheld. To ensure optimal performance on the computer version, Qutest participants are advised to use a touch screen or a mouse as an assistive device for computer usage, and it is not recommended to use a trackpad.

A limitation of this study included control over participants. Even though the platform utilized in current studies has been equipped with a camera proctoring both on phone and desktop/laptop versions,

there is a possibility that participants were distracted by incoming calls or messages. The recommendation for practical use of Qutest regarding this issue is to activate the "don't disturb" mode. The feature will prevent telephone and message distractions so that Qutest participants can focus their attention on the assessment process.

Further development of this test may also consider evaluating test-retest consistency of the results on both test formats and supporting the current test's reliability claims. Steinborn et al. (2018) emphasized the role of intertemporal consistency testing to support internal reliability due to the finding that d2 test performance shifted in the short term but did not affect the test-retest reliability. The importance of intertemporal reliability was also supported by a non-significant practice effect on the d2 test results (Blotenberg & Schmidt-Atzert, 2019). Further development can also consider other demographic attributes such as gender and age. The current study did not control these variables because of the assumption that attention and willpower as basic and universal individual attributes are not affected by those three demographic variables.

Conflict of Interest

There is no conflict of interest in the research. Due to the intellectual property of Qutest, the test is available by request to the corresponding author's email or contacting Talentytica's website.

References

Bates, M. E., & Lemay, Jr., E. P. (2004). The d2 test of attention: Construct validity and extensions in scoring techniques. *Journal of the International Neuropsychological Society, 10*, 392–400, https://doi.org/10.1017/s135561770410307X.

Blotenberg, I, & Schmidt-Atzert, L. (2019). On the locus of the practice effect in sustained attention tests. *Journal of Intelligence, 7*(2), 12-26, https://doi.org/10.3390/jintelligence7020012.

Brickenkamp, R., & Zillmer, E. (1998). *The d2: Test of attention*. Hogrefe & Huber Publishers.

Chadha, N. K. (2009). *Applied psychometry*. SAGE Publications India Pvt Ltd

Cohen, J., LaRue, C., & Cohen, H. H. (2017). Attention interrupted: Cognitive distraction & workplace safety. *Professional Safety, 62*(11), 28-34.

Ghosh, A., Nafalski, A., Tweedale, J., & Dollard, M. (2012). Using hybridized techniques to develop an online workplace risk assessment tool. *Informatyka, Automatyka, Pomiary w Gospodarce i Ochronie Środowiska*, 42-45.

Gomez-Ramirez, M., Hysaj, K., & Niebur, E. (2016). Neural mechanisms of selective attention in the somatosensory system. *Journal of Neurophysiology, 116*(3), 1218-1231, https://doi.org/10.1152/jn.00637.2015.

Grahn, J. A., & Manly, T. (2012). Common neural recruitment across diverse sustained attention tasks. *PLoS ONE, 7*(11), e49556, https://doi.org/10.1371/journal.pone.0049556.

Ha, K. S., Kwon, J. S., & Lyoo, I. K. (2002). Development and standardization of the computerized attention assessment for Korean adults. *Journal of Korean Neuropsychiatric Association*, 335-346.

Hasanah, M. N., & Herliansyah, M. K. (2016). Analisis tes kecepatan dan ketelitian kerja yang representatif pada kondisi ruang kerja ideal. *Undergraduate Thesis*, Industrial Engineering Department, Universitas Gadjah Mada.

Langsrud, Ø. (2003). ANOVA for unbalanced data: Use Type II instead of Type III sums of squares. *Statistics and Computing, 13*(2), 163-167.

Lee, P., Lu, W-S., Liu, C-H., Lin, H-Y., & Hsieh, C-L. (2018). Test–retest reliability and minimal detectable change of the d2 test of attention in patients with schizophrenia. *Archives of Clinical Neuropsychology, 33*(8), 1060–1068, https://doi.org/10.1093/arclin/acx123.

Lord, W., & Clarke, P. (2005). Use of the d2 test of attention as a predictor of driving proficiency. In L. Dorn. (ed.), *Driver behaviour and training (first edition).* (pp. 437-442). Routledge.

Lozano, J. H., Capote, E., & Fernández, M. P. (2015). Convergent validity of the measures of attention and impulsivity in the Trees: Simple Visual Discrimination Test (DiViSA-UAM). *Anales de Psicología/Annals of Psychology, 31*(1), 74–83. https://doi.org/10.6018/analesps.31.1.161371.

Luf, D., & Fichman, N. (2012). Development and use of a computerized test, MATH—CPT, to assess attention. *Perceptual and Motor Skills, 114*(1), 59-74, https://doi.org/10.2466/08.11.22.PMS.114.1.59-74.

Pradipta, R. T. D., Hidayat, R. (2019). Konvergensi dan divergensi tes kraepelin dan tes pauli pada komunitas psikolog. *Master Thesis*, Professional Program of Psychology, Universitas Gadjah Mada.

Raghavendra, B. R., & Telles, S. (2012). Performance in attentional tasks following meditative focusing and focusing without meditation. *Ancient Science of Life, 32*(1), 49–53. https://doi.org/10.4103/0257-7941.113799.

Reyes-Parra, P., Jimenez-Martinez, M., & Lara-Baez, C. (2017). Selective and sustained attention of workers in two different shifts of the day. *Psychologia, 11*(2), 113-120, https://doi.org/10.21500/19002386.3059.

Santacreu, J., & Quiroga, M. A. (2015). The DiViSA's predictive validity by age: An objective online test of attention, *The Spanish Journal of Psychology, 18*, 1–11, https://doi.org/10.1017/sjp.2015.103.

Steinborn, M. B., Langner, R., Flehmig, H. C., & Huestegge, L. (2018). Methodology of performance scoring in the d2 sustained-attention test: Cumulative-reliability functions and practical guidelines. *Psychological Assessment, 30*(3), 339-359, https://doi.org/10.1037/pas0000482.