



Psychological Assessment

Evidence of validity of the Attentional Performance Test (APT)


Jonatas R. Bessa¹

 <http://orcid.org/0000-0002-2918-9666>


Neander Abreu¹

 <http://orcid.org/0000-0001-7636-3666>


Yuri Santana¹

 <http://orcid.org/0000-0002-3000-6564>

Roberto Beirão¹

 <http://orcid.org/0000-0002-3177-2500>

Jamine Cairo¹

 <http://orcid.org/0000-0002-8402-352X>

To cite this article: Bessa, J. R., Abreu, N., Santana, Y., Beirão, R., & Cairo, J. (2021). Evidence of validity of the Attentional Performance Test (APT). *Psicologia: Teoria e Prática*, 23(2), 1–21. doi:10.5935/1980–6906/ePTPPA13231

Submission: 31/03/2020

Acceptance: 16/12/2020



The content of *Psicologia: Teoria e Prática* is distributed under the terms of the Creative Commons Attribution License.

¹ Federal University of Bahia (Ufba), Salvador, BA, Brazil.

Abstract

This study aims to present evidence of validity of the Attentional Performance Test (APT). The search for evidence based on content, internal structure and reliability was carried out. The content's evidence was run with the expert judgement ($n = 07$; $k = 0.84$) and semantic analysis ($n = 12$, $k = 0.77$). Their results suggested an adequate agreement with the content of the assessment of sustained attention (S.A), verbal commands and the images of the instrument. An analysis of the factorial structure ($n = 1086$) resulted on 2 main factors, 4 dimensions and 12 measures (RMSR = 0.08). An analysis of internal consistency ($n = 1086$) of the APT showed adequate values ($\alpha > 70$). This study indicated that the APT presents evidence for content, construct and reliability. The present results contribute to APT evidences to confirm it as an adequate psychometric instrument to assess attention.

Keywords: validity; attention; psychological test; factor analysis; reliability.

EVIDÊNCIAS DE VALIDADE DO TESTE DO DESEMPENHO ATENCIONAL (TDA)

Resumo

Este estudo tem por objetivo apresentar evidências de validade do Teste do Desempenho Atencional (TDA). Buscaram-se evidências baseadas no conteúdo, na estrutura interna e na fidedignidade. As evidências de conteúdo foram realizadas por intermédio da análise de juízes ($n = 7$; $k = 0,84$) e análise semântica ($n = 12$, $k = 0,77$), cujos resultados sugeriram concordância adequada quanto ao conteúdo de avaliação da atenção sustentada, aos comandos verbais e às imagens utilizadas no instrumento. A análise da estrutura fatorial ($n = 1.086$) do instrumento resultou em dois fatores principais, quatro dimensões e 12 medidas (RMSR = 0,08). A análise da consistência interna ($n = 1.086$) do TDA apresentou valores adequados ($\alpha > 70$). Em conjunto, os resultados deste estudo mostram que o TDA apresenta evidências suficientes para o conteúdo, o construto e a fidedignidade. Os atuais resultados contribuem para confirmar o TDA como instrumento psicométrico adequado para avaliação da atenção.

Palavras-chave: validade; atenção; teste psicológico; análise fatorial; confiabilidade.

EVIDENCIA DE VALIDEZ DEL TEST DE DESEMPEÑO ATENCIONAL

Resumen

Este estudio tiene como objetivo presentar evidencia de validez para el Test de Desempenho Atencional (TDA). Se realizó la búsqueda de evidencia basada en contenido,

estructura interna y fiabilidad. Las evidencias de contenido se realizaron a través del análisis de expertos ($n = 7$; $k = 0,84$) y análisis semántico ($n = 12$, $k = 0,77$) y sus resultados sugirieron una adecuada concordancia con el contenido de evaluación de la atención sostenida, los comandos verbales y las figuras del instrumento. El análisis de la estructura factorial ($n = 1.086$) del instrumento se manifestó por dos factores principales, cuatro dimensiones y 12 medidas ($RMSR = 0,08$). El análisis de consistencia interna ($n = 1.086$) del TDA arrojó valores adecuados ($\alpha > 70$). Este estudio sugiere que el TDA presenta evidencia suficiente de contenido, constructo y fiabilidad. Los resultados confirman que el TDA es un instrumento psicométrico adecuado para evaluación de la atención.

Palabras clave: validez; atención; teste psicológico; análisis factorial; fiabilidad.

1. Introduction

Attention allows performing daily, physical and cognitive tasks and activities, as well as serves as an important basis for the activity of other cognitive functions such as executive functions and memory (Gillsoul, Simon, Hogge, & Collette, 2018). Deficits in attentional processes can imply difficulties in cognition in the functioning of everyday tasks. The literature suggests that these damages are due to clinical conditions, among them, attention deficit hyperactivity disorder (Berger, Slobodin, & Cassuto, 2017), bipolar disorder and borderline personality disorder (Gvirtz et al., 2015); or developmental conditions, when aging is observed, such as decline in execution and quality of response of different attentional processes.

Attention can be categorized as a mechanism for selecting information. The model of neural networks associated with the attentional system proposed by Petersen and Posner (2012) suggests that attention can be arranged in three main components: alerting (referring to automatic responsiveness); orientating (directing the attentional focus) and executive attention (processing and sensory modulation voluntarily directed to the stimulus). The activation of these networks would influence the manifestation of the operationalized attentional processes, as proposed by Van Zomeren & Brouwer (1994), resulting in selective, divided, alternating and sustained attention (S.A).

The alerting component is related to the central nervous system's ability to turn attention to non-specific stimuli and it has been documented as a survival mechanism of multiple species, being present even in neonates as a reflex behavior, producing physical, cognitive and emotional effects that assist in the response

readiness (Geva, Zivan, Warsha, & Olchik, 2013). Unlike what happens in the attentional alerting mechanism, the orientating network articulates to direct attention to a specific stimulus, improving the quality of the response. The orientating response is considered a product of a distributed neural network, which includes, among other neuroanatomical components, the frontal eye fields (Geva et al., 2013). The executive attention network, on the other hand, involves the recruitment of a mental and cognitive apparatus, which can deal with the maintenance of attentional targeting for a given task, being associated with error monitoring, being involved in the stimulus detection process (Van Steenbergen & Band, 2013). Therefore, executive attention is a fundamentally regulatory component.

The S.A process involves concentrating on a stimulus or performing a task for at least three minutes (Lin et al., 2018). In everyday life, S.A is used for activities such as attending a conference, reading a book, playing a musical instrument, communicating socially, cooking, directing and learning new content. S.A can also impact the storage of content in memory, academic performance and the implications for personal or collective security (Fortenbaugh, DeGutis, & Esterman, 2017; Lin et al., 2018; Esterman & Rothlein, 2019). S.A is associated with the executive attention network from Petersen & Posner's (2012) neural model. Fortenbaugh et al. (2017) suggested that S.A can be conceptualized through a process of voluntary modulation and directed to a stimulus, being required when there is a need to maintain the attention focus on a task for a long period of time, and for this reason, some authors also call it vigilance.

The use of the name vigilance can be discriminated as a process of S.A in itself. For example, studies suggest that vigilance performance declines with time spent on a task (Thomson, Besner, & Smilek, 2015; Fortenbaugh et al., 2017; Esterman & Rothlein, 2019) as long as the process is maintained, i.e., vigilance it would be part of the S.A process. One of the main hypotheses about this phenomenon involves overload, proposing that tasks involving the vigilance process are monotonous and reduce the ability to sustain the attention focus on the task, resulting in its directing to other external or internal stimuli, such as thoughts and emotional states. For this reason, tasks of S.A / vigilance are of great difficulty for exceeding the limit of information processing (Thomson et al., 2015).

The measurement of S.A involves the time of attentional focus on a specific object in a task. The development of new instruments that use paradigms involving

the investigation of this construct seek to go beyond the classic instruments that propose the detection or discrimination of random targets (Fortenbaugh et al., 2017). In this context, the continuous computerized tests of attention have been indicated for having a greater accuracy of the proposed measures (Cannavò, Conti, & Di Nuovo, 2016).

In the Brazilian context, the tests approved for use in the context of psychological testing in Brazil related to the cognitive function of attention in adults are mostly paper and pencil. According to the Psychological Testing Assessment System (SATEPSI, 2020), only the Computerized Attention Test, visual version, is a computerized instrument of S.A developed for testing the adult population and is favorable for use (SATEPSI, 2020). The use of instruments for assessing attention based on the paper and pencil modality implies ecological limitations of an instrument, that is, not presenting conditions that can capture daily situations and may be less effective due to lack or limitation of control of measurement of errors, less accuracy in the time of application of the instrument and in the measurement of motor response (Canini et al., 2014).

The Attentional Performance Test (APT – *Teste do Desempenho Atencional*, TDA in Brazilian Portuguese) was developed to evaluate adults in a quick test (about 6 minutes). It is a computerized test, simulating a daily scene which seeks to detect the functioning of attentional processes associated with S.A. The ATP is an instrument that simulates a scene in which the tester virtually drives on a road and sees a scene of a road through the car's windshield. Flashing images are displayed over time. Some images are of living beings (target stimuli), to which the tester must detect and respond as quickly as possible, pressing the space bar of the computer. The images of living beings are alternated randomly with images of geometric shapes (distractors) that appear on the screen. In this condition, no response execution should be performed, i.e., the response of pressing the space bar should be inhibited.

The APT consists of 12 measures, being manifested by sums (Total) and variances in performance throughout the test (Vigilance). The instrument's measurements are:

1) *Total Hit* (T.Hit): sum of correct responses to target stimuli, i.e., after stimulus detection, the tester presses the space bar before presenting a new stimulus;

2) *Vigilance of Hit* (V.Hit): standard deviation of the number of hits in the 6 blocks. The Vigilance of Hit allows to analyze the variations in the number of correct answers during the test;

3) *Total of Omission Errors* (T.O.E): sum of detection errors of the target stimulus, that is, absence of motor response in the tester's space bar when the target stimulus is presented on the screen;

4) *Vigilance of Omission Errors* (V.O.E): this measure is the standard deviation of the number of omission errors throughout the six blocks, and allows to observe the variations in the number of omissions during the test;

5) *Average Total of Reaction Time* (A.T.R.T.): the sum of detection and reaction time to target stimuli;

6) *Vigilance of the Average Total of Reaction Time* (V.A.T.R.T): the standard deviation of the reaction time, providing information on the cognitive fluctuation throughout the test;

7) *Total of Commission Errors* (T.C.E): sum of pressure responses in the space bar when the target stimulus is not presented;

8) *Vigilance of Commission Errors* (V.C.E): the standard deviation of the number of commission errors throughout the six blocks. It is possible to observe with this score whether the appraiser has a variation in the number of commission errors during the execution of the instrument;

9) *Total of Anticipation* (T.A): the sum of pressure responses in the space bar before 100 milliseconds, while the target stimulus is in evidence;

10) *Vigilance of Anticipation* (V.A): it is the standard deviation of Anticipation in the six blocks. This measure indicates the Anticipation variation throughout the test;

11) *Total of Motor Perseveration* (T.M.P): this measure is the sum of pressure responses in the space bar more than once for the same target stimulus. It indicates the capacity for motor inhibition;

12) *Vigilance of Motor Perseverance* (V.M.P): this measure is calculated using the standard deviation of motor perseveration throughout the six blocks; it allows you to observe the variation in motor perseverance throughout the test.

Structurally, the APT measures two main factors, named continuous attention (C.A) and impulse control (I.C). The first factor is associated with the

dimensions of: stimulus detection (S.D), measured by the measures of successful hits, omissions and their variances (vigilances); and the process of S.A, estimated by the Average Total of Reaction Time and its fluctuation (variance). The I.C factor is related to the dimensions of cognitive impulsivity (C.I), in which its measures are anticipation and its variance; and motor impulsivity (M.I), whose measures are motor perseverance, commission errors and their variances.

The APT has a short duration when compared to other computerized tests that assess S.A. The study by Lin et al. (2018) suggested that the administration of tests that measure S.A usually lasts 10 or more minutes. However, the assessment of S.A can occur in tasks with a minimum duration of 3 minutes. Similarly, Esterman & Rothlein (2019) suggested that it is possible to analyze the support of attention and its fluctuations in 5 minutes of task. Based on these data, the criterion of time of six minutes was adopted for the APT. The adoption of the assessment of S.D and the process of sustaining attention by performance moment to moment, i.e., on a continuous basis, is seen as the gold standard in the discrimination of attentional deficits (Esterman & Rothlein, 2019).

APT follows the perspective of computerized testing of sustained attention, i.e., it is a simple and repetitive task that requires C.A and greater cognitive effort dispatched by the tester for its execution (Langner & Eickhoff, 2013). The basic paradigm is that of continuous detection of stimuli, which allows not only to assess the ability to sustain attention, but also the I.C to the presentation of distracting stimuli that lead the tester to inhibit the automatic response automation of the target stimuli, avoiding commission errors, anticipations and motor perseverance in responding to the instrument's stimuli (Langner & Eickhoff, 2013). For this purpose, the test was divided into six continuous blocks, without pauses, in which the time interval between exposure of stimuli, 1, 2 and 4 seconds, is modified for each block, i.e., every minute of application. The order of the intervals between stimuli and their changes were previously programmed in the instrument with the intention of avoiding response biases related to the interval time between stimuli, requiring the tester's attention to be sustained throughout the application. At the end of the test, the instrument automatically calculates all measurements, generating a performance report.

According to a document created by the American Educational Research Association (AERA), the American Psychological Association (APA) and the National

Council on Measurement in Education – NCME, 2014), an important step in the process of developing a test or instrument is validity. From there, arguments and sources of evidence will be analyzed, allowing for a greater accuracy in the interpretations of the scores obtained. Despite being a unitary process, validity presents different aspects that can be called sources of evidence, which can be obtained by analyzing the relationship between the content promoted by the test and the construct on which it is desired to measure (AERA, APA, & NCME, 2014).

Evidence of validity based on the content of the test refers to the analysis of the themes, format of the items, questions, images, words, commands of the test administration and formats for obtaining the score that make up the instrument. It can be performed through the agreement of expert judgement or through logical and empirical analyzes of the adequacy of the test content with the domains in which it is desired to analyze (AERA, APA, & NCME, 2014). Also, according to the document (AERA, APA, & NCME, 2014), the evidence that will support the appropriation of the content with the desired construct, in this case the cognitive domain of attention, will be related to the inferences that will be made of the scores of the test.

It is also important to highlight that the analysis of the instrument's internal structure can be configured as another means of obtaining evidence of validity, proposing a one-dimensional criterion based on the homogeneity of the test items, which depends on the number of items and how they are related (AERA, APA, & NCME, 2014). One way to perform this procedure is by factor analysis, which is often used in the evaluation and refinement of instruments and can be defined as multivariate techniques that allow to find an underlying structure in a data matrix and present a quantity of factors that best represent a set of observed variables (Damásio, 2012). After the execution of this procedure, it is suggested the analysis of the degree of interrelation between the variables, carried out by the Cronbach's alpha index, allowing the understanding of how reliable the factorial structure found is (Damásio, 2012).

The present study aims to assess whether the APT measures what is proposed, taking appropriate psychometric measures for clinical use, as well as analyzing its factor structure and reliability. The first hypothesis of the study is that the expert judgement will indicate that the test is presented according to the literature and is suitable for use; hypothesis two estimates that the participants in

the pilot test will suggest that the APT presents figures and instructions suitable for understanding and application; hypothesis three postulates that factor analysis will generate four dimensions, associated with two main factors; and hypothesis four expects the test to show internal reliability / consistency within acceptable psychometric standards for an attention assessment instrument.

2. Method

2.1 Participants

2.1.1 Evidence of validity based on content

2.1.1.1 Analysis Process of Expert Judgement

Seven specialist psychologists (four of them had master's degree and three had a postgraduate degree in psychological assessment) participated in the analysis process of expert judgement, in which they assessed whether the test is in accordance with the constructs to be assessed, i.e., C.A (detection of stimuli and S.A) and I.C (C.I and M.I). The inclusion criteria for this study were to present a minimum experience in neuropsychological assessment and testing of attention of 5 years and to be a professor in the area of cognition or neuropsychology. It should be noted that all the experts who participated in this procedure were recruited in the northeast region of Brazil.

2.1.1.2 Semantic Analysis

For the semantic analysis of the instrument, 12 people ($M = 42.91$; $SD = 20.25$) recruited from the Northeast region of Brazil and aged between 18 and 80 years old, participated in the semantic analysis, evaluating whether the test, items, figures and commands were understandable and suitable for adults of different age groups. 5 participants were between 18 and 30 years old, 5 participants were between 31 and 69 years old and 2 participants were over 70 years old. It should be noted that 62.5% of the sample that responded to the semantic analysis was composed of female individuals. The inclusion criterion for participation in the research was to be an adult between the ages of 18 and 95 and have their authorization granted. The exclusion criterion for participating in this procedure was the self-reported presentation of diagnosed neuropsychiatric, motor or visual acuity disorders that prevented the understanding and manipulation of the instrument.

2.1.1.3 Evidence of Validity Based on Internal Structure and Reliability

The analysis of validity evidence based on internal structure and reliability was performed with a sample of 1086 individuals from the Northeast (79% of the sample), North (10% of the sample) and Southeast (11% of the sample) regions of Brazil. The ages of the participants ranged from 18 to 95 years old ($M = 41.77$, $SD = 20.16$) and 61.41% of the sample was formed by female test subjects. The inclusion criterion for participation in the research was to be an adult between the ages of 18 and 95 and have their authorization granted.

2.2 Instrument

The APT is a computerized instrument that presents a daily scene, in this case, the view of a road through the front window of an automobile, simulating an environmental situation to be tested in a clinical or cognitive context. On the computer screen, a road is presented on which 108 target and 36 non-target stimuli are presented (distractors). The target stimuli are composed of figures of living beings, while the non-targets are geometric figures. (Figure 2.2.1). The evaluator must detect the target stimuli (boy, horse, cat, and dog) and press the space bar once, as soon as possible after detecting a target stimulus. On the other hand, when distractors are detected, i.e., geometric figures (circle, triangle), a pressure response should not be emitted to the space bar of the computer. To perform the test, the test subject is instructed to remain attentive to respond when a target stimulus appears again. The instrument assesses the attentional processes of stimuli detection, S.A, cognitive and M.I of people aged 18 to 90 years old and lasts for 6 minutes. The test is divided into six blocks in which the appearance time of the stimuli modulates throughout the application, avoiding bias and automatism in the reaction time.

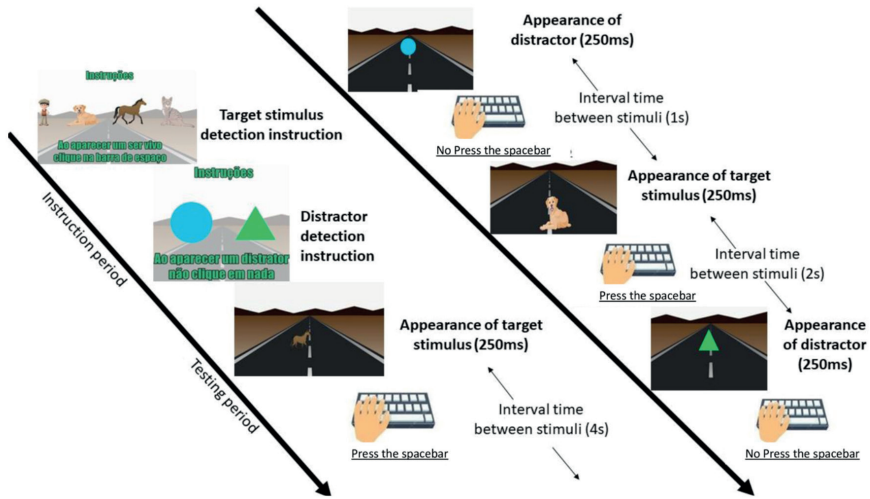


Figure 2.2.1. Stimulus and Application Structure of the Attentional Performance Test.

2.2.1 Data collection procedures

The research participants were invited and tested at universities, nursing homes and at home. In all conditions, it was required to apply the instrument in a room, with a minimum of auditory and visual stimuli around in order to reduce the effect of distractors. To minimize possible interference in the participants' performance, whenever possible, the brightness and temperature of the room were controlled. An anamnesis was performed to obtain general information from the participants and an analysis of the inclusion criteria, age 18 to 95, and the exclusion criteria, the report of the presence of some neuropsychiatric disorder, use of drugs, stimulants and psychiatric medications. Individuals who accepted the invitation to participate signed an informed consent form attesting to the permission to conduct the study. The participant's needs or inconveniences that could be solved by the researcher, such as thirst and / or adjusting the brightness of the computer screen, were investigated before the beginning of testing.

For the expert judgement, participants with expertise in the area of psychological assessment were asked to answer the APT. This procedure was performed with an experimenter inside the room in order to describe the test instructions and consignments. At the end of the application, a questionnaire on

the adequacy of the APT to the literature and to the target audience was given to the experts and also a space for suggesting changes. For this stage, the applicator was removed from the room so that the expert could answer his analysis of the test. The same procedure adopted for the expert judgement was performed for the semantic analysis, with the exception of the questionnaire, which was replaced by a new one aimed at investigating the understanding and adequacy of the figures, execution, instructions and commands of the test to the target audience.

Data collection for confirmatory factor analysis was performed in rooms with minimal distracting stimuli. For this stage, no questionnaire was used. After accepting and signing the informed consent form, the participants received instructions to answer the test. All data collection procedures were carried out with the proper authorization of the Ethics Committee in Psychology of the Institute of Psychology (IPS) of the Federal University of Bahia – UFBA (5686) under the number of opinion with CAE approval: 21717719.6.0000.5686.

2.2.2 Data Analysis Procedures

The software used to perform the statistical analyzes was the R studio version 3.6.1 program for the Macbook Pro® 2010. To calculate the Fleiss Kappa coefficient, the raters package was used, with the concordance command being requested; the Monte Carlo criterion was used, a bootstrap of 1000 and a confidence interval of 0.05 (Falotico & Quatto, 2015). The Landis and Koch (1977) criterion was used for the Kappa index, in which indexes above 0.60 are considered acceptable.

The Kaiser–Meyer–Olkin (KMO) sample fit measure was applied in order to assess the adequacy of factor analysis. The KMO value criterion used was that of the study by Hair, Anderson, & Tatham (1987), which suggests that values above 0.5 are adequate to be factored. Concomitantly, Bartlett’s sphericity test was calculated. The indicators evaluated for internal consistency were standardized by z score, since they present different types of measurement. The value criterion used for Cronbach’s alpha was $\alpha > 0.70$ (Damásio, 2012).

The procedures for the parallel analysis, KMO measurement, Bartlett’s sphericity test and Cronbach’s alpha were performed using the Psych package. Confirmatory factor analysis was performed to test the model according to the number of factors that would be suggested by the parallel analysis. Therefore, the Lavaan packages, the CFA command, and the semPlot were used for the test and

statistical / graphical visualization of the suggested model. The `fa.parallel` command was used to generate the scree plot of the simulated eigenvalues and the database, considering a 95% confidence interval (Damásio, 2012).

3. Results

The expert judgement and semantics were performed to obtain evidence of content validity. The Fleiss Kappa index was used to assess the degree of response agreement between samples of respondents greater than two respondents. The data indicated agreement between the experts for the content ($k = 0.84$; I.C = $0.71 - 0.95$; $p < 0.01$). The data of the target audience participants, in order to analyze the suitability for applying the test, was $k = 0.77$ (I.C = $0.60 - 0.84$, $p < 0.01$). According to the criteria adopted for the evidences of validity based on content (Landis & Koch, 1977), the APT has a strong agreement between the experts who indicated that the test evaluates what is proposed, as well as the semantic analysis indicated that the verbal commands, figures and applicability are suitable for the target audience.

For the factor analysis, the Kaiser–Meyer–Olkin sample fit measure, showed a result of 0.755, while the result of Bartlett's sphericity test was $X^2(66) = 12416.131$; $p < 0.01$. These results indicated that it would be possible to do the factoring of the data (Hair, Anderson, & Tatham, 1987). The parallel analysis performed with the 12 measures of APT suggested the retention of four factors, which were named dimensions. The model of factor structure suggested for the APT was of two main factors that are associated with four dimensions and that branched into 12 measures.

The confirmatory factor analysis showed a good fit of the Root Mean Square Residual (RMSR = 0.08), Akaike Information Criterion (AIC) = 58711.295, and Bayesian Information Criterion (BIC) = 58856.198. As can be seen in Figures 3.1 and 3.2, the analysis of the model suggested two main factors that were named C.A and I.C. There was a covariance of 0.89 between them. The C.A factor was associated with the dimensions S.A (0.93), and S.D (0.70). Similarly, the I.C factor was associated with the M.I (0.35), and C.I (0.59), dimensions.

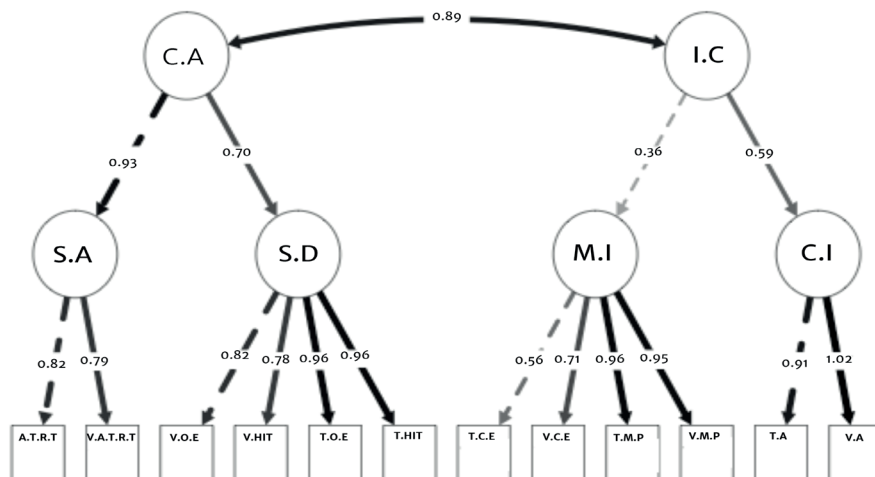


Figure 3.1. Factorial Structure of the Attentional Performance Test.

The first dimension, S.A, was related to the measures A.T.R.T (0.82), and V.A.T.R.T (0.79) of the APT. The S.D was a dimension with the measures of T.Hit (-0.96), V.Hit (0.78), T.O.E (0.96), and V.O.E (0.82). M.I was a dimension composed of the T.C.E (0.56), V.C.E (0.71), T.M.P (0.96), and V.M.P (0.95). The fourth dimension was named C.I and was related to the T.A (0.91), and V.A (1.02) measures.

The total internal consistency of the instrument showed a result of $\alpha = 0.88$. When observing the Cronbach's alpha coefficient specific to the main factors and dimensions, it can be noted that the values ranged from 0.79 to 0.95, which suggests acceptable reliability indexes. The main I.C factor showed $\alpha = 0.81$, while the C.A factor obtained $\alpha = 0.90$. If we take into account the dimensions, it was found that the S.A presented $\alpha = 0.79$; S.D $\alpha = 0.94$; Cognitive impulsivity $\alpha = 0.95$; and M.I $\alpha = 0.88$. Although the instrument has different forms of measurements and measurements, the APT was reliable and consistent, both in specific terms, among its main factors and dimensions, and in general, in the analysis of total consistency.

Figure 3.2. Factor loads of the measures of the APT.

Measures	Factors/Dimensions			
	Continuous Attention		Impulse Control	
	S.A	S.D	M.I	C.I
Average Total Reaction Time (A.T.R.T)	0,93	0.70	0.36	0.59
Vigilance Average Total Reaction Time (V.A.T.R.T)	0.79			
Total Hit (T.Hit)		-0.96		
Vigilance of Hit (V.Hit)		0.78		
Total of Omission Errors (T.O.E)		0.96		
Vigilance of Omission Errors (V.O.E)		0.82		
Total of Commission Errors (T.C.E)			0.56	
Vigilance of Commission Errors (V.C.E)			0.71	
Total of Motor Perseveration (T.M.P)			0.96	
Vigilance of Motor Perseveration (V.M.P)			0.95	
Total of Anticipation (T.A)				0.91
Vigilance of Anticipation (V.A)				1.02

4. Discussion

Validity is an important and methodical process that allows the testing of arguments and making precise inferences about whether the instrument adequately evaluates what it is intended to measure. Despite its unitary character, the validation presents different sources in which evidence is sought to support the statements about the instrument. The data from this study suggest that the expert judgment showed an almost perfect agreement ($k > 0.80$) on the adequacy of the APT, which suggests that in the conception of individuals with expertise in the area of attention testing, the APT measures what it proposes, i.e., the attentional processes of C.A, linked to stimulus detection and the sustained attention, as well as I.C processes, both in impulsivity in the motor and in the anticipation of responses. Another result to be highlighted was the participants' semantic analysis about the suitability of the instrument's figures and commands to the target

audience. The kappa value obtained by this analysis ($k = 0.77$, $p < 0.01$) suggests a strong agreement between the participants, which can be interpreted as an evidence that the items, the understanding and the execution of the instrument are adequate for the adult population of different age groups (Landis & Koch, 1977).

Based on these results, it can be suggested that, according to the expertise of the judges and the semantic analysis with the target audience, that the interpretation of the ATP, its constructs to be evaluated, verbal commands, figures and layout are adequate, which allows to suggest that, based on the content, the APT assesses S.A, simulating an everyday scene. Therefore, the first and second hypotheses of this study had their null hypotheses rejected.

These results indicate that ATP is suitable for the computerized testing of S.A and I.C in adults. In the literature, in general, it is observed that the computerized tests that evaluate S.A usually use measures of response time, commission and omission errors. According to the favorable tests approved to use in testing by SATEPSI (2020), only the Computerized Attention Test, Visual Version, is a computerized instrument that assesses the function of S.A in adults. It has a duration of 15 minutes, evaluates central visual attention, motor impulsivity, visual reaction time and variability in visual reaction time (S.A). In the international literature, there are computerized instruments which assess S.A, such as the Conners Performance Test and the Sustained Attention Task – SART (Langner & Eickhoff, 2013; Fortenbaugh et al., 2017). Although these national and international instruments adopt paradigms and measures similar to the Attention Performance Test, it can be noted that they have an application time of over 10 minutes and do not simulate everyday scenes, not using ecological resources in their testing. The Computerized Digit Vigilance Test (Lin et al., 2018) is a short-term test, i.e., its total application lasts approximately 3 minutes, but it also does not use ecological resources to assess attention. The gradual Conners Performance Test was an instrument developed with a new proposal to evaluate S.A in 8 minutes. In this instrument, the tester must detect the fading of one stimulus image to another, with exogenous clues that signal changes of stimulus from attempt to attempt (Esterman & Rothlein, 2019). Although the application time is reduced in relation to other computerized tests of S.A, its presentation / application structure related to the gradual fading of images followed by replacement for a new figure rarely occurs in the daily lives of individuals.

The APT is a computerized instrument with a short application, of approximately 6 minutes, and that simulates a daily scene, which can reproduce a tester's performance in real day-to-day activity. Confirmatory factor analysis suggested that the APT has two dimensions, cognitive and motor impulsivity, to assess the I.C factor; and its evaluation is not restricted to the measurement of commission errors, i.e., when pressing the space bar at the moment with the non-presentation of a target stimulus. It should be noted that the instrument has two measures of S.A that allow the analysis of this attentional process and its variance throughout the application. The APT also includes the evaluation of the detection and variance of the S.D throughout the application. Studies suggest that these measures may support the analysis of possible attentional deficits in an accurate and specific way (Esterman & Rothlein, 2019; Langner & Eickhoff, 2013; Fortenbaugh et al., 2017).

As for the evidence related to the construct and internal consistency of the APT, it can be noted that the model generated by the confirmatory factor analysis was adequate, in which it can be observed that the APT has two main factors, named C.A and I.C. These main factors are related to four dimensions that are called S.A, S.D, cognitive impulsivity and motor impulsivity.

The instrument's structural model was based on the type of evaluation structure that the test proposes to follow, i.e., the paradigm of continuous stimulus detection (Langner & Eickhoff, 2013). This paradigm is characterized by the large flow of stimuli being presented in short intervals, by the monotony and cognitive effort required in a test of S.A. It involves the automation of responses, which forces the tester to keep an eye on distractors as well and from these to avoid commission errors, anticipations and motor perseverations in responses. These measures are associated with I.C (Langner & Eickhoff, 2013). Concomitantly, the tester must detect the stimuli correctly, avoid omitting the target stimuli and sustain his/her attention, remaining attentive at all times. The total of Hit (correct stimulus), Omission Errors, Average Reaction Time and its vigilances (variances) are the gold standard in the assessment of C.A. In the APT, these measures were grouped into the dimension of C.A. (Esterman & Rothlein, 2019). All measures and dimensions of the instrument related to the named factors are related to criteria mentioned by Langner & Eickhoff (2013), who described the characteristics that involve different tests that evaluate tests of vigilance / S.A.

It was also observed that the adjustment of the model was $RMSR = 0.08$, $AIC = 58711.295$, and $BIC = 58856.198$. For this reason, it can be seen that the evidence of the instrument's internal structure is adequate, reinforcing the statement of rejection of the null hypothesis of the third hypothesis. The data related to internal consistencies, reliability, from the APT suggested that the test is adequately concise, both for main factors and dimensions, as a whole for the construct of S.A and I.C. This statement is possible, since all alpha coefficients showed values above 0.70, according to the criterion suggested in the study by Damásio (2012). From these results, the null hypothesis of the fourth hypothesis of the study can be considered rejected.

In conclusion, with the results obtained by this study, it is stated that the APT presents evidence that this new instrument assesses what is proposed, as well as its items and measures show an appropriate adjustment. However, the need for other sources of evidence of validity is emphasized in order to increase the argumentative power of the instrument's interpretations regarding the assessment of attentional processes (AERA, APA, & NCME, 2014).

For future studies, it is recommended to obtain clinical samples for the analysis of evidence of criterion validity, as well as to carry out the analysis of evidence based on the relationship with other variables, testing by the correlation of the data of the APT with instruments that measure the same construct for convergence analysis and different constructs for discriminant analysis (AERA, APA, & NCME, 2014).

The APT is a computerized instrument and developed in a Brazilian context, which simulates an everyday scene and with a reduced time for evaluating the attentional processes. It is expected, with the other analyzes of validity sources, to obtain additional evidence about the constructs evaluated by the instrument, as well as to suggest its applicability to the Brazilian adult population, favoring the investigation of attentional cognitive processes and use in different contexts of assessment.

References

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

- Berger, I., Slobodin, O., & Cassuto, H. (2017). Usefulness and validity of continuous performance tests in the diagnosis of attention-deficit hyperactivity disorder children. *Archives of Clinical Neuropsychology*, 32(1), 81–93. doi:10.1093/arclin/acw101
- Canini, M., Battista, P., Della Rosa, P. A., Catricalà, E., Salvatore, C., Gilardi, M. C., & Castiglioni, I. (2014). Computerized neuropsychological assessment in aging: Testing efficacy and clinical ecology of different interfaces. *Computational and Mathematical Methods in Medicine*, 2014, 804723. doi:10.1155/2014/804723
- Cannavò, R., Conti, D., & Di Nuovo, A. (2016). Computer-aided assessment of aviation pilots attention: Design of an integrated test and its empirical validation. *Applied Computing and Informatics*, 12(1), 16–26. doi:10.1016/j.aci.2015.05.002
- Damásio, B. F. (2012). Uso da análise fatorial exploratória em psicologia. *Avaliação Psicológica*, 11(2), 213–228.
- Esterman, M., & Rothlein, D. (2019). Models of sustained attention. *Current Opinion in Psychology*, 29, 174–180. doi:10.1016/j.copsyc.2019.03.005
- Falotico, R., & Quatto, P. (2015). Fleiss' Kappa statistic without paradoxes. *Quality & Quantity*, 49(2), 463–470. doi:10.1007/s11135-014-0003-1
- Fortenbaugh, F. C., DeGutis, J., & Esterman, M. (2017). Recent theoretical, neural, and clinical advances in sustained attention research. *Annals of the New York Academy of Sciences*, 1396(1), 70–91. doi:10.1111/nyas.13318
- Geva, R., Zivan, M., Warsha, A., & Olchik, D. (2013). Alerting, orienting or executive attention networks: Differential patterns of pupil dilations. *Frontiers in Behavioral Neuroscience*, 7, 145. doi:10.3389/fnbeh.2013.00145
- Gilsoul, J., Simon, J., Hogge, M., & Collette, F. (2018). Do attentional capacities and processing speed mediate the effect of age on executive functioning? *Aging, Neuropsychology, and Cognition*, 26(2), 1–36. doi:10.1080/13825585.2018.1432746
- Gvirts, H. Z., Braw, Y., Harari, H., Lozin, M., Bloch, Y., Fefer, K., & Levkovitz, Y. (2015). Executive dysfunction in bipolar disorder and borderline personality disorder. *European Psychiatry*, 30(8), 959–964. doi:10.1016/j.eurpsy.2014.12.009
- Hair, J., Anderson, R. O., & Tatham, R. (1987). *Multidimensional data analysis*. New York: Macmillan.
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33, 159–174. doi:10.2307/2529310

- Langner, R., & Eickhoff, S. B. (2013). Sustaining attention to simple tasks: A meta-analytic review of the neural mechanisms of vigilant attention. *Psychological Bulletin*, 139(4), 870–900. doi:10.1037/a0030694
- Lin, G.-H., Wu, C.-T., Huang, Y.-J., Lin, P., Chou, C.-Y., Lee, S.-C., & Hsieh, C.-L. (2018). A reliable and valid assessment of sustained attention for patients with schizophrenia: The Computerized Digit Vigilance Test. *Archives of Clinical Neuropsychology*, 33(2), 227–237. doi:10.1093/arclin/acx064
- Petersen, S. E., & Posner, M. I. (2012). The attention system of the human brain: 20 years after. *Annual Review of Neuroscience*, 35(1), 73–89. doi:10.1146/annurev-neuro-062111-150525
- Sistema de Avaliação de Testes Psicológicos (2020). Lista completa dos testes. Retrieved from https://satepsi.cfp.org.br/lista_teste_completa.cfm
- Thomson, D. R., Besner, D., & Smilek, D. (2015). A resource-control account of sustained attention: Evidence from mind-wandering and vigilance paradigms. *Perspectives on Psychological Science*, 10(1), 82–96. doi:10.1177/1745691614556681
- Van Steenbergen, H., & Band, G. P. (2013). Pupil dilation in the Simon task as a marker of conflict processing. *Frontiers in Human Neuroscience*, 7, 215. doi:10.3389/fnhum.2013.00215
- Van Zomerén, A. H., & Brouwer, W. H. (1994). *Clinical neuropsychology of attention*. Oxford, UK: Oxford University Press.

Authors' notes

Jonatas R. Bessa, Institute of Psychology, Federal University of Bahia (Ufba); **Neander Abreu**, Institute of Psychology, Federal University of Bahia (Ufba); **Yuri Santana**, Institute of Psychology, Federal University of Bahia (Ufba); **Roberto Beirão**, Institute of Psychology, Federal University of Bahia (Ufba); **Jamine Cairo**, Institute of Psychology, Federal University of Bahia (Ufba).

The authors would like to thank Fundação de Amparo à Pesquisa do Estado da Bahia (FAPESB) for the financial support in the form of a master's Scholarship to Jonatas Reis Bessa (number BOL0525 / 2018), which allowed the execution of this study.

Correspondence concerning this article should be addressed to Neander Abreu, Psychology Department, Federal University of Bahia, Salvador, BA, Brazil. Postal Code: 40170-010.

E-mail: neandersa@hotmail.com

EDITORIAL BOARD

Editor-in-chief

Ana Alexandra Caldas Osório

Section editors

Psychological Evaluation

Alexandre Serpa

Luiz Renato Rodrigues Carreiro

Vera Lúcia Esteves Mateus

Psychology and Education

Cristiane Silvestre de Paula

Carlo Schmidt

Social Psychology

Bruna Suguagy do Amaral Dantas

Enzo Banti Bissoli

Clinical Psychology

Eduardo Fraga Almeida Prado

Marina Monzani da Rocha

Carolina Andrea Ziebold Jorquera

Human Development

Maria Cristina Triguero Veloz Teixeira

Rosane Lowenthal

Technical support

Letícia Martinez

Camila Fragoso Ribeiro

EDITORIAL PRODUCTION

Editorial coordination

Ana Cláudia de Mauro

Editorial intern

Júlia Lins Reis

Language editor

Daniel Leão

Layout

Acqua Estúdio Gráfico