UC/FPCE 2018

Universidade de Coimbra
Faculdade de Psicologia e de Ciências da Educação

Toulouse-Piéron Cancellation Test: Validation and normalization studies in cognitive decline

Marisa Pedroso de Lima Marta Neves
(e-mail: marisalima5@hotmail.com)

Supervisors:

Professor Doutor Mário R. Simões (FPCE-UC, CINEICC, PsyAssessmentLab)
Professora Doutora Isabel Santana (FM-UC, CHUC)
Professora Doutora Sandra Freitas (CINEICC, PsyAssessmentLab)

Dissertação de Mestrado em Psicologia, área de especialização em Psicologia Clínica e da Saúde, subárea de especialização em Psicogerontologia Clínica.

This project was supported by BIAL Foundation project 2014/495

Coimbra, 2018

Título da dissertação – *Teste da Barragem de Toulouse-Piéron: Estudos de validação e normalização no declínio cognitivo*

RESUMO

INTRODUÇÃO: O teste da Barragem de Toulouse-Piéron (TP) é um instrumento psicométrico clássico que avalia a atenção seletiva e sustentada, e a velocidade de processamento, sendo igualmente sensível à resistência à fadiga. O teste apresenta três índices principais: Rendimento de trabalho (RT), índice de dispersão (ID) e resultado total (T).

OBJECTIVOS: Estudo de validação psicométrica e clínica com pacientes no espectro da Doença de Alzheimer (DA), incluindo um grupo com Défice Cognitivo Ligeiro (DCL) e um grupo com DA ligeira. Exploração das propriedades psicométricas do teste, estabelecimento de pontos de corte e de dados normativos para a população portuguesa.

METODOLOGIA: A amostra do Estudo 1 - validação do TP para DCL e DA, é composta por 120 participantes (Grupo Controlo: *n*=40, Grupo DCL: *n*=40; Grupo DA: *n*=40). Os grupos clínicos cumprem os respetivos critérios de diagnóstico internacionais estandardizados. O grupo de controlo (Estudo 1) foi retirado do Estudo 2 (*n*=234) (elaboração de dados normativos preliminares do TP para a população portuguesa), e é constituído por participantes cognitivamente saudáveis residentes na comunidade.

RESULTADOS: A amostra do Estudo 1 apresentou uma média de idades de 71.44 ± 9.53 anos e de 6.66 ± 4.31 anos de escolaridade, sendo 61.7% dos sujeitos do sexo feminino. A pontuação total nos índices do TP difere significativamente entre os três grupos (p<.001: Controlo>DCL>DA) quando controlado o efeito das covariáveis (idade e escolaridade). O tamanho do efeito foi maior para o grupo DA em todos os índices do TP comparativamente ao grupo DCL, isto é: η^2 =.54 (TP-RT), η^2=.26 (TP-ID) e η^2=.20 (TP-T), no grupo com DA. No grupo com DCL: η^2 =.17 (TP-RT), η^2=.05 (TP-ID) e η^2=.18 (TP-T). O TP apresentou boa acuidade diagnóstica para o grupo de DA: para o RT o ponto de corte foi <49 pontos (AUC=0.981), para o ID >26 pontos (AUC=0.921) e para o Total foi de <4 pontos (AUC=0.977).

CONCLUSÕES: Os resultados sugerem que o TP é um instrumento psicométrico sensível à presença de défices atencionais nos desempenhos de pacientes com DCL e DA. Os pontos de corte e normas calculadas são de grande utilidade para uso na prática clínica e na investigação.

Palavras chave: atenção sustentada, teste de cancelamento de Toulouse e Piéron, testes neuropsicológicos, Defeito Cognitivo Ligeiro, Doença de Alzheimer.

Title of dissertation – *Toulouse-Piéron Cancellation Test: Validation and Normalization studies in Cognitive Decline*

ABSTRACT

INTRODUCTION: The Toulouse-Piéron Cancellation Test (TP) is a psychometric tool that assesses selective and sustained attention, as well as processing speed and fatigue resistance. It presents three main indexes: Work Efficiency (WE), Dispersion Index (DI) and Total Result (TR).

OBJECTIVES: Clinical validation and psychometric study with patients in the spectrum of Alzheimer's Disease (AD), including a group with Mild Cognitive Impairment (MCI) and a group with mild AD. Exploratory analysis of its psychometric properties, establishment of discriminant cut-off points and of normative data for the Portuguese population.

METHODOLOGY: Study 1 (Validation of the TP for MCI and AD) sample's is composed by 120 subjects (Control Group (CG): n=40; MCI group: n=40; AD group: n=40). The clinical group fulfil the standard international diagnostic criteria. The CG (Study 1) was selected from the Study 2 (n=234) (preliminary normative data of the TP for the Portuguese population) and is constituted by cognitively healthy community-dwelling subjects.

RESULTS: The mean age of the sample was 71.44 ± 9.53 years, the education mean was 6.66 ± 4.31years, and the sample was composed by 61.7% females. The TP total scores significantly differ between three groups (p<.001: Control>MCI>AD), controlling the covariables effect (age and education). The effect sizes were bigger for the AD group comparatively to the MCI's: η^2 =.54 (TP-WE), η^2=.26 (TP-DI) e η^2=.20 (TP-TR), for the AD group. For the MCI group, η^2 =.17 (TP-WE), η^2=.05 (TP-DI) e η^2=.18 (TP-TR). The TP presented good diagnostic accuracy for AD group. The cut-off point for Work Efficiency (WE) was < 49 points (AUC=0.981), for Dispersion Index (DI) was >26 points (AUC=0.921) and for TR was <4 points (AUC=0.977).

CONCLUSIONS: The results showed the sensitivity of the TP to detect attentional deficits in MCI and AD patients, disclosing relevant cut-off points and normative values for use in clinical and research contexts.

Key Words: sustained attention, Toulouse-Piéron Cancellation Test neuropsychological tests, Mild Cognitive Impairment, Alzheimer's disease.

Agradecimentos

Em primeiro lugar, agradecer à Professora Doutora Isabel Santana por toda a sabedoria e conhecimento transmitidos e por todas as oportunidades que me deu para crescer enquanto pessoa e enquanto profissional. Por todas as qualidades quer humanas quer profissionais e pela prontidão em auxiliar nas minhas dúvidas. Sem dúvida um exemplo de rigor e mestria. Um orgulho imenso ter trabalhado a seu lado.

Ao Professor Doutor Mário Simões por todos os ensinamentos, pela disponibilidade e prontidão para me ajudar e por todo o saber sempre disponível para partilhar. Por sempre ter mostrado confiança no meu trabalho e por sempre nos motivar a trabalhar mais e melhor. Por todos os valores éticos e profissionais que nos transmitiu desde a primeira aula. Um modelo a seguir.

À Professora Doutora Sandra Freitas, por todo o conhecimento partilhado nas aulas (e fora delas), por me ter feito ter a certeza de que o Serviço de Neurologia seria a escolha certa para mim, por todas as dúvidas que me esclareceu e acima de tudo por sempre ter estado disponível para me ouvir e aconselhar. Pelas sábias e carinhosas palavras de conforto e de incentivo e pela confiança no meu trabalho. Muito obrigada.

À Dra. Diana Duro, as palavras nunca serão suficientes para expresar a minha gratidão e admiração. Pela disponibilidade e preocupação, pela ajuda, pela motivação e pelo gosto pela Neuropsicologia. Por todo o tempo dispendido comigo e por tudo o que me ensinou. Por sempre confiar nas minhas capacidades e me apoiar. Todos deveriam ter o privilégio de trabalhar e aprender a seu lado.

À Professora Doutora Manuela Vilar, pelo apoio e incentivo que lhe são característicos e todas as horas que, desde o segundo ano deste ciclo que agora cessa, a vi dispender em prol quer da minha, quer da aprendizagem dos meus colegas. Por todas as dúvidas que me esclareceu fora de horas e por ter mostrado sempre interesse no meu trabalho. Por ter dispertado em mim o "bichinho" da avaliação psicológica.

Ao Profesor Valentim Alferes, pelas largas horas que passou tentando explicar-nos a importancia da componente de investigação bem como da sua rigorosa metodologia.

À Professora Mª do Rosário Pinheiro, uma amiga, quase "família", por tudo o apoio e suporte que me deu desde o inicio ao fim deste ciclo.

A todos os profissionais de excelência com quem tive o privilégio de privar durante o último ano, pela forma como me acolheram e confiaram no meu trabalho.

A todos os pacientes da Consulta de Demência do Serviço de Neurologia do CHUC, bem como aos seus familiares e acompanhantes. Um profundo agradecimento por me terem ajudado a crescer enquanto pessoa e enquanto profissional.

À minha colega Bianca por tudo o que este ano vivemos juntas, todas as oportunidades que nos foram dadas e por nos ajudarmos e apoiarmos mutuamente. Pela partilha conjunta de todos os momentos com que este ciclo nos premiou.

Aos meus avós, Joaquim e Benvinda, por todos os esforços que fizeram por mim durante toda a vida, por sempre acreditarem em mim e por todo o amor que sempre me deram, a vocês dedico este trabalho. Aos meus pais, nunca conseguirei agradecer-vos o suficiente. Obrigada pelo apoio, pela confiança, pelos esforços e sacrifícios que fizeram para eu poder chegar até aqui. Obrigada pela preocupação e pelo interesse e por nunca duvidarem de mim. Obrigada por todo o vosso amor.

À tia Bela e ao tio Fernando, e aos “avós” Guida e Jojó, por todo o carinho e amor que sempre me deram, pela preocupação constante. À tia Elsa e ao tio Pedro por estarem sempre do meu lado e me apoiarem.

A todos os meus primos, por termos crescido tanto, juntos, e por sempre estarmos ao lado uns dos outros. Ao Francisco, ao Gonçalo e ao Rodrigo, o meu profundo agradecimento. À avó Isabel, por todos os ensinamentos que tão bem me deu, e por fazer com que a nossa familia se mantenha unida.

Ao Dr. Vitor e à D. Neli, por estarem presentes em todos os momentos da minha vida, por sempre acreditarem em mim e por todo o carinho com que sempre me receberam. Seremos sempre familia. À Inês, por te orgulhares de mim e de me dares força, retribuo em dobro todo o carinho. À Ana, à Xana e aos meninos.

Às minhas “madrinhas” de curso, Sousa, Aguiar, ‘Tina e Mariana, por me terem acolhido desde o primeiro dia e me terem ajudado tanto ao longo destes cinco anos. Obrigada por serem incríveis.

À Carolina, por sempre estar ao meu lado, e por caminharmos juntas, passo a passo nesta caminhada para o sucesso. À Maria, por sermos tão parecidas e sempre nos termos apoiado. À Magda e à Márcia, por tudo o que vivemos juntas, e por tudo o que fizeram por mim.

À Mónica, amiga “de sempre” e de todas as horas, pela amizade e pela cumplicidade. Por todos os momentos que passamos juntas e todas as histórias que guardarei sempre com muito carinho.

À Ana Rita, o meu “tesourinho alentejano”, por me fazer viver a tradição de Coimbra com tanto amor. Serás sempre o orgulho da ‘madrinha.

À Inês Costa e à Inês Figueiredo, pelo “trio maravilha”, pelo suporte, pela amizade e pela partilha.

Aos “Arruínados”, os do costume, os de sempre, os que nunca me falham. Obrigada pela amizade, pelas histórias, por todos os momentos que vivemos juntos. Coimbra sem vocês não seria a mesma. Ao David Brás e ao David Valença, pelos 10 anos de suporte e partilha, por sempre me apoiarem e acreditarem em mim. Obrigada por tudo.

À Tânia, amiga de sempre, os anos não passam por nós. Obrigada pela presença em todos os momentos e pelo carinho e conforto de há tantos anos.

À Rute, à Mafalda, à Marina e à Catarina. Não há palavras para vocês. Pela amizade, cumplicidade, diversão, união e confiança.

Ao Rogério, amigo, companheiro, amor. Por todas as horas que passaste ao meu lado durante estes cinco anos. Por nunca desistires de mim, me dares confiança e força. Por estares sempre do meu lado, em todos os momentos. Pelo carinho e pela cumplicidade. Pelo vagar, tempo e disposição. Pelo apoio incondicional. Por me ajudares a crescer todos os dias. Pelo respeito e sobretudo pela sempre presença.

Content

Introduction

The current increasing of life expectancy composes a relevant demographic change in the last few decades in the developed countries. Those changes provide a worrying trend to an ageing population. Alongside, age is a main risk factor for the development of cognitive impairment and dementia and particularly of Alzheimer's disease.

One such challenge is to validate and normalize neuropsychological assessment instruments that allow to evaluate and monitor the evolution of the ageing process as well as of cognitive decline and dementia, and consequently ensure effective and timely prevention and treatment. So, the main goal of the present thesis is the clinical validation and normalization of the Toulouse-Piéron Cancellation Test for the Portuguese population. This is of prime importance since there are very few studies concerning sustained attention abilities in the elderly.

We will start by making a survey of what has been the development of the study of attention over the time, having in consideration the several processes that integrate it, as well as reflecting how an attentional deficit can occur in the context of a neurodegenerative disease (e.g. Alzheimer's disease), as well as how the presence of other factors (e.g. depressive symptomatology) will affect attentional processes. We will then describe the procedures and materials that were used to validate and normalize this psychometric test. This section will be followed by the Results section, where we will outline the statistical analysis used and the results obtained. After that, in the Discussion section we will study and analyse the results in the context of the current literature. In the end, we will discuss the strenghts and limitations of the present work, and summary propose a set of future research that is essential to be developed.

I - Background

The human brain presents intrinsic limitations related to the extent of information that can be processed in a certain period of time (Banich, 2004). According to the amount of data that it receives the concept of attention plays an essential role and encompasses multiple complementary definitions from different areas of psychology. Since the first studies of William James, Wilhelm Wundt, Edward Titchner and W. B. Pillsburg in the nineteenth century, attention is considered a main cognitive function. At the time, William James, the pioneering investigator, pointed out that attention made humans perceive, discriminate and recall more efficiently (Cohen, 2014). After a period of several investigations in this area (combining experimental techniques, psychopsysiological speculation and self-observation) a lot of progresses were made in the characterization of the attentional phenomena and a period of thirty to fifty years emerged where the study of this concept disappeared from psychological research – it was the age of behaviourism. It was only in the middle of the twentieth century, with the emergence of cognitive psychology that attention regained a central position.

After that, a lot of research has been developed concerning its relevance. For Strauss, Sherman & Spreen (2006) the majority of developed models come up with a multifaceted system that permits the subjects to select important information in place of non-relevant material, such as the Broadbent's filter model, the Treisman's model of attenuation and the Deutsch & Deutsch's pertinence model (Strauss et al., 2006).

> "Attention can be conceptualized as comprising several basic processes. These include sensory selection (filtering, focusing, automatic shifting), response selection (response intention, initiation and inhibition, active switching, and executive supervisory control), attentional capacity (structural and energetic capacity, arousal, effort), and sustained performance (fatigability, vigilance)" (Strauss et al., 2006, p. 546).

According to Benczik and Casella (2007, p.31) it implies the "privileged allocation of resources to relevant events", what generically means that it is characterized by the extent of neural mechanisms that ensure the selection of some stimuli instead of others, as well as the more adequate behaviours to a specific situation. Attention was also defined as a basic and organized mental process that helps the brain to receive, process and respond to internal and external stimuli as well as the global input of information that enters in the brain in a moment (Cohen, 2011). More recently, Fuster (2017, p.7) defined it as the "optimal and selective allocation of limited resources to information processing in the central nervous system that operates in the two major sectors of neurocognitive function: perceptual and executive".

In 2012, Lezak, Howieson, Bigler and Tranel affirmed that attention involves different features, suggesting that many of them are overlapping with each other and with other cognitive domains.

In addition to the multiple definitions of attention, there is also a lack of consensus regarding the subdivision of its different parameters that are commonly divided into spontaneous (involuntary) and controlled (voluntary). In a first moment, individuals are not wittingly involved in the attentional process and there is a quick issuing of unintentional responses (Montiel & Capovilla, 2007). In a second moment, there is a deliberated effort to pay attention to what is going on, being a slowest and more sequenced process (Fuster, 2017).

Despite the numerous definitions and subdivisions of attention, this cognitive ability is commonly divided in sub-domains according to the type of processing that is involved. There are four main sub-domains that are **alertness** (or **arousal)**, (**concentrated** or) **selective attention**, **divided attention** and (**vigilance** or) **sustained attention**. Despite the controversy regarding the lack of agreement on the precise meaning of each one of these terms, some of them end up overlapping. There is also a problem with the fact that the majority of attention tests can not singly evaluate a specific domain of attention. Also, there are many attention tests which require other cognitive abilities such as motor speed, verbal responding or information processing speed (Strauss et al., 2006). For better understanding attention as a

multifaceted system, its four main above-mentioned sub-domains will be described bellow.

Arousal or **alertness** is the propensity to process all the information from external stimuli (Wilding, Munir, & Cornish, 2001). It can be divided in two different modes of function: phasic – defined as the quick organization of inputs to process an expected stimulus; and tonic (or intrinsic) – which is conceptualized as the minimal change that occurs in the state of alert and that is not related with external demands (Alberto, 2003; Weinbach & Henik, 2012).

Selective (or concentrated) attention is the ability of selecting an information instead of another, based in the exclusion of non-relevant distractors. This field of attention allows to respond adequately to one target at a time. One possible explaination for this is the restricted capacity of the human brain, which cannot process all the existing information simultaneously, besides everyday practice shows that subjects are capable of doing more than one task simultaneously (McCallum, 2015).

Divided attention is a kind of shared selective attention that allows the subjects to process different sources of information and successfully perform several tasks at the same time (Levitt & Johnstone, 2001; Herff & Czernochowski, 2017).

Finally, **sustained attention (or vigilance)** is a sub-domain widely dependent of mental control and working memory. It is also related with the ability to maintain attention to the same stimuli for long periods of time, executing a tiring and monotonous task. Long, uninteresting, and for the most part uneventful tasks origine weaker results concerning either speed and accuracy in perceiving hoped-for events. In opposite, if the task is stimulating the subject will be better capable of sustaining attention and maintaining performance (Montiel et al., 2007; McCallum, 2015). This last component of attention is of extreme importance since several attentional disorders such as Attention-Deficit Hyperactivity Disorder or Alzheimer's disease (AD) only manifest themselves after longer testing periods (Benczik et al., 2007; Saunders & Summers, 2011).

So, since the ninetheen century until today, attention has been defined by several authors from different areas of psychology. From William James to Fuster, all of them aimed to attain a more complete and useful definition of the concept due to its main relevance as a primary cognitive function.

The relevance of studying attention

Attention is one of the basic processes of neurocognitive functioning and its complexity demands protracted research. Attentional deficits have a personal and social impact that justifies the development of specific and accurate evaluation tests allowing the assessment of its different sub-domains (Alberto, 2003; Muller, Rothermund & Wentura, 2016). Impaired attention can make it difficult to process information at various levels and explain some performance handicaps in aging and dementia. In patients with mild AD attentional deficits also have a negative impact in other domains such as

memory and executive functions (Rizzo, Anderson, Dawson, Myers, & Ball, 2000).

The study of attention is also very important insofar as it is characterized by an inherent process in all cognitive activity. The deficits in this area lead to changes and influence the results obtained in the neuropsychological evaluations with direct implications in all the tasks. It is thus extremely important to characterize the different attentional deficits since the losses obtained in the tests can be confounded with hyphotetically existent deficits in other cognitive domains. It is needed to be sure that the results in the tests are not being masked by attentional deficits. So, attention is an abilility to direct behaviour to the spatial and temporal charactheristics of the situation (Mesulam, 1985; Cohen, 2014; Chauvin, Gillebert, Rohenkohl, Humphreys, & Nobre, 2016). Moreover, is also important to know which substrates are under each one of the attentional processes for futher understanding both the location and the implications of the deficits found in the neuropsychological tests.

Neurological substrates of attention

The modulation of attention is done by networks related to arousal and located in the brainstem that projects to the thalamus and neocortex where the most complex features are processed (Mesulam, 1985).

Petersen and Posner (2012) updated their research of 1990, highlighting the three main attentional components of the attentional system of the human brain, responsible by stimulus recognition and response initiation mechanisms: alerting, orienting and executive attention. According to their own features these networks are divergent on neuroanatomical substrates. The first one, **alerting**, is related to frontal and parietal dorsolateral areas and with thalamic activities and is defined by the maintenance of a state of alertness (sustained vigilance). The second one, **orienting**, is linked to superior and inferior areas of the parietal cortex as well as frontal eye fields, superior colliculus, pulvinar and reticular thalamic nuclei and is conceptualized as the allocation of the attentional mechanisms to the inputs and new stimuli. The last one, **executive attention**, is related to areas such as the anterior cingulate cortex, medial frontal cortex and lateral prefrontal cortex, responsible for the recognition of struggles between different brain regions while responding accurately to the *in vivo* situations (Petersen et al., 2012).

Executive components are considered the highest levels of cognition and include dimentions such as planning and sequential organization (known as main cognitive functions, in which attention is frequently included). These domains seem to be impaired early in AD and those deficits are considered a part of frontal lobe lesions in addition to difficulties in judgment and social cognition (Kirova, Bays, & Lagalwar, 2015).

Similarly, to what occurs with executive functions, attention and working memory are inter-related, due to the fact that information is stored in working memory through selective attention. At the same time, working memory has been defined as a system involved in processing and storage of information for short periods of time. Besides that, it is responsible by a set of

different tasks like reasoning, retention and reading. This complex system can be divided into four main components: phonological loop, visuospatial sketchpad, episodic buffer and central executive. The central executive is a common component to both working memory and attention, due to its function of coordinating the use of the limited capacity of processing that seems to be common to both the systems. There still remaining some discussions related to the presence of the same substrate for central executive and executive attentional system, especially in view of the similarities and relations between this two components and executive functions (Baddeley, 2012; Karbach & Verhaeghen, 2014).

Neuropsychological instruments for attention evaluation

Cancellation tests

Cancellation tests are traditionally used in neuropsychological assessment of attention and processing speed and their main components are based on target marking. The stimuli may differ according with the test requirements and with the dimension that is being evaluated. So, the targets may be language-based (letters or numbers), auditory (tones), spatial (locations) or visual (squares or circles) types. Thus, the answers can be required in a different modality of the one in presentation. An example of the tasks required is to cut a specific target, usually a letter or a symbol, selecting it from a number of different stimuli (Alberto, 2003; Lunven & Bartolomeo, 2017). The most known cancellation tests are the Toulouse-Pieron Cancellation Test (TP) (do Amaral, 1967), the D2 Test of Attention (Brickenkamp, 1981; 2007) and the Bells Test (Gauthier, Dehaut & Joanette, 1989). This kind of psychological tests are mostly used to evaluate selective and sustained attention. They are also used to detect deficits in visual search in individuals that present neurological problems such as unilateral neglect. If there is no indication to follow a certain order, subjects typically use organized strategies such as left-to-right or top-to-bottom search patterns. Is also frequent that older adults take additional time to execute the required tasks when compared to younger adults (Dalmaijer, Van der Stigchel, Nijboer, Cornelissen & Hussein, 2015). Cancellation tests yet allow the extraction of some indexes like hits, errors and omissions.

Toulouse-Piéron Cancellation Test

Toulouse-Piéron Cancellation Test (TP; do Amaral, 1967; see also Simões et al., 2016), was first developed in 1904 by Édouard Toulouse and is the most known and used psichometric test that evaluates perceptual and attentional skills.

In Portugal, according with a recent study (Almeida, Simões, Almiro, & Seabra-Santos, 2018), the TP is the twenty-third most used test and the 3rd most used neuropsychological test after the Rey Complex Figure Test (Rey, 1959, 1998; Espírito-Santo et al., 2015; see also Simões et al., 2016) and the D2 Test of Attention (Brickencamp, 2007). It is a paper-and-pencil test and encompasses visuo-perceptual and graph abilities. It also demands high concentration levels and fatigue resistance and was one of the pioneering

instruments developed to assess sustained attention and processing speed. It consists of a blank sheet of paper with twenty-five lines and forty small squares per line. The squares are distinguished from each other by the orientation of the rows on the outer surface: in each square the stroke is oriented in eight possible directions, similarly to the wind rose. So, there are eight types of squares being that in each line are randomly arranged five of each of these types and the subject is able to mark them according to the three proposed models in the header. Another important aspect is the application time (10 minutes) and the setting of hits, errors and omissions (do Amaral, 1967). The TP presents two main outcomes and the fisrt one is the Work Efficiency (WE) (Baeta, 2002). This result sets up a measure of both the attentional and perceptual abilities of the subject. It is related to the total score of hits (H), errors (E) and omissions (O) and it is calculated by: $WE = H -$ (E+O). It is a measure of the subject's work quality. The second outcome is the Dispersion Index (DI) (Baeta, 2002), that allows to understand if the result obtained on the WE was influenced by a response pattern of impulsivity or innatention. This index is the percentage of errors and omissions divided by the number of hits obtained by the subject during the test and can be calculated by: DI = $\frac{\text{E+O}}{\text{H}}$ x 100. Beyond the indexes described above, there are three more relevant scores that will be described below. The first one is **Hits**, which are the number of items correctly selected by the subject. The second one is **Errors** that shows the number of items wrongly selected (false positives) by the subjects and the last one is **Omissions**, which are the number of correctly items that the subject had not selected (false negatives). According to do Amaral (1967), there is another index designated by **Total Result** (TR) that allows not only to evaluate the total hits per minute but also establishing a different evaluation of errors and omissions, through the discount of two points for each of the first ones and one point for each of the second ones, so TR is a measure of how much the subject effectively produces per minute. The TR can be calculated by TR = $\frac{H-(O+E\text{x}2+1)}{10}$.

In addition to the above described indexes, this test allows to elaborate the curve of mental effort (Thorndike, 1912), for assessing intraindividual variability in sustained attention. He found three main sources of performance fluctuations over time: accumulating fatigue, effort variations and practice effects. The curve is constructed by the cross of number of hits (obtained in each one of the 10 minutes testing) with each one of those minutes. In this curve each of the minutes of the test is inserted in the axis of the abscissas while the number of hits that corresponds to it is inserted in the axis of the ordinates thus forming the final curve. This provides information related to the variation of the performance of the subjects across time and also related with attentional changes.

It is also known that weak performances in the TP-WE may be interpreted as inattention and global response slowing. On the other hand, a high TP-DI can be explained by distractibility or impulsivity. When there is a disturbance in the flexibility, difficulty in information processing or motor integration, the two indexes are usually affected (Baeta, 2002). Better results in the TP are characterized by high puctuations in both TP-WE and TP-TR

together with lower values of TP-DI. A brief summary of the TP's description will be presented in section II – materials.

Other attentional measures

Currently, there are many psychological tests to evaluate attentional sub-domains such as the Trail Making Test (A and B) and the D2 Test of Attention (that will be described in section II – materials).

Besides these two measures there are some other relevant tests that should be mentioned. Some of them are part of the Wechsler Memory Scale (WMS; Wechsler, 2008) such as mental control, spatial location, sequence of numbers and letters and digit span. Further, the Digit Symbol test is also useful to assess processing speed, attention and working memory (Kaplan & Fein, 2004). Other relevant attentional measures are the Continuous Performance Test (CPT) (Shalev, Ben-Simon, Mevorach, Cohen, & Tsal, 2011), the Paced Auditory Serial Addition Test (PASAT) (Strauss et al., 2006) and the Test of Everyday Attention (TEA) (Robertson, Ward, Ridgeway & Nimmo-Smith, 1994; 1996; Strauss et al., 2006).

Age related disorders with a compromise of attentional processes

Neurodegenerative disorders

Alzheimer's Disease

AD is the most common cause of dementia being responsible for at least 60% of the cases (Lezak et al., 2012). It is the most common neurodegenerative disease, affecting 5 to 7% of people over the age of sixty (Prince et al., 2013; Santana & Duro, 2014). It is characterized by an insidious onset of memory deterioration and progressive cognitive decline together with functional impairment and the emergence of neuropsychiatric symptoms. There is an early deterioration of the episodic memory that is associated with the cortical atrophy that characterizes AD, particularly on the hippocampus and adjacent regions of the temporal lobe (Cunha, Guerreiro, Mendonça, Oliveira, & Santana, 2012). As the disease progresses, prefrontal and parietal regions also become involved and the cognitive and functional compromise worsens (Hannay, Howieson, Loring, Fischer, & Lezak *cit in* Lezak et al., 2012).

Attentional deficits can make it difficult to process information at various levels, and also explain possible functional impairments in both normal aging and in the presence of dementia (Langa & Levine, 2014). Similarly, attention is considered to be a cognitive domain specifically damaged in the presence of neurodegenerative conditions such as MCI and AD what shows the importance of its evaluation when in the presence of a neurodegenerative condition (Saunders et al., 2011). In these pathologies, attentional deficits have been identified as one of the neuropsychological features (even in the earliest stages - mild AD) (Kaiser, Kuhlmann & Bosnjak, 2018). At this point, it is important to enhance the relevance of evaluating attention habilities through well validated psychological tests. The TP allows to assess two main sub-domains that are frequently early impaired in AD: sustained and selective attention as well as was showed by Kaiser et al. (2018).

They concluded that patients with AD had the worst results in divided, sustained, selective and processing speed measures and that their performance was not related to differences in age or education when compared with healthy older adults.

So, the presence of deficits in several cognitive domains (memory, executive functions, attention and language) should be objectively assessed by a comprehensive neuropsychological, emotional and functional assessment enabling to measure the evolution of the clinical picture in order to help the establishment of the final diagnosis (Penã-Casanova et al., 2012). The international AD criteria (Mckhan et al., 2011) will be described in section II (Participants).

Mild Cognitive Impairment

Mild cognitive impairment is defined as a state of cognitive functioning that lays between healthy aging and dementia. More recently MCI was also purposed as a prodromal stage of AD (Petersen et al., 1999; Petersen, 2016). The existent deficits are higger than estimated for the individual's age and education but do not prevent the execution of daily life activities regardless a slight impairment in instrumental daily activities (what distinguishes it from dementia). The prevalence of this clinical condition commonly increases with age (Langa et al., 2014) and after the age of 65 this condition covers between 12 to 15% of the population (Petersen et al., 2010; Roberts et al., 2012). More frequently, MCI patients have memory complains and present an episodic memory impairment on formal testing with a global cognitive function relatively preserved. Although, there is extent clinical and etiological heterogeneity with well known sub-types. Petersen et al. (2014) defined four types of MCI patients according to the involvement of memory (amnestic versus non-amnestic) and the number of cognitive domains affected (single versus multidomain): amnestic MCI (a-MCI), single or multidomain; and nonamnestic (na-MCI), single or multidomain. The amnestic subtype is directly related with the progression to AD, so it can be considered a prodromal stage of the disease. The international criteria (Petersen et al., 1999; Albert et al., 2011) will be described on section II (Participants). Furthermore, as well as in AD, the evaluation of attention in addition with other cognitive domains in MCI patients is crucial for the establishment of the final diagnosis once that the deficits in this particular domain may also impair other cognitive functions as well as its assessment (Sharma, Kaur, Tripathi, Talwar & Sharma, 2017). According to Sharma et al. (2017), evaluation of attention through different tasks can identify AD at a prodromal stage showing that there is a gradual impairment in tests of working memory and control of attention in patients with a-MCI (multidomain) and AD. Similarly, other authors found that MCI patients could present deficits related with response inhibition, switching and cognitive flexibility besides memory (Petersen, 2016). These findings demonstrate the utility of the attention tests like the TP in evaluating MCI conditions once that it is a task that involves not only psychomotor abilities but also attentional control, switching, cognitive flexibility and perception abilities. Furthermore, in 2017, Sherman, Mauser, Nuno and

Sherzai, characterized attentional deficits among individuals with MCI and found that in some cases there is an impairment in both visual attention and visual information processing, highlighting this feature as an AD characteristic even in earliest stages. They also concluded that poorer visual attention was directly associated with worst global cognitive performance.

Attention and Depressive Symptomatology

The presence of depression in late life is frequently seen as a paradox once that the majority of older adults are not depressed itself. It is the combination of the advanced age with a depressive mood (revealing the presence of depressive symptomatology) that are wittingly confounded as being both cause and effect. As a disease it is, it goes hand to hand with physical disability and is related to brain changes and cognitive dysfunction (Kennedy, 2015). According to Rock, Roiser, Riedle and Blackwell (2014) impairment in attentional processing, memory and executive functioning are essential mediators of the depressive cognitive profile together with poor psychosocial functioning. Previously, in 2000, Nebes and collaborators had already showed that patients with depressive symptomatology present cognitive deficits related to a compromised performance in tasks that demand processing resources at highlevels of cognitive functioning like working memory, processing speed and attention. Besides that, depressive symptomatology was responsible by a significant extent of the variance on the neuropsychological measures including episodic memory, visuospatial performance and processing speed. Moreover, Wang and Blazer (2015), also concluded that the presence of depressive symptomatology leads to difficulties in concentration, attention and decision-making and also that memory complains in those patients were due to impaiments in attentional abilities.This was also showed by Eysenck and Keane (2015), who argued that subjects with depressive symptomatology have an impaired inhibitory control (which is an executive process that encompasses working memory resources) that leads to impaired attentional disengagement.

II - Objectives

The present work includes two studies. Study 1 focused on the validation of the TP for the Portuguese population (≥ 43 years old) for use with clinical groups with MCI and mild AD. For this, we compare and analyse the performances of these patients with healthy older adults and establish optimal cut off points for both the clinical groups. We also explore the psychometric properties of the TP, includind convergent validity comparing the results obtained by the three groups on the TP with those obtained on the TMT and the D2 Test of Attention. In Study 2, we evaluated the influence of sociodemographic variables on the TP performance and establish preliminary normative data based on a control group (CG) that comprised cognitively healthy subjects aged from 25 to 84 years old.

III – Methodology

Participants and procedures

This work comprises a convenience and transversal sample. The CG was composed by 234 community-dwelling cognitively healthy subjects without neurological or psychiatric pathology. The recruitment was made in Senior Universities (Universidade Sénior da Nazaré – USN), associations for elderly people (National Association of Elderly Support – ANAI), Day-Care Centers (Quinta das Camélias from Vila Nova de Poiares, Coimbra), as well as caregivers of ambulatory patients with dementia followed at the Neurology Department of the Centro Hospitalar e Universitário de Coimbra (CHUC). All participants in the CG were community volunteers aged between 25 and 84 years. The inclusion criteria included: 1) written informed consent; 2) Portuguese as native language and formal education ≥ 1 year; 3) normal scores according to the normative values defined for the Portuguese Population on the Mini Mental State Examination (MMSE; Freitas et al., 2015) and Montreal Cognitive Assessment (MoCA; Freitas et al., 2011); 4) preserved independence and functionality on activities of daily living; 5) no medication intake that could interfere with normal cognitive functioning; 6) absence of neurological or psychiatric disorders; 7) no significant motor, visual or auditory deficits with a possible negative influence in cognitive performance; 8) no present or past history of alcoholism or drug abuse. The exclusion criteria included: 1) illiteracy; 2) functional deficits with a recognized influence in daily living autonomy; 3) clinical significant depressive symptomatology (assessed by Geriatric Depression Scale – GDS-30: score ≥11 points) according to the normative values defined for the Portuguese Population (Simões et al., 2013; 2017). These neuropsychological tests (MMSE, MoCA, GDS-30 and SMC) will be described in section II - Materials.

The subjects of both clinical groups (MCI and AD) were recruited from the Memory Clinic of the Neurology Department of the CHUC. All patients underwent a cognitive screening (MMSE and MoCA) and a comprehensive neuropsychological assessment with the *Bateria de Lisboa para Avaliação da Demência* (BLAD; Guerreiro, 1998), and a medical exam (by a neurologist). Patients also carried out complementary diagnostic exams (e.g.: genotype study of Apolipoprotein E, APOE, and Cerebrospinal fluid – CSF – analysis through Lumbar Puncture). The final diagnosis was established by the agreement of a multidisciplinary teamwork based on the MCI international criteria: a) presence of memory complains, preferably confirmed by an informant; b) objective memory impairment established by normalized neuropsychological tests (according to age and educational level); c) global cognitive functioning generally preserved; d) daily life activities essentially maintained, even if there are some difficulties in more complex tasks; e) absence of dementia (Petersen et al., 1999; Albert et al., 2011). The MCI's group incorporated the two subtypes of the amnestic domain: a-MCI single-domain and a-MCI multidomain.

Dementia is diagnosed when there are cognitive or behavioural symptoms that cause: a) interference with the capacity to function at work or

at usual activities; b) represent a decline from previous levels of functioning and acting; c) is not explained by delirium or other major psychiatric disorder; d) cognitive impairment is detected and identified over a combination of (1) history-taking from the patient and a knowledgeable informant and (2) an objective cognitive assessment, either a "bedside" mental status examination or neuropsychological testing; e) the cognitive or behavioral impairment encompasses a minimum of two of the subsequent domains: impaired capacity to attain and remember new information; impaired reasoning and management of complex tasks; poor judgment; impaired visuospatial abilities; impaired language functions; or changes in personality or behaviour. Additionally, to the existence of the above-mentioned criteria, to diagnose AD, the following characteristics should be present: A) insidious onset. Symptoms have a gradual onset over months to years, not sudden over hours or days; B) clear-cut history of worsening of cognition by report or observation; C) the initial and most prominent cognitive deficits are evident on history and examination in one of the subsequent categories: 1) Amnestic presentation, that it is the most common syndromic presentation of AD dementia. The deficits should include impairment in learning and recall of recently learned information. There should also be evidence of cognitive dysfunction in at least one other cognitive domain, as defined earlier in the text; 2) Nonamnestic presentations: i) Language presentation: The most prominent deficits are in word-finding, but deficits in other cognitive domains should be present; ii) Visuospatial presentation: The most prominent deficits are in spatial cognition, including object agnosia, impaired face recognition, simultanagnosia, and alexia. Deficits in other cognitive domains should be present; iii) executive dysfunction: The most prominent deficits are impaired reasoning, judgment, and problem solving. Deficits in other cognitive domains should be present; D) The diagnosis of probable AD dementia should not be applied when there is evidence of (a) substantial concomitant cerebrovascular disease, defined by a history of a stroke temporally related to the onset or worsening of cognitive impairment; or the presence of multiple or extensive infarcts or severe white matter hyperintensity burden; or (b) core features of Dementia with Lewy bodies other than dementia itself; or (c) prominent features of behavioral variant frontotemporal dementia; or (d) prominent features of semantic variant primary progressive aphasia or nonfluent/agrammatic variant primary progressive aphasia; or (e) evidence for another concurrent, active neurological disease, or a non-neurological medical comorbidity or use of medication that could have a substantial effect on cognition (McKhann et al., 1984; 2011). In this study, the AD group included only mild dementia patients since subjects in moderate to severe stages were no longer able to complete the required tasks. The final diagnosis was established based on the international criteria described above.

Initially, the objectives of the study were explained to each participant and the informant consent was performed. AD patients performed an oral consent followed by a written consent of the caregiver. The sample groups will be described below.

Control Group

To evaluate this group we established a short protocol composed by the Informed Consent Form, the MMSE (Folstein at al., 1975; Guerreiro et al., 1994; Freitas et al., 2015), the MoCA (Nasreddine et al., 2005; Simões et al., 2008; Freitas et al., 2011), the TP (do Amaral, 1967), the TMT A & B (Cavaco, 2013), the D2 Test of Attention (Brickenkamp, 2007), the GDS-30 (Barreto et al., 2008; Simões et al., 2013) and the Scale of Subjective Memory Complains (SMC; Ginó et al., 2008; 2015). The evaluation procedure was implemented individually within one-hour session through the established fixed order of neuropsychological tests described above.

MCI Group

MCI patients are evaluated once a year (longitudinal follow-up) by a more comprehensive and integrative protocol, including the caregivers. This protocol is composed by: the MMSE (Folstein et al., 1975; Guerreiro et al., 1994; Freitas et al., 2015), the MoCA (Nasreddine et al., 2005; Simões et al., 2008; Freitas et al., 2011), the Alzheimer's Disease Assessment Scale – Cognitive (ADAS-Cog; Mohs et al., 1983; ; Guerreiro et al., 2008; Nogueira et al., 2018), the TMT A & B (Cavaco et al., 2013), the TP (do Amaral, 1967), the GDS-30 (Yesavage et al., 1983; Barreto et al., 2008; Simões et al., 2013), the SMC (Schmand, Jonker, Hooijer & Lindeboom, 1996; Ginó et al., 2008), the Escala de Classificação da Ansiedade de Hamilton (Hamilton, 1959), the Clinical Dementia Rating (CDR; Hughes, Berg, Danziger, Coben & Martin, 1982; Garrett et al., 2003; Santana, 2015b), the Neuropsychiatric Inventory (NPI; Cummings et al., 1994; Cummings, 1997; Santana et al., 2015), the Blessed Dementia Scale (Blessed, Tomlinson, & Roth, 1968; Garcia, 2008) and the Disability Assessment for Dementia Scale (DAD; Gelinas, Gauthier, McIntyre & Gauthier, 1999; Leitão, 2008; Galhardo, 2012).

AD Group

We included only mild AD patients (MMSE ≥ 18) that were previously assessed with the BLAD (Guerreiro, 1998), the MMSE (Folstein et al., 1975; Guerreiro et al., 1994; Freitas et al., 2015) and the MoCA (Nasreddine et al., 2005; Simões et al., 2008; Freitas et al., 2011). As part of this particular study, all patients completed the TMT (A/B), the TP, the GDS-30 and the SMC.

In order to verify the psychometric properties of the TP scores, performances obtained in the TP-WE and the TP-DI were correlated with the total scores of both the MMSE and the MoCA. To determinate the convergent validity first was applied the TP and the TMT A/B followed by the D2 Test of Attention. Besides Stroop Colour Test (colour and colour-word) (Trenerry et al., 1995; Fernandes, 2013; Espírito-Santo et al., 2015) already have normative data for the Portuguese population (in opposite to the CPT, the PASAT and the TEA), it was decided to use both the TMT and the D2 Test of Attention as measures of convergent validity with TP once for the Stroop Test the normative data is only available until 45 years old and our sample was in

majority older and also due to the fact that is already known that patients with AD show clear deficits in inhibitory-control tasks such as the TMT (Kaiser, et al., 2018).

Materials

At the present section, we will describe the seven main neuropsychological tests (the TP, the MMSE, the MoCA, the TMT (A/B), the D2 Test of Attention, the GDS-30 and the SMC) statistically analysed for the purposes of the present study. The protocol of this thesis starts with cognitive neuropsychological measures and ends with psychopathological tests.

Toulouse-Pieron Cancellation Test

The TP (do Amaral, 1967), the main instrument of the present thesis, is a timed cancellation test (dependant on the frontal lobe) that is capable of evaluate sustained attention, processing speed, visuo-perceptive and inhibition abilities. Its administration can be done individually or in group and takes exactly 10 minutes.

Mini Mental State Examination

The MMSE (Folstein, Folstein & McHugh, 1975; Guerreiro et al., 1994; Guerreiro, 1998; Freitas et al., 2015), is a brief cognitive screening instrument used to assess the presence of cognitive impairment. It is composed by 30 items, which evaluate six different domains of cognitive functioning (orientation; repetition; verbal recall; attention and calculation; language and visual construction) (Freitas, Simões, Alves, & Santana, 2015a). The administration takes 5 to 10 minutes with a maximum score of 30 points, meaning that higher results illustrate better performances (Freitas et al., 2015b).

Montreal Cognitive Assessment

The MoCA (Nasreddine et al., 2005; Simões et al., 2008) is a brief cognitive screening instrument developed to evaluate the milder forms of cognitive impairment (Freitas et al., 2011). The scores obtained allow to distinguish between cognitive changes derived from the normal aging process and those that occur due to pathological cognitive deficits (Freitas et al., 2011). This instrument evaluates six cognitive domains which are executive functions, visuospatial ability, language, attention, concentration and working memory, memory and temporal/spatial orientation (Freitas et al., 2011; Freitas, Simões, Alves, & Santana, 2015c). The administration takes 10 to 15 minutes and the punctuations oscillate between 0 to 30 points, meaning that higher scores correspond to better performances (Freitas et al., 2015b).

Trail Making Test A & B

The TMT (forms A and B) (Reitan & Wolfson, 1993; Cavaco et al., 2013), was developed for adults from 18 years old with suspected cognitive dysfunction. The instrument is composed by two tasks (A and B), which are separately timed. It provides information about attention, visual search, eye-

hand coordination, processing speed, sequencing capability and cognitive flexibility. Additionally, task B also evaluates executive functions, namely the ability to switch between sequences (Cavaco et al., 2013). The administration takes approximately 5 minutes and the test scoring is made on the basis of the normative values established for the Portuguese population (including sex, gender and educational level) (Cavaco et al., 2013).

D2 Test of Attention

The D2 Test of Attention (Brickenkamp, 2007) is a timed cancellation test that allows the assessment of selective attention from 9 to 60 years old. It assesses processing speed, rule compliance and quality of performance in response to the discrimination of similar visual stimuli thereby providing an estimation of the individual's attention and concentration abilities. The administration can be done individually or collectively and takes about 8 minutes. The quotation and interpretation are based on the normative values established for the Portuguese population through the number of characters marked and error rates (omission and commission) (Brickenkamp, 2007). In the present study, we only calculated the following scores: E% (Percentage of Errors) that is a variable that measures the qualitative aspect of the subject's performance and TN-E (Total Performance) which is a measure of the quantity of work completed after a single correction for errors and omissions, in order to correlate the results obtained in the application of the D2 Test with those obtained with the TP.

Geriatric Depression Scale

The GDS-30 (Yesavage et al., 1983; Barreto et al., 2008; Simões & Firmino, 2013; Simões et al., 2017) was settled as a straightforward screening measure for depression symptomatology in elderly populations. It sets up a self-report instrument with 30 items (referent to affective and cognitive domains), and the scores range between 0 to 30 points. The administration takes 10 to 15 minutes (answering Yes/No), where higher scores mean a higher level of depression symptomatology (0-10 points: "no depression"; 11-20: "mild depression"; 21-30: "severe depression") (Yesavage et al., 1982; Yesavage et al., 1983).

Subjective Memory Complains Scale

The SMC scale (Schmand et al., 1996; Ginó et al., 2008; 2015), aims to characterize memory complains. It is composed by 10 items with multiple levels of response, according with the complains severity. There are 5 specific questions focused on memory, one about language, one about orientation, one about concentration, and two referring to slow thinking. It can be applied to all subjects aged 49 and over. The total score ranges between 0 and 21 points (maximum level of complains).

Statistical Analysis

Statistical analyses were developed using the *Statistical Package for the Social Sciences* (SPSS, version 22.0 for Windows; IBM Corp., Armonk, N.Y., USA, 2013). Descriptive statistics were used for sample's characterization followed by independent-samples t-test for group comparisons. To assess test-retest reliability we calculated the correlations between the scores at baseline and at follow-up six months later (only for the control group). The convergent validity was performed using Pearson correlations coefficients between the TP indexes and MMSE, MoCA, TMT A/B, and D2 Test of Attention scores. The group differences were examined using two-sample t-test and analysis of variance (ANOVA).

The diagnostic accuracy of the TP for the prediction of the clinical diagnosis of MCI and AD was evaluated by the receiver operating characteristics (ROC) curve analysis. In this analysis, the areas under the curves (AUC) can range from 0 to 1 with bigger AUC meaning better diagnostic accuracy. The optimal cut-off points for each index of the TP that generated the highest Youden value were selected, with higher values meaning the maximization of both sensitivity and specificity. For the analysis of the predictive values of these indexes were performed, for each of the cut-off points, the sensitivity (the probability of subjects with disease to have a positive result), specificity (the probability of subjects with no cognitive impairment to have a negative result), positive predictive value (PPV; the probability of disease in subjects who have a positive result), negative predictive value (NPV; the probability of the absence of disease in subjects who have a negative result) and the classification accuracy (the probability to correctly classify subjects with or without cognitive impairment). Multiple Linear Regression (MLR) analysis was performed to evaluate the significance of age and education as influencing factors on the TP as well of depressive symptomatology and of subjective memory complains.

IV – Results

Study 1: Validation of the TP for MCI and AD

Sociodemographic characterization of the study sample

The present study's sample was composed by 120 subjects, subdivided into 40 participants of the control group and 80 participants of the clinical group (MCI group: n=40; AD group: n=40), matched according to age and educational level.

The features of the study sample and of each subgroup are presented in Table 1. For the following description, we considered the sample size, gender, age and educational level.

Table 1. Sociodemographic characterization of the study sample.

	Total sample	Control Group	Clinical Group	MCI Group	AD Group
N	120	40	80	40	40
Education					
M±SD	6.66±4.31	6.75±4.33	6.64±4.33	6.97±4.35	6.30±4.33
[Min-Max]	[1 - 15]	[4 - 15]	[1 - 15]	[3 - 15]	[1 - 15]
Age					
M±SD	71.44±9.53	71.10±9.53	71.61±9.56	71.40±9.29	71.83±9.97
[Min-Max]	[43 - 87]	[43 - 84]	[43 - 87]	[45 - 85]	[43 - 87]
Gender					
Female (%)	74 (61.7%)	22 (55.0%)	52 (65.0%)	23 (57.5%)	29 (72.5%)

Abbreviations: MCI = Mild Cognitive Impairment; AD = Alzheimer's Disease; M = Mean; SD = Stadard deviation; Min. = minimum value; Max. = maximum value.

We previously selected the participants of the control group from the database of TP's preliminary normative study for the Portuguese population (Study 2) in order to match the patients of the clinical group in age and educational level. As expected, there were no differences between the three groups in age or educational level, respectively, $F_{(1,119)}$=.077, p=.782 and $F_{(1,119)}$=.011, p=.917 (Control vs MCI vs AD), confirmed by the post-hoc analysis.

To further characterize the global cognitive performance of the study sample, the results of each group in the MMSE, MoCA, D2, TMT A/B, GDS-30 and SMC scale are presented in Table 2.

We found statistically significant differences between Control, MCI and AD groups in the MMSE [$F_{(2, 119)}$=92,936, p<.001, η^2=.614], the MoCA [$F_{(2,105)}$=92,086, p<.001, η^2=.674], the TMT A [$F_{(2,98)}$=20,06, p<.001, η^2=.295], the TMT B [$F_{(2, 71)}$=6,14, p<.01, η^2=.151], the GDS-30 [$F_{(2,105)}$=49,005 p<.001, η^2=.272)], and the SMC scale [$F_{(2,105)}$=21,880 p<.001, η^2=.488]. Post-hoc test analysis confirmed that the control group had a better cognitive performance than both clinical groups in all tests except TMT A (where there were no satistically significant differences between controls and MCI's); and that the MCI group revealed significant higher results than the AD group in all described measures.

In order to characterize the influence of the presence of depressive symptomatology and of memory complains in the results of the TP, we correlated both the SMC and the GDS-30 with the six indexes of the TP. We found that the MCI group scored higher than controls and AD patients in the GDS-30 and that the GDS scores were negativelly correlated with the total scores on the TP-WE (r=-.267, p<.01) and TP-H (r=-.260, p<.01). GDS-30 scores were also positively correlated with SMC (r=.748, p<.01) and with the TP-E (r=.214, p<.05). Concerning SMC, we observed that it was statistically negatively correlated with the TP-WE (r=-.188, p<.05) and positively correlated with the TP-E (r=.223, p<.05).

After that, we performed a MLR analysis (enter method) to analyse the influence of these two variables in the TP results of the total sample and for each one of the three groups (Tables 3 to 6).

Table 2. Cognitive characterization of groups.

	Total sample	Control Group	Clinical Group	MCI Group	AD Group
N	120	40	80	40	40
MMSE					
M±SD	26.00±3.73	28.89±1.41	25.29±9.53	28.08±1.54	22.37±3.17
[Min-Max]	[17 - 30]	[26 - 30]	[17 - 30]	[24 - 30]	[17 - 29]
MoCA					
M±SD	16.72±6.42	25.40±3.27	15.46±5.76	20.00±3.51	11.54±4.23
[Min-Max]	[4 - 30]	[20 - 30]	[4 - 26]	[12 - 26]	[4 - 19]
TMT A					
M±SD	112.39±71.35	80.21±34.80	128.48±79.31	89.22±37.99	165.44±90.18
[Min-Max]	[32 - 375]	[32 - 167]	[33 - 375]	[33 - 173]	[35 - 375]
TMT B					
M±SD	243.25±136.47	210.50±118.91	268.45±145.73	233.28±98.92	364.82±204.02
[Min-Max]	[55 - 914]	[55 - 570]	[66 - 914]	[66 - 420]	[174 - 914]
GDS-30					
M±SD	8.25±6.10	2.67±2.88	8.75±6.70	11.13±6.61	6.69±4.77
[Min-Max]	[0 - 26]	[0 - 7]	[0 - 26]	[2 - 26]	[0 - 19]
SMC					
M±SD	7.06±4.42	5.00±5.93	7.25±4.27	8.17±4.27	6.46±4.17
[Min-Max]	[1 - 20]	[0 - 13]	[1 - 20]	[1 - 16]	[1 - 20]
D2 – TN-E					
M±SD	-	103.70±67.91	-	-	-
[Min-Max]		[-7 - 208]			
D2 – E%					
M±SD	-	30.87±41.54	-	-	-
[Min-Max]		[0 - 160]			

Abbreviations: MCI = Mild Cognitive Impairment; AD = Alzheimer's Disease; MMSE = Mini-Mental State Examination; MoCA = Montreal Cognitive Assessment; TMT = Trail Making Test form A and B; GDS = Geriatric Depression Scale; SMC = Subjective Memory Complaints. M = Mean; SD: Standard Deviation; Min = minimum value; Max = maximum value.

Table 3. MLR analysis for depressive symptomatology (DS) and subjective memory complains (SMC) (Total Sample).

Variable	B	SEB	β
TP-WE $_1$			
DS	-41.682	16.448	-.237
SMC	33.698	14.244	.221
TP-DI $_2$			
DS	16.833	16.743	.096
SMC	-25.536	14.500	-.169
TP-TR $_3$			
DS	-3.923	2.905	-.129
SMC	4.976	2.516	.189

$_1R^2$=.073, $F_{(2,119)}$=4.599, p=.012; $_2R^2$=.028, $F_{(2,119)}$=1.669, p=.193; $_3R^2$=.037, $F_{(2,119)}$=2.259, p=.109

Abbreviations: WE = Work Efficiency; DI = Dispersion Index; TR = Total Result; SEB = standard error of B.

Table 4. MLR analysis for depressive symptomatology (DS) and subjective memory complains (SMC) (Control group).

Variable	B	SEB	β
TP-WE $_1$			
DS	-15.333	56.603	-.044
SMC	32.556	18.868	.280
TP-DI $_2$			
DS	-6.848	17.215	-.067
SMC	-.895	5.738	-.026
TP-TR $_3$			
DS	-1.075	5.998	-.029
SMC	2.934	1.999	.241

$_1R^2$ = .075, $F_{(2,39)}$=1.497, p=.237; $_2R^2$ =.006, $F_{(2,39)}$=.112, p=.895; $_3R^2$ =.056, $F_{(2,39)}$=1.090, p=.347

Abbreviations: WE = Work Efficiency; DI = Dispersion Index; TR = Total Result; SEB = standard error of B.

Table 5. MLR analysis for depressive symptomatology (DS) and subjective memory complains (SMC) (MCI group).

Variable	B	SEB	β
***TP-WE*$_1$**			
DS	-29.334	13.837	-.344
SMC	37.051	18.186	.330
***TP-DI*$_2$**			
DS	13.013	7.096	.308
SMC	-9.270	9.326	-.167
***TP-TR*$_3$**			
DS	-3.006	1.398	-.348
SMC	3.350	1.838	.339

$_1R^2$=.147, $F_{(2,39)}$=3.197, p=.05; $_2R^2$=.086, $F_{(2,39)}$=1.750, p=.188; $_3R^2$=.153, $F_{(2,39)}$=3.332, p=.047

Abbreviations: WE = Work Efficiency; DI = Dispersion Index; TR = Total Result; SEB = standard error of B.

Table 6. MLR analysis for depressive symptomatology (DS) and subjective memory complains (SMC) (AD group).

Variable	B	SEB	β
***TP-WE*$_1$**			
DS	-10.849	23.152	-.077
SMC	10.476	21.674	.079
***TP-DI*$_2$**			
DS	24.991	38.246	.107
SMC	-15.511	35.240	-.071
***TP-TR*$_3$**			
DS	1.648	6.666	.040
SMC	4.670	6.641	.122

$_1R^2$ =.012, $F_{(2,39)}$=.232, p=.801; $_2R^2$ =.016, $F_{(2,39)}$=.302, p=.741; $_3R^2$ =.017, $F_{(2,39)}$=314, p=.733

Abbreviations: WE = Work Efficiency; DI = Dispersion Index; TR = Total Result; SEB = standard error of B.

The results presented on table 3 showed that both depressive symptomatology and subjective memory complains are significant contributors to the prediction of the TP scores but only for the TP-WE (total sample). To this model, the adjusted R^2 value was .073, which means that 7,3% of the variance on the TP-WE was explained by the obtained scores on the GDS-30 and the SMC scale. After that, the same MLR was performed for each one of the three groups. The results showed that the same two variables only revelead significant influences on the MCI group, more specifically for TP-WE and TP-TR indexes. To this model, the adjusted R^2 value was .147, meaning that 14,7% of the variance on the TP-WE is derived from the results obtained on the GDS-30 and the SMC scale. For the TP-TR, the adjusted R^2 value was .153, which means that 15,3% of the variance is explained by the two above-mentioned variables.

Psychometric properties of the TP

Convergent validity

The convergent validity was performed through the Pearson correlations between the TP and the MMSE, the MoCA, the TMT and the D2 Test of Attention. We observed significant negative correlations between the total scores of the TP (DI, E, O) and the MMSE (r= -.617; r= -.434; r= -.333, p<.01) and MoCA (r= -.620; r= -.404; r= -.304, p<.01)..The results also showed significant negative correlations between the total scores of the TP-WE and the TMT A (r= -.565, p<.01) and B (r= -.524, p<.01), the TP-TR and the TMT A (r= -.578, p<.01) and B (r= -.503, p<.01) as well as between the TP-H and the TMT A (r= -.582, p<.01) and B (r= -.433, p<.01). We also found significant positive correlations between the total scores of the TP-DI and the TMT A (r= .486, p<.01) and B (r= .422, p<.01), the TP-E and TMT A (r= .387, p<.01) and between the D2(E%) and the TP-O (r= .950, p<.01). Regarding D2, we observed significant positive correlations between the D2 (TN-E) and the TP-WE (r= .959, p<.01), the D2 (TN-E) and the TP-TR (r= .942, p<.01). As expected, we also found significant negative correlations between the D2 (TN-E) and the TP-DI (r= -.920, p<.01) and the

Reliability – Test-Retest

The test-retest reliability was measured through the Pearson's correlation coefficient between the baseline and the six-month follow-up data. This analysis was performed only for a sub-sample (n=30) in the control group. The value obtained was r=.877 (TP-DI), r=.903 (TP-WE) and r=.854 (TP-TR), p<.01.

Differences between groups in TP performance

We conducted an ANOVA in order to analyze the differences between the groups on the several TP indexes. In Table 7 we presented in detail the TP performances for all subgroups and for all the six indexes. Results from group comparisons are presented in Table 8. The results confirmed the existence of significant differences between groups: Control Group vs Clinical Group, Control Group vs MCI, Control Group vs AD and MCI vs AD in all the indexes of the TP excluding in Control Group vs MCI (TP-DI and omissions) and in Control Group vs Clinical Group (in omissions).

Table 7. Characterization of the performance on TP.

	Total sample	Control Group	Clinical Group	MCI Group	AD Group
N	120	40	80	40	40
WE					
M±SD	66.81±76.16	124.48±55.50	66.81±76.16	79.50±43.14	-3.55±64.06
[Min-Max]	[-234 - 295]	[35 - 245]	[-234 - 295]	[0 - 190]	[-234 – 79]
n	120	40	80	40	40
DI					
M±SD	53.48±75.98	19.20±16.16	53.48±75.98	27.32±21.37	113.91±106.04
[Min-Max]%	[0 – 460]%	[0 – 66]%	[0 – 460]%	[2 – 100]%	[7.14 – 460]%
n	120	40	80	40	40
TR					
M±SD	3.46±14.22	12.61±7.48	3.46±14.22	7.59±4.37	-3.87±18.49
[Min-Max]	[-110.0 – 24.3]	[2.5 – 24.3]	[-110.0 – 24.3]	[-1.0 – 18.9]	[-110.0 – 7.8]
n	120	40	80	40	40
Hits					
M±SD	90.40±56.16	163.64±74.95	90.40±56.16	102.13±39.94	53.05±23.60
[Min-Max]	[9 - 303]	[59 - 303]	[9 - 303]	[23 - 208]	[9 - 126]
n	120	40	80	40	40
Errors					
M±SD	4.84±10.74	0.93±1.39	4.84±10.74	1.70±3.92	9.35±14.91
[Min-Max]	[0 - 80]	[0 - 5]	[0 - 80]	[0 - 22]	[0 - 80]
n	120	40	80	40	40
Omissions					
M±SD	34.94±46.58	34.71±27.45	34.94±46.58	21.88±14.75	48.08±66.00
[Min-Max]	[2 - 295]	[6 - 112]	[2 - 295]	[2 - 63]	[3 - 295]
n	120	40	80	40	40

Abbreviations: MCI = Mild Cognitive Impairment; AD = Alzheimer's Disease; WE = Work Efficiency; DI = Dispersion Index; TR= Total Result; H = Hits; E = Errors; O = Omissions; M = Mean; SD = Standard Deviation; Min. = minimum value; Max. = maximum value.

Table 8. Group differences in cognitive indexes of the TP.

	Control Group vs Clinical Group	Control Group vs MCI	Control Group vs AD	MCI vs AD
WE	$F_{(1, 119)}$=47.98, p<.001, η_2=.29	$F_{(1, 79)}$=16.38, p<.001, η_2=.17	$F_{(1, 79)}$=91.28, p<.001, η_2=.54	$F_{(1, 79)}$= 46.26, p<.001, η_2=.37
DI	$F_{(1, 119)}$=13.49, p<.001, η_2=.10	$F_{(1, 79)}$=3.68, p=.059, η_2=.05	$F_{(1, 79)}$=30.68, p<.001, η_2=.29	$F_{(1, 79)}$=25.64, p<.001, η_2=.25
TR	$F_{(1, 119)}$=18.94, p<.001, η_2=.14	$F_{(1, 79)}$=16.71, p<.001, η_2=.18	$F_{(1, 79)}$=27.48, p<.001, η_2=.26	$F_{(1, 79)}$=14.54, p<.001, η_2=.16
H	$F_{(1, 119)}$=56.90, p<.001, η_2=.33	$F_{(1, 79)}$=16.10, p<.001, η_2=.17	$F_{(1, 79)}$=84.39, p<.001, η_2=.52	$F_{(1, 79)}$=44.77, p<.001, η_2=.37
E	$F_{(1, 119)}$=8.123, p<.01, η_2=.06	$F_{(1, 79)}$=4.67, p=.03, η_2=.06	$F_{(1, 79)}$=14.20, p<.001, η_2=.16	$F_{(1, 79)}$=9.85, p<.01, η_2=.11
O	$F_{(1, 119)}$=1.71, p=.194, η_2=.01	$F_{(1, 79)}$=0.35, p=.558, η_2=.004	$F_{(1, 79)}$=4.78, p=.03, η_2=.06	$F_{(1, 79)}$=6.01, p=.016, η_2=.07

Abbreviations: MCI = Mild Cognitive Impairment; AD = Alzheimer's Disease; WE = Work Efficiency; DI = Dispersion Index; TR = Total Result; H = Hits; E = Errors; O = Omissions;

Concerning the group differences presented above, we will now describe the respective mean differences. So, for TP-WE, the mean differences were respectively 88.00 ± 12.46 (Control vs Clinical Group), 46.48 ± 11.06 (Control vs MCI), 129.53 ± 13.36 (Control vs AD) and 83.05 ± 12.21 (MCI vs AD). For TP-DI, the mean differences were respectively -51.42 ± 13.99 (Control vs Clinical Group), -8.11 ± 4.24 (Control vs MCI), -94.72 ± 16.96 (Control vs AD), and -86.60 ± 17.10 (MCI vs AD). For the TP-TR, the mean differences were respectively 10.41 ± 2.39 (Control vs Clinical Group), 4.68 ± 1.15 (Control vs MCI), 16.14 ± 3.06 (Control vs AD) and 11.45 ± 3.00 (MCI vs AD). For TP-H, the mean differences were respectively 70.79 ± 9.39 (Control vs Clinical Group), 46.25 ± 11.53 (Control vs MCI), 95.33 ± 10.34 (Control vs AD) and 49.08 ± 7.33 (MCI vs AD). For TP-E, the mean differences were respectively -5.20 ± 1.83 (Control vs Clinical Group), -1.38 ± .636 (Control vs MCI), -9.03 ± 2.36 (Control vs AD) and -7.65 ± 2.44 (MCI vs AD). Lastly, for TP-O, the mean differences were respectively -10.68 ± 8.17 (Control vs Clinical Group), 2.43 ± 4.13 (Control vs MCI), -23.76 ± 10.97 (Control vs AD) and -26.20 ± 10.69 (MCI vs AD). Through the above presented results, we can conclude that the control group presented better scores than both the MCI and AD groups in all indexes. The only exception was omissions where there were no statistically significant differences between controls and MCI patients. The MCI group also performed better than the AD group in all measures of TP.

Validity and Diagnostic Accuracy of TP – Cut-off points

Receiver operating characteristics (ROC) curve analysis were calculated to measure the diagnostic accuracy of the TP to distinguish MCI and AD patients from healthy older adults. Graphic representations of the ROC curves are delivered in figures 1 to 12.

Figure 1. ROC curve analysis of the TP-WE to detect MCI and AD, respectively.

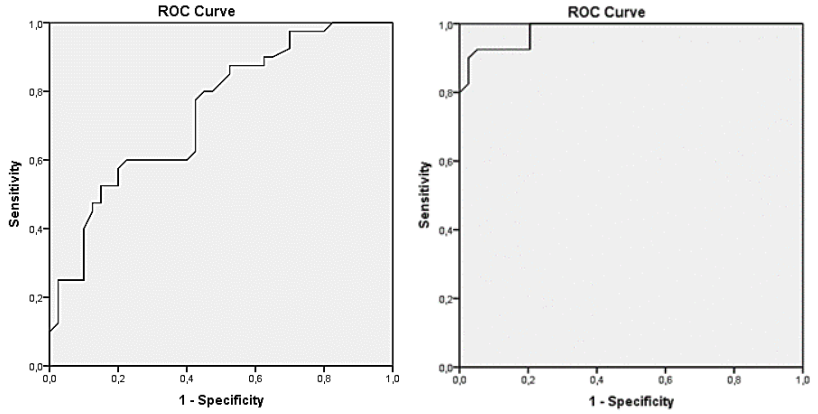


Figure 2. ROC curve analysis of the TP-DI to detect MCI and AD, respectively.

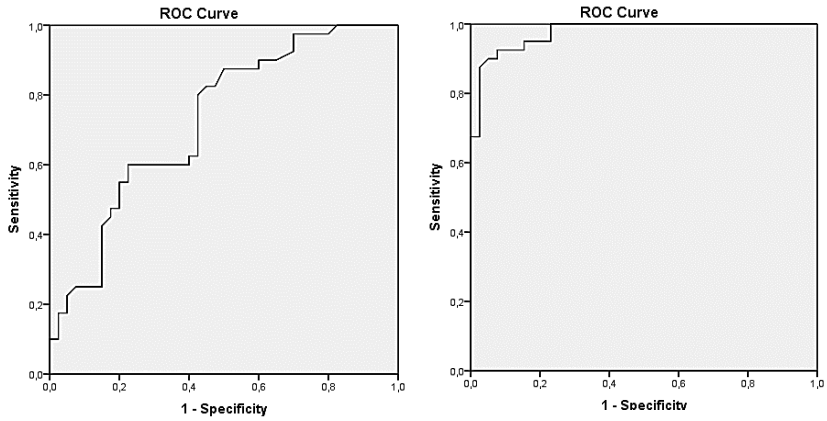


Figure 3. ROC curve analysis of the TP-TR to detect MCI and AD, respectively.

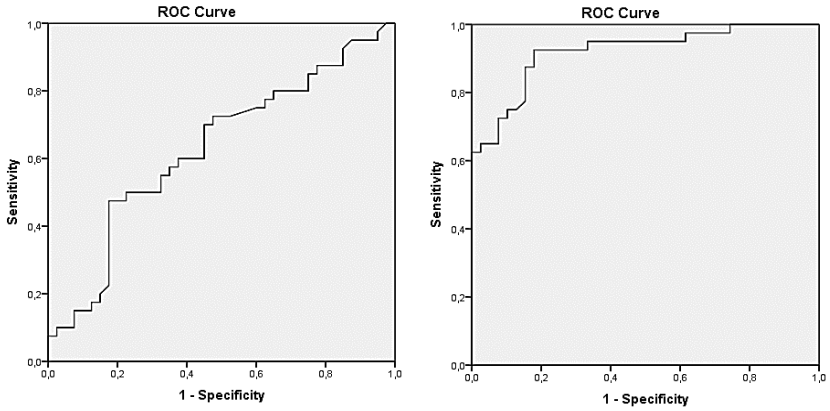


Figure 4. ROC curve analysis of the TP-Hits to detect MCI and AD, respectively.

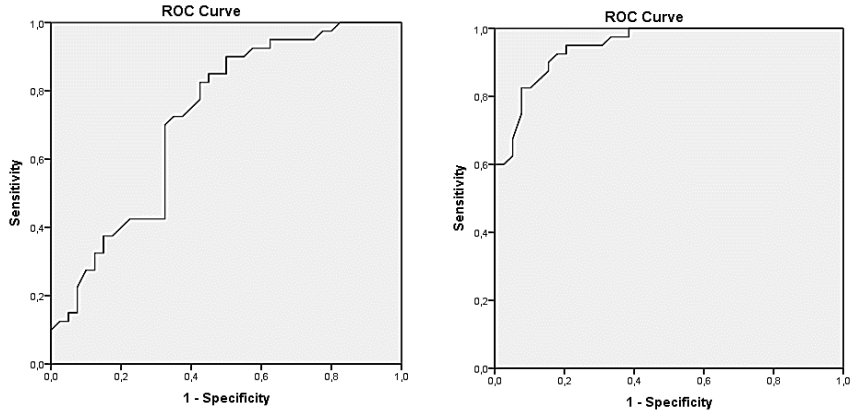


Figure 5. ROC curve analysis of the TP-Errors to detect MCI and AD, respectively.

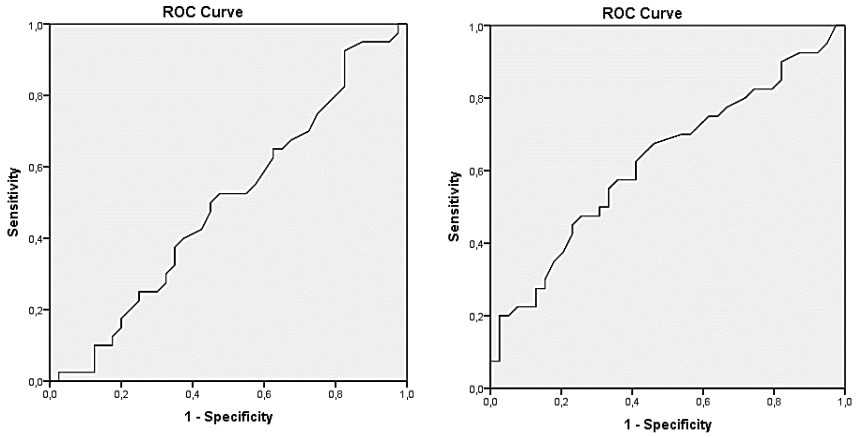
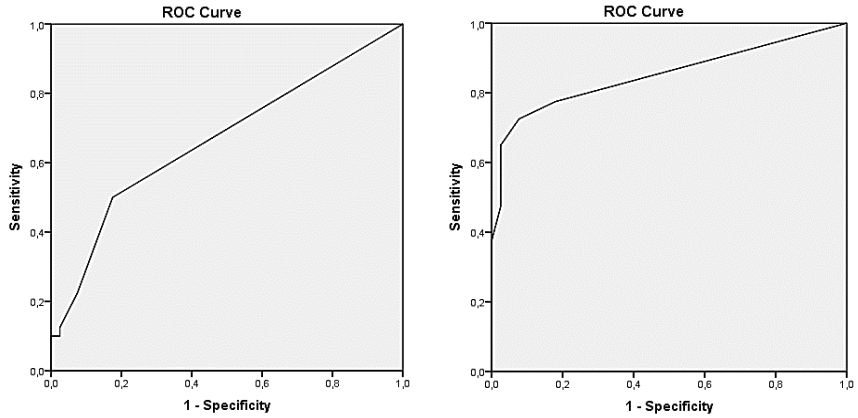


Figure 6. ROC curve analysis of the TP-Omissions to detect MCI and AD respectively.



The AUC for the MCI in the three main indexes of the TP was 0.739 [95% confidence interval (CI) = 0.631-0.847] for the TP-WE, 0.629 [95% confidence interval (CI) = 0.506-0.753] for the TP-DI and 0.839 [95% confidence interval (CI)=0.801-0.876] for the TP-TR, respectively. For AD the AUC's were 0.981 [95% (CI) = 0.960-1.000] for the TP-WE, 0.921 [95% confidence interval (CI) = 0.861-0.980] for the TP-DI and 0.977 [95% confidence interval (CI) = 0.951-1.000] for the TP-TR.

On table 9 we presented the optimal cut-off points for maximum accuracy (according to Youden's index) and the respective values of sensitivity, specificity, PPV, NPV, and classification accuracy.

Table 9. ROC curve analysis.

	TP	Cut-off	AUC	Sensitivity	Specificity	PPV	NPV	Classifcation Accuracy
MCI	WE	< 73	0.739	53	85	78	64	69
	DI	> 27	0.629	48	83	73	61	65
	TR	< 8	0.726	60	78	73	66	69
	H	< 135	0.720	83	58	66	77	71
	E	> 1	0.666	50	83	75	62	67
	O[1]	> 7	0.496	93	18	53	72	56
AD	WE	< 49	0.981	93	95	95	93	94
	DI	> 26	0.921	93	82	84	92	88
	TR	< 4	0.977	90	95	95	90	93
	H	< 70	0.949	83	93	92	85	88
	E	> 2	0.851	73	93	91	78	83
	O	> 34	0.626	45	77	66	58	61

Abbreviations: AUC = area under the curve; PPV = positive predictive value; NPV = negative predictive value; TP = Toulouse-Pièron; WE = work efficiency; DI = dispersion index; TR = total result; H = hits; E = errors; O = omission.

Note: Values of sensitivity, specificity, PPV, NPV, and classification accuracy were expressed in percentage. Cut-off points indicate the minimum score required for presence of signal.

[1] We decided to maintain the values of sensitivity, specificity, PPV, NPV, and classification accuracy for omissions even in the absence of statiscally significant differences between Control and MCI group: $F_{(1, 79)}=0.35$, *ns*, $\eta^2=.004$.

In the MCI group the optimal cut-off point for the TP (WE below 73, DI above 27 and TR below 8), allowed to identify MCI patients and was capable of discriminating them from the healthy older adults. Based on this cut-off points, TP presented a sensitivity of 53% (WE), 48% (DI) and 60% (TR), a specificity of 85% (WE), 83% (DI) and 78% (TR), a PPV of 78% (WE) and 73% (DI and TR), a NPV of 64% (WE), 61% (DI) and 66% (TR), and lastly a classification accuracy of 69% (WE and TR) and 65% (DI).

To AD patients, the optimal cut-off points for the TP (WE below 49, DI above 26 and TR below 4), permited to distinguish AD patients from healthy older adults. Based on these cut-off points, TP presented a sensitivity of 93% (WE and DI) and 90% (TR), a specificity of 95% (WE and TR) and 82% (DI), a PPV of 95% (WE and TR) and 84% (DI), a NPV of 93% (WE), 92% (DI) and 90% (TR), and finally a classification accuracy of 94% (WE), 88% (DI) and 93% (TR).

Figure 7. ROC curve analysis of the TP-WE to distinguish MCI from AD.

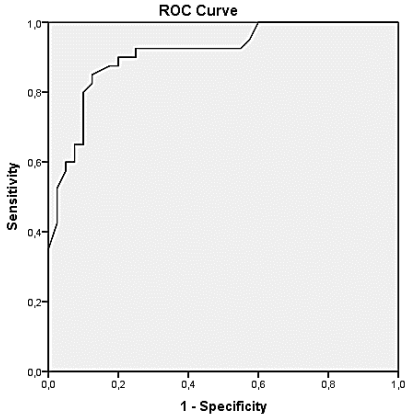


Figure 8. ROC curve analysis of the TP-DI to distinguish MCI from AD.

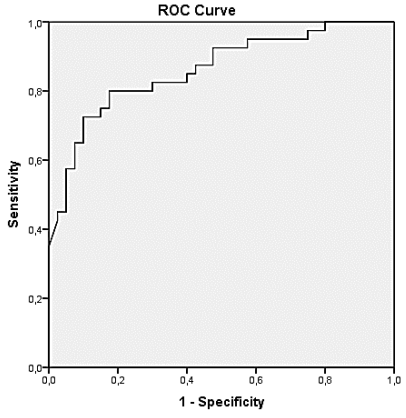


Figure 9. ROC curve analysis of the TP-TR to distinguish MCI from AD.

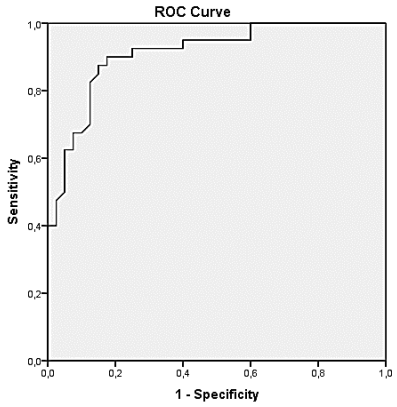


Figure 10. ROC curve analysis of the TP-Hits to distinguish MCI from AD.

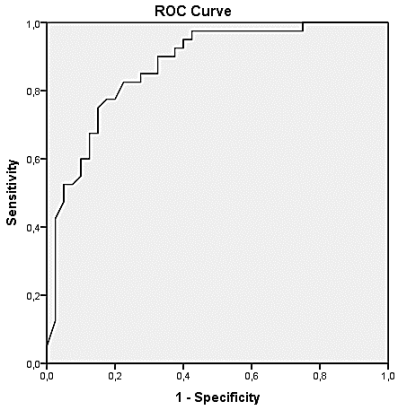


Figure 11. ROC curve analysis of the TP-Errors to distinguish MCI from AD.

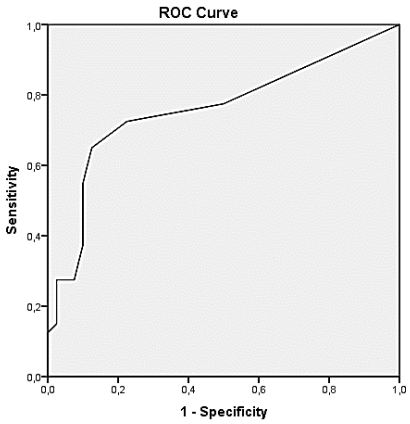
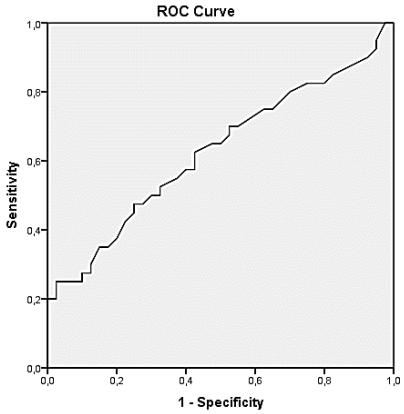


Figure 12. ROC curve analysis of the TP-Omissions to distinguish MCI from AD.



The AUC for the MCI vs AD in the three main indexes of the TP was 0.910 [95% confidence interval (CI) = 0.897-0.974] for the TP-WE, 0.862 [95% confidence interval (CI) = 0.781-0.942] for the TP-DI and 0.911 [95% confidence interval (CI) = 0.848-0.974] for the TP-TR, respectively.

On table 10 we presented the optimal cut-off points for maximum accuracy (according to Youden's index) and the respective values of sensitivity, specificity, PPV, NPV, and classification accuracy.

Table 10. ROC curve analysis.

	TP	Cut-off	AUC	Sensitivity	Specificity	PPV	NPV	Classifcation Accuracy
MCI vs AD	WE	< 37	0.910	85	88	88	85	87
	DI	> 38	0.862	80	83	82	81	82
	TR	< 3	0.911	88	85	85	88	87
	H	< 66	0.873	75	85	83	77	80
	E	> 3	0.759	65	88	84	72	77
	O	> 32	0.626	48	75	66	59	62

Abbreviations: AUC = area under the curve; PPV = positive predictive value; NPV = negative predictive value; TP = Toulouse-Pièron; WE = work efficiency; DI = dispersion index; TR = total result; H = hits; E = errors; O = omission.
Note: Values of sensitivity, specificity, PPV, NPV, and classification accuracy were expressed in percentage. Cut-off points indicate the minimum score required for presence of signal.

For distinguish MCI from AD patients the optimal cut-off point for the TP were TP-WE below 37, TP-DI above 38 and TR below 3. Based on this cut-off points, TP presented a sensitivity of 85% (WE), 80% (DI) and 88% (TR), a specificity of 88% (WE), 83% (DI) and 85% (TR), a PPV of 88% (WE) and 82% (DI) and 85% (TR), a NPV of 85% (WE), 81% (DI) and 88% (TR), and lastly a classification accuracy of 87% (WE and TR) and 82% (DI).

Study 2 – Preliminary normative study of the TP for the Portuguese population

We collected data from 320 healthy control subjects. Of the 320 total subjects, 48 (15%) were excluded due to their low performance on the brief cognitive screening tests according to the normative data of the MMSE (Freitas et al., 2015) and 38 (12%) were excluded following the initial interview due mostly to the personal history of neurological or psychiatric disorders. The sample selection and stratification process culminated with a final sample of 234 participants (mean age = 50.15 ± 16.26, age range = [25-84]; mean education = 8.76 ± 4.41, education range = [1-15]).

The sociodemographic features of the total sample are described in Table 11 and the subject's performance on the other psychological assessment instruments are found in Table 12.

Table 11. Sociodemographic characterization of the study population.

	Levels	Total Sample n (%)
Age	25-49	120 (51.3)
	50-64	57 (24.4)
	≥ 65	57 (24.4)
Gender	Male	98 (41.9)
	Female	136 (58.1)
Education	Primary (1-4)	43 (18.4)
	Middle (5-10)	63 (26.9)
	High (≥ 11)	128 (54.7)

Table 12. Performance of the study subjects in complementary psychological measures.

Age Groups	**MMSE** (M ± SD)	**TMT A** (M ± SD)	**TMT B** (M ± SD)	**D2 (TN-E)** (M ± SD)	**D2 (E%)** (M ± SD)
25-49	29,58 ± 0.76	41,45 ± 13.24	92.81 ± 34.57	142.43 ± 46.36	12.00 ± 9.43
50-64	29,30 ± 0.88	49.77 ± 18.26	109.27 60.28	165.14 ± 41.22	15.00 ± 6.48
≥ 65	29.30 ± 0.76	68.80 ± 27.45	186.95±82.19	70.62 ± 65.36	44.00 ± 46.8

Abbreviations: MMSE = Mini Mental State Examination; TMT A = Trail Making Test A; TMT B = Trail Making Test B; D2 = D2 Test of Attention.

We found statistically significant correlations between the total scores of the TP indexes (WE and DI) with age (r=-.598 and r=.153, p<.01) respectively, education (r=.426, p<.01 and r=-.128, p=.051) respectively, but not with gender (r=.092 and r=-.009, p=.896), respectively (Table 13).

MLR analysis (stepwise method) was performed to examine the contribution of relevant sociodemographic variables to the indexes of the TP and to assess the added contributions of their interactions. The intercorrelations of the statiscally significant predictive variables can be found in Table 14.

The regression model ($F_{(2,233)}$=81.44, p<.001, TP-WE; $F_{(2,233)}$=3,253, p=.040, TP-DI) included the two variables together: age (ß=-.511, t=-9.552, p<.001, TP-WE; ß=.111, t=1.618, p=.107, TP-DI) and education (ß=.256, t=4.779, p<.001, TP-WE; ß=-.091, t=-1.316, p=.189, TP-DI). As described on Table 14, beta weights indicated that age was the most significant contributor to the prediction of the TP scores. To this model, the adjusted R^2 value was .414, which means that 41.4% of the variance on the TP-WE was explained by the model (education explains 18,2% and age 23,2%) and on the TP-DI the value was 3% (R^2=.03).

Table 13. Correlations between TP (WE and DI) and Predictor Variables.

Variable	**WE**	**DI**	**Age**	**Education**
WE	-	-.226**	-.598**	.426**
DI		-	.153**	-.128
Age			-	-.360**
Education				-

Abbreviations: WE = Work Efficiency; DI = Dispersion Index; ** p<.01

Table 14. MLR analysis for age and education

Variable	**B**	**SEB**	**β**
WE			
Education	19.801	4.143	.256
Age	-37.032	3.877	-.511
DI			
Education	-.936	.711	-.091
Age	1.076	.665	.111

R^2 = .414; $F_{(2,233)}$ = 81.44, p<.001; R^2 = .03; $F_{(2,233)}$ = 3,253, p=.040

Abbreviations: WE = Work Efficiency; DI = Dispersion Index; SEB = standard error of B.

Preliminary normative data

Following the results of MLR, we considered age and education for the development of normative data of the TP for the Portuguese population. The normative data were calculated and stratified according to the distributional properties of each variable. We considered the subjects' education, divided into three levels (1–4 years, 5–10 years and ≥ 11 years) and three age groups: 25–49, 50–64, and over 65 years.

The TP scores are expressed as the mean ± S.D.); values below (TP-WE) and above (TP-DI) 1 S.D., 1.5 S.D. and 2 S.D. can be indicated as cut-off points for possible cognitive impairment (Tables 15 and 16).

Table 15. Preliminary normative data of the TP-WE scores according to age and education.

Work Efficiency (WE)		Education			
Age		**Primary (1-4)**	**Middle (5-10)**	**High (≥ 11)**	**All education**
25-49	***n***	9	35	76	120
	M±SD	174 ± 51	193 ± 52	208 ± 45	200.85±48.01
	SD	123, 98, 72	141, 115, 89	163, 141, 118	153, 129, 105
50-64	***n***	9	15	33	57
	M±SD	140 ± 47	168 ± 47	181 ± 49	170.84±49.94
	SD	93, 70, 46	121, 98, 74	132, 108, 83	121, 96, 71
≥ 65	***n***	25	13	19	57
	M±SD	93 ± 44	106 ± 27	140 ± 47	111.58±46.16
	SD	49, 27, 5	79, 66, 52	93, 70, 46	65, 42, 19
All age	***n***	43	63	128	234
	M±SD	119.81±56.60	169.25±57.29	190.51±51.88	171.79±60.09
	SD	63, 35, 7	112, 83, 55	139, 113, 87	112, 82, 52

Note: WE values are presented below 1 *SD*, 1.5 *SDs*, and 2 *SDs*, respectively.

Table 16. Preliminary normative data of the TP-DI scores according to age and education

Dispersion Index (DI) (%)		Education			
Age		**Primary (1-4)**	**Middle (5-10)**	**High (≥ 11)**	**All education**
25-49	***n***	9	35	76	120
	M±SD	11.30 ± 5.16	10.19 ± 5.27	11.02 ± 5.16	10.80±5.16
	SD	17, 19, 22	16, 18, 21	16, 19, 21	153, 129, 105
50-64	***n***	9	15	33	57
	M±SD	9.89 ± 6.29	13.07 ± 10.84	8.87 ± 5.71	10.13±7.54
	SD	18, 19, 23	24, 29, 35	15, 17, 20	121, 96, 71
≥ 65	***n***	25	13	19	57
	M±SD	14.88 ± 12.73	15.92 ± 17.40	11.44 ± 4.41	13.97±11.99
	SD	28, 34, 40	33, 42, 51	16, 18, 20	26, 32, 38
All age	***n***	43	63	128	234
	M±SD	13.09±10.49	12.06±10.28	10.53±5.26	11.41±8.01
	SD	24, 29, 34	22, 28, 33	16, 18, 21	19, 23, 27

Note: DI values are presented above 1 *SD*, 1.5 *SDs*, and 2 *SDs*, respectively.

V – Discussion

Generally, slight research has been developed to directly assess sustained attention in older adults. At present, it seems that sustained attention is impaired in ageing, although it may be dependent of task difficulty (Zanto & Gazzaley, 2014).

The main objective of the study 1 was to validate the TP as a measure of sustained and selective attention for MCI and AD. The results confirmed its great potential and offer strong evidence that the TP is a good instrument to distinguish the performance between healthy older adults and patients with AD and also between MCI and AD. Concerning the results obtained with the MCI group, the TP was capable of discriminating between controls and MCI patients, but its sensitivity and specificity were better for the AD group. In this case, better effect sizes for AD were also reported, corroborating the proper use of this psychological tool for evaluating this clinical condition.

Therefore, it is important to evaluate the psychometric properties of the test in order to ensure an imperative methodological rigor, to guarantee adequate results and an optimal use of the TP in clinical field (Strauss et al., 2006). We observed significant correlations between the TP and both the MMSE and the MoCA total scores, as well as with the TMT A/B and the D2 scores, which is highly suggestive of convergent validity (Cohen, 1988). Regarding test-retest reliability, the TP indexes showed high correlations between the baseline and a six-month follow-up.

Given the observed results of each index per group, an increased (or decreased, according to the specific index) tendency of means could be perceived. These results showed the expected tendency to obtain worse performances in AD group in comparision with both MCI and CG and are congruent with Perry and Hodges (1999) and Prince et al. (2013) studies: they found evidences that after an early amnesic stage in AD, attention is one of the first non-memory domains which becomes impaired, even before the emergence of deficits in language and visuospatial skills. These results are also consistent with the possibility that the problems with daily live activities, which are common in mild AD patients, may be related to attentional deficits.

Concerning the absence of significant differences between Control and MCI groups in TP-DI and omissions, this could be explained by the influence that the number of omissions have on the total score of TP-DI, once that this number is always bigger than the number of errors in both groups. The total number of omissions can yet be explained by external factors such as fatigue, the time of the day in which the assessment occurred, impulsivity, anxiety factors, distractibility and testing session, once that the MCI patients are already familiarized with testing procedures in opposite to healthy older subjects. This can be a possible explanation for the bigger number of omissions in CG comparatively with MCI's (Eysenck, 2015).

We were particularly interested in exploring the diagnostic validity of the TP in our study sample. As estimated, the results of ROC curve analysis showed a higher discriminant capacity of the TP for AD than for the MCI group. The optimal cut-off points for the MCI group were, respectively, above 27 for the TP-DI total score and below 73 for the TP-WE and below 8 for the

TP-TR. For the AD group, the optimal cut-off points were above 26 for the TP-DI and below 49 for the TP-WE and below 4 for the TP-TR, respectively. The results of ROC curve analysis also showed a high discriminat ability of the TP discriminating between MCI and AD. So, the optimal cut-off points were, respectively, above 38 for the TP-DI total score and below 37 for the TP-WE and below 3 for the TP-TR. These obtained values cannot be compared because there are no references for any international study including cut-off points for the TP concerning MCI and AD patients.

For the MCI group, the three main indexes of the TP showed a poor sensitivity of 63% (TP-WE), 48% (TP-DI) and 60% (TP-TR) but a good specificity of 85% (TP-WE) 83% (TP-DI) 78% (TP-TR). Consequently, they showed a poor classification accuracy of respectively 69% (TP-WE and TP-TR) and 65% (TP-DI). However, in AD patients, the three main indexes of TP showed a high sensitivity of 93% (TP-WE and TP-DI) and 90% (TP-TR) and a high specificity of 95% (TP-WE and TP-TR) and 83% (TP-DI). The results also showed a high classification accuracy of respectively 94% (TP-WE), 88% (TP-DI) and 93% (TP-TR). The same occurred in MCI vs AD, in which the three main indexes of TP showed a sensitivity of 85% (TP-WE), 80% (TP-DI) and 88% (TP-TR) and a high specificity of 88% (TP-WE), 83% (TP-DI) and 85% (TP-TR). The results also showed a good classification accuracy of respectively 87% (TP-WE and TP-TR), and 82% (TP-DI).

So, regarding the MCI group, the observed sensitivity and classification accuracy should be viewed as an indicator of high likelihood to have false-negative cases and the obtained cut-off points should be used carefully. Further studies should be conducted regarding the division of the MCI group in a-MCI single-domain and a-MCI multidomain in order to understand if sustained attention may be affected in multidomain MCI, which could be a marker of future conversion to AD.

Through the acquired results, we can conclude that the TP is useful for the identification of attentional deficits in AD but has less discriminative utility with milder forms of cognitive impairment.

We also evaluated the influence of the presence of depressive symptomatology in the subject's performance as well as of subjective memory complains. Our results showed that higher scores in both the SMC scale and on the GDS-30 influenced the subject's performance on some of the TP indexes. Concerning the total sample, the SMC scores are responsible for 4,4% of the variance of the TP-WE, and the GDS scores for 2,9%. Together, they explain 7,3% of this index variance. Regarding the three groups separately, the influence of these two variables was only significant on the performances of the MCI group. For the TP-WE, the GDS scores explained 5,2% of the variance and the SMC scores were responsible for 9,5% (the two variables together are responsible for 14,7% of the TP-WE results). For the TP-TR, subjective memory complains (R^2=.101) and depressive symptomatology (R^2=.052) are responsible for 15,3% of this index variance.

Subjects who revealed higher scores in the GDS-30 always performed worse in attentional measures compared with those who had non-significant scores on the depression scale (<11). The same occurs with the obtained

correlations between GDS-30 and SMC. Subjects that have higher results in GDS-30 also present several more memory complains and perform worse in the TP comparively with those who did not show raised punctuations in the depressive scale. These results are congruent with those of Nebes (2000), Peckam, McHugh & Otto (2010) and Nogueira (2016) that had already reported that patients with moderate to severe depressive symptomatology presented slight to moderate impairments in processing speed, working memory and attention abilities.

During the establishment of normative data based on the more influent sociodemographic variables, we observed that age was the most significant predictor of the results in the TP when compared with education. Together, they explain 41.4% of the variance of the results in the TP-WE and 3% of the variance of the TP-DI, which are, respectively a large and a small effect according to Cohen (1988).

As expected, our results showed that the TP-WE total scores increased with the educational level and decreased with age progression (similarly with what was pointed out by Baeta in 2002), and that the inverse occured with the TP-DI.

In another study, Montiel, Figueiredo, Lustosa and Dias (2006) found no significant differences between subjects according to the age variable on the TP-DI, and the differences found in the TP-WE were not corroborated by Tukey post-hoc tests. In addition to this, Alchieri, Lunkes and Zimmer (2002) and Araújo (2011) also found no significant differences in the TP scores according to age. These results could possibly be explained by the mean ages and the *n*'s of the age sub-groups in each of these studies. In Montiel et al. (2006) the values of mean age and *S.D.* were respectively *M*=26.76 and *S.D.*=9.04. In Alchieri et al. (2002) only 8.7% of the subjects had more than 40 years old and in Araújo (2011) only 3.4% of the sample was older than 50 years old. According to our results, the differences in sustained attention were only significant after the age of 49, so it is probable that the previous described studies did not found statistically significant differences between groups due to their young age and low percentages of older adults.

Regarding the effects of gender, according to what was pointed out by Baeta (2002) and Rebollo and Montiel (2006), this variable also did not contribute significantly to our data distribution. Finally, we determined the means and S.D for each sub-group, intersecting the several education levels and age groups. Additionally, the cut-off points of 1 S.D., 1.5 S.D.s, and 2 S.D.s were presented.

Limitations

However, some limitations must be mentioned. First, in study 1, only the amnestic subtype of MCI (single and multidomain) was included, so the generalization of the results for other types of MCI should be done carefully. Thus, in Study 2, our data can not be corroborated because the existent international studies determined their normative values based on percentiles and for non-equivalent age and education sub-groups (Alchieri et al., 2002; Araújo, 2011; Maureira et al., 2014). Furthermore, in our country there are no studies with national representative samples using the TP after the age of 45

(Baeta, 2002; do Amaral, 1967). Moreover, besides the rate of the total explained variance was 41.4% in the TP-WE, in the TP-DI was only 3%. So, we highlight the need to increase the total sample, allowing for a better stratification according to the different groups of age and education in order to provide an adequate balance of the study sample. Moreover, in the last few decades, the educational setting has quickly changed as an outcome of the restructuring of the school system and the definition of higher required educational highlands. These changes are already reflected in the younger levels of the study sample. However, the older group continues to be characterized by a lower education level. The sample distribution could not fully remove this disagreement, nevertheless, the attained partition is relatively close to the real one.

Strengths

Besides these limits, these two studies have an important set of strengths. Regarding Study 1, we believe that its added value is the rigorous and meticulous methodology used. It included homogeneous clinical groups (including an MCI group that allows to afford the knowledge about the discriminant ability of the TP within the spectrum of AD), equivalent sample-sizes reducing the probability of occurrence of biases in statistical analysis, and the overall matching between the three groups regarding age and education (the most significant sociodemographic variables influencing the TP's results). Concerning Study 2, it addresses the importance of considering age and education influences on test performances as well as provides preliminary normative data for the Portuguese population. These above-mentioned normative values are useful in both clinical and research contexts, where the TP has been increasingly used to assess attentional deficits (as well as processing speed) and for evaluating attentional abilities in the renewal of driving licenses. Moreover, although we chose to not include illiterate participants, one of the advantages of the TP is its easy application to this population group due to the absence of a verbal component.

Suggested directions for future research include further examination of the impact of test-taking strategies on overall performance measures and validation studies with other clinical groups such as patients with multiple sclerosis, epilepsy or other forms of dementia (as Frontotemporal dementia, Dementia with Lewy Bodies, Parkinson's disease dementia, or vascular dementia). Also, it is important to continue with studies focused on its psychometric properties (increasing the number of subjects in test-retest reliability), as well as to increase the normative sample in order to achieve representativeness (based on balanced matched groups according to age and education) allowing also to establish normative data based on percentiles.

VI - Conclusions

The present work demonstrates the importance of the administration of the TP in MCI and AD patients, according to its role in evaluating sustained and selective attention and processing speed deficits. The main results showed the clinical validity of the test to distinguish between healthy older adults and AD patients, presenting the cut-off points to classify the attained results.

Our main goal was to contribute to turn the TP a more useful tool in the psychological and psychometric field. Based on our results, we may consider changes in its indexes (regarding the original study of do Amaral, 1967) and add new ones, aiming to obtain a better diagnostic accuracy. Furthermore, we obtained preliminary normative data in order to guarantee a better use of the TP scores that are currently used in both clinical and research contexts, extending its utility to older adults.

References

Albert, M. S., DeKosky, S. T., Dickson, D., Dubois, B., Feldman, H. H., Fox, N. C., … Phelps, C. H. (2011). The diagnosis of mild cognitive impairment due to Alzheimer's disease: Recommendations from the National Institute on Aging and Alzheimer's Association workgroup. *Alzheimer's & Dementia*, *7*(3), 270-279. doi: 10.1016/j.jalz.2011.03.008

Alberto, I. (2003). Atenção, por favor (!) à avaliação da atenção [Attention, please (!) to the assessment of the attention]. *Psychologica, 34*, 231-243.

Alchieri, J. C., Lunkes, C. L. K., & Zimmer, D. (2002). Toulouse-Piéron: Atualizações de resultados para o estado do Rio Grande do Sul [Toulouse-Piéron: Results updates for the state of Rio Grande do Sul]. *Avaliação Psicológica*, *1*(2), 111-118.

Almeida, L. S., Simões, M. R., Almiro, P., Santos, M. J. S. (2018, Setembro). *Práticas de avaliação psicológica em Portugal: Resultados gerais de inquérito internacional realizado sobre a utilização de testes e outros instrumentos* [Psychological assessment practices in Portugal: General results of an international survey carried out on the use of tests and other instruments]. Conferência plenária apresentada no 10º Congresso da Associação Ibero-Americana de Diagnóstico e Avaliação Psicológica (AIDAP/AIDEP). Coimbra, Portugal.

Araujo, R. S. (2011). *Teste de atenção concentrada Toulouse-Pierón: Atualização dos estudos de padronização, validade e precisão* [Toulouse-Pierón attention test: Updating of standardization, validity and reliability studies] (Doctoral dissertation). Universidade de São Paulo, São Paulo, Brasil.

Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology*, *63*, 1-29. doi: 10.1146/annurev-psych-120710-100422

Baeta, Élia. (2002). Bateria para avaliação neuropsicológica de adultos com epilepsia [Battery for neuropsychological assessment of adults with epilepsy]. *Psicologia*, *16*(1), 79-96.

Banich, M. T. (2004). *Cognitive neuroscience and neuropsychology*. Boston: Houghton Mifflin College Division.

Barreto, J., Leuschner, A., Santos, F., & Sobral, M. (2008). Escala de Depressão Geriátrica [Geriatric Depression Scale]. In Alexandre de Mendonça, Manuela Guerreiro, & Grupo de Estudos de Envelhecimento Cerebral e Demências (Eds.), *Escalas e Testes na Demência* [Scales and Tests in Dementia] (2.ª ed.; pp. 69-72). Lisboa: Novartis.

Benczik, E., & Casella, E. (2007). Atenção [Attention]. In P. W. Schelini (Ed.). *Alguns domínios da avaliação psicológica* [Some domains of psychological assessment] (pp.31-58). Campinas: Editora Alínea.

Blessed, G., Tomlinson, B. E., & Roth, M. (1968). The association between quantitative measures of dementia and of senile change in the cerebral grey matter of elderly subjects. *British Journal of Psychiatry, 114*, 797–811. doi: 10.1192/bjp.114.512.797

Brickenkamp R. (1981). Test d2 Aufmerksamkeits-Belastungs-Test (7th ed.). Gottingen: Hogrefe.

Brickenkamp, R. (2007). *Teste de Atenção D2: Manual técnico* [D2 Attention Test: Technical manual] (adaptação portuguesa: Carla Ferreira e António M. Rocha). Lisboa: Cegoc.

Cavaco, S., Gonçalves, A., Pinto, C., Almeida, E., Gomes, F., Moreira, I., Fernandes, J., & Teixeira-Pinto, A. (2013b). Trail Making Test: Regression-based norms for the Portuguese population. *Archives of Clinical Neuropsychology, 28*(2), 189-198. doi:10.1093/arclin/acs115

Cavaco, S., Pinto, C., Gonçalves, A., Gomes, F., Pereira, A., & Malaquias, C. (2008). Trail Making Test: Dados normativos dos 21 aos 65 anos [Trail Making Test: Normative data from 21 to 65 years]. *Psychologica, 49*, 221-238.

Chauvin, J., Gillebert, C., Rohenkohl, G., Humphreys, W., & Nobre, C. (2016). Temporal orienting of attention can be preserved in normal aging. *Psychology and Aging, 31*(5), 442. doi: 10.1037/pag0000105

Cohen, R. J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cohen, R. A. (2011). Attention. In J. S. Kreutzer, J. DeLuca, B. Caplan (Eds.), *Encyclopedia of Clinical Neuropsychology* (pp. 282-292). New York: Springer.

Cohen, R. A. (2014). Historical and philosophical antecedents. In R. A. Cohen (Ed.), *The neuropsychology of attention* (pp. 11-17). New York, Springer.

Cunha, C., Guerreiro, M., Mendonça, A., Oliveira, P. E., & Santana, I. (2012). Serial position effects in Alzheimer's disease, mild cognitive impairment, and normal aging: Predictive value for conversion to dementia. *Journal of Clinical and Experimental Neuropsychology, 34*(8), 841-852. doi: 10.1080/13803395.2012.689814

Dalmaijer, E. S., Van der Stigchel, S., Nijboer, T. C., Cornelissen, T. H., & Husain, M. (2015). Cancellation Tools: All-in-one software for administration and analysis of cancellation tasks. *Behavior Research Methods, 47*(4), 1065-1075. doi: 10.3758/s13428-014-0522-7

do Amaral, J. R. (1967). *O teste da barragem de Toulouse e Piéron: Elementos de aferição para a população portuguesa* [The cancellation test of the Toulouse and Piéron: Standardization elements for the Portuguese population]. Lisboa: Fundação Calouste Gulbenkian.

Espírito-Santo, H., Lemos, L., Fernandes, D., Cardoso, D., Neves, C. S., Caldas, L., … Daniel, F. (2015). Teste de Stroop [Stroop test]. In Mário R. Simões, Isabel Santana, & Grupo de Estudos de Envelhecimento Cerebral e Demência (GEECD) (Coords.), *Escalas e Testes na Demência* [Scales and Tests in Dementia] (3ª ed.; pp. 110-115). Lisboa: Novartis.

Espírito-Santo, H., Lemos, L., Ventura, L., Moitinho, S., Pinto, A. L., Rodrigues, F., … Daniel, F. B. (2015). Teste da Figura Complexa de Rey-Osterrieth [Rey-Osterrieth Complex Figure Test]. In Mário R. Simões, Isabel Santana, & Grupo de Estudos de Envelhecimento Cerebral e Demência (GEECD) (Coords.), *Escalas e Testes na Demência* [Scales and Tests in Dementia] (3ª ed.; pp. 100-107). Lisboa: Novartis.

Eysenck, M. W., & Keane, M. T. (2015). Cognition and emotion. In M. W. Eysenck & M. T. Keane (Eds.), *Cognitive psychology: A student's handbook* (7th ed., pp. 648-667). New York: Guildford.

Fernandes, S. (2013). *STROOP: Teste de Cores e Palavras. Manual* [Stroop test: Manual]. Lisboa: CEGOC-TEA.

Folstein, M., Folstein, S., & McHugh, P. (1975). Mini-Mental State: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric*

Research, 12(3), 189–198. https://doi.org/10.1016/0022-3956(75)90026-6

Freitas, S., Simões, M. R., Alves, L., & Santana, I. (2011). Montreal Cognitive Assessment (MoCA): Normative study for the Portuguese population. *Journal of Clinical Experimental Neuropsycholgy, 33*(9), 989-996. doi: 10.1080/13803395.2011.589374.

Freitas, S., Simões, M. R., Alves, L., & Santana, I. (2014). Mini Mental State Examination (MMSE): Normative study for the Portuguese population in a community stratified sample. *Applied Neuropsychology: Adults, 22*(4), 311-319. doi: 10.1080/23279095.2014.926455.

Freitas, S., Simões, M.R., Alves, L., & Santana, I. (2015a). Mini Mental State Examination (MMSE). In Mário R. Simões, Isabel Santana e Grupo de Estudo de Envelhecimento Cerebral e Demência (Eds.), *Escalas e Testes na Demência* [Scales and Tests in Dementia] (3ª ed.; pp. 18-23). Porto Salvo: Novartis.

Freitas, S., Simões, M. R., Alves, L., & Santana, I. (2015b). Montreal Cognitive Assessment (MoCA). In Mário R. Simões, Isabel Santana e Grupo de Estudos de Envelhecimento Cerebral (Eds.), *Escalas e Testes na Demência* [Scales and Tests in Dementia] (3.ª ed.; pp. 24-31). Porto Salvo: Novartis.

Fuster, J. M. (2017). Prefrontal executive functions predict and preadapt. In E. Goldberg (Ed.), *Executive functions in health and disease* (pp. 7-15). New York: Academic Press.

Galhardo, J. (2012). *Estudo de validação da versão portuguesa da Escala de Avaliação da Incapacidade Funcional na Demência (Disability Assessment for Dementia – DAD)*. (Tese de Mestrado não publicada). Faculdade de Psicologia e de Ciências da Educação da Universidade de Coimbra, Coimbra.

Garcia, C. (2008). Escala de Demência de Blessed [Blessed Dementia Scale]. In Alexandre de Mendonça, Manuela Guerreiro e Grupo de Estudos de Envelhecimento Cerebral e Demência (Eds.), *Escalas e Testes na Demência* [Scales and Tests in Dementia] (2ª ed.; pp. 103-106). Lisboa: Novartis.

Garret, C., Santos, F., Tracana, I., Barreto, J., Sobral, M., & Fonseca, R. (2003). Avaliação Clínica da Demência. In A. de Mendonça, M. Guerreiro, & Grupo de Estudos de Envelhecimento Cerebral e Demência (Eds.), *Escalas e testes na demência* [Scales and Tests in Dementia] (pp. 11-26). Lisboa: Novartis.

Gauthier, L., Dehaut, F., & Joanette, Y. (1989). The bells test: A quantitative and qualitative test for visual neglect. *The International Journal of Clinical Neuropsychology, 11*(2), 49-54.

Gauthier, S., Reisberg, B., Zaudig, M., Petersen, R. C., Ritchie, K., Broich, K., ... & Cummings, J. L. (2006). Mild cognitive impairment. *The Lancet, 367*(9518), 1262-1270. doi: 10.1016/S0140-6736(06)68542-5

Gelinas, I., Gauthier, L., McIntyre, M., & Gauthier, S. (1999). Development of a functional measure for persons with Alzheimer's disease: The disability assessment for dementia. *American Journal of Occupational Therapy, 53*, 471-481. doi:10.5014/ajot.53.5.471

Ginó, S., Guerreiro, M., & Garcia, C. (2008). Escalas de Queixas de Memória (SMC). In Alexandre de Mendonça, Manuela Guerreiro e Grupo de Estudos de Envelhecimento Cerebral e Demência (Eds.), *Escalas e Testes na Demência* [Scales and Tests in Dementia] (2ª ed.; pp. 117-120). Lisboa: Novartis.

Ginó, S., Mendes, T., Mendonça, A., & Guerreiro, M. (2015). Escala de Queixas Subjectivas de Memória (QSM). In Mário R. Simões, Isabel Santana e Grupo de Estudo de Envelhecimento Cerebral e Demência (Eds.), *Escalas e Testes na Demência* [Scales and Tests in Dementia] (3ª ed; pp. 44-49). Porto Salvo: Novartis.

Guerreiro, M., Silva, A. P., Botelho, M., Leitão, O., Castro-Caldas, A., & Garcia, C. (1994). Adaptação à população portuguesa da tradução do Mini Mental State Examination [MMSE: Adaptation to Portuguese population]. *Revista Portuguesa de Neurologia, 1*, 9.

Guerreiro, M. (1998). *Contributo da neuropsicologia para o estudo das demências*. Tese de doutoramento não publicada [Contribution of neuropsychology to the study of dementia. Unpublished PhD]. Faculdade de Medicina de Lisboa, Lisboa, Portugal.

Guerreiro, M., Fonseca, S., Barreto, J., & Garcia, C. (2008). Escala de Avaliação da Doença de Alzheimer (Alzheimer Disease Assessment Scale [ADAS]). In A. de Mendonça, M. Guerreiro, & Grupo de Estudos de Envelhecimento Cerebral e Demência (Eds.), *Escalas e Testes na Demência* [Scales and Tests in Dementia] (2nd ed., pp. 41-68). Lisboa: Novartis.

Hamilton, M. (1959). The assessment of anxiety stats by rating. *British Journal of Medical Psychology, 32*, 50-55. https://doi.org/10.1111/j.2044-8341.1959.tb00467.x

Heaton, R. K. (1981). *A manual for the Wisconsin Card Sorting Test*. Odessa, FL: Psychological Assessment Resources.

Herff, S. A., & Czernochowski, D. (2017). The role of divided attention and expertise in melody recognition. *Musicae Scientiae*, 1-18. doi: 1029864917731126.

Hughes, C. P., Berg, L., Danziger, W. L., Coben, L. A., & Martin, R. L. (1982). A new clinical scale for the staging of dementia. *The British Journal of Psychiatry, 140*, 566–572. https://doi.org/10.1192/bjp.140.6.566

IBM Corp. (Released 2013). *IBM SPSS Statistics for Windows (Version 22.0).* Armonk, NY: Author.

Kaiser, A., Kuhlmann, B. G., & Bosnjak, M. (2018). A meta-analysis of inhibitory-control deficits in patients diagnosed with Alzheimer's dementia. *Neuropsychology, 32*(5), 615-633. doi: doi.org/10.1037/neu0000460

Karbach, J., & Verhaeghen, P. (2014). Making working memory work: a meta-analysis of executive-control and working memory training in older adults. *Psychological Science, 25*(11), 2027-2037. doi: 10.1177/0956797614548725

Kennedy, G. J. (2015). The problem of depression in late life. In G. J. Kennedy (Ed.), *Geriatric depression: a clinical guide* (pp. 2-18). New York: Guilford.

Kirova, A. M., Bays, R. B., & Lagalwar, S. (2015). Working memory and executive function decline across normal aging, mild cognitive impairment, and Alzheimer's disease. *BioMed Research International*, 1-9. doi: 10.1155/2015/748212

Langa, K. M., & Levine, D. A. (2014). The diagnosis and management of mild cognitive impairment: A clinical review. *JAMA, 312*(23), 2551-2561. doi: 10.1001/jama.2014.13806

Leitão, O. R. (2008). Avaliação da Incapacidade Funcional na Demência (Disability Assessment for Dementia Scale [DAD]). In A. de Mendonça, M. Guerreiro, & Grupo de Estudos de Envelhecimento Cerebral e Demência (Eds.), *Escalas e Testes na Demência* [Scales and Tests in Dementia] (2nd ed., pp. 77-98). Lisboa: Novartis.

Levitt, T., & Johnstone, B. (2001). The assessment and rehabilitation of attention disorders. In B. Johnstone & H. Stonnington (Eds). *Rehabilitation of Neuropsychological disorders: A practical guide for rehabilitation professionals* (pp. 27-52). Philadelphia: Psychology Press.

Lezak, M. D., Howieson, D. B., Bigler, E. D., & Tranel, D. (2012). Neuropsychological Assessment (5th ed.). Oxford University Press: USA.

Lunven, M., & Bartolomeo, P. (2017). Attention and spatial cognition: Neural and anatomical substrates of visual neglect. *Annals of Physical and Rehabilitation Medicine*, *60*(3), 124-129. doi: 10.1016/j.rehab.2016.01.004

Mahoney, J. R., Verghese, J., Goldin, Y., Lipton, R., & Holtzer, R. (2010). Alerting, orienting, and executive attention in older adults. *Journal of the International Neuropsychological Society*, *16*(5), 877-889. doi: 10.1017/S1355617710000767

Maureira, F., Trujillo, H., & Flores, E. (2014). Propiedades psicométricas y datos normativos del test de atención Toulouse-Piéron y del test de memoria visual de Benton forma D en estudiantes de educación física de Chile [Psychometric properties and normative data of the Toulouse-Piéron Attention Test and the Benton Form D Visual Memory Test in physical education students in Chile]. *Gaceta de Psiquiatría Universitaria, 10*(2), 238-245.

McCallum, C. (2015). Attention. *Encyclopaedia Britannica*. Retrieved from www.britannica.com/science/attention

McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., & Stadlan, E. M. (1984). Clinical diagnosis of Alzheimer's disease: Report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology*, *34*, 939–944. doi: 10.1212/WNL.34.7.939

McKhann, G. M., Knopman, D. S., Chertkow, H., Hyman, B. T., Jack Jr, C. R., Kawas, C. H., ... & Mohs, R. C. (2011). The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia*, *7*(3), 263-269. doi: 10.1016/j.jalz.2011.03.005

Mesulam, M. M. (Ed.). (1985). *Principles of behavioral neurology*. New York: Oxford University Press.

Mohs, R. C., Rosen, W. G., & Davis, K. L. (1983). The Alzheimer's disease assessment scale: An instrument for assessing treatment efficacy. *Psychopharmacology Bulletin, 19*, 448-450.

Montiel, J., Figueiredo, E., Lustosa, D., & Dias, N. (2006). Evidência de validade para o teste de atenção concentrada Toulouse-Piéron no contexto de trânsito. *Psicologia: Pesquisa & Trânsito, 2*(1), 19-27.

Montiel, J. & Capovilla, A. (2007). Avaliação da atenção: Teste de Atenção por Cancelamento [Assessment of Attention: Cancellation Attention Test]. In A. Capovilla & C. Capovilla (Eds). *Teoria e pesquisa em avaliação neuropsicológica* [Theory and research in neuropsychological assessment] (pp. 114-118). São Paulo: Memmon.

Muller, S., Rothermund, K., & Wentura, D. (2016) Relevance drives attention: Attentional bias for gain- and loss-related stimuli is driven by delayed disengagement. *The Quarterly Journal of Experimental Psychology*, *69*(4), 752-763. doi: 10.1080/17470218.2015.1049624

Nasreddine, Z., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., … Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: A Brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society, 53*(4), 695–699. doi: 10.1111/j.1532-5415.2005.53221.x

Nebes, R. D., Butters, M. A., Mulsant, B. H., Pollock, B. G., Zmuda, M. D., Houck, P. R., & Reynolds, C. F. (2000). Decreased working memory and processing speed mediate cognitive impairment in geriatric depression. *Psychological Medicine, 30*(3), 679-691. doi: 10.1017/S0033291799001968

Nogueira, J., Freitas, S., Duro, D., Tábuas-Pereira, M., Guerreiro, M., Almeida, J., & Santana, I. (2018). Alzheimer's Disease Assessment Scale-Cognitive Subscale (ADAS-Cog): Normative Data for the Portuguese Population. *Acta Medica Portuguesa*, *31*(2), 94-100, doi: 10.20344/amp.8859

Nogueira, V. (2016). Depressão e suicídio [Depression and suicide]. In H. Firmino, M. R. Simões, & J. Cerejeira (Eds.), *Saúde mental das pessoas mais velhas* [Mental health of older people] (pp. 257-266). Coimbra: Lidel.

Peckham, A. D., McHugh, R. K., & Otto, M. W. (2010). A meta-analysis of the magnitude of biased attention in depression. *Depression and anxiety*, *27*(12), 1135-1142. doi: 10.1002/da.20755

Peña-Casanova, J., Sánchez-Benavides, G., Sola, S., Manero-Borrás, R. M., & Casals-Coll, M. (2012). Neuropsychology of Alzheimer's Disease. *Archives of Medical Research, 43*, 686–693. doi: 10.1016/j.arcmed.2012.08.015

Perry, R. J., & Hodges, J. R. (1999). Attention and executive deficits in Alzheimer's disease: A critical review. *Brain*, *122*(3), 383-404. doi: 10.1.1.667.8869

Petersen, R. C., Smith, G. E., Waring, S. C., Ivnik, R. J., Tangalos, E.G., & Kokmen, E. (1999). Mild cognitive impairment: Clinical characterization and outcome. *Archives of Neurology*, *56*(3), 303-308. doi: 10.1001/archneur.56.3.303

Petersen, R. C., Roberts, R. O., Knopman, D. S., Geda, Y. E., Cha, R. H., Pankratz, V. S., … Rocca, W. A. (2010). Prevalence of mild cognitive impairment higher in men: The Mayo Clinic Study of Aging. *Neurology*, *75*, 889-897. doi: 10.1212/WNL.0b013e3181f11d85.

Petersen, S. E., & Posner, M. I. (2012). The attention system of the human brain: 20 years after. *Annual Review of Neuroscience*, *35*, 73-89. doi: 10.1146/annurev-neuro-062111-150525

Petersen, R. C., Caracciolo, B., Brayne, C., Gauthier, S., Jelic, V., & Fratiglioni L. (2014). Mild Cognitive Impairment: A concept in evolution. *Journal of Internal Medicine, 275*, 214-228. doi: 10.1111/joim.12190

Petersen, R. C. (2016). Mild cognitive impairment. *CONTINUUM: Lifelong Learning in Neurology*, *22*(2), 404-418. doi: 10.1212/CON.0000000000000313

Prince, M., Bryce, R., Albanese, E., Wimo, A., Ribeiro, W., & Ferri, C. P. (2013). The global prevalence of dementia: A systematic review and metaanalysis. *Alzheimer's & Dementia, 9*(1), 63–75. doi:10.1016/j.jalz.2012.11.007

Reitan, R. M., & Wolfson, D. (1995). Category Test and Trail Making Test as measures of frontal lobe functions. *The Clinical Neuropsychologist, 9*(1), 50-56. https://doi.org/10.1080/13854049508402057

Rebollo, M. A., & Montiel, S. (2006). Attention and the executive functions. *Revista de Neurologia*, *42*, S3-7.

Rey, A. (1959). *Manuel: Test de Copie d`une Figure Complexe*. Paris, EU: Centre de Psychologie Appliqueé.

Rey, A. (1998). *Teste de Cópia de Figuras Complexas* [Complex Figure Test] (adaptação portuguesa: A. Menezes Rocha, M. Helena Coelho). Lisboa: Cegoc.

Rizzo, M., Anderson, S. W., Dawson, J., Myers, R., & Ball, K. (2000). Visual attention impairments in Alzheimer's disease. *Neurology*, *54*(10), 1954-1959. doi: doi.org/10.1212/WNL.54.10.1954

Roberts, R. O., Geda, Y. E., Knopman, D. S., Cha, R. H., Pankratz, V. S., Boeve, B. F., … Petersen, R. C. (2012). The incidence of MCI differs by subtype and is higher in men: The Mayo Clinic Study of Aging. *Neurology, 78*, 342-351. doi: 10.1212/WNL.0b013e3182452862

Robertson, I. H., Ward, T., Ridgeway, V., & Nimmo-Smith, I. (1994). *Test of Everyday Attention*. San Antonio, TX: PsychCorp/Pearson.

Robertson, I. H., Ward, T., Ridgeway, V., & Nimmo-Smith, I. (1996). The structure of normal human attention: The Test of Everyday Attention. *Journal of the International Neuropsychological Society*, *2*, 525-534. https://doi.org/10.1017/S1355617700001697

Rock, P. L., Roiser, J. P., Riedel, W. J., & Blackwell, A. D. (2014). Cognitive impairment in depression: a systematic review and meta-analysis. *Psychological Medicine*, *44*(10), 2029-2040. doi: 10.1017/S0033291713002535

Santana, I., & Duro, D. (2014). Envelhecimento cerebral e demência [Brain aging and dementia]. In M. T. Veríssimo (Ed.), *Geriatria fundamental* [Geriatrics fundamental] (pp. 159-166). Lisboa: Lidel – Edições Técnicas.

Santana, I., Cunha, C., Duro, D., Santiago, B., Simões, M. R., Ferreira, A.R., …. Fernandes, L. (2015). Inventário Neuropsiquiátrico (NPI) [Neuropsychiatric Inventory, NPI]. In Mário R. Simões, Isabel Santana e Grupo de Estudos de Envelhecimento Cerebral e Demência (Eds.), *Escalas e Testes na Demência* (3ª edição; pp. 120-127) [Scales and Tests in Dementia, 3rd edition]. Porto Salvo: Novartis.

Saunders, N. L., & Summers, M. J. (2011). Longitudinal deficits to attention, executive, and working memory in subtypes of mild cognitive impairment. *Neuropsychology*, *25*(2), 237. doi: 10.1037/a0021134.

Schmand, B., Jonker, C., Hooijer, C., & Lindeboom, J. (1996). Subjective memory and memory complaints may announce dementia. *Neurology*, *46*, 121-125. doi: https://doi.org/10.1212/WNL.46.1.121

Shalev, L., Ben-Simon, A., Mevorach, C., Cohen, Y., & Tsal, Y. (2011). Conjunctive Continuous Performance Task (CCPT) — A pure measure of sustained attention. *Neuropsychologia*, *49*(9), 2584-2591. doi: 10.1016/j.neuropsychologia.2011.05.006

Sharma, S., Kaur, S., Tripathi, M., Talwar, A., & Sharma, R. (2017). Differential Deficits in Attention, Working and Semantic Memory Discriminates Between Mild Cognitive Impairment and Alzheimer's Disease. *Indian Journal of Physiology and Pharmacology*, *61*(4), 348-356.

Sherman, D. S., Mauser, J., Nuno, M., & Sherzai, D. (2017). The efficacy of cognitive intervention in mild cognitive impairment (MCI): A meta-analysis of outcomes on neuropsychological measures. *Neuropsychology Review*, *27*(4), 1-45. doi: 10.1007/s11065-017-9363-3

Simões, M. R., Albuquerque, C. P., Pinho, M. S., Vilar, M., Pereira, M., Lopes, A. F., Seabra-Santos, M. J., Alberto, I., Lopes, C., Martins, C., & Moura, O. (2016). *Bateria de Avaliação Neuropsicológica de Coimbra (BANC): Manual de administração e cotação* [Coimbra Neuropsychological Assessment Battery (BANC): Administration and scoring manual]. Lisboa: Cegoc.

Simões, M. R., Freitas, S., Santana, I., Firmino, H., Martins, C., Nasreddine, Z., & Vilar, M. (2008). Montreal Cognitive Assessment (MoCA). *Coimbra: Serviço de Avaliação Psicológica, Faculdade de Psicologia e de Ciências da Educação da Universidade de Coimbra*.

Simões, M. R., Prieto, G., Pinho, M. S., Sobral, M., Firmino, H., & Grupo de Estudos de Envelhecimento Cerebral e Demência. (2013). Geriatric Depression Scale (GDS-30). *Coimbra: Laboratório de Avaliação Psicológica e Psicometria, Faculdade de Psicologia e de Ciências da Educação da Universidade de Coimbra*.

Simões, M.R., Sousa, L.B., Pinho, M. S., Vilar, M., Prieto, G., & Firmino, H. (2017). Escala de Depressão Geriátrica (GDS) [Geriatric Depression Scale]. In Miguel M. Gonçalves, Mário R. Simões, Leandro S. Almeida (Coords.), *Psicologia Clínica e da Saúde: Instrumentos de Avaliação* [Clinical and Health Psychology: Assessment Tools] (pp. 219-233). Lisboa: Pactor.

Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). Attention. In E. Strauss, E. Sherman & O. Spreen (Eds). *A Compendium of Neuropsychological Tests: Administration, Norms, and Commentary* (3rd ed., pp. 546-559). New York: Oxford University Press.

Thorndike, E. L. (1912). The curve of work. *Psychological Review, 19*(3), 165-194. doi.org/10.1037/h0073541

Trenerry, M. R., Crosson, B., DeBoe, J., & Leber, W. R. (1995). *Stroop neuropsychological screening test: Manual*. Tampa: Psychological Assessment Resources.

Wang, S., & Blazer, D. G. (2015). Depression and cognition in the elderly. *Annual Review of Clinical Psychology, 11*, 331-360. doi: 10.1146/annurev-clinpsy-032814-112828

Wechsler, D. (2008). *Escala de Memória de Wechsler – Terceira edição (WMS-III): Manual de administração e cotação [Wechsler Memory Scale – (WMS-III): Administration and scoring manual]*. Lisboa, Portugal: Cegoc-Tea.

Weinbach, N., & Henik, A. (2012). The relationship between alertness and executive control. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(6), 1530. doi: 10.1037/a0027875

Wilding, J., Munir, F., & Cornish, K. (2001). The nature of attentional differences between groups of children differentiated by teacher ratings of attention and hyperactivity. *British Journal of Psychology, 92*(2), 357-371. https://doi.org/10.1348/000712601162239

Yesavage, J. A., Brink, T. L., Rose, T. L., Lum, O., Huang, V., Adey, M., & Leirer, V. O. (1983). Development and validation of a geriatric depression screening scale: A preliminary report. *Journal of Psychiatric Research, 17(*1), 37-49, doi: 10.1016/0022-3956(82)90033-4

Zanto, T., & Gazzaley, A. (2014). Attention and ageing. In A. C. Nobre & S. Kastner (Eds.), *The Oxford handbook of attention* (pp. 927-971). Oxford: Oxford University Press.