

**PRODUÇÃO DE EVIDÊNCIA DE VALIDADE PARA UM GAME-BASED
ASSESSMENT PARA AVALIAÇÃO DE SINTOMAS DE TDAH**

Universidade Presbiteriana Mackenzie

Coordenadoria de Fomento à Pesquisa - Pró-reitoria de Pesquisa e
Pós-Graduação

Julia Takieddine ¹
Apoio: PIBIC Mackenzie

Dr. Alexandre Luiz de Oliveira Serpa ²

Centro de Ciências Biológicas e da Saúde (CCBS) – Curso de Psicologia¹

Programa de pós-graduação em Ciências do Desenvolvimento Humano ²

10396144@mackenzista.com.br

alexandre.serpa@mackenzie.br

2025

RESUMO

O Transtorno de Déficit de Atenção e Hiperatividade (TDAH) é um transtorno do neurodesenvolvimento prevalente na infância e adolescência, caracterizado por desatenção, hiperatividade e impulsividade, com impactos significativos no funcionamento acadêmico, social e emocional. A avaliação neuropsicológica é fundamental para compreender os déficits cognitivos associados, como os de atenção sustentada; entretanto, métodos tradicionais tendem a ser longos, cansativos e pouco motivadores para crianças, podendo comprometer a validade dos resultados. Nesse cenário, avaliações baseadas em jogos surgem como alternativas promissoras, ao aliarem maior engajamento, brevidade e validade ecológica. O presente estudo teve como objetivo investigar propriedades psicométricas do Airship Repair, desenvolvido com base no paradigma do Teste de Desempenho Contínuo (CPT), para avaliação da atenção sustentada em crianças brasileiras com e sem sintomas de TDAH de forma lúdica e digital. Trata-se de um estudo quantitativo, transversal, composto por 09 participantes, entre 07 e 11 anos, em processo de investigação de queixas relacionadas ao TDAH. Os instrumentos utilizados foram: Airship Repair, tarefa CPT Flex (SAFE), escala MTA-SNAP-IV (pais) e questionário sociodemográfico. Os resultados preliminares não indicaram diferenças estatisticamente significativas entre os grupos definidos pela pontuação no MTA-SNAP-IV. Contudo, observou-se tendência de diferenciação nos escores do Airship Repair em relação aos grupos com sintomas de desatenção e hiperatividade/impulsividade, sugerindo potencial sensibilidade do jogo para captar variações no desempenho atencional. Conclui-se que, embora limitados pelo tamanho reduzido da amostra, os achados reforçam a viabilidade do Airship Repair como instrumento complementar na avaliação neuropsicológica infantil, apontando a necessidade de estudos com amostras mais amplas.

Palavras-chave: atenção sustentada; avaliação digital; avaliação neuropsicológica; jogos sérios; transtorno de déficit de atenção e hiperatividade

ABSTRACT

Attention Deficit Hyperactivity Disorder (ADHD) is a neurodevelopmental disorder prevalent in childhood and adolescence, characterized by inattention, hyperactivity, and impulsivity, with significant impacts on academic, social, and emotional functioning. Neuropsychological assessment is essential to understand the associated cognitive deficits, such as sustained attention; however, traditional methods tend to be long, tiring, and less engaging for children, which may compromise the validity of the results. In this context, game-based assessments emerge as promising alternatives by combining greater engagement, brevity, and ecological validity. The present study aimed to investigate the psychometric properties of *Airship Repair*, developed based on the Continuous Performance Test (CPT) paradigm, for the evaluation of sustained attention in Brazilian children with and without ADHD symptoms in a playful and digital format. This is a quantitative, cross-sectional study composed of 09 participants, aged between 7 and 11 years, undergoing neuropsychological evaluation for the investigation of complaints related to ADHD. The instruments used were: *Airship Repair*, CPT Flex task (SAFE), MTA-SNAP-IV (parents), and a sociodemographic questionnaire. Preliminary results did not indicate statistically significant differences between groups defined by MTA-SNAP-IV scores. However, a trend of differentiation was observed in *Airship Repair* scores in relation to groups with symptoms of inattention and hyperactivity/impulsivity, suggesting the game's potential sensitivity to capture variations in attentional performance. It is concluded that, although limited by the small sample size, the findings reinforce the feasibility of *Airship Repair* as a complementary instrument in child neuropsychological assessment, highlighting the need for studies with larger samples.

Keywords: sustained attention; digital assessment; neuropsychological assessment; serious games; attention deficit hyperactivity disorder

1. INTRODUÇÃO

O Transtorno de Déficit de Atenção e Hiperatividade (TDAH), conforme definido pelo Manual Diagnóstico e Estatístico de Transtornos Mentais (AMERICAN PSYCHIATRIC ASSOCIATION, 2013), caracteriza-se por um padrão persistente de desatenção e/ou hiperatividade-impulsividade que interfere significativamente no funcionamento e desenvolvimento do indivíduo. As consequências funcionais do TDAH são amplas, acarretando prejuízos em múltiplos âmbitos da vida, como maior risco de rejeição social, redução do desempenho escolar e sucesso acadêmico e, a longo prazo, maiores taxas de desemprego e propensão a comportamentos de risco, como o abuso de substâncias (GARCIA et al., 2019).

O TDAH apresenta alta prevalência na infância e adolescência, estimada entre 3% e 8% globalmente e em 7,6% para a faixa etária de 6 a 17 anos no Brasil (COMISSÃO NACIONAL DE INCORPORAÇÃO DE TECNOLOGIAS NO SUS [CONITEC], 2022). Em geral, os sintomas primários se manifestam antes dos 12 anos (AMERICAN PSYCHIATRIC ASSOCIATION, 2013). Embora usualmente associado a esse período do desenvolvimento, o transtorno frequentemente persiste na idade adulta, intensificando os desafios sociais, emocionais e profissionais associados à sua avaliação, diagnóstico e tratamento (AMERICAN PSYCHIATRIC ASSOCIATION, 2013). Diante desse cenário e das dificuldades impostas por um diagnóstico tardio, ressalta-se a importância crucial da detecção precoce e do manejo adequado como estratégias fundamentais para mitigar os impactos negativos prolongados na vida dos indivíduos.

Para um diagnóstico e tratamento eficazes do TDAH, é fundamental compreender o funcionamento neuropsicológico do indivíduo, ou seja, a interação entre seu desenvolvimento cognitivo, psicossocial e comportamental (GONÇALVES et al., 2011). Essa compreensão é frequentemente buscada por meio da avaliação neuropsicológica, processo que utiliza entrevistas e testes psicológicos para investigar a repercussão de possíveis disfunções na cognição e no comportamento, comparando o desempenho do indivíduo com normas para sua idade, sexo e escolaridade (ARGIMON; LOPES, 2017; MIRANDA, BORGES & ROCCA, 2010).

Contudo, a aplicação de instrumentos e testes neuropsicológicos tradicionais apresenta desafios significativos, especialmente em crianças e adolescentes com sintomas de TDAH. Frequentemente compostas por longas baterias de tarefas, como evidenciado por Chaytor e Schmitter-Edgecombe (2003), essas avaliações podem induzir cansaço e desmotivação,

comprometendo a qualidade e a validade dos resultados obtidos. A extensão e o formato dos instrumentos podem ser pouco ecológicos e não refletir adequadamente o funcionamento da criança em seu cotidiano. Diante dessas limitações, buscam-se abordagens alternativas que otimizem o processo avaliativo, tornando-o mais eficaz, engajador e menos exaustivo para essa população. Nesse contexto, a avaliação neuropsicológica digital e, mais recentemente, o uso de jogos sérios (*Serious Games*) e avaliações baseadas em jogos (*Game-Based Assessments - GBA*) têm emergido como alternativas promissoras (DAMAŠEVIČIUS et al., 2023; SEELOW, 2019).

Entre os domínios cognitivos frequentemente comprometidos no TDAH, a atenção sustentada destaca-se pela sua relevância clínica e pelo impacto direto em contextos acadêmicos e sociais. Tendo em vista essa importância, o presente estudo estabeleceu a atenção sustentada como foco principal de investigação. Para isso, utilizou-se o jogo digital *Airship Repair*, desenvolvido a partir do paradigma do Continuous Performance Test (CPT), amplamente empregado em avaliações neuropsicológicas, como alternativa para a mensuração desse construto em crianças.

Considerando a importância do diagnóstico precoce do TDAH, os desafios inerentes à avaliação tradicional e o potencial promissor dos *game-based assessment* (GBA), este estudo justifica-se pela necessidade de investigar a adequação do *Airship Repair* como ferramenta de avaliação complementar para crianças com suspeita ou queixa relacionada ao TDAH. Portanto, o objetivo do estudo é produzir evidências de validade e fidedignidade na avaliação baseada em jogo *Airship Repair* para avaliação de atenção sustentada em uma população de crianças com sintomas de Transtorno de Déficit de Atenção e Hiperatividade.

Objetivo geral

Produzir evidências de validade e fidedignidade na avaliação baseada em jogo *Airship Repair* para avaliação de atenção sustentada em uma população de crianças com sintomas de Transtorno de Déficit de Atenção e Hiperatividade.

2. REFERENCIAL TEÓRICO

O Transtorno de Déficit de Atenção e Hiperatividade é caracterizado por padrões de (1) desatenção, manifestada por divagação em tarefas, desorganização e dificuldade de

concentração; (2) hiperatividade, refere-se, principalmente, ao excesso de atividade motora e (3) impulsividade, relacionada a ações precipitadas sem premeditação. Diante disso, o transtorno pode se apresentar em três subtipos, sendo eles (1) Tipo Predominantemente Desatento; (2) Tipo Predominantemente Hiperativo-Impulsivo e (3) Tipo Combinado, com a ocorrência tanto dos sintomas de desatenção como sintomas de hiperatividade/impulsividade (AMERICAN PSYCHIATRIC ASSOCIATION, 2013). No que concerne aos processos cognitivos, é possível notar que o TDAH está associado diretamente a alterações nas funções executivas, aspecto mais complexo da cognição; memória e atenção (CHARACH, 2020; ROHDE et al., 2000).

O processo diagnóstico desse transtorno deve ser conduzido por uma equipe multiprofissional, adotando uma abordagem integral que esteja alinhada aos critérios estabelecidos no Manual Diagnóstico e Estatístico de Transtornos Mentais – 5ª edição (DSM 5; AMERICAN PSYCHIATRIC ASSOCIATION, 2013). Para o diagnóstico, é necessário que a criança demonstre “seis ou mais sintomas persistentes por pelo menos seis meses em um grau que é inconsistente com o nível de desenvolvimento e ter impacto negativo diretamente nas atividades sociais e acadêmicas” (AMERICAN PSYCHIATRIC ASSOCIATION, 2013), cuja presença acarreta uma série de prejuízos ao comparar a amostra clínica com aquelas tipicamente observadas em indivíduos com o mesmo nível de desenvolvimento. Além disso, o transtorno deve se manifestar em dois ou mais ambientes distintos, como escolar e familiar.

Considerando a presença de alterações cognitivas no TDAH, destaca-se a importância da avaliação neuropsicológica, definida por Cunha (1993) como o procedimento que tem por objetivo investigar as funções cognitivas e comportamentais do paciente, buscando avaliar alterações cognitivas específicas, bem como seu diagnóstico diferencial. No contexto do TDAH, a avaliação neuropsicológica contribui na caracterização do perfil atencional, executivo e de memória, permitindo compreender como esses domínios impactam a vida social, acadêmica e emocional da criança (GONÇALVES et al., 2011; MALLOY-DINIZ et al., 2018). Entretanto, instrumentos tradicionais, embora amplamente utilizados, frequentemente se mostram extensos e pouco atrativos para crianças, o que pode comprometer a validade ecológica e a motivação durante a aplicação (CHAYTOR; SCHMITTER-EDGEcombe, 2003). Diante dessas limitações, abordagens digitais e gamificadas surgem como alternativa ou complementos promissores ao processo avaliativo.

Entre os domínios cognitivos avaliados no TDAH, a atenção sustentada, definida como a capacidade de manter o foco em estímulos ou tarefas por períodos prolongados (DIAMOND,

2013; COUTINHO; MATTOS; MALLOY-DINIZ, 2018), constitui uma das funções executivas mais frequentemente afetadas. Tradicionalmente, a atenção sustentada é avaliada por meio de tarefas do tipo Teste de Desempenho Contínuo (Continuous Performance Test – CPT), reconhecidas por sua sensibilidade, mas que também apresentam limitações quanto à validade ecológica e ao engajamento dos participantes (CALLAN et al., 2024). Nesse sentido, investigações que utilizem paradigmas baseados no CPT em formatos digitais ou gamificados, como o Airship Repair, tornam-se particularmente promissoras para avaliar esse construto de maneira mais atrativa e contextualizada ao cotidiano infantil.

A utilização de jogos, sejam eles físicos ou digitais, é muito associada ao entretenimento, principalmente no mundo contemporâneo, onde o acesso à internet e a aparelhos smartphones atinge grande parte da sociedade de forma direta ou indireta. Os jogos sérios são uma classe de jogos criados com o propósito de desenvolver uma determinada habilidade, ensinar um conteúdo, realizar uma avaliação, promover uma atividade de recuperação física ou, de modo geral, possuir um objetivo implícito (WANG et al., 2022; WATTANASOONTORN et al., 2013; XAVIER et al., 2022). Eles são usados há décadas na educação e nos últimos anos o interesse pela sua adoção da área da saúde vem aumentando, servindo tanto como ferramenta auxiliar no monitoramento, na detecção de determinadas condições, no tratamento/terapia e na reabilitação, em pacientes de todas as faixas etárias (DAMAŠEVIČIUS et al., 2023; HORNE-MOYER et al., 2014; KOO; VIZER, 2019).

Os jogos digitais possuem algumas vantagens em relação aos métodos tradicionais de avaliação e intervenção tanto para o paciente quanto para o profissional. Do ponto de vista do paciente, os jogos oferecem uma abordagem que facilita o engajamento, diversão, favorecem a repetição, permitem que os indivíduos testem escolhas, risco e outros comportamentos sem consequência pessoal imediata e, por fim, a experimentação de estratégias comportamentais que possam impactar positivamente na saúde individual (BARANOWSKI et al., 2016). Do ponto de vista dos profissionais, eles permitem maior independência e padronização na aplicação de técnicas de avaliação e intervenção, o uso de informações complementares, oriundas de aparelhos de biofeedback, bem como relatórios detalhados dos resultados (BARANOWSKI et al., 2013; L'YVONNET et al., 2021).

As avaliações baseadas em jogos (*Game-Based Assessment* - GBA) são atividades que fazem parte do design de um jogo para realizar avaliações de habilidades ou competências do jogador em um contexto mais ecológico, permitindo o uso de narrativas e ambientes simulados que favorecem o processo de aprendizado; motivador, na medida em que o jogo é uma atividade

mais interessante para o indivíduo do que marcar respostas em uma folha de papel; e engajadoras para os jogadores (BAPTISTA; OLIVEIRA, 2019; BARANOWSKI et al., 2018; MISLEVY et al., 2012). Essa abordagem oferece potencial para aumentar a motivação, o engajamento e a precisão dos dados coletados em um contexto mais lúdico e ecológico (BARANOWSKI et al., 2018; MISLEVY et al., 2012; VALLADARES-RODRÍGUEZ et al., 2016).

Um exemplo de GBA desenvolvido com esse propósito é o jogo Airship Repair (SERPA, 2023). O Airship Repair está baseado no clássico protocolo do teste de desempenho contínuo (Continuous Performance Test – CPT). Esse paradigma é amplamente utilizado na avaliação psiquiátrica e neuropsicológica para mensurar a atenção sustentada, com foco na característica de vigilância. A atenção sustentada é definida como a capacidade de manter o foco por um período prolongado sem perder o desempenho na tarefa. Enquanto a vigilância é a habilidade de sustentar a atenção para detectar estímulos-alvo, mantendo a atenção mesmo em contextos de baixa estimulação (LEZAK et al., 2012). O CPT também é usado como medida para avaliar, além desses construtos, o controle inibitório (DIAMOND, 2013), que é uma habilidade presente nas funções executivas e se baseia na autorregulação para executar ações de acordo com os estímulos apresentados ou inibir respostas concorrentes.

A literatura sobre o uso do CPT no diagnóstico e na diferenciação de Transtorno de Déficit de Atenção e Hiperatividade revela que, embora o teste tenha mostrado potencial para identificar o TDAH, sua utilidade clínica geral ainda é questionada (CALLAN et al., 2024; VARELA et al., 2024). Uma ferramenta frequentemente utilizada para avaliar alterações nos processos atencionais é o Conners CPT-3, que auxilia na identificação de déficits relacionados à atenção e pode ser útil na avaliação de indivíduos com TDAH. De acordo com a revisão realizada por Callan et al. (2024), dois estudos encontraram que o teste Conners CPT-3 poderia diferenciar com eficácia entre crianças com TDAH e controles saudáveis. Já o estudo de Wang et al. (2021) concluiu que, embora tanto o Conners CPT-3 quanto o Conners Auditory Test of Attention (CATA) individualmente previsse o diagnóstico de TDAH, juntos eles forneceram uma previsão melhor do que cada um isoladamente. A combinação do Conners CPT-3 com entrevistas clínicas semi-estruturadas, como sugerido por Pettersson et al. (2018), pode melhorar a sensibilidade e a especificidade do diagnóstico de TDAH, uma vez que, por si só, o Conners CPT-3 não possui validade ecológica suficiente para diagnosticar o transtorno com precisão.

De maneira geral, os estudos indicam que os testes no formato CPT podem ser mais eficazes como parte de uma avaliação abrangente, incluindo entrevistas clínicas e outros instrumentos diagnósticos, do que como ferramentas isoladas para o diagnóstico de TDAH. Com isso, sugere-se que tarefas baseadas no paradigma CPT não devam ser usadas como ferramenta de diagnóstico, nem mesmo em um ambiente clínico, onde a prevalência do TDAH é mais alta. Tais assertivas são também válidas para tarefas com formato inovador, como o Airship Repair.

Para que ferramentas inovadoras como esta sejam integradas com segurança e eficácia à prática clínica, é indispensável que suas propriedades psicométricas sejam rigorosamente investigadas. A validade, definida como o grau em que evidências apoiam a interpretação dos escores de um teste para um propósito específico (AMERICAN EDUCATIONAL RESEARCH ASSOCIATION et al., 2014), é um aspecto central nesse processo. É necessário acumular evidências que demonstrem que as pontuações provenientes do jogo medem adequadamente o construto de interesse, seguindo recomendações de diretrizes internacionais como as da AERA e da ITC (INTERNATIONAL TEST COMMISSION & ASSOCIATION OF TEST PUBLISHERS, 2022).

METODOLOGIA

A presente pesquisa tratou-se de um estudo transversal quantitativo, voltado à investigação das propriedades psicométricas do instrumento de avaliação baseado em jogo, Airship Repair. O estudo foi desenvolvido a partir da montagem e aplicação de uma bateria de avaliação de sintomas e prejuízos de funcionalidades associados ao TDAH em crianças. A bateria foi composta pelo Airship Repair em conjunto com o Continuous Performance Test Flex (CPT-Flex), pelo MTA-SNAP-IV, bem como um questionário de informações sociodemográficas.

O estudo respeitou todos os preceitos éticos aplicáveis às pesquisas com seres humanos, conforme a Resolução nº 466/2012 do Conselho Nacional de Saúde. O projeto foi submetido e aprovado pelo Comitê de Ética em Pesquisa da Universidade Presbiteriana Mackenzie (CAAE: 87854125.4.0000.0084; Parecer nº 7.586.712). A participação foi voluntária, mediante assinatura do Termo de Consentimento Livre e Esclarecido (TCLE) pelos responsáveis e do Termo de Assentimento Livre e Esclarecido (TALE) pelas crianças, assegurando

confidencialidade, privacidade e o direito de desistência em qualquer etapa do estudo, sem prejuízos.

3.1 Participantes

A amostra foi composta por 09 crianças, com idades entre 7 e 10 anos ($M = 8,6$, $DP = 1,1$), que estavam em processo de avaliação neuropsicológica para a investigação de queixas relacionadas ao Transtorno de Déficit de Atenção e Hiperatividade (TDAH). Destes, 55,6% eram do sexo masculino, com todos os participantes pertencentes à classe social B e C, com maior frequência dos estratos B2 (33,3%) e B1 e C1 (22,2%, cada). A seleção dos participantes ocorreu por meio do contato com os pais de crianças que buscavam a avaliação neuropsicológica em clínicas parceiras. A coleta foi realizada em ambiente clínico controlado, em sessões individuais com duração aproximada de 30 minutos. Enquanto os pais responderam aos questionários, as crianças realizaram as tarefas computadorizadas Airship Repair e CPT-Flex. Os critérios de inclusão da amostra foram crianças com idade igual ou superior a 06 anos e menor ou igual a 11 anos de idade, com suspeita ou diagnóstico de TDAH. Os critérios de exclusão da amostra englobaram os participantes com menos de 06 anos e mais de 11 anos, que declararam, via relato dos pais ou responsáveis, possuir diagnóstico prévio de alguma condição neurológica que impacte a capacidade cognitiva (por exemplo, epilepsia, convulsões, tumores cerebrais, hidrocefalia, entre outros) ou que possuíam comorbidades físicas, psicológicas ou psiquiátricas que inviabilizassem a aplicação das tarefas ou dos testes (ex. ataques de pânico em vigência, esquizofrenia, entre outros).

3.2 Instrumentos

3.2.1 Questionário Sociodemográfico

O questionário sociodemográfico, elaborado para esta pesquisa, coletou informações básicas dos participantes. A classificação socioeconômica foi obtida com base no Critério de Brasil da Associação Brasileira de Empresas de Pesquisa - ABEP (2024), que avalia o poder de consumo das famílias a partir da posse de bens, nível de escolaridade do chefe da família e acesso a serviços públicos. A classificação socioeconômica seguiu os pontos e cortes estabelecidos pela ABEP (2024), permitindo enquadrar as famílias nas classes A, B1, B2, C1, C2 e DE.

3.2.2 Airship Repair

O Airship Repair é um game-based assessment cujo propósito é avaliar o funcionamento da habilidade de atenção sustentada dos jogadores. Nele, o examinando deve ajudar a mecânica Tray a consertar um dirigível. Para isso, ela solicita que o examinando clique na barra de espaço sempre que a ferramenta que ela precisa aparecer no centro da tela. Essa ferramenta é o estímulo alvo e fica indicada em um quadro do lado esquerdo superior da tela durante toda a tarefa. A duração da tarefa é de 5 minutos e nela são utilizadas quatro figuras simples (mochila, martelo, chave de fenda e serrote) como estímulo. Em cada aplicação um dos estímulos é definido aleatoriamente como estímulo alvo. Os estímulos são apresentados aleatoriamente, de modo que o total de estímulos alvos em cada sessão não é conhecido previamente. Do mesmo modo, o intervalo de apresentação dos estímulos é variável, sendo definido pelo algoritmo no momento da aplicação dentro de um intervalo definido no momento do desenvolvimento da tarefa. Essa configuração está baseada no protocolo clássico Continuous Performance Test (CPT), comumente utilizado para avaliação da atenção concentrada. A avaliação do desempenho é realizada por meio do tempo de resposta médio dos examinandos e diferentes padrões de respostas estão associados a condições ou transtornos clínicos, como TDAH ou impulsividade. Em cada sessão do Airship Repair, o jogador deve clicar na tela ou no botão do mouse assim que o estímulo alvo aparecer na tela e não deve clicar quando outro estímulo ou nenhum estímulo estiver apresentado. O Airship Repair possui evidências de validade relacionadas à critério, com idade, escolaridade e classificação socioeconômica e evidências de validade convergente com o teste Stroop. As medidas produzidas pela avaliação são o número de estímulos apresentados, o número de estímulos-alvo apresentados, o número de estímulos clicados corretamente, incorretamente, não clicados e os cliques quando nenhum estímulo estava na tela.

3.2.3 Bateria Soluções para Avaliação das Funções Executivas (SAFE, ESQUARANTE-BARBOSA et al., 2025):

Desenvolvida no Brasil e possui sete tarefas para avaliação das funções executivas, em formato digital. Para o presente estudo se propõe o uso da tarefa CPT-Flex. Ele consiste na apresentação de diferentes figuras geométricas (quadrado, triângulo, círculo e um X), exibidas uma de cada vez, em intervalos variáveis de 1, 2 ou 4 segundos. O examinando deve pressionar

a barra de espaço do teclado sempre que o estímulo-alvo surgir na tela, e não pressionar a mesma barra quando outro estímulo surgir ou quando a tela estiver vazia. O estímulo-alvo é determinado pela cor das figuras: se a cor estiver cinza, as figuras “quadrado”, “triângulo” e “círculo” são o alvo; se laranja, o alvo é a letra “X”. As cores foram escolhidas por serem distinguíveis por examinandos com qualquer dos três tipos de daltonismo (protanopia, deutanopia e tritanopia), tornando o teste acessível a essas pessoas. O CPT-Flex avalia a atenção sustentada, a vigilância, o controle inibitório e a flexibilidade cognitiva, habilidades que abrangem as funções executivas. A confiabilidade teste-reteste foi moderada, com coeficientes de correlação variando de 0,51 a 0,74. A validade convergente foi confirmada através de correlações positivas com outros testes de atenção, como o AOL e o FDT. O teste é aplicado via plataforma VOL Vetor Online (<https://vol.vetoreditora.com.br/>) e leva cerca de 15 minutos.

3.2.4 MTA-SNAP-IV (MATTOS; SERRA-PINHEIRO; ROHDE; PINTO, 2006):

O MTA-SNAP-IV é a versão curta da escala SNAP-IV, desenvolvida para avaliar os sintomas centrais de TDAH conforme estabelecido pelo DSM-IV-TR. Possui 26 itens em formato de escala Likert com 4 opções. É respondida pelos pais da criança e permite a produção de medidas para hiperatividade-impulsividade, desatenção e comportamentos desafiadores de oposição. Costa et al. (2019) encontraram evidências favoráveis em relação a estrutura fatorial e fidedignidade de escala para a população brasileira. Os resultados indicaram ainda evidências favoráveis de validade por meio das correlações com a escala K-SADS-PL e ressaltaram ainda a capacidade da escala na predição de diagnóstico de TDAH.

3.3 Análise de dados

Os dados foram submetidos a análises estatísticas descritivas para caracterizar a amostra e as variáveis dos instrumentos. Devido ao tamanho da amostra clínica, não foi possível realizar comparações com propósito inferencial entre eles. Contudo, as diferenças de médias foram calculadas com o propósito de avaliar tendências potenciais entre os participantes. Assim, foram avaliadas diferenças de médias nas medidas dos instrumentos, por meio do teste t, comparando os grupos com indicativos de depressão e hiperatividade/impulsividade definidos com base na pontuação no MTA-SNAP-IV, e também diferenças sociodemográficas, em relação ao sexo.

As análises estatísticas foram conduzidas no software Jamovi, versão 2.6 (THE JAMOV PROJECT, 2024).

3. RESULTADO E DISCUSSÃO

Devido ao número de crianças com suspeita de TDAH recrutadas, serão apresentadas as análises preliminares realizadas com as medidas do jogo sério Airship Repair e da tarefa Continuous Performance Test Flex. Os dados descritivos das medidas dos instrumentos para toda a amostra estão apresentados na Tabela 1. Em relação ao MTA-SNAP-IV, foi encontrada uma maior proporção de crianças com indicativos de desatenção e/ou hiperatividade/impulsividade do que o esperado, por se tratar de uma amostra clínica, com grande variabilidade de pontuação na amostra para ambas medidas. Da amostra, 66,7% apresentaram indicação para Desatenção e 55,6% para Hiperatividade/Impulsividade. No Airship Repair os acertos tiveram uma média de 47,3 pontos, variando entre 21 e 73 acertos. Isso indica que os participantes tiveram uma grande variabilidade no desempenho na atividade. Além disso, os erros de omissão e comissão foram altos, com média de 21,3 e 42,2 pontos, respectivamente. No CPT-Flex, os resultados indicaram também um alto número de omissões e comissões, com média de 27,0 e 10,8 pontos, respectivamente, um tempo médio de reação de 489,5 milissegundos e um acerto médio de 162,2 pontos.

Tabela 1 – Estatísticas descritivas das medidas dos instrumentos na amostra total

Instrumento/Medida	N = 9
MTA-SNAP-IV	
Desatenção	
Média (DP)	6.3 (2.3)
Amplitude	2.0 - 9.0
Hiperatividade/Impulsividade	
Média (DP)	5.3 (3.0)
Amplitude	1.0 - 9.0
Indicativo Desatenção	
Sim	6 (66.7%)
Não	3 (33.3%)
Indicativo Hiperatividade/Impulsividade	
Sim	5 (55.6%)
Não	4 (44.4%)
Airship Repair	
Total Hits	

Média (DP)	47.3 (22.7)
Amplitude	21.0 - 73.0
Target Stimuli Shown	
Média (DP)	79.2 (10.4)
Amplitude	62.0 - 92.0
Targets Not Clicked	
Média (DP)	21.3 (18.4)
Amplitude	1.0 - 42.0
Total Stimuli Shown	
Média (DP)	314.7 (10.4)
Amplitude	304.0 - 328.0
Total Clicks Wrong Stimuli	
Média (DP)	42.2 (25.4)
Amplitude	14.0 - 72.0
Total Clicks Stimuli Off	
Média (DP)	4.3 (3.1)
Amplitude	0.0 - 8.0
CPT_Flex	
Right Answers	
Média (DP)	162.2 (19.2)
Amplitude	135.0 - 180.0
Omission Errors	
Média (DP)	27.0 (18.7)
Amplitude	10.0 - 53.0
Commission Errors	
Média (DP)	10.8 (3.0)
Amplitude	7.0 - 14.0
Median Reaction Time	
Média (DP)	489.5 (56.6)
Amplitude	425.0 - 563.0

Na Tabela 2, há a análise dos resultados por medida do instrumento Airship Repair e CPT-Flex para a variável sexo. Os resultados sugerem que não houve diferenças significativas de desempenho entre meninos e meninas na maioria das medidas dos instrumentos. A taxa de acertos, a quantidade de estímulos exibidos e os erros de comissão foram muito semelhantes entre os grupos. No entanto, uma diferença estatisticamente significativa foi observada na variável "Total Clicks Stimuli Off" do Airship Repair, onde a média de erros cometidos por meninos foi maior que a das meninas. Isso indica uma tendência a que os meninos pressionem a tela em momentos em que não há estímulos, em comparação com as meninas, o que poderia refletir um perfil de maior impulsividade.

Tabela 2 – Estatísticas descritivas e teste t das medidas do Airship Repair e do CPT-Flex em relação ao sexo

Variável Dependente: Sexo		M	F	Total	p
		N (%)	5 (55.6)	4 (44.4)	9
Airship Repair: Total Hits	Média (DP)	47.5 (20.2)	47.0 (36.8)	47.3 (22.7)	0.983
Airship Repair: Targets Not Clicked	Média (DP)	23.8 (19.4)	16.5 (21.9)	21.3 (18.4)	0.698
Airship Repair: Total Stimuli Shown	Média (DP)	314.0 (9.1)	316.0 (17.0)	314.7 (10.4)	0.852
Airship Repair: Total Clicks Wrong Stimuli	Média (DP)	43.5 (25.2)	39.5 (36.1)	42.2 (25.4)	0.878
Airship Repair: Total Clicks Stimuli Off	Média (DP)	6.2 (1.3)	0.5 (0.7)	4.3 (3.1)	0.004
CPT-Flex: Right Answers	Média (DP)	168.0 (NA)	160.3 (23.0)	162.2 (19.2)	0.800
CPT-Flex: Omission Errors	Média (DP)	18.0 (NA)	30.0 (21.7)	27.0 (18.7)	0.679
CPT-Flex: Commission Errors	Média (DP)	14.0 (NA)	9.7 (2.5)	10.8 (3.0)	0.274
CPT-Flex: Median Reaction Time	Média (DP)	484.0 (NA)	491.3 (69.2)	489.5 (56.6)	0.935

Ao considerar a divisão dos participantes em grupos a partir das subescalas do MTA-SNAP-IV (desatenção e hiperatividade/impulsividade), nota-se variações nos padrões de desempenho (Tabela 3). Crianças com maiores escores de desatenção apresentaram uma tendência a maior taxa de acertos e menor número de omissões no jogo, isto é, foram mais assertivas em identificar o estímulo-alvo e menos propensas a falhar na resposta. Em avaliações baseadas no paradigma do CPT, erros de omissão são associados a dificuldades em manter a atenção de forma sustentada por todo o período da atividade. Portanto, no Airship Repair, a menor frequência de omissões no grupo com indicativo de desatenção é um achado que merece mais investigação, pois contraria os lapsos atencionais típicos desse perfil.

Já o grupo com indicação predominante de hiperatividade/impulsividade apresentou menor frequência de erros de comissão, ou seja, respostas emitidas em momentos inadequados - como pressionar a tecla quando um estímulo não era o alvo ou quando nenhum estímulo estava presente. Esse padrão também se contrapõe à literatura sobre o desempenho em tarefas derivadas do CPT, onde se esperaria maior impulsividade e, consequentemente, mais erros de comissão.

Nesse sentido, os dados preliminares sugerem que o Airship Repair, com base nestes achados, não reproduz os padrões esperados de desempenho, mesmo em uma amostra pequena. As análises inferenciais, entretanto, não revelaram diferenças estatisticamente significativas

entre os grupos, resultado que deve ser compreendido diante do tamanho amostral reduzido. Com apenas nove participantes, o estudo não apresenta poder estatístico suficiente para detectar efeitos sutis entre os grupos. No entanto, os achados apontam para uma tendência de diferenciação entre perfis atencionais em crianças, indicando o potencial do Airship Repair como ferramenta complementar de avaliação.

Tabela 3 – Estatísticas descritivas e teste t das medidas do Airship Repair e do CPT-Flex para cada indicativo do MTA-SNAP-IV

Variável Dependente: Indicativo de desatenção		Sim	Não	Total	p
	N (%)	6 (66.7)	3 (33.3)	9	
Airship Repair: Total Hits	Média (DP)	53.3 (25.1)	41.3 (23.5)	47.3 (22.7)	0.578
Airship Repair: Targets Not Clicked	Média (DP)	16.3 (20.0)	26.3 (19.1)	21.3 (18.4)	0.566
Airship Repair: Total Stimuli Shown	Média (DP)	316.7 (11.0)	312.7 (11.7)	314.7 (10.4)	0.689
Airship Repair: Total Clicks Wrong Stimuli	Média (DP)	47.0 (24.6)	37.3 (30.6)	42.2 (25.4)	0.692
Airship Repair: Total Clicks Stimuli Off	Média (DP)	4.7 (4.2)	4.0 (2.6)	4.3 (3.1)	0.826
CPT-Flex: Right Answers	Média (DP)	150.5 (21.9)	174.0 (8.5)	162.2 (19.2)	0.293
CPT-Flex: Omission Errors	Média (DP)	40.0 (18.4)	14.0 (5.7)	27.0 (18.7)	0.196
CPT-Flex: Commission Errors	Média (DP)	9.5 (3.5)	12.0 (2.8)	10.8 (3.0)	0.517
CPT-Flex: Median Reaction Time	Média (DP)	524.5 (54.4)	454.5 (41.7)	489.5 (56.6)	0.286
Variável Dependente: Indicativo de Hiperatividade/Impulsividade		Sim	Não	Total	p
	N (%)	5 (55.6)	4 (44.4)	9	
Airship Repair: Total Hits	Média (DP)	23.0 (2.8)	59.5 (16.3)	47.3 (22.7)	0.041
Airship Repair: Targets Not Clicked	Média (DP)	35.5 (4.9)	14.2 (18.8)	21.3 (18.4)	0.210
Airship Repair: Total Stimuli Shown	Média (DP)	305.0 (1.4)	319.5 (9.3)	314.7 (10.4)	0.107
Airship Repair: Total Clicks Wrong Stimuli	Média (DP)	16.5 (3.5)	55.0 (20.3)	42.2 (25.4)	0.065
Airship Repair: Total Clicks Stimuli Off	Média (DP)	4.5 (4.9)	4.2 (2.9)	4.3 (3.1)	0.938
CPT-Flex: Right Answers	Média (DP)	160.3 (23.0)	168.0 (NA)	162.2 (19.2)	0.800
CPT-Flex: Omission Errors	Média (DP)	30.0 (21.7)	18.0 (NA)	27.0 (18.7)	0.679

CPT-Flex: Commission Errors	Média (DP)	9.7 (2.5)	14.0 (NA)	10.8 (3.0)	0.274
CPT-Flex: Median Reaction Time	Média (DP)	491.3 (69.2)	484.0 (NA)	489.5 (56.6)	0.935

4. CONCLUSÃO

O presente estudo buscou investigar as propriedades psicométricas do Airship Repair, um game-based assessment voltado à avaliação da atenção sustentada em crianças com suspeita ou diagnóstico de TDAH. Os resultados preliminares, embora não tenham indicado diferenças estatisticamente significativas entre os grupos, em virtude do tamanho amostral reduzido, sugerem uma tendência de diferenciação nas pontuações do jogo em relação aos perfis de desatenção e de hiperatividade/impulsividade.

As crianças com maiores escores de desatenção apresentaram indícios de maior frequência de omissões, enquanto aquelas com predominância de hiperatividade/impulsividade demonstraram tendência a mais erros de comissão. Esses padrões, consistentes com o que é relatado na literatura acerca do paradigma do Continuous Performance Test (CPT), reforçam que o Airship Repair é capaz de replicar os padrões observados em tarefas do tipo CPT, diferenciando lapsos de atenção sustentada e de respostas impulsivas.

Dessa forma, mesmo em caráter preliminar, os achados apontam para o potencial do Airship Repair como instrumento complementar de avaliação neuropsicológica, capaz de captar manifestações cognitivas associadas ao TDAH de forma lúdica e motivadora. Para estudos futuros, recomenda-se a ampliação da amostra e o refinamento das análises estatísticas, de modo a confirmar a validade e a fidedignidade do instrumento em diferentes contextos clínicos e educacionais.

5. REFERÊNCIAS

- AERA - AMERICAN EDUCATIONAL RESEARCH ASSOCIATION; AMERICAN PSYCHOLOGICAL ASSOCIATION; NATIONAL COUNCIL ON MEASUREMENT IN EDUCATION. *Standards for educational and psychological testing*. Washington, DC: AERA, 2014. Disponível em: <https://www.testingstandards.net/uploads/7/6/6/4/76643089/9780935302356.pdf>
- APA - AMERICAN PSYCHIATRIC ASSOCIATION. *Diagnostic and statistical manual of mental disorders*. 5. ed. Washington, DC: American Psychiatric Publishing, 2013.
- ARGIMON, I. D. L.; LOPES, R. M. F. Avaliação neuropsicológica Infantil: aspectos históricos, teóricos e técnicos. In: TISSER, L. (org.). *Avaliação neuropsicológica infantil*. Novo Hamburgo: Sinopsys, 2017.
- ASSOCIAÇÃO BRASILEIRA DE EMPRESAS DE PESQUISA – ABEP. Critério de Classificação Econômica Brasil. São Paulo: ABEP, 2024. Disponível em: https://abep.org/wp-content/uploads/2024/09/01_cceb_2024.pdf. Acesso em: ago. 2025.
- BAPTISTA, G.; OLIVEIRA, T.. Gamification and serious games: A literature meta-analysis and integrative model. *Computers in human behavior*, v. 92, p. 306-315, 2019. <https://doi.org/10.1016/j.chb.2018.11.030>
- BARANOWSKI, M. T. et al.. Business models for successfully maintaining games for health. *Games for Health Journal*, v. 2, n. 2, p. 64-69, 2013. <https://doi.org/10.1186/s40064-016-1837-x>
- BARANOWSKI, T. et al. Games for Health for Children—Current Status and Needed Research. *Games for Health Journal*, v. 5, n. 1, p. 1–12, 2016.
- BARANOWSKI, M. T. et al. Playing for real: Video games and stories for health-related behavior change. *American Journal of Preventive Medicine*, 34(1), 74–82.e10, 2018. <https://doi.org/10.1016/j.amepre.2007.09.027>
- BITTENCOURT, R.; ALMEIDA, S. O. et al. Adoption of technological products and services for different consumer profiles: an analysis using the technology adoption propensity scale. *Review of Business Management*, v. 21, n. 01, p. 86-102, 2019. <https://doi.org/10.7819/rbgn.v0i0.3955>.
- BRASIL. Conselho Nacional de Saúde. Resolução nº 466, de 12 de dezembro de 2012. Aprova as diretrizes e normas regulamentadoras de pesquisas envolvendo seres humanos. Brasília, DF, 2012.
- CALLAN, P. D. et al. Diagnostic utility of conners continuous performance test-3 for attention deficit/hyperactivity disorder: a systematic review. *Journal of Attention Disorders*, v. 28, n. 6, p. 992-1007, 2024. <https://doi.org/10.1177/10870547231223727>
- CHARACH, A. *Children with Attention Deficit Hyperactivity Disorder: Epidemiology, Comorbidity and Assessment*. In: TREMBLAY, R. E.; BOIVIN, M.; PETERS, R. DeV. (eds.). SCHACHAR, R. (topic ed.). *Encyclopedia on Early Childhood Development* [online], 2020. Disponível em <https://www.child-encyclopedia.com/hyperactivity-and-inattention-adhd/according-experts/children-attention-deficit-hyperactivity>
- CHAYTOR, Naomi; SCHMITTER-EDGECOMBE, Maureen. The ecological validity of neuropsychological tests: A review of the literature on everyday cognitive skills. *Neuropsychology review*, v. 13, n. 4, p. 181-197, 2003.

- COMMISSION, I. T.; PUBLISHERS, A. OF T. *Guidelines for Technology-Based Assessment*, 2022. Disponível em: <https://www.intestcom.org/page/16>
- COSTA, D. S. et al. Parent SNAP-IV rating of attention-deficit/hyperactivity disorder: accuracy in a clinical sample of ADHD, validity, and reliability in a Brazilian sample. *Jornal de pediatria*, v. 95, n. 6, p. 736-743, 2019.
- COUTINHO, G. et al. Atenção. In: MALLOY-DINIZ, L. F.; FUENTES, D.; MATTOS, P.; ABREU, N. (Orgs.). *Avaliação Neuropsicológica 2*. ed. Porto Alegre: Artmed, 2018.
- CUNHA, J. Psicodiagnóstico-R. 4. ed. Porto Alegre, RS: Artes Médicas, 1993
- DAMAŠEVIČIUS, R.; MASKELIŪNAS, R.; BLAŽAUSKAS, T. Serious games and gamification in healthcare: a meta-review. *Information*, v. 14, n. 2, p. 105, 2023.
- DIAMOND, Adele. Executive functions. *Annual review of psychology*, v. 64, n. 1, p. 135-168, 2013.
- ESQUARANTE-BARBOSA, R. M.; SERPA, A. L. O.; SOARES, G. P.; SILVA, L. E.; MALLOY-DINIZ, L. F. *CPT-Flex: Manual técnico*. São Paulo: Vetor Editora, 2025.
- GARCIA T, et al. Perfis psicossociais de adultos com TDAH: um estudo comparativo de amostras psiquiátricas prisionais e ambulatoriais. *Revista Europeia de Psicologia Aplicada ao Contexto Jurídico*, v. 11, n. 41, p. 49, 2019.
- GONÇALVES, H. A.; PUREZA, J. R.; PRANDO, M. L. Transtorno de déficit de atenção e hiperatividade: breve revisão teórica no contexto da neuropsicologia infantil. *Revista Neuropsicologia Latinoamericana*, v. 3, n. 3, p. 20-24, 2011.
- HORNE-MOYER, H. L. et al. The use of electronic games in therapy: a review with clinical implications. *Current psychiatry reports*, v. 16, n. 12, p. 520, 2014.
<https://doi.org/10.1007/s11920-014-0520-6>
- INTERNATIONAL TEST COMMISSION; ASSOCIATION OF TEST PUBLISHERS (Orgs.). *Guidelines for Technology-Based Assessment*. 2022. p. 1–173. .
- THE JAMOVI PROJECT. *jamovi*. (Version 2.6) [Computer Software]. 2024. Retrieved from <https://www.jamovi.org>.
- KOO, B. M.; VIZER, L. M. Mobile technology for cognitive assessment of older adults: a scoping review. *Innovation in aging*, v. 3, n. 1, p. igy038, 2019.
<https://doi.org/10.1093/geroni/igy038>
- LEZAK, M. D.; HOWIESON, D. B.; BIGLER, E. D.; TRANEL, D. *Neuropsychological assessment*. 5. ed. New York: Oxford University Press, 2012.
- L'YVONNET, T. et al. Serious mini-games for cognitive function assessment. *Alzheimer's & Dementia*, v. 17, p. e057775, 2021.
- MALLOY-DINIZ, L. F. et al. *Avaliação Neuropsicológica-2*. ed. Porto Alegre: Artmed, 2018.
- MATTOS, P. et al. Apresentação de uma versão em português para uso no Brasil do instrumento MTA-SNAP-IV de avaliação de sintomas de transtorno do déficit de atenção/hiperatividade e sintomas de transtorno desafiador e de oposição. *Revista de psiquiatria do Rio Grande do Sul*, v. 28, p. 290-297, 2006.

MIRANDA, M. C.; BORGES, M.; ROCCA, C. C. de A. Avaliação neuropsicológica infantil. In: MALLOY-DINIZ, L. F.; FUENTES, D.; MATTOS, P.; ABREU, N. (Orgs.). *Avaliação neuropsicológica*. Porto Alegre: Artmed, 2010.

MISLEVY, R. J. et al. Three things game designers need to know about assessment. In: *Assessment in game-based learning: Foundations, innovations, and perspectives*. New York, NY: Springer New York, 2012.

PETTERSSON, R.; SÖDERSTRÖM, S.; NILSSON, K. W. Diagnosing ADHD in adults: An examination of the discriminative validity of neuropsychological tests and diagnostic assessment instruments. *Journal of attention disorders*, v. 22, n. 11, p. 1019-1031, 2018. <https://doi.org/10.1177/1087054715618788>

ROHDE, L. A. et al. Transtorno de déficit de atenção/hiperatividade. *Brazilian Journal of Psychiatry*, v. 22, p. 07-11, 2000.

SEELow, D.. The art of assessment: Using game based assessments to disrupt, innovate, reform and transform testing. *Journal of Applied Testing Technology*, p. 1-16, 2019.

VALLADARES-RODRÍGUEZ, S. et al. Trends on the application of serious games to neuropsychological evaluation: A scoping review. *Journal of biomedical informatics*, v. 64, p. 296-319, 2016.

VARELA, J. L. et al. A systematic review of the utility of continuous performance tests among adults with ADHD. *The Clinical Neuropsychologist*, v. 38, n. 7, p. 1524-1585, 2024.

WANG, L. J. et al. Validity of visual and auditory attention tests for detecting ADHD. *Journal of attention disorders*, v. 25, n. 8, p. 1160-1169, 2021.

WANG, Y. et al. Application of serious games in health care: scoping review and bibliometric analysis. *Frontiers in Public Health*, v. 10, p. 896974, 2022.

WATTANASOONTORN, V. et al. Serious games for health. *Entertainment Computing*, v. 4, n. 4, p. 231-247, 2013. <https://doi.org/10.1016/j.entcom.2013.09.002>

XAVIER, A. et al. Children's perspectives on using serious games as a complement to promoting their social-emotional skills. *International Journal of Environmental Research and Public Health*, v. 19, n. 15, p. 9613, 2022.