TALi DETECT: a serious game for the assessment of early childhood attention

SIMONE GINDIDIS & AZADEH FEIZPOUR

Attention encompasses an individual's capacity to sustain, direct, maintain, control their focus and behaviour. Appraisal of attentional skills are an important component of comprehensive childhood assessments where concerns for learning or behaviour are present as attention is strongly associated with the ability to learn, regulate emotions, socialise, and achieve (Diamond, 2013; Pondé et al., 2012; Ribner, 2020). Serious games offer clinicians an innovative method with which to assess attention development in young children that can increase motivation and engagement (Fleming et al., 2017; Khaleghi et al., 2021; Lumsden et al., 2016). Despite significant potential, adoption of serious games by clinicians is complicated by the need for tools that are informed by robust theory, have strong psychometric properties, are designed for the correct population with recent norms, and that are supported by rigorous research (American Psychological Association, 2020, 2021; Australian Psychological Society, 2018; Pratt et al., 2018). This benchmark for design and evidence is rarely met; resulting in few digital tools qualifying for clinical use (Byambasuren et al., 2019; Lui et al., 2017; Marshall et al., 2020; Torous et al., 2018; Wang et al., 2018).

The potential impact of digital health tools is more evident now in the wake of the COVID-19 pandemic than ever before. Digital health tools

such as serious games offer interactive, scalable resources that can be accessed in-person or remotely (Torous et al., 2020). In addition to the lack in availability of scientifically robust tools, real-world engagement with digital health tools such as mental health apps remains a challenge (Baumel & Kane, 2018; Gilbody et al., 2015; Torous et al., 2020). Given emerging literature indicates the quality and design of digital tools influence real-world user engagement (Baumel & Kane, 2018), high quality and evidence-based serious games may enable more accurate monitoring of client progress as well as adherence to intervention (Patel & Butte, 2020).

TALi DETECT (DETECT) is Software as a Medical Device (SaMD) successfully navigating the research-practice divide. TALi DETECT is a scientifically validated serious game designed to assess cognitive attention development in children aged 4-7 years currently used by Australian clinicians. The program assesses children's abilities across three core attention networks: selective, sustained, and executive attention (Fan et al., 2002). Children complete a series of gamified subtests (games) within set time or level limits and the games are designed based on established, standardised assessments of attention and executive function. The games are embedded in an imaginative storyline narrated by a character (avatar) that children select at the start of the game. Children are guided through the TALi world by their avatar and together they embark on an adventure that involves helping various animals and creatures in the game to complete tasks. Assessing children using these gamified, engaging stories is a basic form of stealth assessment (Georgiadis et al., 2018; Wang et al., 2015). Compared with traditional assessments, children lose sight of the fact they are being assessed; instead immersed in a fun gaming environment resulting in covert measurement of cognitive attention skills in a naturalistic setting.

Some of the TALi DETECT games are simple and easy, tapping into less sophisticated constructs such as one's visual search ability. Other games are more difficult to complete due to increased cognitive demand on more sophisticated constructs, such as one's executive function skills (inhibitory control and cognitive flexibility).

Image 1: Welcome to TALi DETECT

Feed Whizz is one of the TALi DETECT games that has been designed to assess a child's efficiency in visual information processing. This is accomplished by measuring response speed to randomly appearing visual targets. *Feed Whizz* is a gamified task based on the Simple Reaction Time paradigm in cognitive neuroscience literature (Stebbins, 2007). While children are playing this game, their responses enable exploration of the psychological factors underlying their stimulus-response behaviour and motor speed. Translating this task to a serious game enables young children to both understand the requirement of the test, and to complete the test without adult involvement with help from the avatar narrating the story.

The second and third games —*Look for Lobsters* and *Scan the Sky* — are designed to assess how quickly a child can locate targets in a crowded visual scene. Children search for a particular object amongst a complex array of objects enabling the evaluation of attention skills related to a child's visual search efficiency.
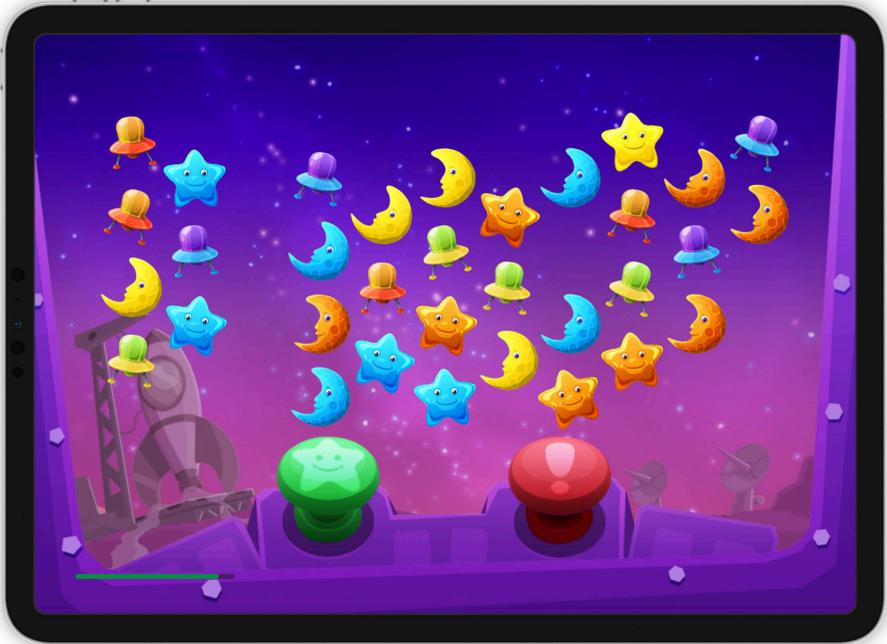
Image 2: TALi DETECT Visual search subtest: Scan the Sky.

Find the Frog is the fourth TALi DETECT game and is designed to assess how effectively a child can monitor tasks and divide their attention. Children monitor two events at the same time and make a considered response when prompted by the game. Performance on this task enables evaluation of visual search, vigilance, and divided attention skills.
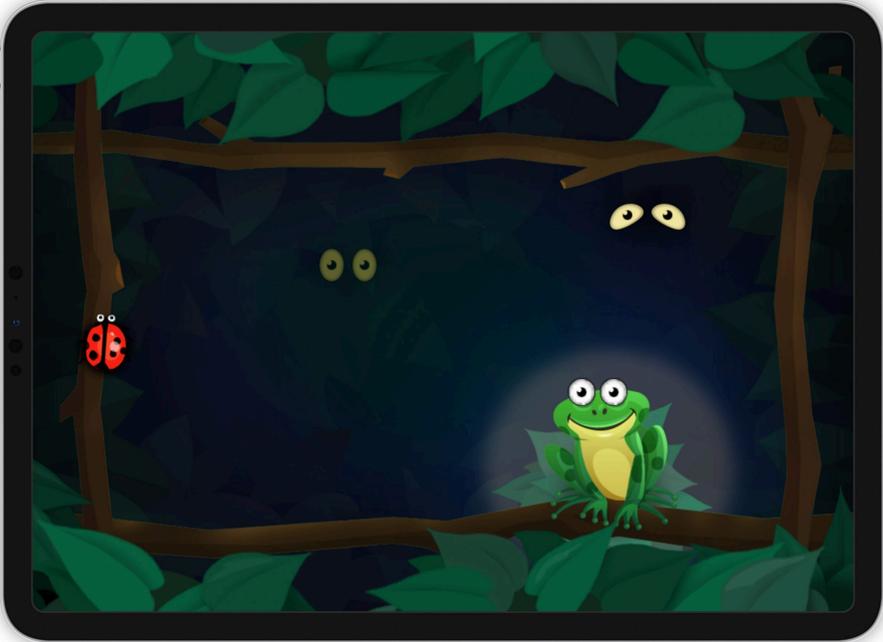
Image 3: TALi DETECT Divided attention subtest: Find the Frog

The fifth game —*Don't Pat the Pig* — is designed to assess how long children can maintain their attention on a task. A series of animals (objects) are presented one at a time in the centre of the screen in a fixed order. Children are instructed by their avatar to watch the objects as they appear on the screen and respond by pressing a button for every object, except one. Performance on this game is an evaluation of attention skills related to vigilance and inhibition. This game is based on Go/No-go paradigm in cognitive neuroscience literature which when delivered in its original format, would be an arduous test for young children to complete. Gamification of this cognitive task in TALi DETECT has aided participation and compliance from children with what would otherwise be laborious neurocognitive test requirements.

Image 4: TALi DETECT Sustained Attention to Response Task: Don't Pat the Pig.

The sixth game —*Make Jam* —is designed to assess how effectively a child can inhibit an already initiated motor response or suppress an impulsive behaviour. Children are asked to respond based on the instructions presented by their avatar, pressing a left or right button if a target appears or withholding a response to the target if it is rapidly followed by a stop cue. This game is based on a very well-established cognitive paradigm called Stop Signal Task (Logan et al., 2014) which measures inhibitory ability.

The final TALi DETECT game is *Sort the Blocks* where children are asked to sort objects based on a particular sorting rule and switch from one sorting rule to another when prompted. This game is designed to assess how quickly a child can switch between different tasks and is based on a very well-established cognitive paradigm called the Dimensional Change Card Sort (Zelazo, 2006) which measures children's cognitive flexibility.

TALi DETECT – and serious games in general – offer two potential benefits: a uniform user experience, and a platform where a child's responses are

automatically recorded and scored. Although automatic response recording and scoring are foundational elements of most games (Ávila-Pesántez et al., 2017; Hanneghan & Tang, 2011), the impact of an engaging, objective, and accurate assessment provides a promising benefit to clinicians. Serious games therefore not only limit the likelihood of human error (Mrazik, Martin et al., 2015; Styck & Walsh, 2016) impacting test interpretation, the objective data collected during gameplay can aid clinical conceptualisation of development. The current version of TALi DETECT is not designed as a standalone tool for diagnosing developmental delays or neurodevelopmental disorders, instead providing additional data to a clinician's formal assessment process along with relevant clinical interviews, rating scales, cognitive and academic assessments.

To be used as a clinical-grade tool and counter the lack of evidence underpinning many digital health tools, a validation study was conducted. The validation study served two purposes. The first, was to evaluate the psychometric properties related to latent factor structure of TALi DETECT as a cognitive measure of attention. The study involved 340 neurotypical children and structural equation modelling results confirm that six of the games (cognitive subtests) provide foundational assessment of the three key attentional domains. This means that TALi DETECT features a comprehensive battery of diverse attention assessment tasks that facilitate estimation of a child's strengths and weaknesses across major domains of attention. This includes the ability to orient attention, referred to as selective attention in neuroscientific literature; the ability to maintain attention, referred to as sustained attention; and the ability to switch/regulate attention, referred to as executive attention (Raz & Buhle, 2006).

The second purpose of the validation study was to determine the convergent validity, reliability, and developmental sensitivity of TALi DETECT as an effective psychometric tool. Validity is a term used in psychology which is evaluated by investigating what psychological qualities or constructs are measured by a test. Validity can be assessed in several ways, and convergent validity for TALi DETECT was determined using correlations between TALi DETECT and a well-established measure of the same construct (attention); in this case, the Test of Everyday Attention, Second Edition Junior (TEA-Ch2J) (Manly et al., 2016). The TEA-Ch2J is a

measure of selective attention and sustained attention in children aged 5-7 years. Correlation analyses between five TALi DETECT games and comparable gamified subtests from the TEA-Ch2J confirmed convergent validity in the satisfactory range (~0.3-0.7; Carlozzi et al., 2017).

Reliability is another important psychometric property for a test. Reliability is defined as the stability of scores obtained by a person when re-examined with the same test multiple times. Stability of test scores across time is important for longitudinal studies. High reliability in this context would normally be considered at around or larger than 0.7 and an intermediate reliability at around 0.6 (Fan et al., 2002). In testing the reliability of TALi DETECT, two attentional domain scores showed acceptable reliability (Selective and Sustained), while one showed lower reliability (Cognitive Flexibility).

Associations of test scores with age reflect the validity of the tests for measuring cognitive development during childhood. The effects of age on the three latent variables were examined. Age was positively associated with performance on selective attention, sustained attention, and cognitive flexibility; indicating TALi DETECT has good developmental sensitivity. Moreover, using the data collected through the study as a normative sample, TALi DETECT provides clinicians with an index for a child's attention-related cognitive capabilities relative to the reference group. This is clinically important, as it allows for identification of children with performance outside the typical range.

Overall, TALi DETECT provides an objective, performance-based measure of cognitive attention during a child's early years to detect attention vulnerabilities earlier in life. TALi DETECT illustrates the impetus for serious games developers to engage in ongoing research that is informed by robust scientific and clinical input to ensure high quality games are scalable, safe, and show real-world utility. Although avenues for future development of serious games in the digital therapeutic space are infinite (Fleming et al., 2017; Floryan et al., 2020), a lack of investment in evidence will likely result in the continued proliferation of digital health tools that are not fit for purpose.

REFERENCES

American Psychological Association. (2020). *APA Guidelines for psychological assessment and evaluation*. APA Taskforce on Psychological Assessment and Evaluation Guidelines. https://www.apa.org/about/policy/guidelines-psychological-assessment-evaluation.pdf

American Psychological Association. (2021). *APA Guidelines on evidence-based psychological practice in health care*. Workgroup of the Committee on Professional Practice and Standards COPPS) and the Board of Professional Affairs (BPA). https://www.apa.org/about/policy/psychological-practice-health-care.pdf

Australian Psychological Society. (2018). *Ethical guidelines for psychological assessment and the use of psychological tests*.

Ávila-Pesántez, D., Rivera, L. A., & Alban, M. S. (2017). *Approaches for Serious Game Design: A Systematic Literature Review*. *8*(3), 11.

Baumel, A., & Kane, J. M. (2018). Examining Predictors of Real-World User Engagement with Self-Guided eHealth Interventions: Analysis of Mobile Apps and Websites Using a Novel Dataset. *Journal of Medical Internet Research*, *20*(12), e11491. https://doi.org/10.2196/11491

Byambasuren, O., Beller, E., & Glasziou, P. (2019). Current Knowledge and Adoption of Mobile Health Apps Among Australian General Practitioners: Survey Study. *JMIR MHealth and UHealth*, *7*(6), e13199. https://doi.org/10.2196/13199

Carlozzi, N. E., Tulsky, D. S., Wolf, T. J., Goodnight, S., Heaton, R. K., Casaletto, K. B., Wong, A.K., Baum, C. M., Gershon, R. C., & Heinemann, A. W. (2017). Construct Validity of the NIH Toolbox Cognition Battery in Individuals With Stroke. *Rehabilitation Psychology*, *62*(4), 443–454. https://doi.org/10.1037/rep0000195

Diamond, A. (2013). Executive functions. In S. T. Fiske (Ed.), *Annual review of psychology* (Vol. 64, pp. 135–168). Annual Reviews. Fan, J., Mccandliss, B., Sommer, T., Raz, A., & Posner, M. (2002). Testing the Efficiency and

Independence of Attentional Networks. *Journal of Cognitive Neuroscience*, *14*, 340–347. https://doi.org/10.1162/089892902317361886

Fleming, T. M., Bavin, L., Stasiak, K., Hermansson-Webb, E., Merry, S. N., Cheek, C., Lucassen, M., Lau, H. M., Pollmuller, B., & Hetrick, S. (2017). Serious Games and Gamification for Mental Health: Current Status and Promising Directions. *Frontiers in Psychiatry*, *7*. https://doi.org/10.3389/fpsyt.2016.00215

Floryan, M., Chow, P. I., Schueller, S. M., & Ritterband, L. M. (2020). The Model of Gamification Principles for Digital Health Interventions: Evaluation of Validity and Potential Utility. *Journal of Medical Internet Research*, *22*(6), e16506. https://doi.org/10.2196/16506

Georgiadis, K., Van Lankveld, G., Bahreini, K., & Westera, W. (2018). Accommodating Stealth Assessment in Serious Games: Towards Developing a Generic Tool. *2018 10th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games)*, 1–4. https://doi.org/10.1109/VS-Games.2018.8493409

Gilbody, S., Littlewood, E., Hewitt, C., Brierley, G., Tharmanathan, P., Araya, R., Barkham, M., Bower, P., Cooper, C., Gask, L., Kessler, D., Lester, H., Lovell, K., Parry, G., Richards, D. A., Andersen, P., Brabyn, S., Knowles, S., Shepherd, C., … White, D. (2015). Computerised cognitive behaviour therapy (cCBT) as treatment for depression in primary care (REEACT trial): Large scale pragmatic randomised controlled trial. *BMJ*, h5627. https://doi.org/10.1136/bmj.h5627

Hanneghan, M., & Tang, S. (2011). Game content model: An ontology for documenting serious game design. *Developments in E-Systems Engineering*, 431–436. https://doi.org/10.1109/DeSE.2011.68

Khaleghi, A., Aghaei, Z., & Mahdavi, M. A. (2021). A Gamification Framework for Cognitive Assessment and Cognitive Training: Qualitative Study. *JMIR Serious Games*, *9*(2), e21900. https://doi.org/10.2196/21900

Logan, G. D., Van Zandt, T., Verbruggen, F., & Wagenmakers, E.-J. (2014). On the ability to inhibit thought and action: General and special theories of an

act of control. *Psychological Review*, *121*(1), 66–95. https://doi.org/10.1037/a0035230

Lui, J. H. L., Marcus, D. K., & Barry, C. T. (2017). Evidence-based apps? A review of mental health mobile applications in a psychotherapy context. *Professional Psychology: Research and Practice*, *48*(3), 199–210. https://doi.org/10.1037/pro0000122

Lumsden, J., Edwards, E. A., Lawrence, N. S., Coyle, D., & Munafò, M. R. (2016). Gamification of Cognitive Assessment and Cognitive Training: A Systematic Review of Applications and Efficacy. *JMIR Serious Games*, *4*(2), e11. https://doi.org/10.2196/games.5888

Manly, T., Anderson, V., Crawford, J., George, M., Underbjerg, M., & Robertson, I. (2016). *Test of Everyday Attention for Children Second Edition (TEA-Ch2)*. London, UK: Pearson.

Marshall, J. M., Dunstan, D. A., & Bartik, W. (2020). Clinical or gimmickal: The use and effectiveness of mobile mental health apps for treating anxiety and depression. *Australian & New Zealand Journal of Psychiatry*, *54*(1), 20–28. https://doi.org/10.1177/0004867419876700

Mrazik, Martin, Janzen, Troy M, Dombrowski, Stefan C, Barford, Sean W, & Krawchuk, Lindsey L. (2015). Administration and scoring errors of graduate students learning the WISC-IV: Issues and controversies. *Canadian Journal of School Psychology*, *27*(4), 279–290. https://doi.org/10.1177/0829573512454106

Patel, N. A., & Butte, A. J. (2020). Characteristics and challenges of the clinical pipeline of digital therapeutics. *Npj Digital Medicine*, *3*(1), 159. https://doi.org/10.1038/s41746-020-00370-8

Pondé, M. P., Cruz-Freire, A. C., & Silveira, A. A. (2012). Relationship Between Learning Problems and Attention Deficit in Childhood. *Journal of Attention Disorders*, *16*(6), 505–509. https://doi.org/10.1177/1087054711423625

Pratt, C., Baker, J., Kinsella, S., Banks, G., Macqueen, P., Dear, G., O'Keefe, V., Hagiliassis, N., Power, M., Howe, W., Senior, G., Kendall, I., & Mathews,

R. (2018). *Practice guide for the use of psychological tests and instruments*. Australian Psychological Society.

Raz, A., & Buhle, J. (2006). Typologies of attentional networks. *Nature Reviews. Neuroscience.*, *7*(5), 367–379. https://doi.org/10.1038/nrn1903

Ribner, A. D. (2020). Executive function facilitates learning from math instruction in kindergarten: Evidence from the ECLS-K. *Learning and Instruction*, *65*, 101251. https://doi.org/10.1016/j.learninstruc.2019.101251

Stebbins, G. T. (2007). Chapter 27—Neuropsychological Testing. In C. G. Goetz (Ed.), *Textbook of Clinical Neurology (Third Edition)* (pp. 539–557). W.B. Saunders. https://doi.org/10.1016/B978-141603618-0.10027-X

Styck, K. M., & Walsh, S. M. (2016). Evaluating the prevalence and impact of examiner errors on the Wechsler scales of intelligence: A meta-analysis. *Psychological Assessment*, *28*(1), 3–17. https://doi.org/10.1037/pas0000157

Torous, J., Jän Myrick, K., Rauseo-Ricupero, N., & Firth, J. (2020). Digital Mental Health and COVID-19: Using Technology Today to Accelerate the Curve on Access and Quality Tomorrow. *JMIR Mental Health*, *7*(3), e18848. https://doi.org/10.2196/18848

Torous, J., Luo, J., & Chan, S. R. (2018). Mental health apps: What to tell patients. *Current Psychiatry*, *17*(3), 21–25.

Wang, K., Varma, D. S., & Prosperi, M. (2018). A systematic review of the effectiveness of mobile apps for monitoring and management of mental health symptoms or disorders. I*ournal of Psychiatric Research*, *107*, 73–78. https://doi.org/10.1016/j.jpsychires.2018.10.006

Wang, L., Shute, V., & Moore, G. R. (2015). Lessons Learned and Best Practices of Stealth Assessment: *International Journal of Gaming and Computer-Mediated Simulations*, *7*(4), 66–87. https://doi.org/10.4018/IJGCMS.2015100104

Zelazo, P. D. (2006). The Dimensional Change Card Sort (DCCS): A method of assessing executive function in children. *Nature Protocols*, *1*(1), 297–301. https://doi.org/10.1038/nprot.2006.46