Assessment of attention through user's performance in a virtual reality game

David Mendez-Encinas[1], Aaron Sujar[2], David Delgado-Gómez[1,*]

[1]Departamento de Estadística, Universidad Carlos III de Madrid, Leganés, Spain
[2]Departamento de Ciencias de la Computación, Universidad Rey Juan Carlos, Móstoles, Spain

Abstract

Serious games can help traditional evaluation methods by creating an objective system. There are computerised tests that measure user performance through reaction time, hits, and misses but they are focused on discrete variables. A virtual reality game has been developed where the user must maintain attention to get a target fish among non-target fishes. The game has been divided into sequences where speeds and interstimulus times vary uniformly. Distractors have been included in some sequences. The game through the device records variables related to the user's movement, eyes movement, and performance. The variables are recorded continuously. Random forest regressor was used to infer the Attention Control Scale Test with the variables recorded. Although the sample is small, it has been found that errors made with and without distractors, together with reaction time are predictors related to the score obtained in the test. Other variables like eye gaze also suggest a correlation with the attention control scale score. Virtual reality applications and new devices can help in the assessment of psychological variables.

Keywords

Virtual reality, Regression, Attention, CPT, Video games

1. Introduction

According to a recent review by Francés et al. [1], the most common psychiatric disorder in children and adolescents, with an estimated prevalence between 5% and 11%, is Attention-deficit/hyperactivity disorder (ADHD). The main symptoms are difficulty in maintaining attention and impulsive behaviour [2]. These symptoms are related to impairments in executive functions, which are the mental activities that define a person's behaviour [3]

The psychological diagnosis of this disorder is based on caregivers filling out questionnaires together a clinical interview with a psychiatrist. However, this method is criticised for its lack of precision[4], the subjectivity of clinicians and caregivers[5], the lack of precision of the scales [6], and the possibility of faking [7]. There are computerised tests that try not to rely on human judgement for diagnosis. The most commonly used is Conners' Continuous Performance Test (CPT) [4]. Most computerized tests are based on the go/no-go paradigm. Participants must be

I Congreso Español de Videojuegos, December 1–2, 2022, Madrid, Spain
*Corresponding author.
mendezencinasdavid@gmail.com (D. Mendez-Encinas); aaron.sujar@urjc.es (A. Sujar); ddelgado@est-econ.uc3m.es (D. Delgado-Gómez)
https://www.linkedin.com/in/david-m (D. Mendez-Encinas); https://aaronsujar.com/ (A. Sujar)
0000-0002-7891-6137 (A. Sujar); 0000-0002-2976-2602 (D. Delgado-Gómez)

attentive to perform an action when certain stimulus appear on the screen and control their impulses on the stimuli to be inhibited. These tests usually record the number of hits, omissions, commissions and reaction times as predictors of inattention and hyperactivity [8]. In addition, some alternatives incorporate tracking devices to capture the participant's motion. For example, Reh et al. [9] registered the movements with an infrared camera following a reflective marker attached to a headband while performing the QbTest. O'Mahony et al. [10] measured the subject's wrist and ankle motion with inertial measurement units (IMUs) while performing the TOVA test. Delgado-Gomez et al. [11] developed an adapted Continuous Performance Test, using Microsoft Kinect.

Furthermore, serious games are applications designed as commercial video games but with a non-playful purpose. Traditionally, serious games have been aimed at acquiring new skills, either in education or in training [12]. In recent years, new research has studied the use of video games in the health environment, as a treatment tool [13], rehabilitation environment[14], or evaluation and assessment [15]. The use of serious games is intended to facilitate assessment through a stimulating activity [16] in an objective way[17]. The GroundsKeeper [18] is a game based on the well-known *"Whac-a-Mole"* but using *Sifteo Cubes*®. In order to play, users must take the mallet cube to hit the mole when it appears in one of the three cubes over the table. Faraone et al. [18] found that their game has similar diagnostic accuracy to CPT. Closer to our proposal, the Virtual Classroom [19] is a virtual reality (VR) solution where users perform a CPT on the blackboard of a virtual classroom. Through the measurement of head movement (motor activity), Muhlberger et al. [20] found that motor activity correlated with the hyperactivity questionnaire scores ($r = .32, p < .01$). These results have been replicated by the AULA software team [21], registering omissions, commissions, response times, and motor activity, concluding that these data help to discriminate the different ADHD subtypes [22]. More examples can be found in a recent systematic review [23].

In this article, we present a prototype to evaluate the attention capacity of users. A game has been created in which the user has to maintain attention during the execution of a task. This task is based on thee go/no go paradigm in which users should respond or inhibit depending on the appearance of target/non-target fish. While the user plays the game, the device collects a set of variables. Through a regression analysis, the aim is to find out which are the most significant variables that can predict the user's attentional capacity.

2. Method

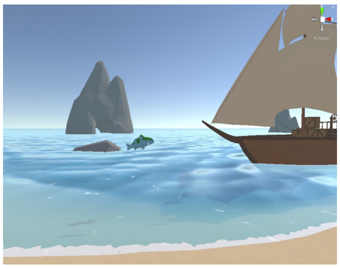
2.1. Participants

A group (N = 34) composed of students from the Carlos III and IES Jose Hierro High School , Spain, (age: Mean = 19, SD = 3, 35% women) participated in the study. Participants had ages between 14 and 26. Four participants stated that they were diagnosed with ADHD. Students were informed that their participation was voluntary and was under no circumstances considered in their academic evaluation. Subjects had the possibility to withdraw from the test at any time if they felt some kind of sickness derived from virtual reality. An identification number was given to each participant to anonymize the data. All participants were informed of the study and signed the required informed consent form.

2.2. Fish Attention Game

Before starting, a previous tutorial where the VR headset and sensors were calibrated along with a short explanation about how to use the controllers was provided before starting the fish game.

(a) Introduction to the game



(b) A fish as a stimulus

(c) Distractors of the game

(d) Models of fishes

Figure 1: Images of the game

The experiment consisted of a virtual reality game in which participants, facing the sea, should press the button when they see a fish. They should inhibit when the target fish appears (See Figure 1a). Before starting, the target fish will be shown to the user among the 5 possible ones (Fig. 1d). To start, the player needs to press a button within the virtual environment. Once the test is started, the fish will come out one by one in sequence (Fig. 1b). This sequence is common for all the participants in order to measure the same way all the results.

The complete game was formed by 16 sequences some with distractors and some without them with also different speeds and numbers of fish (See Table 1). The total duration of the whole test was around 10 minutes. However, the sequences are not identical to each other, instead, variations are introduced in order to make the users pay attention. Also, the number of stimuli varies, as well as the fish velocity. The speed is measured in Unity units per second which is equivalent to meters per second. In addition, the interstimulus time, i.e. the time between one fish and the next, is also varied. If the participant does not answer within this time, it is recorded as an error. Furthermore, in some sequences there will appear some distractors such as moving crabs, seagulls and a boat (See Figure 1c), also the same pattern for all the people. The objective of these distractors is to measure if people with attention problems get more distracted than others.

The first sequence was used as a tutorial to get comfortable with the objective of the game and was discarded once analysing the variables. All variables gathered by the game can be observed in Table 2.

The game has been designed to be played in virtual reality, but we have selected the HP

Table 1
Game Sequences

Sequence	Nº of fishes	Nº of Target Fishes	Speed	Interstimulus time (Seconds)	Has distractors
1	10	2	5	2	No
2	10	2	8	2	No
3	10	2	8	4	Yes
4	15	3	10	1	Yes
5	5	1	10	4	No
6	10	2	12	1	No
7	20	4	12	Mix of (1,2 and 4)	Yes
8	20	4	12	Mix of (1,2 and 4)	No
9	10	2	8	2	Yes
10	10	2	8	4	No
11	15	3	10	1	No
12	5	1	10	4	Yes
13	10	2	12	1	Yes
14	20	4	12	Mix of (1,2 and 4)	No
15	20	4	12	Mix of (1,2 and 4)	Yes
16	10	2	10	2	Yes

Table 2
Variables gather by the game

Variable Name	Type	Range
Percentage looking seagull	float	[0-1]
Percentage looking crab	float	[0-1]
Percentage looking boat	float	[0-1]
Percentage looking fish	float	[0-1]
Cognitive load[24]	float	[0-1]
Errors with distractions	int	positive
Errors without distractions	int	positive
Reaction time average	float	(0,10)
Average blinks	int	positive
Eye dilation normalized variance	float	positive
Gaze variance in x axis	float	positive
Gaze variance in y axis	float	positive
Gaze variance in z axis	float	positive
Eye vergence mean	float	positive
Hands velocity total energy	float	positive

Reverb G2 Omnicept Edition goggles, because it has eye tracking built-in and pupillometry. It also includes heart rate measurement and face cam that is not registered in this experiment.

2.3. Attention Control Scale questionnaire

After the game is completed, the users are invited to fill out the Attention Control Scale (ATTC) questionnaire [25]. It is designed to measure two main components of attention (focusing of attention and shifting of attention). The ATTC consists of 20 items that are assessed on a four-point Likert scale from 1 (rarely) to 4 (always). The test produces a total scale and two subscales, each of which measures one of the two main components of attention. Scales scores are calculated as the sum of the respective items with some items having reverse scoring.

Also, participants fill out a demographic questionnaire. The main variables gathered by this questionnaire can be found in table 3. Both questionnaires were presented to the subjects via Google Forms.

2.4. Statistics

Table 3
Variables gather by the test

Variable Name	Type	Range
Age	int	positive
Gender	bool	male/female
Wearing glasses	bool	true/false
Has tried VR	bool	true/false
Habitual VR player	bool	true/false
Habitual videogames player	bool	true/false
Color blind	bool	true/false
Vision problems	bool	true/false
ADHD diagnosis	bool	true/false
ATTC total mark	int	positive

This paper aims to estimate the relationship between the variables collected by the device and the target variable (ATTC score). We have chosen a regression analysis using Random Forest techniques due to its interpretability. We can obtain the feature importance of the model variables, and also it is a reliable model for this many variables. Using classification for that many values will not be a good idea. We have too many classes for this problem and there are no classes as such.

The random forest regressor was used to infer the ATTC questionnaire mark. The procedure was the following: split the N=34 set into 20% test 80% train, make five-fold cross-validation in order to get the best parameters of the random forest. Get the feature importance along with the train and test scores with the trained model, and repeat this process 1000 times in order to get all the possible test/train splits. The model was trained by minimising the R^2 score (Eq. 1). Other scores used to assess the performance of the model were the RMSE (Eq. 2) and the MAPE (Eq. 3). All the final scores are the mean of the 1000 iterations giving a more complete view of this small sample as all the combinations should have been performed.

$$R^2 = \frac{\sum_{i=1}^{n}(Predicted_i - Mean(Real))^2}{\sum_{i=1}^{n}(Real_i - Mean(Real))^2} \tag{1}$$

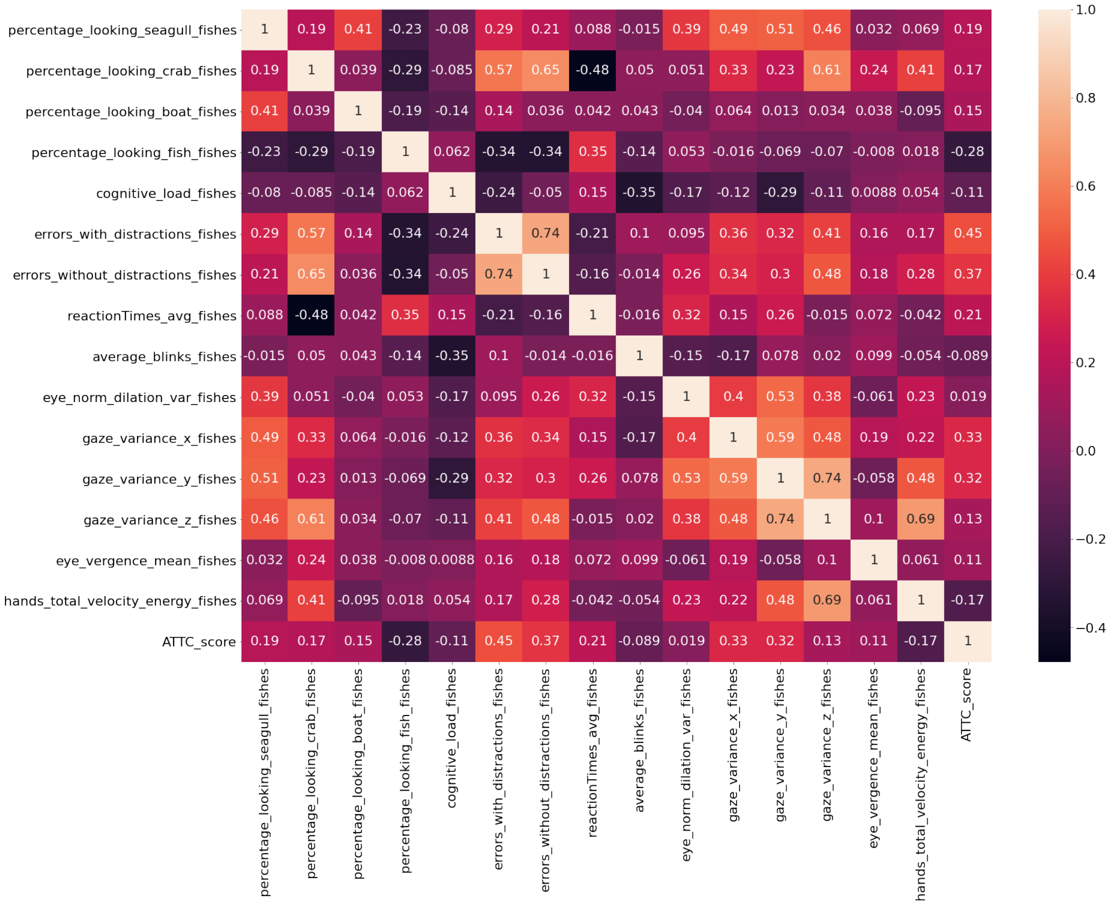


Figure 2: Covariance between variables including the target variable ATTC Score

$$RMSE = \sqrt{\frac{1}{n}\Sigma_{i=1}^{n}\Big(Predicted_i - Real_i\Big)^2} \tag{2}$$

$$MAPE = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{Real_i - Predicted_i}{Real_i}\right| \tag{3}$$

The correlation between variables can be seen in Figure 2 and it tells that there are some variables the target variable.

3. Results

The covariance table (Fig. 2) is very illustrative. We can check the correlation of the ATTC score with some of the variables. The highest correlation is found with the errors in the sections that have distractors, that is, the higher the score in the test, the more errors they make when there are distractors present. However, the next variable that has a high correlation is the errors made, but this time without distractors. This makes sense, as those with a lower attention span will be less attentive to the target fish. This is also consistent with the correlations of time looking at distractors (seagulls (0.19), crab(0.17) and boat(0.15)). These distractions seem to affect also the reaction time which has a correlation of 0.21 with a high test score. Another interesting variable is the time spent looking at fishes. This variable has an inverse correlation to a high test score (-0.28). This may mean that more attentive people look more at fishes than at distractors.

On the other hand, the correlations of eye movement are also interesting. Users with higher ATTC scores are those who move their eyes more (Variance of x(0.33), y(0.32), z(0.13)). Finally, it is worth pointing out the negative correlation of hand speed (-0.17). This may be due to the fact that the more attentive users were able to choose the fish faster, while the inattentive users may hesitate.

On the Figure 3 can be seen the variable importances. We see for example that the model has decided that the variable that represents the most important is the errors made in the presence of distractors. In second place are eye movement and hand speed. And finally, response time, time looking at the fish, or distractors also are important.

In order to verify how good our regressor is, we are going to compare the training and test performance against a baseline. To do this, we have run the model by assigning the value of the ATTC test variable to the mean obtained from the whole sample, which was 47 points. The baseline RMSE and MAPE can give us an idea about the performance of the regressor. Although we have a small sample the correlation of the test is clearly positive which means that it has predictive potential.

Table 4
Experiment results

Partition	RMSE	MAPE	Correlation
Baseline	9.06	0.1556	N/A
Train	3.82	0.0698	0.9306
Test	9.05	0.1761	0.2509

4. Conclusion

In this paper, a prototype of a game to measure a player's attention is presented. As well as solutions such as Virtual Classroom [19] or Aula Nesplora [21], statistical analysis shows that the model has a positive correlation with the attention test (ATTC). However, this claim should be tested with a larger sample, since the sample size is not enough. The variables selected as important by the model (see Figure 3) are quite significant, and the most important variables are in line with what multiple studies previously suggested as important regressors in order to assess inattention [8]. For example, the variable that measures the errors made is more significant when distractors are present. This is relevant since incorporating environmental distractors improved the test's utility [26]. In the same way, the response time can also be a predictor of the user's attention. The hand's total velocity energy, i.e., the amount of user

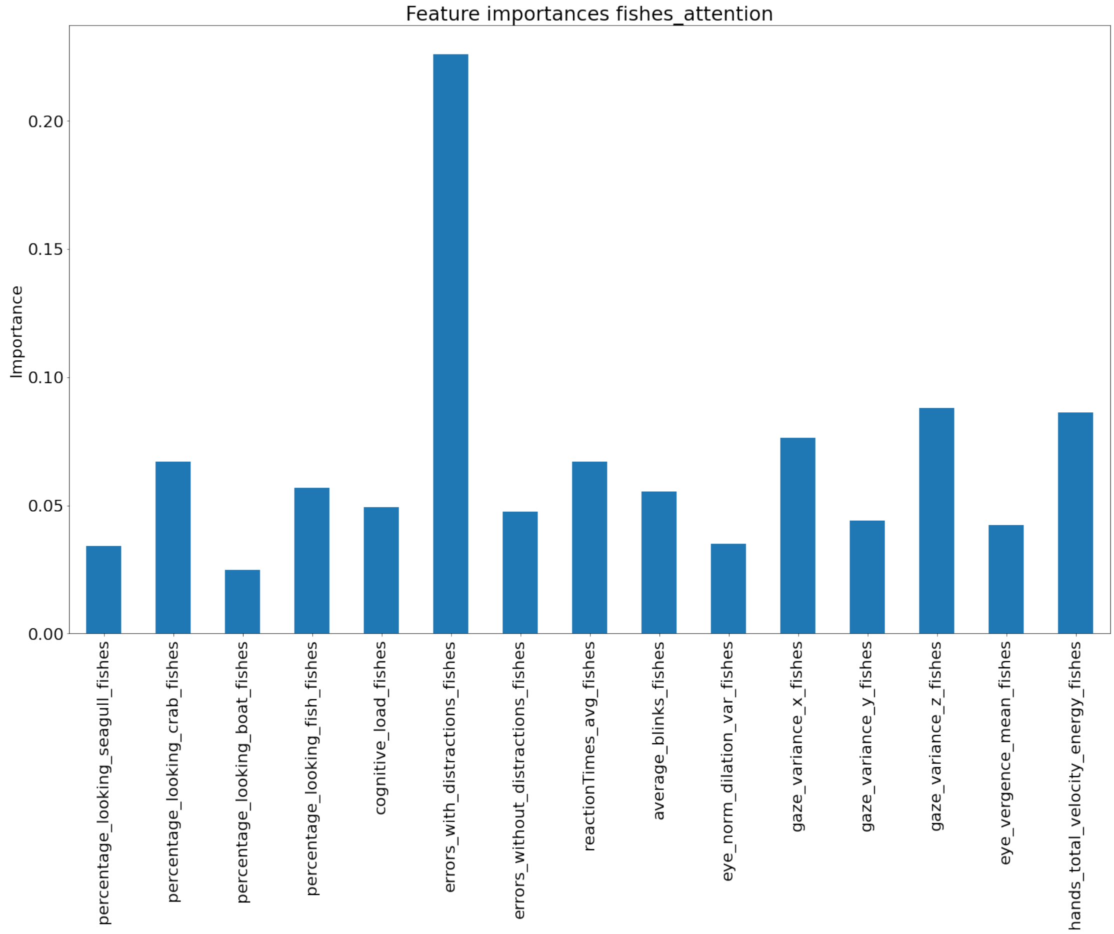


Figure 3: Variable mean importance of the random forest models after 1000 iterations of different random partitions

movement in the hands, has also been tested in previous work [10, 11]. However, we differ from the aforementioned work in that we collect all the variables on a continuous basis and that the regression technique can tell us which variables are more important than others.

Other variables identified as important are the gaze variance in the y and z axis, the pupil dilation variance, and the cognitive load. This has also been tested according to the advances in eye tracking technology. Some studies are basing the diagnosis of ADHD on eye analysis. For example, pupil size has been shown to be a predictor of ADHD [27], or Varela Casal et al. [28] developed a go/no-go test where users must discriminate cartoon images of a tadpole from a fish. They found that attention-related eye vergence is different between healthy controls, clinical controls, and children with ADHD.

However, limitations in sample size reduce the power of the conclusions. Future work should expand the sample, and include a healthy control group and another group with ADHD-diagnosed participants. This will allow the regressor to be trained on a larger sample. In addition, other psychological tests should be administered and could even be compared against the CPT.

It is important to mention that the test lasts more than 10 minutes. This may cause sickness in users who are not used to using VR on a regular basis. Future studies should check if it is possible to reduce the total play time while obtaining the same results.

Finally, future work could be to make other games to assess other psychological areas. Hyperactivity is another symptom of ADHD, but other game mechanics must be designed. Also, plannification, working memory and other executive functions could be considered. Although the VR goggles used in this experiment are expensive and could be heavy, maybe some years from now this technology will become lighter and available to the common user.

Acknowledgments

This research was partially funded by: Ministerio de Ciencia e Innovación, Proyectos de Transición Ecológica y Transición Digital TED2021-130980B-I00, UC3M (Grant for the requalification of permanent lectures, David Delgado-Gómez), Instituto Salud Carlos III, grant number DTS21/00091.

References

[1] L. Francés, J. Quintero, A. Fernández, A. Ruiz, J. Caules, G. Fillon, A. Hervás, C. V. Soler, Current state of knowledge on the prevalence of neurodevelopmental disorders in childhood according to the dsm-5: a systematic review in accordance with the prisma criteria, Child and Adolescent Psychiatry and Mental Health 16 (2022) 27. doi:10.1186/s13034-022-00462-1.

[2] American Psychiatric Association, Diagnostic and statistical manual of mental disorders: DSM-5, 5th ed. ed., Autor, Washington, DC, 2013.

[3] T. E. Brown, Add/adhd and impaired executive function in clinical practice, Curr Atten Disord Rep 1 (2009) 37–41. doi:10.1007/s12618-009-0006-3.

[4] M. C. Edwards, E. S. Gardner, J. J. Chelonis, E. G. Schulz, R. A. Flake, P. F. Diaz, Estimates of the validity and utility of the conners' continuous performance test in the assessment of inattentive and/or hyperactive-impulsive behaviors in children, Journal of Abnormal Child Psychology 35 (2007) 393–404. doi:10.1007/s10802-007-9098-3.

[5] T. C. Chi, S. P. Hinshaw, Mother–child relationships of children with adhd: The role of maternal depressive symptoms and depression-related distortions, Journal of abnormal child psychology 30 (2002) 387–400. doi:10.1023/a:1015770025043.

[6] M. R. Basco, J. Q. Bostic, D. Davies, A. J. Rush, B. Witte, W. Hendrickse, V. Barnett, Methods to improve diagnostic accuracy in a community mental health setting, American Journal of Psychiatry 157 (2000) 1599–1605. doi:10.1176/appi.ajp.157.10.1599.

[7] R. A. Sansone, L. A. Sansone, Faking attention deficit hyperactivity disorder, Innovations in Clinical Neuroscience 8 (2011) 10.

[8] T. D. Parsons, T. Duffield, J. Asbee, A comparison of virtual reality classroom continuous performance tests to traditional continuous performance tests in delineating adhd: a meta-analysis, Neuropsychology review 29 (2019) 338–356. doi:10.1007/s11065-019-09407-6.

[9] V. Reh, M. Schmidt, W. Rief, H. Christiansen, Preliminary evidence for altered motion tracking-based hyperactivity in ADHD siblings, Behavioral and Brain Functions 10 (2014) 7. doi:10.1186/1744-9081-10-7.

[10] N. O'Mahony, B. Florentino-Liano, J. J. Carballo, E. Baca-García, A. A. Rodríguez, Objective diagnosis of ADHD using IMUs, Medical Engineering & Physics 36 (2014) 922–926. doi:10.1016/j.medengphy.2014.02.023.

[11] D. Delgado-Gomez, I. Peñuelas-Calvo, A. E. Masó-Besga, S. Vallejo-Oñate, I. Baltasar Tello, E. Arrua Duarte, M. C. Vera Varela, J. Carballo, E. Baca-García, Microsoft Kinect-based Continuous Performance Test: An objective attention deficit hyperactivity disorder assessment, J Med Internet Res 19 (2017) e79. doi:10.2196/jmir.6985.

[12] D. R. Michael, S. L. Chen, Serious games: Games that educate, train, and inform, Muska & Lipman/Premier-Trade, 2005.

[13] L. Jiménez-Muñoz, I. Peñuelas-Calvo, P. Calvo-Rivera, I. Díaz-Oliván, M. Moreno, E. Baca-García, A. Porras-Segovia, Video games for the treatment of autism spectrum disorder: A systematic review, Journal of Autism and Developmental Disorders 52 (2022) 169–188. doi:10.1007/s10803-021-04934-9.

[14] O. Czech, A. Wrzeciono, L. Batalík, J. Szczepańska-Gieracha, I. Malicka, S. Rutkowski, Virtual reality intervention as a support method during wound care and rehabilitation after burns: A systematic review and meta-analysis, Complementary Therapies in Medicine 68 (2022) 102837. doi:https://doi.org/10.1016/j.ctim.2022.102837.

[15] T. M. Connolly, Psychology, pedagogy, and assessment in serious games, Igi Global, 2013.

[16] K. C. Bul, I. H. Franken, S. Van der Oord, P. M. Kato, M. Danckaerts, L. J. Vreeke, A. Willems, H. J. Van Oers, R. Van den Heuvel, R. Van Slagmaat, et al., Development and user satisfaction of "plan-it commander," a serious game for children with adhd, Games for health journal 4 (2015) 502–512. doi:10.1089/g4h.2015.0021.

[17] D. Delgado-Gómez, A. Sújar, J. Ardoy-Cuadros, A. Bejarano-Gómez, D. Aguado, C. Miguelez-Fernandez, H. Blasco-Fontecilla, I. Peñuelas-Calvo, Objective assessment of attention-deficit hyperactivity disorder (adhd) using an infinite runner-based computer game: A pilot study, Brain Sciences 10 (2020). doi:10.3390/brainsci10100716.

[18] S. V. Faraone, J. H. Newcorn, K. M. Antshel, L. Adler, K. Roots, M. Heller, The groundskeeper gaming platform as a diagnostic tool for attention-deficit/hyperactivity disorder: sensitivity, specificity, and relation to other measures, Journal of child and adolescent psychopharmacology 26 (2016) 672–685. doi:10.1089/cap.2015.0174.

[19] Y. Pollak, P. L. Weiss, A. A. Rizzo, M. Weizer, L. Shriki, R. S. Shalev, V. Gross-Tsur, The utility of a continuous performance test embedded in virtual reality in measuring ADHD-related deficits, Journal of Developmental & Behavioral Pediatrics 30 (2009). doi:10.1097/DBP.0b013e3181969b22.

[20] A. Mühlberger, K. Jekel, T. Probst, M. Schecklmann, A. Conzelmann, M. Andreatta, A. A. Rizzo, P. Pauli, M. Romanos, The influence of methylphenidate on hyperactivity and attention deficits in children with ADHD: A virtual classroom test, Journal of Attention Disorders 24 (2020) 277–289. doi:10.1177/1087054716647480.

[21] U. Díaz-Orueta, C. Garcia-López, N. Crespo-Eguílaz, R. Sánchez-Carpintero, G. Climent, J. Narbona, AULA virtual reality test as an attention measure: Convergent validity with conners' continuous performance test, Child Neuropsychol. 20 (2014) 328–342.

doi:10.1080/09297049.2013.792332, pMID: 23638628.

[22] D. Areces, C. Rodríguez, T. García, M. Cueli, P. González-Castro, Efficacy of a continuous performance test based on virtual reality in the diagnosis of ADHD and its clinical presentations, J. Atten. Disord. 22 (2018) 1081–1091. URL: https://doi.org/10.1177/1087054716629711. doi:10.1177/1087054716629711.

[23] I. Peñuelas-Calvo, L. K. Jiang-Lin, B. Girela-Serrano, D. Delgado-Gomez, R. Navarro-Jimenez, E. Baca-Garcia, A. Porras-Segovia, Video games for the assessment and treatment of attention-deficit/hyperactivity disorder: a systematic review, European Child & Adolescent Psychiatry (2020). doi:10.1007/s00787-020-01557-w.

[24] E. Siegel, J. Wei, A. Gomes, M. Oliviera, P. Sundaramoorthy, K. Smathers, M. Vankipuram, HP Omnicept Cognitive Load Database (HPO-CLD)–Developing a Multimodal Inference Engine for Detecting Real-time Mental Workload in VR, Technical Report, Technical report, HP Labs, Palo Alto, 2021.

[25] D. Derryberry, M. A. Reed, Anxiety-related attentional biases and their regulation by attentional control., Journal of Abnormal Psychology 111 (2002) 225–236. doi:10.1037/0021-843X.111.2.225.

[26] I. Berger, O. Slobodin, H. Cassuto, Usefulness and validity of continuous performance tests in the diagnosis of attention-deficit hyperactivity disorder children, Archives of Clinical Neuropsychology 32 (2017) 81–93. doi:10.1093/arclin/acw101.

[27] W. Das, S. Khanna, A novel pupillometric-based application for the automated detection of ADHD using machine learning, in: 11th ACM Int. Conf. Bioinformatics, Comput. Biol. Heal. Informatics, BCB ’20, Association for Computing Machinery, New York, NY, USA, 2020. doi:10.1145/3388440.3412427.

[28] P. Varela Casal, F. L. Esposito, I. M. Martínez, A. Capdevila, M. S. Puig, N. de la Osa, L. Ezpeleta, A. P. i Lluna, S. V. Faraone, J. A. Ramos-Quiroga, H. Supèr, J. Cañete, Clinical validation of eye vergence as an objective marker for diagnosis of ADHD in children, J. Atten. Disord. 23 (2019) 599–614. doi:10.1177/1087054717749931.