




OPEN

Attention and impulsivity assessment using virtual reality games

David Mendez-Encinas¹, Aaron Sujar¹ , Sofia Bayona² & David Delgado-Gomez¹

The assessment of cognitive functions is mainly based on standardized neuropsychological tests, widely used in various fields such as personnel recruitment, education, or health. This paper presents a virtual reality game that allows collecting continuous measurements of both the performance and behaviour of the subject in an immersive, controllable, and naturalistic experience. The application registers variables related to the user's eye movements through the use of virtual reality goggles, as well as variables of the game performance. We study how virtual reality can provide data to help predict scores on the Attention Control Scale Test and the Barratt Impulsiveness Scale. We design the application and test it with a pilot group. We build a random forest regressor model to predict the attention and impulsivity scales' total score. When evaluating the performance of the model, we obtain a positive correlation with attention (0.434) and with impulsivity (0.382). In addition, our model identified that the most significant variables are the time spent looking at the target or at distractors, the eye movements variability, the number of blinks and the pupil dilation in both attention and impulsivity. Our results are consistent with previous results in the literature showing that it is possible to use data collected in virtual reality to predict the degree of attention and impulsivity.

One of the main current problems in child psychiatry is obtaining an accurate diagnosis of attention-deficit/hyperactivity disorder (ADHD). It has been shown that medical professionals commit about 20% false positives and 20% false negatives¹. Many of the games focused on ADHD diagnosis are developed targeting children and adolescents, so there remains a need to develop therapeutic tests for adults². In addition, high impulsivity and inattention can also occur in other mental disorders and conditions. For example, high impulsivity is associated with suicidal ideation³ or addictions⁴. On the other hand, inattention is a factor present in schizophrenia, anxiety, depression, post-traumatic stress disorder, epilepsy, among others⁵. The importance of being able to evaluate these two constructs is not restricted only to the health field. Several studies have shown their relationship with academic performance^{6,7}. In other fields, such as personnel recruitment, inattention or impulsivity are characteristics that are taken into account in the selection of workers such as air traffic controllers or policemen^{8,9}. All these facts illustrate the importance of obtaining accurate assessments of a person's impulsivity and inattention.

Inattention and impulsivity are frequently assessed using the completion of scales and questionnaires. It is frequently conducted using the completion of scales and questionnaires. However, some studies point out the disadvantages of these methods, such as subjective interpretation, lack of precision¹⁰, or malingering to obtain a benefit (economic, academic, access to drugs, etc.)¹¹.

An alternative to the scales and questionnaires is the continuous performance tasks (CPT), originally created to prove that brain-damaged patients performed tasks worse than healthy controls¹². These evaluations typically consist of the participant trying to distinguish between target/non-target stimuli in order to conduct a given action, more commonly known as the go/no-go paradigm¹³. These stimuli are presented on a computer so that reaction time and errors can be measured without human supervision.

Recently, with the advances in computer science, the trend is towards the creation of more advanced assessments through video games to support diagnosis and cognitive assessment¹⁴. In general, these games are mostly, but not exclusively, oriented to child and adolescent related disorders^{15,16}. By appearing more closely related to everyday-life hobbies and offering pleasant and fun environments, they improve participation and achieve a more realistic performance¹⁷. If necessary, the game and tasks can often be adapted to adults by modifying the appearance and mechanics for a more mature version¹⁸. On the one hand, many of the video games proposed are gamified versions of CPTs^{19,20} or versions based on the go/no-go paradigm seeking a direct comparison with an

¹Departamento de Estadística, Universidad Carlos III de Madrid, Leganes, Spain. ²Departamento de Ciencias de la Computación, Universidad Rey Juan Carlos, Móstoles, Spain. ✉email: aaron.sujar@urjc.es

already established instrument^{21,22}. On the other hand, assessment can also be done with other types of games that mimic the look and feel of other commercial games, providing a more natural and familiar experience^{23,24}.

In this paper, we present an approach to support cognitive assessment using a virtual reality (VR) application that tries to predict the degree of attention and impulsivity of the participant. The application gathers variables related to the user's movements through the use of VR goggles and in-game performance. Variables are recorded in a continuous manner during the entire time the player is in the game. Our hypothesis is that applying machine learning techniques to these variables will allow us to predict the degree of impulsivity and attention of the participant.

The remaining of the article is structured as follows. In the next section, we will cover some of the most relevant psychometrical tests, starting with self-report scales and evolving to VR and video game-based assessments. Then, in Section “[The treasure hunt VR application](#)”, we will describe in detail the proposed VR application, the registered variables and the analytical methods applied. Section “[Experimental results](#)” describes the experiment conducted on the model's performance and the results obtained. Finally, we will discuss the implications and limitations of our results in Section “[Discussion](#)”, and conclude and propose future work in Section “[Conclusion](#)”.

Related work

Attention and impulsiveness are frequently measured through self-reporting scales. A self-report scale widely used to measure attention control²⁵ is the Attention Control Scale (ATTC) questionnaire²⁶. In the case of impulsivity, the Barratt Impulsiveness Scale (BIS-11) is one of the most commonly administered self-report scales for the assessment of impulsiveness in both research and clinical settings²⁷. In this work, we will use our VR application to predict the scores of these two well-validated self-report scales.

Other approaches, instead of relying on questionnaires, focus on performing tasks. An example of this is the d2 Test of Attention²⁸ which measures attention and concentration processes. The task consists of detecting target characters (the character 'd' with two marks above it, around it or below it) that are interspersed with non-target characters (character 'p' or character 'd' with more or less than two marks). The Familiar Figures Matching Test (MFF20)²⁹ is a matching task where the participant must select a target, from a set of six perceptually similar pictures, that exactly matches the sample. This test is able to discriminate reflective from impulsive children.

With the advance of technology, computerized tests emerged in which, for example, a series of visual and/or auditory stimuli are presented over a period of time. Those tests allow all subjects to be exposed to the same tasks and stimuli while recording variables automatically. The CPTs have been used in a variety of settings, for instance, to test the attention of brain-damaged patients¹², or to test the skills of radar operators³⁰. Two of the most commonly used CPTs are the Conners CPT³¹ and the Test of Variables of Attention (TOVA)³². These tests require the participant to maintain attention on a changing stimulus displayed on the screen, and they must react when a target stimulus appears. However, they differ on some key points³³, where Conners CPT uses a “response inhibition” paradigm (90% target and 10% non-target), TOVA uses a mix of this and a “rare target” paradigm (22.5% targets and 77.5% non-targets). Furthermore, TOVA is not based on letters, but uses the positions of a square (up and down) as the target stimuli. In this way, no linguistic demands or left-right discrimination are required.

With the aim of finding new predictors, other authors have incorporated new technologies into computerized tests to take into account other variables like the patient's movement. For example, O'Mahony et al.³⁴ used wrist and ankle trackers while performing the TOVA test. Delgado-Gomez et al.³⁵ adapted the CPT using the Microsoft Kinect by replacing key presses with arm movements. Valera Casal et al.³⁶ found that eye vergence can be used to discriminate attention-deficit/hyperactivity disorder (ADHD) in children. Snyder et al.³⁷ affirm that electroencephalographic (EEG)-based assessment obtained while participants conduct CPTs can be used to improve the accuracy of ADHD diagnosis.

Some works included the CPT in VR applications. The Virtual Classroom²⁰ is a VR solution that introduces users into a classroom to perform a CPT on the blackboard. Rodriguez et al.³⁸ found that Aula Nesplora, another VR-CPT solution within a virtual class, is a better predictor of ADHD than TOVA.

In order to improve participation in the assessment process and increase their ecological validity³⁹, serious games and gamification have been exploited to enhance existing CPTs or to create new assessment alternatives. For example, MOXO-CPT¹⁹ uses pictures instead of letters. In comparison to the use of letters that are clearly distinctive target/non-target stimuli, similar pictures that share shapes and colours without being entirely equal are used. Also, it incorporates auditory and visual distractors as external interfering stimuli.

Other solutions more similar to commercial video games have also been developed for assessment or treatment. One of the most known examples is Akili's EndeavourRx²² which was the first therapeutic game cleared by the U.S. Food and Drug Administration (FDA) intended to relieve and treat ADHD. The game Antonyms²¹ contains different activities within a storyline, where each activity is inspired by classic neuropsychological tasks and players have to fill a backpack or move along a route based on the go/no-go paradigm. Timó's Adventure⁴⁰ is another computer-based game, which through a storyline with different mini-games, is able to measure executive functions (EFs) like attention or inhibition. Delgado-Gómez et al.²³ developed an infinite runner-based computer game, where a raccoon has to jump between platforms to reach the end, to assess attention capacity in ADHD patients. The authors found a link between the jump distance and omissions with the Strengths and Weaknesses of ADHD-Symptoms and Normal-Behavior (SWAN) scale⁴¹.

Our proposal is based on creating a VR application that includes a game that allows continuous data collection. We record information such as errors or reaction time and the VR device allows us to obtain eye movement information. We then use all these data to create a model to predict the scores obtained on the ATTC and BIS-11 tests, related to attention and impulsivity.

The treasure hunt VR application

In this section, we will explain in detail the VR application developed. The application is composed of two modules. The first consists of the game itself, which allows for the collection of the variables related to the user's performance and behaviour. The second module implements the design of the analytical model to predict the degree of attention and impulsivity.

Game description. The game for the VR application consists of performing a sustained attention task where distractors appear and the user must resist the impulse of diverting the gaze in order not to lose sight of the target.

The setting of the game is an island. On the sandy seashore, three treasure chests are lined up. Two of these chests are empty, while inside the other chest, the player can see that there is a treasure (see Fig. 1). The chests close and begin to alternate their position with each other, similarly to the three shells and a pea game. Players must keep their focus on the closed chest containing the treasure to be able to select it once the movement of the chests ceases. Once selected, the chests are opened and players can see if they made the right choice. After some time, the chests are closed again, starting a new round of movement. Between rounds, there is a fixed time when the chests will not move to create the need to keep watching them until they close and start moving again.

The sequence of movements of each chest is fixed and will be common to all players. In addition, from time to time distractors are introduced into the scene, such as seagulls flying or crabs moving. The complete game is made up of twenty sequences in which both the number of movements of the chests and the speed of the movements vary (it can be found as Supplementary Fig. S1 online). The total duration of the entire test is about ten minutes. This time varies slightly depending on how long each participant took to choose the chest, as we allowed a maximum of ten seconds for each selection.

The application requires a VR headset and hand controllers to gather the following variables during the game playing time: eye gaze (left/right eye 3D vector), pupil dilation, blinks, eye vergence, and hand movement. Throughout the execution of the game, the application stores the player's performance through the following variables: the reaction time, the time spent looking at the correct chest, looking at a wrong chest or at each distractor, the number of successes, the number of errors, (either because no chest was selected within the 10 seconds or because a wrong chest was selected).

Analytical methods. In order to find out how the variables collected by our application can help to predict the scores of the scales administered, we use a regression analysis. The chosen machine learning technique was random forest because it works well with high-dimensional data and the interpretation of the model variables is easier to understand than in other machine learning techniques. Concisely, the Random Forest technique constructs several decision trees from randomly obtained subsets of the available observations and predictors. Once the Random Forest is constructed, it classifies each new observation by averaging the predictions made by each of these trees⁴².

Experimental results

In this section, we describe the experiment developed to evaluate the performance of the proposed application. We start by explaining the sample, the set-up and then, we detail the results obtained.

Experiment description. We run a pilot study to evaluate our analytical model. We recruited students from the Carlos III University and the IES Jose Hierro High School, Spain. The sample was composed of 34 students with ages between 15 and 26 (mean = 19.24, SD = 2.985 and 37% of them were women). None of them had vision problems or epileptic disorders.

At the beginning, students were informed that their participation was voluntary and under no circumstances was it considered in their academic evaluation. Subjects had the possibility to withdraw from the experiment at any time if they felt some kind of sickness derived from VR or for no reason. An identification number was

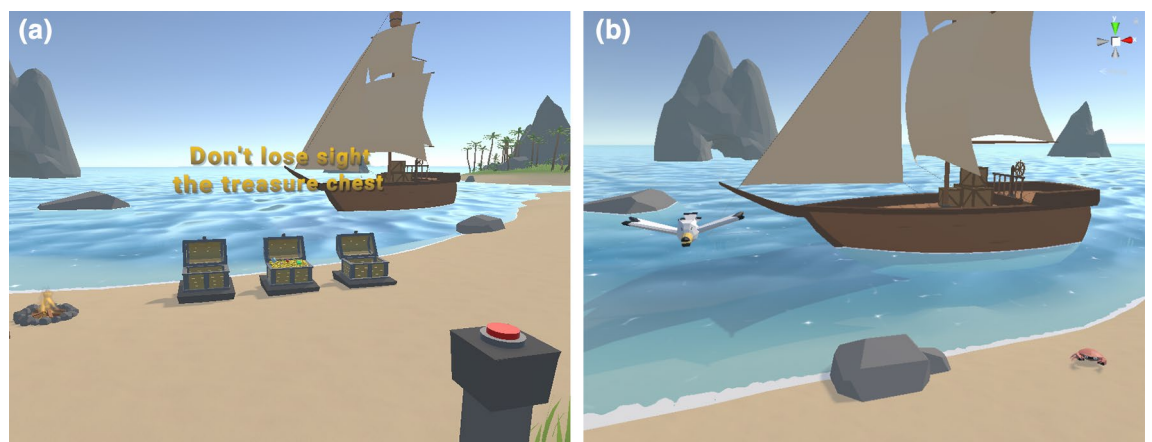


Figure 1. Image of the game's beach scenario.

assigned to each participant to anonymize the data. All participants were informed of the study and signed the required informed consent form.

Fig. 2 shows the different steps that compose the following experiment. Participants were assisted to put on the headset ensuring a good position and the calibration process was started. The players were given the hand controls. Before playing the game, participants went through a practical tutorial in which the basic VR controls were explained so that they would be comfortable with the VR experience. After this familiarization, participants played the game. During the game, the previously mentioned variables were collected by the VR headset at a rate of 90 times per second. However, we observed that several occlusion errors had been produced when registering the hand movements, so we discarded this information for our model.

After the game is over, and in order to determine how well our application is capable of predicting the degree of attention and impulsivity, we administered the Attention Control Scale (ATTC) questionnaire²⁶ and the Barratt Impulsiveness Scale (BIS-11)²⁷. The first scale is composed of 20 items that are assessed on a four-point Likert scale from 1 (rarely) to 4 (always). The BIS-11 scale is designed to assess the impulsivity of a person. It has 30 items that are also assessed on a four-point Likert scale with some reverse scoring. On average, the whole procedure (familiarization, playing the game and filling the questionnaires) lasted about 20 minutes per user.

As explained, the VR and game related variables, together with the answers to the questionnaires were used to evaluate the performance of our application using a random forest in a repeated validation set-up. For each repetition, the sample is divided into 80% train and 20% test. In addition, a five-fold cross-validation was conducted on the training data to obtain the best parameters of the random forest model in each repetition. The number of repetitions was set to 1000 times. The score metrics used to measure the model performance were the Root Mean Squared Error (RMSE) and the correlation.

In our work, we used the HP Reverb G2 Omnicept goggles with its two-hand motion controllers. The main reason for this choice was that this headset has built-in eye tracking and pupillometry sensors. However, any VR goggles that include these features can serve as an alternative. Additionally, the HP Reverb G2 Omnicept goggles have high-quality 2160x2160 LCD panels per eye. The interpupillary distance (IPD) can be adjusted and four cameras are built into the headset to track the movement of the controllers.

Results. Before moving on to the results of the repeated validation experiment described above, and in order to obtain a first insight, we present the correlation matrix between the predictors and the scores of attention and impulsivity (see Fig. 3).

Regarding the ATTC score, the highest absolute correlation value can be observed in the time looking at the correct chest. The value -0.52 indicates that the relationship is inverse, the longer players are able to keep their attention on the correct chest, the lower the ATTC score (the more attentive). There were also high positive correlations between ATTC score and the time spent looking at the distractors (boat 0.41, seagulls 0.36, fish 0.31, and crabs 0.3). This shows that the higher the ATTC score (more inattentive), the more time spent looking at the distractors, or at the wrong chest (correlation 0.41). We observe a correlation of 0.39 in the errors made in those sequences that did not have distractors. Interestingly, looking at the eye gaze correlation values (x 0.32, y

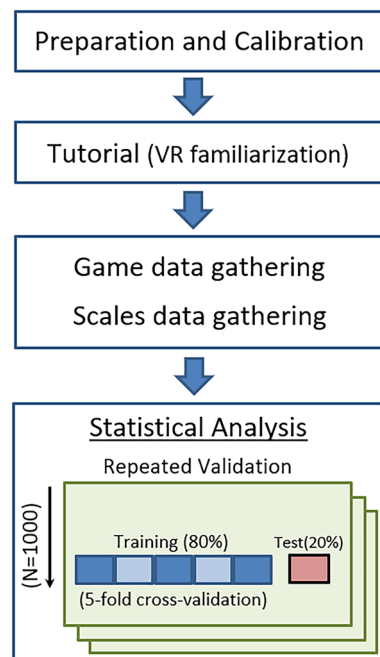


Figure 2. Summary of the assessment process using machine learning.

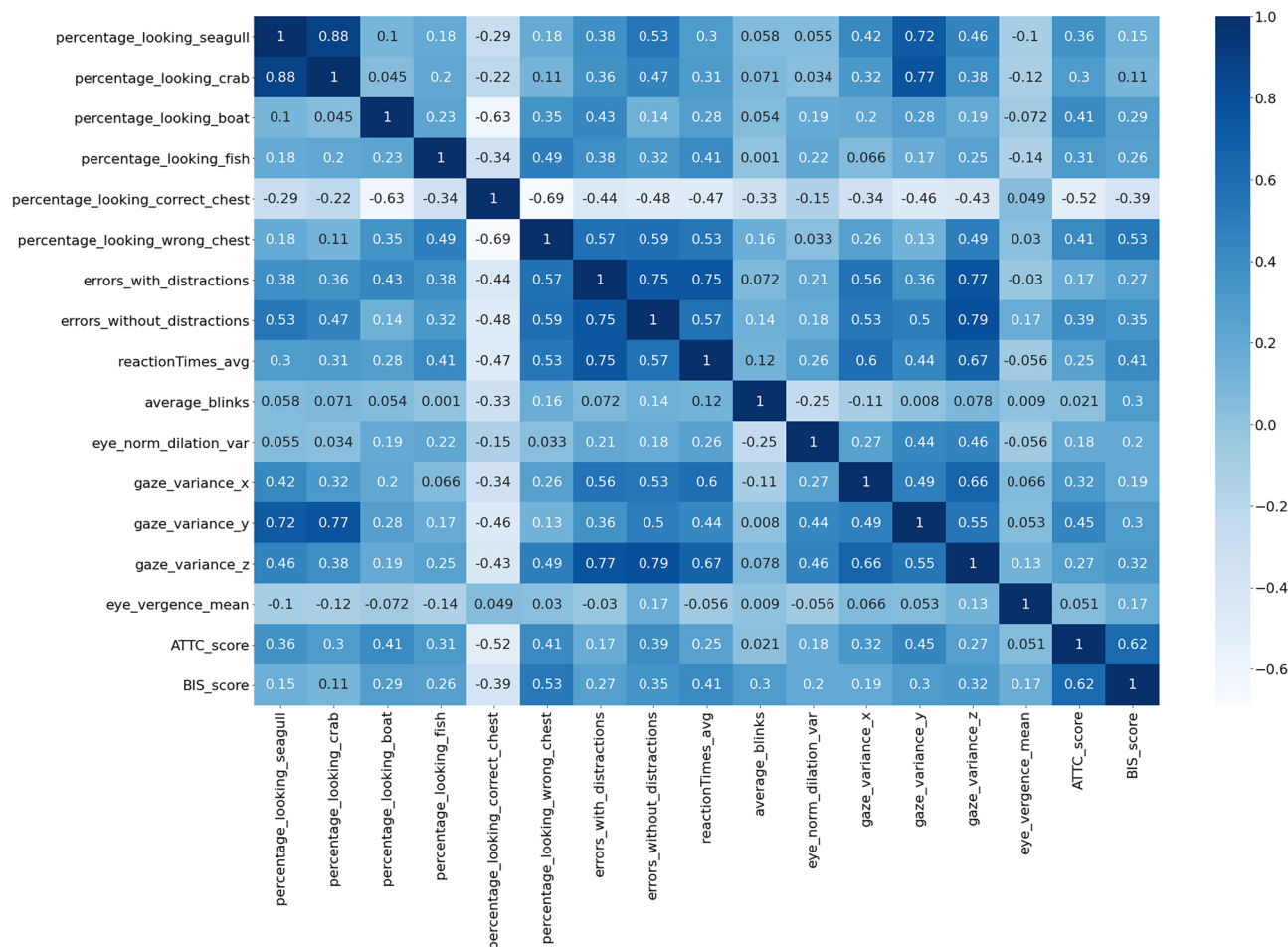


Figure 3. Correlation matrix between the predictors and the ATTC and BIS-11 scale scores.

0.45, z 0.27), there is a positive correlation between the variability of eye movement in those directions and the ATTC score. Additionally, the eye dilation correlation is 0.18.

In terms of impulsivity, we see a correlation of 0.53 between a BIS-11 score with the time spent looking at the wrong chest. Besides, the correlation between the time looking at the correct chest and the BIS-11 score is inverse -0.39 (less impulsive subjects tend to look more to the correct chest). Errors made with or without distractors have a positive correlation with impulsivity (0.27 and 0.35 respectively). Like it happened with attention, there is a positive correlation between eye movement (x 0.19, y 0.3, z 0.32). Pupil dilation, eye vergence and blinks also show a positive correlation with impulsivity (0.2, 0.17 and 0.3, respectively). A positive correlation between impulsivity and inattention with respect to reaction time (0.41 and 0.25 respectively) was also found.

Next, we present the model's predictive performance. In order to obtain a reference baseline, the average of the target scores of the training group was used as an estimate of the participants' scores. The results are summarized in Table 1. Regarding the results for attention, the baseline RMSE was 8.946, whereas the test RSME was 8.269. The test correlation was 0.434. Regarding the results for impulsivity, the baseline RMSE (11.203) was again higher than the test RSME (10.538). The test correlation was 0.382.

The models also output the features' importance; this can be understood as the weight each model gives to the variables. In the case of attention (see Fig. 4), the model indicates as the most important variables the time spent looking at the boat (distractor), together with the eye movements variability in the vertical axis. Regarding

	Attention	Impulsivity
Baseline RMSE	8.946	11.203
Test RMSE	8.269	10.538
Test correlation	0.434	0.382

Table 1. RMSE and correlations obtained by the model for attention and impulsivity. The test RMSE was better than the baseline for both constructs.

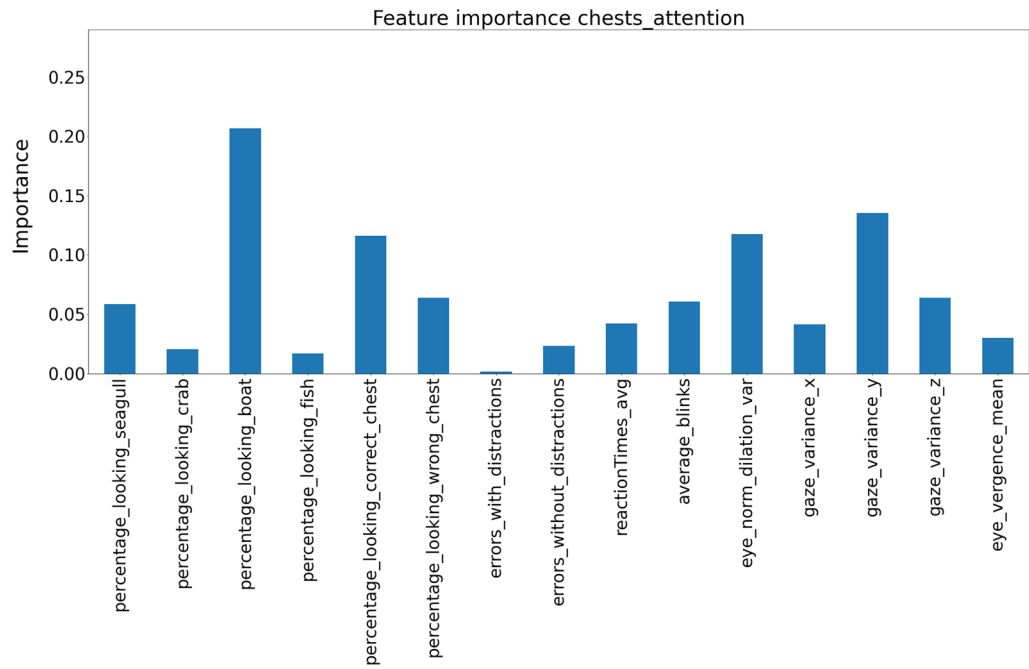


Figure 4. Features importance between the chest game and the attention score.

the most important variables for predicting impulsivity according to the model (see Fig. 5), the time looking at an incorrect chest was the most important variable and then the number of blinks.

In the next section, we will discuss these results and compare them with previous results in the literature.

Institutional review. The study was conducted in accordance with the Declaration of Helsinki, and approved by the Ethics Committee of Universidad Autónoma de Madrid (protocol code CEI-125-2575 and date of approval 15/07/2022).

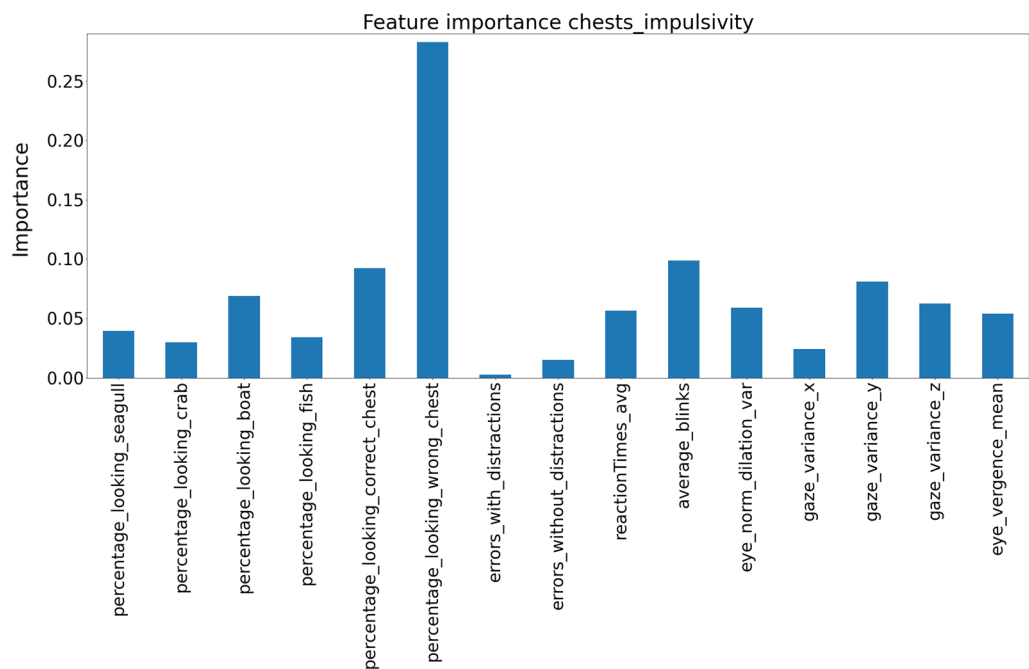


Figure 5. Features importance between the chest game and the impulsivity score.

Informed consent. Informed consent was obtained from all subjects involved in the study.

Discussion

In this work, we have proposed the Treasure Hunt VR application that gathers performance and behavioural data that are then used to accurately predict the scores of the ATTC and BIS-11 self-report scales. One of the advantages of the model is that it allows us to know which are the most informative variables (see Figs. 4 and 5). Our results are generally consistent with what multiple studies suggest as important regressors for assessing attention and impulsivity.

First of all, errors are the most common predictors in the literature⁴³. Both high impulsivity and inattention make the participant more likely to make mistakes while performing a task. Previous studies showed that there is a clear relationship between the number of errors and the level of inattention and impulsivity^{21,40,44}. The results of our model corroborate this.

Also of interest are the results related to the distractors. For example, some studies claim that distractors help to assess ADHD^{19,45}. In our case, there is a positive correlation between the percentage of time looking at the distractors and the test scores. This correlation is higher with the attention score than with the impulsivity score. It should be noted that the different distractors have different influences on the user. The most relevant distractor seems to be the boat, which is also the biggest. In comparison, the crab distractor, perhaps because it is smaller or because its movement is limited to a single small area, has less influence.

Another interesting variable is the reaction time. In the literature, we can find that reaction time can be a predictor of the user's attention⁴⁴. An inattentive person will take longer to react, and an impulsive person will have more variability in reaction time³⁸. In our case, although the reaction time is not the most important variable according to the model, we can observe that it correlates positively with ATTC and BIS-11 scores. This could be due to the fact that players who had followed the impulse to look at the distractors, at the moment when the movement of the chests stopped, could take longer to choose the treasure chest.

Our model shows that eye predictors are also related to inattention and impulsivity. According to our results, it is remarkable the positive correlation between eye vergence with impulsivity, agreeing with the results of Varela et al.³⁶. Our results show the importance of the number of blinks for impulsivity assessment. This is consistent with the work of Ray Li et al.⁴⁶ where they detected a greater number of blinks in participants with higher impulsivity scores. Also, our results show that gaze variance is related to both inattentiveness and impulsiveness. Lev et al.⁴⁷ found that ADHD patients spent more time gazing at irrelevant regions.

Furthermore, our model shows that the variance in pupil dilation is an important variable. However, the relationship is direct (more inattention, more dilation variability). This result seems at odds with some others in the literature^{48,49} but this can be due to the fact that the task in those studies was different since it was a visuo-spatial working memory task.

It has been observed that it is possible to obtain information about the player in an objective way, so that it can be used to predict the degree of attention and impulsivity. This is relevant because it minimizes the problem of subjectivity or malingering when answering scales or questionnaires. Our proposal differs from previous works because it has a game-like appearance and uses virtual reality devices, which allows us to continuously collect a lot of information to get a more global view of player behaviour. In addition, rather than looking at the number of errors made while distractors are present, as most studies do, we analyze the entire time spent looking at the distractors. These findings contribute to the field by highlighting the importance of registering temporal information, instead of simply using, for instance, the number of errors. Additionally, this work shows that it is possible to use machine learning techniques to help in the analysis and prediction of constructs, in this case, the degree of attention and impulsivity. The potential of these techniques could also be used to help validate new scales or questionnaires that measure the same constructs.

Our approach sheds light on which of the variables collected are more relevant and correlate more with the score on the scales measuring the degree of impulsivity and inattention. Additionally, it opens the field to other studies that could focus on predicting other constructs. However, we recommend developers of therapeutic, diagnostic or assessment games to be very specific about which constructs or executive functions they wish to measure or improve¹⁸. In this way, the game can be designed as simply as possible to focus only on the target constructs or executive functions to minimize interference of other constructs or EFs.

In general, the analysis of the data shows that our results are consistent with previous findings in the literature, which suggests that using VR applications including games like the one presented can be useful as an additional way of information to assess attention and impulsivity.

Limitations. The main limitation of our work is the small sample size which may not be enough to obtain a better generalization of the model. Another limitation is that occlusions when registering hands movements with this particular version of the hardware led us to discard the hands' data for the analysis. Our study was also limited by the variables the VR device was able to register. As discussed in the related work, other devices could provide new information that could be included to improve the predictions. Additionally, some subjects in our study suggested that the duration of 10 minutes per game was too long and that they might feel somewhat overwhelmed. Further work should study the optimal duration of the games depending on the target sample.

Conclusion

In this work, we developed a VR application including a game to capture variables that were then used to create a machine learning model to predict the level of attention and impulsivity of the players.

A theoretical contribution of our work is that we have shown the usefulness of recording objective variables continuously. This allows us to study user behavior, seeing not only the things the user has done, but also

observing the time spent on each action. Our method provides a way of knowing which variables are the most important when generating a predictive model for inattention and impulsivity and we plan to study how this could assist in the diagnosis of ADHD in future work. The random forest model can help us to improve the effectiveness of new VR applications since it can tell which variables or game mechanics are less relevant and therefore susceptible of being eliminated. For example, in our case, the model indicates that the distractor that affects the least is the crab.

Finally, the hardware currently used (HP Reverb G2 Omnicept and Laptop with an NVidia 3060 Graphics card) is relatively expensive, somewhat heavy and, the version used in this study suffered from hand capture occlusions. For future studies, we recommend the most recent version of the hardware which incorporates two additional cameras to track the hands and hence palliates this problem. Also, those kinds of VR devices may not be suitable for all individuals, which would make it difficult to apply, for the time being, the proposed approach as a general assessment technique. Hopefully, as this technology is in a maturing phase, better, cheaper and lighter solutions will become available to the common user in the coming years, enabling this type of VR application to become more widely used. Other forms of interaction could also be used in future work. For example, other motion capture devices also record the movement of other parts of the body or even brain-computer interactions.

Data availability

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Received: 10 January 2023; Accepted: 10 August 2023

Published online: 22 August 2023

References

- Bruchmüller, K., Margraf, J. & Schneider, S. Is ADHD diagnosed in accord with diagnostic criteria? Overdiagnosis and influence of client gender on diagnosis. *J. Consult. Clin. Psychol.* **80**, 128. <https://doi.org/10.1037/a0026582> (2012).
- Kooij, J. *et al.* Updated European consensus statement on diagnosis and treatment of adult ADHD. *Eur. Psychiatry* **56**, 14–34. <https://doi.org/10.1016/j.eurpsy.2018.11.001> (2019).
- Klonsky, E. D. & May, A. Rethinking impulsivity in suicide. *Suicide Life-Threat. Behav.* **40**, 612–619. <https://doi.org/10.1521/suli.2010.40.6.612> (2010).
- Dawe, S. L. N. The role of impulsivity in the development of substance use and eating disorder. *Neurosci. Biobehav. Rev.* **28**, 343–351. <https://doi.org/10.1016/j.neubiorev.2004.03.007> (2004) (**Festschrift in Honour of Jeffrey Gray - Issue 1: Anxiety and Neuroticism**).
- Nasiri, E. *et al.* A comprehensive review of attention tests: Can we assess what we exactly do not understand?. *Egypti. J. Neurol. Psychiatry Neurosurg.* **59**, 26. <https://doi.org/10.1186/s41983-023-00628-4> (2023).
- Merrill, C. & Tymms, P. B. Inattention, hyperactivity and impulsiveness: Their impact on academic achievement and progress. *Br. J. Educ. Psychol.* **71**, 43–56. <https://doi.org/10.1348/000709901158389> (2001).
- McClelland, M. M., Acock, A. C., Piccinin, A., Rhea, S. A. & Stallings, M. C. Relations between preschool attention Span-Persistence and age 25 educational outcomes. *Early Child. Res. Q* **28**, 314–324 (2012).
- Čosić, K. *et al.* New tools and methods in selection of air traffic controllers based on multimodal psychophysiological measurements. *IEEE Access* **7**, 174873–174888. <https://doi.org/10.1109/ACCESS.2019.2957357> (2019).
- Próchniak, P. Traits of personality and preferred values among fearless and courageous policemen. *Int. J. Psychol. Stud.* **4**, 113 (2012).
- Flynn, R. M., Colón-Acosta, N., Zhou, J. & Bower, J. A game-based repeated assessment for cognitive monitoring: Initial usability and adherence study in a summer camp setting. *J. Autism Dev. Disord.* **49**, 2003–2014. <https://doi.org/10.1007/s10803-019-03881-w> (2019).
- Hirsch, O. & Christiansen, H. Faking ADHD? Symptom validity testing and its relation to self-reported, observer-reported symptoms, and neuropsychological measures of attention in adults with ADHD. *J. Atten. Disord.* **22**, 269–280. <https://doi.org/10.1177/1087054715596577> (2018).
- Rosvold, H. E., Mirsky, A. F., Sarason, I., Bransome, E. D. Jr. & Beck, L. H. A continuous performance test of brain damage. *J. Consult. Psychol.* **20**, 343. <https://doi.org/10.1037/h0043220> (1956).
- Trommer, B. L., Hoepfner, J.-A.B., Lorber, R. & Armstrong, K. J. The go-no-go paradigm in attention deficit disorder. *Ann. Neurol.* **24**, 610–614. <https://doi.org/10.1002/ana.410240504> (1988).
- Lumsden, J., Edwards, E. A., Lawrence, N. S., Coyle, D. & Munafò, M. R. Gamification of cognitive assessment and cognitive training: A systematic review of applications and efficacy. *JMIR Serious Games* **4**, e11. <https://doi.org/10.2196/games.5888> (2016).
- Tong, T., Chignell, M., Tierney, M. C. & Lee, J. A serious game for clinical assessment of cognitive status: Validation study. *JMIR Serious Games* **4**, e7. <https://doi.org/10.2196/games.5006> (2016).
- Ijaz, K., Ahmadpour, N., Naismith, S. L. & Calvo, R. A. An immersive virtual reality platform for assessing spatial navigation memory in predementia screening: Feasibility and usability study. *JMIR Ment. Health* **6**, e13887. <https://doi.org/10.2196/13887> (2019).
- Bioulac, S. *et al.* Virtual remediation versus methylphenidate to improve distractibility in children with ADHD: A controlled randomized clinical trial study. *J. Atten. Disord.* **24**, 326–335. <https://doi.org/10.1177/1087054718759751> (2020).
- Sújar, A. *et al.* Developing serious video games to treat attention deficit hyperactivity disorder: Tutorial guide. *JMIR Serious Games* **10**, e33884. <https://doi.org/10.2196/33884> (2022).
- Berger, I., Slobodin, O. & Cassuto, H. Usefulness and validity of continuous performance tests in the diagnosis of attention-deficit hyperactivity disorder children. *Arch. Clin. Neuropsychol.* **32**, 81–93. <https://doi.org/10.1093/arclin/acw101> (2017).
- Pollak, Y. *et al.* The utility of a continuous performance test embedded in virtual reality in measuring ADHD-related deficits. *J. Dev. Behav. Pediatr.* **30**, 1–10. <https://doi.org/10.1097/DBP.0b013e3181969b22> (2009).
- Crepaldi, M. *et al.* The use of a serious game to assess inhibition mechanisms in children. *Front. Comput. Sci.* <https://doi.org/10.3389/fcomp.2020.00034> (2020).
- Waltz, E. First video game to treat disease gains fda okay. *Nat. Biotechnol.* **38**, 1224. <https://doi.org/10.1038/s41587-020-0726-6> (2020).
- Delgado-Gómez, D. *et al.* Objective assessment of attention-deficit hyperactivity disorder (ADHD) using an infinite runner-based computer game: A pilot study. *Brain Sci.* <https://doi.org/10.3390/brainsci10100716> (2020).
- Sujar, A. *et al.* Attention deficit hyperactivity disorder assessment based on patient behavior exhibited in a car video game: A pilot study. *Brain Sci.* <https://doi.org/10.3390/brainsci12070877> (2022).

25. Clarke, P. J. & Todd, J. Lessons unlearned: A conceptual review and meta-analysis of the relationship between the attention control scale and objective attention control. *Cogn. Emot.* **35**, 1447–1459. <https://doi.org/10.1080/02699931.2021.1987861> (2021).
26. Derryberry, D. & Reed, M. A. Anxiety-related attentional biases and their regulation by attentional control. *J. Abnorm. Psychol.* **111**, 225–236. <https://doi.org/10.1037/0021-843X.111.2.225> (2002).
27. Stanford, M. S. *et al.* Fifty years of the Barratt impulsiveness scale: An update and review. *Pers. Individ. Diff.* **47**, 385–395. <https://doi.org/10.1016/j.paid.2009.04.008> (2009).
28. Bates, M. E. & Lemay, E. P. The d2 test of attention: Construct validity and extensions in scoring techniques. *J. Int. Neuropsychol. Soc.* **10**, 392–400. <https://doi.org/10.1017/S135561770410307X> (2004).
29. de-la Peña, M. T. C., Otero, J. M. & Romero, E. Comparison among various methods of assessment of impulsiveness. *Percept. Motor Skills* **77**, 567–575. <https://doi.org/10.2466/pms.1993.77.2.567> (1993).
30. Mackworth, J. F. & Taylor, M. M. The d' measure of signal detectability in vigilancelike situations. *Can. J. Psychol.* **17**, 302. <https://doi.org/10.1037/h0092885> (1963).
31. Homack, S. & Riccio, C. A. Conners' continuous performance test (2nd ed.; ccpt-ii). *J. Atten. Disord.* **9**, 556–558. <https://doi.org/10.1177/1087054705283578> (2006).
32. Greenberg, L. M. & Waldmant, I. D. Developmental normative data on the test of variables of attention (t.o.v.a.TM). *J. Child Psychol. Psychiatry* **34**, 1019–1030. <https://doi.org/10.1111/j.1469-7610.1993.tb01105.x> (1993).
33. Edwards, M. C. *et al.* Estimates of the validity and utility of the Conners' continuous performance test in the assessment of inattentive and/or hyperactive-impulsive behaviors in children. *J. Abnorm. Child Psychol.* **35**, 393–404. <https://doi.org/10.1007/s10802-007-9098-3> (2007).
34. O'Mahony, N., Florentino-Liano, B., Carballo, J. J., Baca-García, E. & Rodríguez, A. A. Objective diagnosis of ADHD using IMUs. *Med. Eng. Phys.* **36**, 922–926. <https://doi.org/10.1016/j.medengphy.2014.02.023> (2014).
35. Delgado-Gomez, D. *et al.* Microsoft kinect-based continuous performance test: An objective attention deficit hyperactivity disorder assessment. *J. Med. Internet Res.* **19**, e79. <https://doi.org/10.2196/jmir.6985> (2017).
36. Varela Casal, P. *et al.* Clinical validation of eye vergence as an objective marker for diagnosis of ADHD in children. *J. Atten. Disord.* **23**, 599–614. <https://doi.org/10.1177/1087054717749931> (2019).
37. Snyder, S. M., Rugino, T. A., Hornig, M. & Stein, M. A. Integration of an EEG biomarker with a clinician's ADHD evaluation. *Brain Behav.* **5**, e00330. <https://doi.org/10.1002/brb3.330> (2015).
38. Rodríguez, C., Arecos, D., García, T., Cueli, M. & González-Castro, P. Comparison between two continuous performance tests for identifying ADHD: Traditional vs. virtual reality. *Int. J. Clin. Health Psychol.* **18**, 254–263. <https://doi.org/10.1016/j.ijchp.2018.06.003> (2018).
39. Peñuelas-Calvo, I. *et al.* Video games for the assessment and treatment of attention-deficit/hyperactivity disorder: A systematic review. *Eur. Child Adolesc. Psychiatry* **31**, 5–20. <https://doi.org/10.1007/s00787-020-01557-w> (2022).
40. Peijnenborgh, J. C. *et al.* A study on the validity of a computer-based game to assess cognitive processes, reward mechanisms, and time perception in children aged 4–8 years. *JMIR Serious Games* **4**, e15. <https://doi.org/10.2196/games.5997> (2016).
41. Hay, D. A., Bennett, K. S., Levy, F., Sergeant, J. & Swanson, J. A twin study of attention-deficit/hyperactivity disorder dimensions rated by the strengths and weaknesses of ADHD-symptoms and normal-behavior (swan) scale. *Biol. Psychiatry* **61**, 700–705. <https://doi.org/10.1016/j.biopsych.2006.04.040> (2007).
42. Cutler, A., Cutler, D. R. & Stevens, J. R. *Random Forests* (Springer, 2012).
43. Parsons, T. D., Duffield, T. & Asbee, J. A comparison of virtual reality classroom continuous performance tests to traditional continuous performance tests in delineating ADHD: a meta-analysis. *Neuropsychol. Rev.* **29**, 338–356. <https://doi.org/10.1007/s11065-019-09407-6> (2019).
44. Faraone, S. V. *et al.* The groundskeeper gaming platform as a diagnostic tool for attention-deficit/hyperactivity disorder: Sensitivity, specificity, and relation to other measures. *J. Child Adolesc. Psychopharmacol.* **26**, 672–685. <https://doi.org/10.1089/cap.2015.0174> (2016).
45. Parsons, T. D., Bowerly, T., Buckwalter, J. G. & Rizzo, A. A. A controlled clinical comparison of attention performance in children with ADHD in a virtual reality classroom compared to standard neuropsychological methods. *Child Neuropsychol.* **13**, 363–381. <https://doi.org/10.1080/13825580600943473> (2007).
46. Ray Li, C., Chen, S.-H., Lin, W. & Yang, Y. Attentional blink in adolescents with varying levels of impulsivity. *J. Psychiatr. Res.* **39**, 197–205. <https://doi.org/10.1016/j.jpsychires.2004.06.003> (2005).
47. Lev, A., Braw, Y., Elbaum, T., Wagner, M. & Rassovsky, Y. Eye tracking during a continuous performance test: Utility for assessing ADHD patients. *J. Atten. Disord.* **26**, 245–255. <https://doi.org/10.1177/1087054720972786> (2022).
48. Das, W. & Khanna, S. A novel pupillometric-based application for the automated detection of ADHD using machine learning. In *Proceedings of the 11th ACM International Conference on Bioinformatics, Computational Biology and Health Informatics, BCB '20*. Association for Computing Machinery. <https://doi.org/10.1145/3388440.3412427> (2020).
49. Wainstein, G. *et al.* Pupil size tracks attentional performance in attention-deficit/hyperactivity disorder. *Sci. Rep.* <https://doi.org/10.1038/s41598-017-08246-w> (2017).

Acknowledgements

This research was partially funded by: Grant TED2021-130980B-I00 funded by MCIN/AEI/10.13039/501100011033 and by the “European Union NextGenerationEU/PRTR”. Grant RED2022-134259-T funded by MCIN/AEI/10.13039/501100011033. Project “DTS21/00091”, funded by Instituto de Salud Carlos III (ISCIII) and co-funded by the European Union. Grant for the requalification of permanent lectures, UC3M, David Delgado-Gómez.

Author contributions

Conceptualization, D.D.-G., A.S. and D.M.-E.; methodology, D.D.-G. and D.M.-E.; software, D.M.-E. and A.S.; formal analysis, D.D.-G. and D.M.-E.; resources, D.D.-G. and D.M.-E.; data curation, D.M.-E.; investigation D.D.-G., A.S. and D.M.-E.; writing-original draft preparation, A.S., S.B. and D.D.-G.; writing-review and editing, A.S., S.B. and D.D.-G.; project administration, D.D.-G.; funding acquisition, S.B. and D.D.-G. All authors have read and agreed to the published version of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-40455-4>.

Correspondence and requests for materials should be addressed to A.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023