

RESEARCH

Open Access



# Playing for cognition: investigating the feasibility and user experience of a virtual reality serious game for cognitive assessment in children with congenital heart disease

C. L. van de Wouw<sup>1</sup>, E. Bousché<sup>1</sup>, M. M. J. van Schooneveld<sup>2</sup>, J. M. P. Breur<sup>3</sup>, H. P. Nijenhuis<sup>4</sup>, H. Huygelier<sup>1,5</sup> and T. C. W. Nijboer<sup>1\*</sup>

## Abstract

**Background** In order to facilitate the development and implementation of innovative technology in clinical practice, it is important to understand the user experience of end-users. Virtual Reality (VR) offers the possibility to assess cognitive functioning in a dynamic environment that simulates real-world situations. The purpose of this cross-sectional study was to investigate the feasibility of a VR Serious Game for cognitive assessment in school-aged children with congenital heart disease (CHD). The sub-aims were two-fold: (1) to objectively evaluate the feasibility of the VR Serious Game in children with CHD in comparison to typically developing (TD) children and (2) to explore the user experience of both groups following their interaction with the VR Serious Game.

**Results** A total of 101 children participated in this study; 98 children were included in the final analysis (CHD:  $n = 54$ ; TD:  $n = 47$ ). The VR Serious Game appeared feasible for both children with CHD and TD children, with 88% children completing the innovative VR assessment without encountering any issues. There were no discernible differences in completion rates between groups. Children with CHD reported significantly lower scores than TD children on three user experience scales: *Engagement*, *Flow* and *Presence*. Nonetheless, the scores for *Engagement* and *Flow* were still considered "moderate to good". Both groups reported minimal adverse physiological reactions.

**Conclusions** The findings suggested that the VR Serious Game was feasible for children with CHD and that the user experience was positive. Future research should investigate the effectiveness of the VR Serious Game compared with a conventional or digital neuropsychological assessment, prioritising the development of novel outcome measures that can better estimate and explain the impact of cognitive impairment on daily functioning.

**Keywords** Clinical implementation, Congenital heart disease, Cognition, Serious game, Virtual reality

\*Correspondence:

T. C. W. Nijboer

t.c.w.nijboer@uu.nl

Full list of author information is available at the end of the article



© The Author(s) 2024. **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/>.

## Background

Children with congenital heart disease (CHD) who undergo early-life cardiac surgery are at risk of cognitive impairment, particularly in executive functions (EFs) [1–7]. This can contribute to a developmental cascade, hindering academic achievement and psychosocial outcomes, whilst exacerbating long-term restrictions in daily life [5, 6, 8]. Cognitive impairment may be latent in early development, emerging as children face increased academic demands, a phenomenon referred to as “growing into deficits” [9]. This stresses the importance of assessing cognitive functioning in school-age children to better predict their capabilities in daily life.

Cognitive functioning is typically assessed with a conventional neuropsychological assessment (NPA; i.e., paper-and-pencil tests), which aims to estimate a child's maximum cognitive capacity. The NPA is administered in a quiet room with minimal external distractions, devoid of the complexities of real-life situations [10]. As a result, clinicians face challenges in estimating and explaining the impact of cognitive impairment on daily functioning and providing meaningful recommendations for daily life [11, 12].

To address these limitations, Virtual Reality (VR) has emerged as a promising tool for more sensitive and ecologically valid cognitive assessment, as compared to conventional NPA [13, 14]. VR encompasses computer-generated interactive virtual environments that simulate the real world, ranging from *low-immersive* VR with a limited field of view (e.g., ‘window on the world’ 2D displays) to *high-immersive* VR with a 360° field of view (e.g., head-mounted 3D displays [HMDs]; [15]).

Serious Gaming involves the gamification of tasks, such as goal setting, narrative, feedback and rewards, used in nonentertainment settings [16]. By combining VR game performance with NPA test results, clinicians can obtain isolated scores per cognitive function as well as clinically relevant and objective scores that reflect cognitive functioning in real-life situations [17]. Ultimately, this approach may offer valuable insights into a child's cognitive capabilities, bridging the gap between a child's cognition at the level of functions and at the level of activities [18].

VR simulations demonstrate feasibility in various paediatric clinical populations, including acquired brain injury (see [19–21]) and attention deficit hyperactivity disorder (see [20]). However, there is a paucity of research evaluating the feasibility of Serious Gaming in children with CHD, despite a clear clinical need. Therefore, the overarching aim of the present study was to evaluate the feasibility and user experience of employing a VR Serious Game for cognitive assessment in school-aged children with CHD, compared to typically developing (TD)

children. Feasibility (i.e., ability to complete assessment) is a necessary step for implementing technology in clinical practice. Prioritising user-friendliness enhances the feasibility of technology, increasing the likelihood of successful implementation in clinical practice [22].

## Methods

### Participants

Children with severe CHD, demanding surgical intervention within their first year of life [23], were consecutively enrolled from the cardiology departments of both the Wilhelmina Children's Hospital in Utrecht and the University Medical Centre in Groningen, the Netherlands, under the supervision of paediatric cardiologists. In addition, recruitment of children with CHD was facilitated through announcements on the Hartekind Foundation (*Stichting Hartekind*) website. Eligible children met the following criteria: (1) aged between 10 and 13 years; (2) voluntary participation; and (3) informed consent provided by guardians, as well as by children aged 12 years and older. This age range was decided upon, as clinical follow-up was up to eight years of age and there was a need for follow-up at age 10 and beyond. Following the ‘growing into deficits’ principle—combined with the need to also ensure a homogenous group – the upper age limit was 13 years of age, since environmental demands change especially during this age range.

TD children were recruited by contacting schools and (sport) associations in the Brabant, Groningen, and Utrecht regions of the Netherlands. Inclusion criteria included: (1) aged between 10 and 13 years old; (2) no history of neurological or psychiatric disease; (3) voluntary participation; and (4) informed consent provided by guardians, as well as by children aged 12 years and older.

For both groups, children diagnosed with epilepsy were excluded due to photosensitivity and reflex seizures [24]. The inclusion period spanned from September 2021 to December 2022.

The present study received ethical approval from the Faculty Ethics Review Board of the Faculty of Social and Behavioural Sciences at Utrecht University (UU; protocol numbers 20–648, 21–0136, 21–0369, and 22–0058) and the Medical Research Ethics Committee (protocol number 20–693/C). The study protocol adhered to the Declaration of Helsinki, as well as the requirements of UU and the Faculty Ethics Review Board.

### Procedure, tests, and outcome measures

Children with CHD were identified through examination of the electronic patient database at participating hospitals by paediatric cardiologists. Eligible children and their guardians were provided with a brief study overview during their hospital appointments by their cardiologist. If

guardians expressed interest, a researcher from UU initiated contact, sharing the information letter and consent form through either email or post. In situations where no hospital appointment was scheduled, researchers from UU contacted the guardians by telephone to provide them with study information. For children who registered on the Hartekind Foundation (*Stichting Hartekind*) website, their eligibility was determined by a UU researcher. Subsequently, an information letter and consent form were sent to the guardians. Guardians were required to respond to schedule an appointment, and UU researchers followed up after one week.

TD children were recruited by contacting schools and (sport) associations in the Brabant, Groningen, and Utrecht regions of the Netherlands. Researchers from UU initiated contact by sending an email containing an information letter and an informed consent form. Schools and (sport) associations were required to respond to schedule an appointment, with UU researchers following up after one week.

The informed consent, signed by guardians and children aged 12 years and older, could either be emailed or brought in person to the appointment.

For children with CHD, the test protocol comprised a single 90-min session conducted at participating hospitals. The whole session included a conventional NPA and a VR Serious Game. The session was divided into two blocks, with a short break in between the two blocks. Additional breaks were accommodated as needed. The assessment took place in a quiet room with minimal external distractions and was administered by a trained researcher. The sequence of cognitive assessments (conventional NPA [A; approximately 60 min] and innovative VR Serious Game [B; approximately 30 min]) was counterbalanced, alternating between A and B blocks of tasks. The data from both the NPA and the VR Serious Games were used for the current aims. The innovative VR Serious Game was counterbalanced and divided into two 'quests' (i.e., 1 and 2). In the first quest, the following minigames were administered in this order: Alien Outpost, Nautilus Underwater Memory, Mystical Lake, and Galactic Diamond Belt. The second quest comprised the following minigames in this order: Mystical Lake, Galactic Diamond Belt, Alien Outpost, and Nautilus Underwater Memory. All children interacted with the VR Serious Game whilst seated.

The researcher helped the children with fitting the VR goggles to their heads and gave consecutive instructions per game. This was repeated after the block of NPA for the children with CHD. The virtual environment was streamed to a second device, so the researcher could observe what the children were presented with. For TD children, the test protocol (innovative VR Serious Game)

involved a single 30-min session conducted at participating hospitals, schools and (sport) associations. TD children were assessed in groups of two up to four at the same time. Again, the researcher helped the TD children with fitting the VR goggles to their heads and gave consecutive instructions per game.

Relevant clinical information for the children with CHD was gathered from the electronic patient database. This information included diagnosis (e.g., Atrioventricular Septal Defect, Hypoplastic Left Heart Syndrome, Tetralogy of Fallot, Transposition of the Great Arteries, and Tricuspid Atresia), the age at the time of their first surgery (in days), as well as the total number of surgeries they had undergone before the age of six months.

Prior to assessment, children reported their daily gaming hours. Age and sex of each child were recorded. Following the assessment, children completed a questionnaire to evaluate their VR user experience (see [12]).

#### **VR serious game**

**Apparatus** The Serious Game 'Koji's Quest', developed by NeuroReality [25], was employed. The Serious Game was run on Oculus Meta Quest 2 head-mounted display VR glasses (immersive 360-degree field of view and a fast-switch LCD display with 1832 × 1920 resolution per eye). Children interacted with the VR Serious Game using touch controllers.

**Divided attention** *Alien Outpost* aimed to assess children's divided attention using two four-minute paradigms. In the first paradigm, children had to determine whether an alien appeared in a central spaceship within a space-themed setting (i.e., central attention). They responded by pressing either a green button (alien present) or a red button (no alien). In the second paradigm, an alien always appeared in the central spaceship, but children had to discern whether the alien looked "happy" or "sad" (i.e., central attention). They responded by pressing either a green button ("happy-looking" alien) or a red button ("sad-looking" alien). In both paradigms, children were also required to identify an alien in the surrounding spaceships by selecting the spaceship where the alien appeared (i.e., peripheral attention). The central flash duration (i.e., alien display duration in the central spaceship) was adjusted based on difficulty level, whilst the peripheral flash duration (i.e., alien display duration in the surrounding spaceships) remained fixed. In addition, the number of surrounding spaceships and distractor stimuli were adjusted based on difficulty level. Accuracy and difficulty levels for both central and peripheral attention were recorded separately.

**Executive functions** *Mystical Lake* aimed to assess children's EFs using three paradigms, each lasting for

three, three, and four minutes, respectively. In the first paradigm, children were presented with individual fish jumping out of the water, and their task was to quickly identify the correct food associated with each fish species. Children had to learn and remember three different specific food-fish associations. The second paradigm challenged children with multiple fish jumping out of the water simultaneously, each at a different pace. In addition to recalling the correct food-fish associations, children had to prioritise the order in which the fish should be fed. In the third paradigm, multi-tasking was introduced. Children had to monitor water levels in fishbowls, which gradually decreased over time. To prevent the fish from going hungry, children had to continuously refill the fishbowls whilst also ensuring they fed both the fast- and slow-moving fish. The number of fish jumping out of the water simultaneously, their jumping speed, as well as the rate at which fishbowls ran out of water, were dependent on difficulty level.

**Selective attention** *Galactic Diamond Belt* aimed to measure children's selective attention using a single four-minute visual search paradigm. In this task, children were required to select a single target (i.e., specific gem of colour and shape) within a space-themed setting whilst ignoring irrelevant distractors. Distractors, which included gems of other shapes and colours, were randomly scattered across space. The number of distractors, display duration, and speed at which the gems moved, were dependent on difficulty level.

**Memory** *Nautilus Underwater Memory* aimed to assess children's short-term memory using two three-minute paradigms. In the first paradigm, children were instructed to memorise a sequence of seashells opening and then reproduce the sequence. The sequence was the same across children. In the second paradigm, children were asked to memorise the spatial locations of various objects for a self-determined period (up to 30s) and then reproduce the object placements. The length of sequence and the number of objects that children needed to memorise were dependent on difficulty level.

For all mini-games, the starting difficulty level was set at 500. Difficulty levels varied from 100 to 1000 based on a 3-up 1-down staircase.

#### **User experience**

The user experience questionnaire [12] consisted of 15 items, comprising five categories: *Engagement* (i.e., the feeling of active involvement and enjoyment of the content; [12]), *Flow* (i.e., the mental state characterised by full immersion in an activity, intense focus, and a distorted perception of time; [12]), *Presence* (i.e., the feeling

of full immersion within a virtual environment), *Side effects* (i.e., adverse physiological reactions, such as nausea), and *Transportation* (i.e., the feeling of being transported to an alternate world; [12]). Each item was rated on a six-point Likert scale, ranging from negative (0) to positive (5).

#### **Statistical analyses**

##### **Participant characteristics**

Demographic characteristics (i.e., age and sex) were compared between groups (CHD versus TD children) using non-parametric tests. For the continuous variable, age, a Mann–Whitney U test was performed, whilst for the categorical variable, sex, a Chi-square test was performed. Furthermore, mean and standard deviation (SD) were determined for the number of surgeries. Median and interquartile range (IQR) were determined for age at first surgery.

##### **Feasibility**

*Feasibility* was determined based on the number of children who completed the innovative VR assessment. Logistic regression was conducted to investigate whether groups differed in their rates of completing the innovative VR assessment, with sex as a covariate.

##### **User experience**

To evaluate *user experience*, scale scores were calculated for five categories: *Engagement*, *Flow*, *Presence*, *Side effects*, and *Transportation*. Each scale score was derived by summing the responses from three related items within the respective category, resulting in a scale score ranging from 0 to 15. A score of 0 indicated a very negative experience, and a score of 15 indicated a very positive experience. However, a higher score indicated more side effects. One item (item 7—"I felt present in the virtual environment") was missing from the questionnaire, therefore the scale score for *Presence* was calculated from two related items. To compare the user experience between CHD and TD children, an ANCOVA was conducted, with sex as a covariate. Post hoc analysis was performed with a Bonferroni adjustment.

#### **Results**

A total of 101 children participated in this study. This comprised 54 children with CHD and 47 TD children. However, three children with CHD were subsequently excluded from the analysis for the following reasons: (1) diagnosis of CHARGE syndrome, a rare genetic disorder associated with a wide range of health and physical problems ( $n = 1$ ). This exclusion was made after extracting relevant clinical information from the electronic patient database; children with CHARGE syndrome

often experience cognitive impairment [26]; and (2) age-related discrepancies ( $n=2$ ). Two children were excluded because their age at the time of testing was

14 years, which exceeded the age originally scheduled for the appointment. Subsequently, a total of 98 children were included in the final analysis. This consisted of 51 children with CHD and 47 children with TD.

**Table 1** Demographic and clinical characteristics, and user experience comparison between CHD and TD children

	CHD (n=51)	TD (n=47)	P
Age, years, Mean (SD)	11.20 (1.28)	11.00 (.72)	.34
Sex, male (%)	64.7	38.3	.01**
Age first surgery, days, Median (IQR)	4, 61		
Number surgeries, Mean (SD)	2.41 (1.38)		
Diagnosis, N (%)			
Transposition of the Great Vessels	9 (20.0)		
+Ventricular Septal Defect	3 (6.7)		
+Pulmonary Stenosis	1 (2.2)		
Atrioventricular Septal Defect	5 (11.1)		
Tricuspid Atresia	5 (11.1)		
Hypoplastic Left Heart Syndrome	4 (8.9)		
Tetralogy of Fallot	4 (8.9)		
Other <sup>a</sup>	14 (31.1)		
Gaming, hours, Mean (SD)	1.40 (.97)	1.74 (1.71)	.92
User experience <sup>b</sup>			
Engagement (0 – 15), Mean (SD)	11.43 (3.48)	13.96 (1.43)	
Adjusted Mean (SE)	11.32 (.48)	14.12 (.60)	<.001**
Transportation (0 – 15), Mean (SD)	10.46 (3.61)	12.35 (3.27)	
Adjusted Mean (SE)	10.56 (.59)	12.20 (.73)	.090
Flow (0 – 15), Mean (SD)	9.91 (3.19)	12.04 (1.82)	
Adjusted Mean (SE)	9.99 (.47)	11.93 (.58)	.012*
Presence (0 – 15), Mean (SD)	5.91 (2.73)	7.39 (2.29)	
Adjusted Mean (SE)	5.93 (.44)	7.37 (.55)	.047*
Side effects (0 – 15), Mean (SD)	4.34 (3.70)	6.43 (4.35)	
Adjusted Mean (SE)	4.42 (.68)	6.31 (.84)	.090

<sup>a</sup> Anomalous Left Coronary Artery from the Pulmonary Artery, Aortic Coarctation, Congenital Aortic Stenosis, Congenital complete A-V block, Double Outlet Right Ventricle, Pulmonary Atresia with Ventricular Septal Defect, Single Ventricle Heart, Total Anomalous Pulmonary Venous Return, Ventricular Septal Defect

<sup>b</sup> 58 children completed user experience questionnaire (35 children with CHD; 23 TD children)

\*  $p < .05$ , \*\*  $p < .01$

**Demographic and clinical characteristics**

Table 1 presents an overview of the participant characteristics of both groups (i.e., children with CHD and TD children).

A higher number of males were observed in the CHD group compared to the TD group (CHD:  $n=33$ ; TD:  $n=18$ ;  $\chi^2=6.834$ ,  $p=0.009$ ). Hypoplastic Left Heart Syndrome, Transposition of the Great Arteries, and Tetralogy of Fallot appear more prevalent among males [27, 28]. The children in both groups did not significantly differ in age (CHD:  $M=11.20$  [SD=1.28]; TD:  $M=11.00$  [SD=0.72];  $U=1069.500$ ,  $z=-0.958$ ,  $p=0.338$ ).

Children in both groups spent a similar amount of time playing video games for leisure (CHD: 1.40 h; TD: 1.74 h;  $U=254.000$ ,  $z=-0.107$ ,  $p=0.915$ ; see Table 1).

**Feasibility**

Out of the 98 children, 88% completed the innovative VR assessment without encountering any issues. CHD and TD children exhibited comparable completion rates for the innovative VR assessment (boys:  $n=47$  [55%]), suggesting that the VR Serious Game was feasible for both CHD and TD children. Group (i.e., children with CHD or TD children;  $p=0.389$ ) and sex (i.e., boy or girl;  $p=0.127$ ) had no significant effect on the likelihood of participants completing the innovative VR assessment (see Table 2).

Among the children with CHD, five children (42%) experienced software malfunctions during the innovative VR assessment. In contrast, four TD children (33%) encountered similar software malfunctions. In addition, two children with CHD (17%) displayed signs of disinterest or a short attention span, resulting in incomplete gameplay. In the TD group, one child (8%) reported discomfort due to the size of the VR glasses used.

**Table 2** Logistic regression predicting likelihood of innovative VR assessment completion based on group (i.e., CHD Children or TD Children) and sex (i.e., Boy or Girl)

	B	SE	Wald	df	p	Odds Ratio	95% CI for Odds Ratio	
							Lower	Upper
Group	.565	.656	.743	1	.389	1.760	.487	6.360
Sex	-1.033	.677	2.333	1	.127	.356	.094	1.340
Constant	2.294	.547	17.558	1	<.001	9.913		

Sex is for girls compared to boys; Group is for CHD compared to TD



### User experience

A total of 58 children completed the user experience questionnaire (see Table 1). This comprised 35 children with CHD and 23 TD children. Children with CHD reported significantly lower scores than TD children for *Engagement* (CHD:  $M=11.32$  [SE=0.48], TD:  $M=14.12$  [SE=0.60];  $F[1, 55]=103.582$ ,  $p<0.001$ , partial  $\eta^2=0.192$ ), *Flow* (CHD:  $M=9.99$  [SE=0.47], TD:  $M=11.93$  [SE=0.58];  $F[1, 55]=49.762$ ,  $p=0.012$ , partial  $\eta^2=0.108$ ), and *Presence* (CHD:  $M=5.93$  [SE=0.44], TD:  $M=7.37$  [SE=0.55];  $F[1, 55]=27.587$ ,  $p=0.047$ , partial  $\eta^2=0.070$ ). In addition, children with CHD reported lower scores for *Transportation* (CHD:  $M=10.56$  [SE=0.59], TD:  $M=12.20$  [SE=0.73];  $F[1, 55]=35.721$ ,  $p=0.090$ , partial  $\eta^2=0.051$ ), however this was not significant at the 0.05 level. On the other hand, TD children reported higher scores for *Side effects* (CHD:  $M=4.42$  [SE=0.68], TD:  $M=6.31$  [SE=0.84];  $F[1, 55]=47.236$ ,  $p=0.090$ , partial  $\eta^2=0.051$ ), although this was not significant at the 0.05 level. Side effects encompassed nausea (i.e., "I felt nauseous"; CHD:  $n=3$ ; TD:  $n=2$ ), warmth (i.e., "I felt warm"; CHD:  $n=14$ ; TD:  $n=9$ ), and headaches (i.e., "I had a headache"; CHD:  $n=6$ ; TD:  $n=11$ ). A higher score indicated more side effects, thereby resulting in a more negative experience.

### Discussion

In order to facilitate the development and implementation of innovative technology (i.e., a new approach that has emerged in the published literature over the last two decades) in clinical practice, it is necessary to understand the user experience of end-users (i.e., children with CHD). The overarching aim of the present study was to evaluate the feasibility of a VR Serious Game for cognitive assessment in school-aged children with CHD. The sub-aims were two-fold: (1) to objectively evaluate the feasibility of the VR Serious Game in children with CHD compared to TD children, and (2) to explore the user experience of both groups after their interaction with the VR Serious Game.

The findings indicated that the VR Serious Game was feasible for both children with CHD and TD children, with 88% children completing the innovative VR assessment without encountering any issues. Notably, there were no discernible differences in completion rates between the CHD and TD groups or between male and female participants. This finding is of clinical relevance as it suggests that the VR Serious Game may also be feasible for cognitive training. The primary factor contributing to unsuccessful completion was software malfunctions (75%). It is important to note that children with CHD also performed the NPA tests in between blocks of VR

assessment. This indicates that it may be applicable as part of a larger clinical (cognitive) assessment. Therefore, the application of innovative VR appears to hold promise for various paediatric clinical populations [19–21].

Furthermore, children with CHD reported significantly lower scores for *Engagement*, *Flow* and *Presence* compared to TD children. However, the scores for *Engagement* and *Flow* were still considered "moderate to good" (greater than eight; [21]). It is important to note that the evaluation of *Presence* was conducted using only two items, namely "I was part of the environment" and "I felt integrated in the environment" instead of three items, which may have contributed to the lower reported scores. Presence is an important factor for enhancing the VR user experience, particularly when utilising VR Serious Games in clinical practice to better predict children's capabilities in daily life [21]. Therefore, to explore this further, the use of a user experience questionnaire with greater emphasis on presence should be considered.

Enhanced levels of flow and immersion (i.e., presence) have been associated with increased enjoyment and motivation during play [22]. Therefore, innovative VR assessment has the potential to enhance patient compliance and more effective cognitive monitoring [29].

Moreover, children with CHD and TD children reported minimal adverse physiological reactions. TD children reported a higher frequency of headaches compared to children with CHD. TD children were assessed in groups of two or four simultaneously, therefore, it is possible that they may have experienced increased external distractions, potentially contributing to the occurrence of headaches.

### Strengths and limitations

The present study's strength lies in its extensive inclusion of children (CHD:  $n=54$ ; TD:  $n=47$ ). The CHD cohort represents a diverse range of CHD diagnoses, encompassing the most commonly reported CHD diagnoses in the Netherlands [30]. In addition, children with CHD were recruited from two Centres of Expertise for CHD, enhancing the sample's representativeness and generalisability to the larger, target population. It is important to highlight that none of the children with CHD attended specialist schools; they were all integrated into mainstream education and did not encounter participation restrictions (e.g., at school and in the community). The majority of children with CHD seen in clinical practice attend mainstream education. It is plausible that children with severe CHD, demanding surgical intervention within their first year of life [23], might encounter fewer enduring negative outcomes compared to children with milder CHD who undergo surgery at a later stage. Consequently, the feasibility and user experience

of children with milder CHD operated at a later stage remain uncertain.

### Future research

To further investigate the added value of the VR Serious Game, future research should explore the associations between performance in the VR Serious Game and test results from a conventional or digital NPA. Although a digital NPA offers similar advantages to a VR Serious Game, such as precise and standardised stimulus presentation and automated scoring [31], the VR Serious Game has unique advantages. It allows the development of ecologically valid environments (e.g., a dynamic setting with multisensory distractions), which can elicit more natural behaviour [31], therefore, providing valuable insights into a child's cognitive capabilities. In addition, task difficulty can be tailored to each individual child (e.g., by adjusting stimuli complexity and flash duration), resulting in a more accurate cognitive assessment.

Furthermore, researchers should explore differences in cognitive performance between children with CHD and TD children. To better estimate and explain the impact of cognitive impairment on daily functioning, researchers should prioritise the development of novel outcome measures. For example, performance stability measures (i.e., the number of fluctuations in test performance, [32]). To enhance the interpretability of test results, it is important to understand how a task was executed. It would be informative to not only identify the number of errors a child makes, but also to pinpoint precisely when these errors occurred. This can provide valuable insights into whether the errors are due to comprehension problems, fatigue, or fluctuations in attention.

### Conclusions

The VR Serious Game appeared feasible for children with CHD and TD children (88% completion rate). Children with CHD reported significantly lower scores than TD children on three user experience scales: *Engagement*, *Flow* and *Presence*. Nonetheless, the scores for *Engagement* and *Flow* were still positive. *Presence* is an essential factor for enhancing the VR user experience, therefore, demands further investigation. In addition, future research should investigate the effectiveness of the VR Serious Game compared with a conventional or digital NPA, prioritising the development of novel outcome measures that can better estimate and explain the impact of cognitive impairment on daily functioning.

### Acknowledgements

We would like to extend our gratitude to all individuals who played an essential role in this research project. We are grateful to the dedicated clinicians, M.C.A. Bolland-Sprong and M. Uniken Venema-van Uden, for their contributions. Our thanks also go to the committed students, E. Hoogers, N. Pletting, S. Nieuwenhuizen, and Y. van der Hoorn, for their significant involvement.

Furthermore, we express our profound thanks to NeuroReality for their essential support, both in terms of software (VR Serious Game Koji's Quest) and hardware (Oculus Meta Quest 2 head-mounted display VR glasses). Moreover, we would like to acknowledge and thank the children who participated in this study and their guardians for their invaluable contributions.

### Authors' contributions

T.C.W.N served as the principal investigator. Conventional NPA selection was carried out by C.V.W, E.B, M.M.J.S, and T.C.W.N. Eligible children from participating hospitals were identified by H.P.N and J.M.P.B. Data analysis was supported by H.H and T.C.W.N in collaboration with C.L.W. C.L.W wrote the main manuscript text. T.C.W.N wrote the main revision. All authors contributed by editing and reviewing the manuscript.

### Funding

The research article was financially supported by the Hartekind Foundation (*Stichting Hartekind*; CHD and TD children), Kinderrevalidatiefonds Adriaanstichting (TD children) K.F. Hein Fonds (hardware), and Utrechts Universiteitsfonds.

### Availability of data and materials

No datasets were generated or analysed during the current study.

### Declarations

#### Ethics approval and consent to participate

The present study received ethical approval from the Faculty Ethics Review Board of the Faculty of Social and Behavioural Sciences at Utrecht University (UU; protocol numbers 20–648, 21–0136, 21–0369, and 22–0058) and the Medical Research Ethics Committee (protocol number 20–693/C). The study protocol adhered to the Declaration of Helsinki, as well as the requirements of UU and the Faculty Ethics Review Board. Informed consent was provided by guardians, as well as by children aged 12 years and older.

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare no competing interests.

#### Author details

<sup>1</sup>Experimental Psychology, Helmholtz Institute, Utrecht University, Utrecht, the Netherlands. <sup>2</sup>Department of Pediatric Psychology, Wilhelmina's Children Hospital, University Medical Center Utrecht, Utrecht, the Netherlands. <sup>3</sup>Department of Pediatric Cardiology, Wilhelmina Children's Hospital, Utrecht, the Netherlands. <sup>4</sup>Department of Pediatric Cardiology, University of Groningen, Beatrix Children Hospital, University Medical Center Groningen, Groningen, the Netherlands. <sup>5</sup>Brain and Cognition, Leuven Brain Institute, KU Leuven, Leuven, Belgium.

Received: 24 November 2023 Accepted: 21 February 2024

Published online: 18 April 2024

### References

- Bellinger DC, Newburger JW. Neuropsychological, psychosocial, and quality-of-life outcomes in children and adolescents with congenital heart disease. *Prog Pediatr Cardiol*. 2010;29(2):87–92. <https://doi.org/10.1016/j.ppedcard.2010.06.007>.
- Bellinger DC, Wypij D, Rivkin MJ, DeMaso DR, Robertson RL Jr, Dunbar-Masterson C, et al. Adolescents with d-transposition of the great arteries corrected with the arterial switch procedure: neuropsychological assessment and structural brain imaging. *Circulation*. 2011;124(12):1361–9. <https://doi.org/10.1161/CIRCULATIONAHA.111.026963>.
- Calderon J, Angeard N, Pinabiaux C, Bonnet D, Jambaqué I. Facial expression recognition and emotion understanding in children after neonatal open-heart surgery for transposition of the great arteries. *Dev Med Child Neurol*. 2014;56(6):564–71. <https://doi.org/10.1111/dmnc.12381>.

4. Cassidy AR, White MT, DeMaso DR, Newburger JW, Bellinger DC. Executive function in children and adolescents with critical cyanotic congenital heart disease. *J Int Neuropsychol Soc*. 2015;21(1):34–49. <https://doi.org/10.1017/S1355617714001027>.
5. Feldmann M, Bataillard C, Ehrler M, Ullrich C, Knirsch W, Gosteli-Peter MA, et al. Cognitive and executive function in congenital heart disease: a meta-analysis. *Pediatrics*. 2021;148(4). <https://doi.org/10.1542/peds.2021-050875>.
6. Jackson WM, Davis N, Calderon J, Lee JJ, Feirsen N, Bellinger DC, et al. Executive functions in children with heart disease: a systematic review and meta-analysis. *Cardiol Young*. 2021;31(12):1914–22. <https://doi.org/10.1017/S1047951121001074>.
7. Kasmi L, Calderon J, Montreuil M, Geronikola N, Lambert V, Belli E, et al. Neurocognitive and psychological outcomes in adults with dextro-transposition of the great arteries corrected by the arterial switch operation. *Ann Thorac Surg*. 2018;105(3):830–6. <https://doi.org/10.1016/j.athoracsur.2017.06.055>.
8. van de Wouw CL, Visser M, Gorter JW, Huygelier H, Nijboer TCW. Systematic review of the effectiveness of innovative, gamified interventions for cognitive training in pediatric acquired brain injury. *Neuropsychol Rehabil*. 2023;1–32. <https://doi.org/10.1080/09602011.2023.2174561>
9. Lambregts SA, Smetsers JE, Verhoeven IM, de Kloet AJ, van de Port IG, Ribbers GM, et al. Cognitive function and participation in children and youth with mild traumatic brain injury two years after injury. *Brain Inj*. 2018;32(2):230–41. <https://doi.org/10.1080/02699052.2017.1406990>.
10. Chevignard M, Pillon B, Pradat-Diehl P, Taillefer C, Rousseau S, Le Bras C, et al. An ecological approach to planning dysfunction: script execution. *Cortex*. 2000;36(5):649–69. [https://doi.org/10.1016/S0010-9452\(08\)70543-4](https://doi.org/10.1016/S0010-9452(08)70543-4).
11. Donovan NJ, Kendall DL, Heaton SC, Kwon S, Velozo CA, Duncan PW. Conceptualizing functional cognition in stroke. *Neurorehabil Neural Repair*. 2008;22(2):122–35. <https://doi.org/10.1177/1545968307306239>.
12. Spreij LA, Visser-Meily JM, Sibbel J, Gosselt IK, Nijboer TC. Feasibility and user-experience of virtual reality in neuropsychological assessment following stroke. *Neuropsychol Rehabil*. 2022;32(4):499–519. <https://doi.org/10.1080/09602011.2020.1831935>.
13. Neguț A, Jurma AM, David D. Virtual-reality-based attention assessment of ADHD: ClinicaVR: Classroom-CPT versus a traditional continuous performance test. *Child Neuropsychol*. 2017;23(6):692–712. <https://doi.org/10.1080/09297049.2016.1186617>.
14. Parsey CM, Schmitter-Edgecombe M. Applications of technology in neuropsychological assessment. *Clin Neuropsychol*. 2013;27(8):1328–61. <https://doi.org/10.1080/13854046.2013.834971>.
15. Huygelier H, Mattheus E, Abeele VV, van Ee R, Gillebert CR. The use of the term virtual reality in post-stroke rehabilitation: a scoping review and commentary. *Psychologica Belgica*. 2021;61(1):145. <https://doi.org/10.5334/pb.1033>.
16. Vermeir JF, White MJ, Johnson D, Crombez G, Van Ryckeghem DM. The effects of gamification on computerized cognitive training: systematic review and meta-analysis. *JMIR Serious Games*. 2020;8(3):e18644. <https://doi.org/10.2196/18644>.
17. Parsons TD. Ecological validity in virtual reality-based neuropsychological assessment. In: *Encyclopedia of Information Science and Technology*, Third Edition. IGI Global; 2015. p. 1006–1015. <https://doi.org/10.4018/978-1-4666-5888-2.ch095>
18. World Health Organization. International classification of functioning, disability, and health: Children & youth version: ICF-CY. 2007.
19. Erez N, Weiss PL, Kizony R, Rand D. Comparing performance within a virtual supermarket of children with traumatic brain injury to typically developing children: a pilot study. *OTJR*. 2013;33(4):218–27. <https://doi.org/10.3928/15394492-20130912-04>.
20. Parsons TD, 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*. 2019;29:338–56. <https://doi.org/10.1007/s11065-019-09407-6>.
21. Verheul F, Gosselt I, Spreij L, Visser-Meily A, Te Winkel S, Rentinck I, et al. Can serious play and clinical cognitive assessment go together? On the feasibility and user-experience of virtual reality simulations in pediatric neurorehabilitation. *J Pediatr Rehabil Med*. 2022;15(2):265–74. <https://doi.org/10.3233/prm-200801>.
22. Moizer J, Lean J, Dell'Aquila E, Walsh P, Keary A, O'Byrne D, Sica LS. An approach to evaluating the user experience of serious games. *Comput Educ*. 2019;136:141–51. <https://doi.org/10.1016/j.compedu.2019.04.006>.
23. Qu Y, Lin S, Zhuang J, Bloom MS, Smith M, Nie Z, Liu X. First-Trimester Maternal Folic Acid Supplementation Reduced Risks of Severe and Most Congenital Heart Diseases in Offspring: A Large Case-Control Study. *J Am Heart Assoc*. 2020;9(13):e015652. <https://doi.org/10.1161/jaha.119.015652>.
24. Wolf P, Benbadis S, Dimova PS, Vinayan KP, Michaelis R, Reuber M, Yacubian EM. The importance of semiological information based on epileptic seizure history. *Epileptic Disord*. 2020;22(1):15–31. <https://doi.org/10.1684/epd.2020.1137>.
25. NeuroReality. Koji's Quest. <https://neuro-reality.com/products>.
26. Hsu P, Ma A, Wilson M, Williams G, Curotta J, Munns CF, Mehr S. CHARGE syndrome: a review. *J Paediatr Child Health*. 2014;50(7):504–11. <https://doi.org/10.1111/jpc.12497>.
27. Azakie A, Russell IA. Gender differences in pediatric cardiac surgery: the surgeon's perspective. *J Thorac Cardiovasc Surg*. 2004;128(1):4–6. <https://doi.org/10.1016/j.jtcvs.2004.04.002>.
28. Miller-Hance WC, Tacy TA. Gender differences in pediatric cardiac surgery: the cardiologist's perspective. *J Thorac Cardiovasc Surg*. 2004;128(1):7–10. <https://doi.org/10.1016/j.jtcvs.2004.04.008>.
29. Zygouris S, Tsolaki M. Computerized cognitive testing for older adults: a review. *Am J Alzheimers Dis Other Dement*. 2015;30(1):13–28. <https://doi.org/10.1177/1533317514522852>.
30. VZinfo. Aangeboren afwijkingen van het hartvaatstelsel: leeftijd en geslacht. Ministerie van Volksgezondheid, Welzijn en Sport, Rijksinstituut voor Volksgezondheid en Milieu. <https://www.vzinfo.nl/aangeboren-afwijkingen-van-hartvaatstelsel/leeftijd-en-geslacht>.
31. Spreij LA, Gosselt IK, Visser-Meily JM, Nijboer TC. Digital neuropsychological assessment: feasibility and applicability in patients with acquired brain injury. *J Clin Exp Neuropsychol*. 2020;42(8):781–93. <https://doi.org/10.1080/13803395.2020.1808595>.
32. Spreij LA, Gosselt IK, Visser-Meily JM, Hoogerbrugge AJ, Kootstra TM, Nijboer TC. The journey is just as important as the destination—Digital neuropsychological assessment provides performance stability measures in patients with acquired brain injury. *PLoS ONE*. 2021;16(7):e0249886. <https://doi.org/10.1371/journal.pone.0249886>.

## Publisher's Note

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