Supplementary Material

**Supplementary Methods**

**Participants**

The participants were recruited from Kirkkonummi and Espoo, Finland, (a) by sending email invitations to the parents of typically developing children who had already performed the HMD part (see 2.3 Procedure) when participating in an earlier study (Seesjärvi et al., 2022b), and (b) by sending advertisement letters to parents via schools’ electronic message boards and giving brief educational lectures at schools and online, during which the opportunity to take part in the study was mentioned.

**The differences between HMD- and FSD-EPELI**

In the HMD version, the participants used Oculus Go goggles (2560 x 1440 resolution, 60/72 Hz refresh rate, 16:9 aspect ratio, and 101-degree horizontal FOV) and the related hand controller. To ensure participant safety, the children performed the task while sitting in a revolving chair (see Figure 1A). To look around, they could rotate their head and spin with the chair. The participant interacted with the objects by pointing at them with a ray coming from the virtual hand controller object and pushing a button with a thumb. Movable objects could be picked up and released by pointing at a desired location and pushing the button again. The participant could check the time by raising the hand controller and looking at it, like checking the time from a wristwatch. The drums in the virtual environment were played by swinging the hand controller at them. This was required in only one scenario but could be done spontaneously in others. Navigating was performed by teleporting between waypoint circles on the floor by pointing at them and pressing the button. Motion tracking sensors in the HMD and hand controller recorded movements. Only the hand controller movement data was used in this study, as in the earlier study the head and hand movement measures were strongly correlated (*r* = .84; see Seesjärvi et al. 2022b supplementary material). For audio, the integrated speakers of the Oculus Go headset, located near (~ 3 cm) each ear, were used. The loudness was kept fixed at a predetermined level that allowed the children to hear the instructions clearly. The examiner was instructed to adjust the volume if the participant could not hear the instructions easily during the training phase of the task, but this was not needed for any of the participants.

In the FSD version, the participants used typical laptop/desktop computers and a web browser. The participants who performed the FSD version supervised by an experimenter used a Dell Latitude 7400 laptop computer (Intel i5-8365 @ 1,6 MHz CPU, 16 GB of RAM, Intel UHD Graphics 620 GPU, Windows 10 OS, 14-inch screen, 1920 x 1080 resolution, 60 Hz refresh rate, 16:9 aspect ratio) and a standard mouse (Logitech M100). While the viewing distance was not strictly controlled, it was approximated to be 60 cm on average, which results in a 29-degree horizontal FOV. The participants who performed the FSD version at home used various typical laptop/desktop configurations. Of these participants, 25 used a mouse, four a trackpad, while six did not report this. The screen size range was 12–32 inches, the most popular options being 14 inches (four participants) and 24 inches (five participants). As the viewing distance was not strictly controlled, the exact FOV is not known. However, it can be approximated that the horizontal FOV varied between 25–60 degrees if a typical viewing distance of 60 cm was used. Thus, the FOV was on average markedly narrower in the FSD version compared to the HMD version. In the control system, there were several differences compared to the HMD version. The participants looked around with mouse/trackpad as opposed to natural head movement. There was a crosshairs in the middle of the screen, and interaction with the objects was done by aiming at them by rotating the view until the target object was in the crosshairs, and then pressing the left mouse/trackpad button. The time could be checked by pressing the right mouse/trackpad button, which resulted a white circle in the lower right corner being replaced with a clock. The virtual drums were used in the same way as other objects (aiming with the crosshairs and pressing the left button). Navigation was performed by teleporting between waypoint circles like in the HMD version. For audio, the use of headphones was advised but not required. Forty-four participants used headphones, 11 reported using conventional speakers, and six failed to report this. The correct loudness level was determined individually for each participant by first finding the threshold level of hearing and then setting the loudness to a comfortable level above it. After this, the children could manually adjust the level before the task.

**Procedure**

The HMD-EPELI session started with the Matrix reasoning subtest of the WISC-IV test battery (Wechsler, 2003) and continued with EPELI. After EPELI, the children orally answered several related questionnaires, which were read aloud and filled out by the experimenter. These included a translated version of the Simulator Sickness Questionnaire (Kennedy et al., 1993; see also Seesjärvi et al., 2022a), a gaming experience questionnaire (Seesjärvi et al., 2022a), a shortened version of the Presence Questionnaire 3.0 (Witmer et al., 2005; see also Seesjärvi et al., 2022a), and an object familiarity questionnaire (Seesjärvi et al., 2022a). Also, to probe their familiarity with the task contents, the children were asked “From a scale of 1 (not at all) to 7 (very much), how much have you performed similar tasks in real life?”. After a break, the children performed the Similarities subtest of WISC-IV and other tasks, which were not analyzed in this study. The parents filled out several questionnaires, of which Behavior Rating Inventory for Executive Functions (BRIEF; Gioia et al., 2000) was analyzed in this study (see Seesjärvi et al. 2022b for the list of other questionnaires). From BRIEF, the raw score of Global Executive Composite (GEC) was used.

The FSD-EPELI session began with EPELI. After EPELI, the children orally answered the Simulator Sickness Questionnaire, the shortened version of the Presence Questionnaire 3.0, and a hardware questionnaire (see Supplementary Table 2), as read aloud and filled out by the experimenter (lab group) or the parent (home group). As with the HMD-EPELI session, the parents filled out several questionnaires, and the BRIEF measure of GEC was included in the analyses.

**References**

Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2000). *Behavior Rating Inventory of Executive Function: BRIEF*. Odessa, FL: Psychological Assessment Resources

Kennedy, R. S., Lane, N. E., Berbaum, K. S., & Lilienthal, M. G. (1993). Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness. *The International Journal of Aviation Psychology*, *3*(3), 203–220. <https://doi.org/10.1207/s15327108ijap0303_3>

Seesjärvi, E., Puhakka, J., Aronen, E. T., Hering, A., Zuber, S., Merzon, L., Kliegel, M., Laine, M., & Salmi, J. (2022a). EPELI: A Novel Virtual Reality Task for the Assessment of Goal-directed Behavior in Real-life Contexts. *Psychological Research*. <https://doi.org/10.1007/s00426-022-01770-z>

Seesjärvi, E., Puhakka, J., Aronen, E. T., Lipsanen, J., Mannerkoski, M., Hering, A., Zuber, S., Kliegel, M., Laine, M., & Salmi, J. (2022b). Quantifying ADHD Symptoms in Open-Ended Everyday Life Contexts With a New Virtual Reality Task. *Journal of Attention Disorders*, *26*(11), 1394-1411. <https://doi.org/10.1177/10870547211044214>

Wechsler, D. (2003). *Wechsler Intelligence Scale for Children-Fourth Edition (WISC-IV) administration and scoring manual*. The Psychological Corporation.

Witmer, B. G., Jerome, C. J., & Singer, M. J. (2005). The Factor Structure of the Presence Questionnaire. *Presence: Teleoperators and Virtual Environments*, *14*(3), 298–312. <https://doi.org/10.1162/105474605323384654>

**Supplementary Table 1. Descriptive statistics.**

|  |  |  |
| --- | --- | --- |
|  | Mean (SD) | Home vs. laboratory comparison |
| Variable | All participants (*n*=72) | Laboratory group (*n*=37) | Home group (*n*=35) | Test statistics | *p* |
| Age | 11 yrs. 0 mos. (1 yr. 0 mos.) | 10 yrs. 11 mos.(0 yrs. 11 mos.) | 11 yrs. 1 mo. (1 yr. 0 mos.) | *t* = 0.668 | .507 |
| Handedness (left / right) | 5 / 67 | 3 / 34 | 2 / 33 | *Fisher’s exact test* | > .999 |
| Gender (boy / girl) | 43 / 29 | 22 / 15 | 21 / 14 | *Fisher’s exact test* | > .999 |
| Parental income\* | 4.45 (0.79) | 4.44 (0.91) | 4.46 (0.66) | *t* = 0.068 | .946 |
| Parental education\*\* | 2.89 (0.30) | 2.83 (0.36) | 2.94 (0.20) | *t* = 1.592 | .117 |
| *Note*. \* Per adult before tax; 1 = less than 1500 €/m, 2 = 1500–2200 €/m, 3 = 2200–3000 €/m, 4 = 3000–4000 €/m, 5 = over 4000 €/m. \*\* 1 = Comprehensive school, 2 = High school / Vocational school, 3 = University degree, higher vocational school degree, or equivalent. \*\*\* Standard score. |

**Supplementary Table 2. The hardware questionnaire.**

We ask you to report certain specifications about your hardware. If you do not know the answer, please use “not known”.

1. Computer brand
2. Computer model
3. Processor
4. Amount of RAM
5. Display size (inches)
6. Display driver
7. OS version
8. Web browser version
9. Did you use mouse or trackpad?
10. Did you use headphones or not?
11. Free comments

**Supplementary Table 3. The general linear models with each FSD-EPELI measure as the dependent variable and the place of the assessment (home/laboratory) and time (1st/2nd session) as independent variables.**

|  |  |  |
| --- | --- | --- |
|  | Descriptive Statistics | Linear model test statistics |
|  |
| Dependent variable | Place, mean (SD) | Session, mean (SD) | Home vs laboratory | 1st vs 2nd session |  |
| Home | Laboratory | 1st | 2nd | estimate (SD) | t | *p* | estimate (SD) | t | *p* |  |
| Total score | 56.35(8.12) | 57.86(5.82) | 54.25(7.45) | 60.18(5.05) | 2.026(1.527) | 1.327 | .189 | **6.099(1.527)** | **3.994** | **<.001** |  |
| Task efficacy | 0.44(0.10) | 0.49(0.10) | 0.47(0.10) | 0.46(0.11) | 0.046(0.024) | 1.886 | .064 | -0.009(0.024) | -0.366 | .716 |  |
| Navigation efficacy | 0.09(0.02) | 0.09(0.02) | 0.09(0.02) | 0.09(0.02) | 0.004(0.004) | 0.983 | .329 | 0.005(0.004) | 1.116 | .268 |  |
| Controller motion | 41313.42(8660.43) | 43254.4(7550.03) | 42540.42(8869.9) | 42069.4(7341.39) | 1914.78(1960.098) | 0.977 | .332 | -308.323(1960.098) | -0.157 | .875 |  |
| Total actions | 442.85 (129.28) | 410.28(86.06) | 403.78(117.99) | 449.74(96.15) | -28.879(25.861) | -1.117 | .268 | 43.504(25.861) | 1.682 | .097 |  |
| TBPM | 7.32(3.01) | 8.00(2.33) | 6.83(2.83) | 8.56(2.23) | 0.829(0.61) | 1.358 | .179 | **1.796(0.61)** | **2.942** | **.004** |  |
| Clock checks | 68.68(37.56) | 61.89(30.06) | 65.11(36.12) | 65.26(31.78) | -6.824(8.199) | -0.832 | .408 | -0.426(8.199) | -0.052 | .959 |  |
| Clock check durations | 82.74(61.57) | 71.53(42.07) | 75.04(51.53) | 78.91(53.68) | -10.994(12.725) | -0.864 | .391 | 3.096(12.725) | 0.243 | .809 |  |
| EBPM | 4.15(0.74) | 4.42(0.65) | 4.08(0.73) | 4.5(0.62) | 0.307(0.160) | 1.925 | .058 | **0.443(0.16)** | **2.775** | **.007** |  |
| *Note. N* = 70. The effects that are significant at the level of *p* < .05 are written in bold. |  |

**Supplementary Table 4. The general linear models with perceived presence questions as the dependent variables and the place of the assessment (home/laboratory) and time (1st/2nd session) as independent variables.**

|  |  |  |
| --- | --- | --- |
|  | Descriptive statistics | Linear model test statistics |
|  | Place, mean (SD) | Session, mean (SD) | Home vs laboratory | 1st and 2nd session |
| Question | Home | Laboratory | 1st | 2nd | estimate (SD) | t | *p* | estimate (SD) | t | *p* |
| 1.      How natural did your interactions with the environment seem? | 4.45(1.68) | 4.38(1.67) | 4.66(1.60) | 4.18(1.71) | -0.133 (0.418) | -0.318 | .752 | -0.496 (0.415) | -1.197 | .236 |
| 2.      How much did the environment involve you? | 4.72(1.75) | 4.78(1.73) | 5.06(1.70) | 4.47(1.73) | -0.016 (0.432) | -0.036 | .971 | -0.594 (0.429) | -1.383 | .171 |
| 3.      How natural was the mechanism which controlled movement through the environment? | 3.34(1.61) | 4.05(1.82) | 4.00(1.74) | 3.50(1.76) | 0.656 (0.434) | 1.513 | .135 | -0.418 (0.431) | -0.97 | .336 |
| 4.      How much did your experiences in the virtual environment seem consistent with your real-world experiences? | 4.52(1.55) | 4.30(1.78) | 4.50(1.44) | 4.29(1.88) | -0.250 (0.423) | -0.592 | .556 | -0.237 (0.420) | -0.565 | .574 |
| 5.      How much did the visual display quality interfere or distract you from performing assigned tasks or required activities? | 1.59 (1.05) | 1.78(0.98) | 1.59(0.87) | 1.79(1.12) | 0.227 (0.253) | 0.896 | .374 | 0.229 (0.251) | 0.911 | .366 |
| 6.      How much did the control devices interfere with the performance of assigned tasks or with other activities? | 1.76(1.50) | 1.38(0.83) | 1.66(1.33) | 1.44(1.02) | -0.414 (0.294) | -1.411 | .163 | -0.267 (0.292) | -0.915 | .364 |
| 7.      How well could you concentrate on the assigned tasks or required activities? | 5.38(1.32) | 5.24(1.52) | 5.25(1.44) | 5.35(1.43) | -0.125 (0.361) | -0.346 | .730 | 0.087 (0.359) | 0.243 | .808 |
| 8.      How well could you hear sounds? | 6.76(0.79) | 6.76(0.55) | 6.66(0.83) | 6.85(0.44) | 0.023 (0.165) | 0.142 | .888 | 0.200 (0.164) | 1.218 | .228 |
| 9.      Were there moments during the virtual environment experience when you felt completely focused on the task or environment? | 3.93(2.20) | 4.70(1.93) | 4.25(2.06) | 4.47(2.11) | 0.812 (0.516) | 1.576 | .12 | 0.322 (0.512) | 0.629 | .531 |
| Three additional questions that were not in the original Presence Questionnaire 3.0: |
| 10.    How enthusiastic did you feel about the tasks? | 5.14(1.62) | 5.30(1.54) | 5.56(1.39) | 4.91(1.68) | 0.078 (0.389) | 0.201 | .841 | -0.641 (0.386) | -1.659 | .102 |
| 11.    How interesting did the tasks seem to you? | 4.38(1.86) | 4.70(1.65) | 4.97(1.71) | 4.18(1.70) | 0.227 (0.429) | 0.529 | .599 | -0.764 (0.426) | -1.795 | .077 |
| 12.    How much effort did you put into your performance? | 6.34(0.90) | 6.00(0.94) | 6.19(0.93) | 6.12(0.95) | -0.359 (0.232) | -1.548 | .127 | -0.115 (0.231) | -0.498 | .620 |
| *Note. N* = 66. All questions were answered with a Likert scale with a range of 1–7. |