# © 2018, American Psychological Association. This paper is not the copy of record and may not exactly replicate the final, authoritative version of the article. Please do not copy or cite without authors' permission. The final article will be available, upon publication, via its DOI: 10.1037/neu0000517

# Neuropsychological assessment of visual selective attention and processing capacity with

# head-mounted displays

Rebecca M. Foerster<sup>1,3</sup>, Christian H. Poth<sup>1,3</sup>, Christian Behler<sup>2,3</sup>, Mario Botsch<sup>2,3</sup>, & Werner X.

Schneider<sup>1,3</sup>

<sup>1</sup>Neuro-cognitive Psychology, <sup>2</sup>Computer Graphics and Geometry Processing,

<sup>3</sup>Cluster of Excellence Cognitive Interaction Technology,

Bielefeld University, P. O. Box 100131, D-33501 Bielefeld, Germany

Author note:

Rebecca M. Foerster (corresponding author), Neuro-cognitive Psychology and Cluster of Excellence Cognitive Interaction Technology 'CITEC', Bielefeld University. P. O. Box 100131, D-33501 Bielefeld, Germany. Phone: +49 (0)521 106-4503 or +49 (0)521 106-4505. Fax: +49 (0)521 106-156934. E-mail: <u>rebecca.foerster@uni-bielefeld.de</u>. Christian H. Poth, Neurocognitive Psychology and Cluster of Excellence Cognitive Interaction Technology 'CITEC', Bielefeld University. Christian Behler, Graphics & Geometry Group and Cluster of Excellence Cognitive Interaction Technology 'CITEC', Bielefeld University. Mario Botsch, Graphics & Geometry Group and Cluster of Excellence Cognitive Interaction Technology 'CITEC', Bielefeld University. Werner X. Schneider, Neuro-cognitive Psychology and Cluster of Excellence Cognitive Interaction Technology 'CITEC', Bielefeld University. This research was supported by the Cluster of Excellence Cognitive Interaction Technology 'CITEC' (EXC 277) at Bielefeld University, which is funded by the German Research Foundation (DFG).

#### Abstract

**Objective:** Neuropsychological patients often suffer from impairments in visual selective attention and processing capacity components. Their assessment demands a high standardization of testing conditions, which is difficult to achieve across institutions. Head-mounted displays (HMDs) provide a solution. These virtual reality devices cover the entire visual field in a shielded way and thus keep visual stimulation constant. For neuropsychological assessment with HMDs, sufficient reliability is required. We have previously demonstrated that an early developer version of an HMD can be used to reliably measure components of visual processing capacity. However, it is unclear whether this also holds for the assessment of components of visual selective attention. Moreover, it has yet to be established whether now commercially available HMDs are capable of reliable neuropsychological assessment.

**Method:** We assessed the test-retest reliabilities of several components of visual selective attention and processing capacity of healthy subjects with the commercially available HTC Vive. Using an assessment procedure (combiTVA) derived from the theory of visual attention (Bundesen, 1990), we measured attentional selectivity, lateral bias, processing speed, visual working memory capacity, and the threshold of conscious perception. We compared the reliabilities of these components measured with the HTC Vive with those of a cathode ray tube (CRT)-screen, the gold-standard of visual presentation in the laboratory.

**Results:** Both devices provided comparable reliabilities.

**Conclusions:** Thus, HMDs fulfill the requirement to replace standard screens. With their inherent visual standardization and portability, they offer unprecedented opportunities for neuropsychological assessment, such as computerized bedside testing and comparisons of test values across institutions.

**Keywords:** virtual reality, head-mounted display, visual attention, visual capacity, neuropsychological diagnostics, test-retest reliability

# **Public Significance Statement**

- With their inherent visual standardization and portability, head-mounted displays (HMDs) offer unprecedented opportunities for neuropsychological assessment, such as computerized bedside testing and comparisons across institutions.
- We assessed test-retest reliabilities of five components of visual selective attention and processing capacity by the HMD HTC Vive and a CRT, the current gold-standard of visual presentation.
- Results show comparable reliability of both devices demonstrating that HMDs have the capability to advance the neuropsychological assessment to the next level.

#### **1** Introduction

To act efficiently in a rapidly changing and crowded environment, it is essential to apprehend fast and accurately the visual input reaching the eyes in every moment of time. For instance, to drive safely through heavy traffic, nearby vehicles and traffic signs should be processed preferentially, memorized, and used for action control in the right moment of time, while advertisement signs or houses next to the road should be ignored. This ability of visual apprehension is limited at the level of processing capacity in two ways (Habekost, 2015). First, the speed with which visual information can be processed is limited (Bundesen, 1990) – a fact that also influences the test values of classic neuropsychological tests such as the Trail Making Test (Bowie & Harvey, 2006; Foerster & Schneider, 2015; Reitan & Wolfson, 1994). Second, even with enough time, only a limited number of visual objects can be retained in visual working memory for controlling behavior (Cowan, 2000a, 2000b; Luck & Vogel, 1997; Vangkilde, Bundesen, & Coull, 2011). Note that there are alternative views on how working memory might be limited (Buschman, Siegel, Roy, & Miller, 2011; Ma, Husain, & Bays, 2014; Oberauer, 2009).

In order to handle these severe limitations in visual processing capacity, mechanisms of selective visual attention have been postulated (Bundesen, 1990; Desimone & Duncan, 1995; Foerster & Schneider, 2013; Posner & Petersen, 1990). Selective visual attention ensures that relevant information is preferentially processed, so that it has a high likelihood of winning the competition for the limited access to visual working memory.

The two types of processing limitations (processing speed and working memory capacity) vary among individuals and across the life span and can be impaired differently in various neurological and psychiatric disorders (Finke et al., 2005; Finkel, Reynolds, McArdle, & Pedersen, 2007; Habekost, 2015; Habekost & Starrfelt, 2009; Kraft et al., 2013; Unsworth & Engle, 2007). The exact composition of visual processing impairments is therefore helpful to

diagnose and differentiate neuropsychological disorders and to specify the associated functional deficits. For instance, patients with simultanagnosia have problems to perceive multiple objects simultaneously, pointing to a deficit in visual capacity. By assessing working memory capacity and visual processing speed independently, it had been shown that simultanagnosia is associated with a severe reduction in processing speed rather than working memory capacity (Duncan et al., 2003; Neitzel et al., 2016). Additionally ventral, but not dorsal simultanagnosia was associated with a left-sided attention bias. Such differential characterization of the symptoms of a disease demands accurate and reliable assessment of the specific components of visual selective attention and processing capacities.

An elegant way to measure differential components of visual attention and processing capabilities simultaneously for neuropsychological purposes is accomplished in the combiTVA assessment (Duncan et al., 1999; Vangkilde et al., 2011). Grounding on the theory of visual attention (TVA, Bundesen, 1990; Bundesen & Habekost, 2008), combiTVA measures visual selective attention and visual capacities without reaction-time pressure. Specifically, red and blue or only red letters are presented briefly and their presentation is terminated by pattern masks. Presentation durations vary across trials. In the end of each trial, participants have to report only the red letters without any time limit. The advantage of this accuracy-based rather than reaction time-based procedure is that the estimated visual components are not confounded by motor abilities. Therefore, such paradigms are not only applied in basic research (e.g., Poth, Petersen, Bundesen, & Schneider, 2014; Poth & Schneider, 2018; Vangkilde et al., 2011), but also frequently used to assess patients with motor impairments (Finke et al., 2005; Habekost, 2015; Habekost, Petersen, & Vangkilde, 2014).

Based on the data of the combiTVA (Vangkilde et al., 2011) paradigm, five components (parameters) of visual processing can be estimated (Bundesen, 1990). Firstly, top-down

controlled visual selectivity (visual selective attention) can be estimated based on the performance difference between trials with only red target letters and trials with red target letters among blue distractor letters. Selective attention is only necessary in the latter case in order to ensure that red and not blue letters will be encoded for later report. Secondly, the performance difference for target letters at individual positions in the visual field allows to estimate the spatial distribution of visual attention. This parameter can be used to assess the lateral attentional bias – the processing difference across the two hemifields (left vs. right). Thirdly, the minimum presentation duration that has to be exceeded to allow processing for object recognition provides an estimate of the temporal threshold of conscious perception. Fourthly, the rate at which report accuracy improves with increasing presentation duration allows to estimate visual processing speed. Finally, the maximum number of objects that can be retained for later report at long presentation durations delivers an estimate of the capacity of visual working memory.

These five visual processing components of combiTVA are well-grounded in experimental psychology research and cognitive neuroscience (Bundesen & Habekost, 2014). Importantly, they are not primarily performance descriptions but offer reasonable interpretations in terms of psychological processes (Bundesen, 1990; Bundesen, Habekost, & Kyllingsbaek, 2005) and have been found to be affected differentially in a number of neuropsychological disorders (for a review, see, Habekost, 2015). Deficits in top-down controlled selective attention, for instance, have been observed after parietal and frontal lesions and in patients with spina bifida myelomeningocele, mild cognitive impairment, and Alzheimer's disease (Bublak et al., 2005; Caspersen & Habekost, 2013; Peers et al., 2005; Redel et al., 2012).

CombiTVA allows to measure visual selective attention and processing components simultaneously and free from motoric confounds. However, the estimates of visual processing components of combiTVA and computer-based visual tests in general strongly depend on the

exact visual presentation during assessment. As an example, selection between visual objects that differ greatly in visual similarity is easier and faster than selection between objects that differ only slightly (Duncan & Humphreys, 1989; Estes, 1972; Pashler, 1987; Verghese & Nakayama, 1994). Unfortunately, different testing conditions (e.g., different lighting in different testing rooms) influence the exact visual characteristics of the presented stimuli and can thereby change for instance the contrast range of the presented stimuli which in turn affects the selection performance of the observer as well as the absolute values of the other measured visual person parameters. The exact visual characteristics of the presented stimuli depend on the computer monitor used, the lighting, the head-to-screen distance, and other factors of the testing situation. In general, values of visual components are only comparable if they are assessed within the same setup. It is not proper to compare individual values across varying testing situations and institutions, even if the same test procedure is used. This complicates comparisons of study results obtained by different research institutions. This challenge is a severe problem for clinical diagnostics, for monitoring the progress of a disease, and for the evaluation of a treatment in which patients might pass from an acute clinic over a rehabilitation clinic to a medical practice, all applying unique assessment setups. Only strict standardization of visual presentation makes it possible to compare visual components across different laboratories and clinics and to obtain statistical assessment norms. In addition, the common assessment cabins prevent to test special patient groups, i.e., patients who cannot sit upright or be transported.

A promising solution to enhance standardization and applicability is to utilize headmounted displays (HMDs) for assessing visual processing abilities (Foerster, Poth, Behler, Botsch, & Schneider, 2016). The increasingly available HMDs provide an inherently standardized visual environment and cover the entire visual field. Therefore, stimulus appearance and viewing distance are uniquely determined by the assessment software. Furthermore, the

devices are portable and light of weight. In this way, constant visual presentations across studies and institutions can be guaranteed.

However, the prerequisite of using HMDs to assess components of visual processing in a variety of clinical and applied settings is that HMDs provide sufficient reliability. Reliable assessment with HMDs cannot be regarded as self-evident as HMDs differ in a number of ways from the usually employed cathode-ray-tube (CRT) screens (refresh rate, screen resolution, monitor model, etc.) that are usually employed for visual assessments (see also, Poth et al., 2018). Recently, we have provided first evidence that non-selective visual processing components can be measured reliably with a developer version of the HMD Oculus Rift (Foerster et al., 2016). In that study, we used a whole report TVA assessment (Bundesen, 1990; Bundesen & Habekost, 2014; Duncan et al., 1999; Vangkilde et al., 2011) to assess processing capacity parameters such as visual processing speed, visual working memory capacity, and the threshold of conscious perception. Reliabilities of all components were comparable across Oculus Rift and the CRT.

Yet, it is still unknown whether selective visual attention, another key process of TVA, can be measured similarly reliably with HMDs and whether other devices can be used that are now commercially available and differ from the early research prototypes. In the present study, we therefore investigated whether the commercially available HTC Vive (https://www.vive.com, Figure 1) allows to assess visual selective attention as well as processing capacity parameters as reliable as a standard CRT. Specifically, we used the combiTVA assessment (Vangkilde et al., 2011) by the HTC Vive and a standard CRT screen to measure and compare the test-retest reliabilities of the visual attention parameters top-down controlled selectivity and lateral attentional bias as well as the visual capacity parameters threshold of conscious perception, visual processing speed, and visual working memory capacity.

#### 2. Methods

#### 2.1 Participants

Initially, forty participants were recruited at Bielefeld University, Germany. Three participants had to be excluded due to data loss. Their data was substituted by three additionally recruited participants from Bielefeld University. Of these 40 participants, two were excluded after the TVA fitting procedure (see below) because the fitting procedure did not converge on plausible values. One participant had an implausibly high processing speed parameter deviating by 160 standard deviations from the remaining sample. The other participant had an implausibly high visual selectivity parameter, deviating by 10 standard deviations from the remaining sample. The remaining analyzed sample consisted of 15 males and 23 females with an average age of 25 years ranging from 19 to 38. Thirty-six participants stated to be right-handed and two participants stated to be left-handed. All participants reported normal or lens-corrected visual acuity, were naïve with respect to the purpose of the study, and provided written informed consent before the start of the experiment. Participation was rewarded with 8 €/h and a short 3D HTC Vive experience if desired (either The Cubicle. by Roel van Beek or InMind VR by Nival VR). The study was approved by the Bielefeld University's ethics committee and performed in accordance with the guidelines of the German Psychological Association (DGPs).

# 2.2 Apparatus

The experiment took place in a dimply-lit room. TVA-based assessment was performed on a Dell Precision T3600 (Dell, Round Rock, TX, USA) computer using 64bit Windows 7 (Microsoft, Seattle, WA, USA) and an NVIDIA GeForce GTX 1080 graphics card (NVIDIA, Santa Clara, CA, USA). The HTC Vive (HTC, New Taipei City, Taiwan) as well as a 19-inch

color CRT monitor (G90fB, ViewSonic, Brea, CA, USA) were used for stimulus presentation (Figure 1). The CRT had a refresh rate of 90 Hz and a spatial resolution of 1,024 x 786 pixels extending 36 x 27 cm. Participants' viewing distance was fixed at 71 cm with a chin rest.

Stereoscopic visualization in the HTC Vive is enabled by presenting individual images for the left and the right eye. To this end, the HTC Vive features two low-persistence OLED displays (one for each eye) with a refresh rate of 90 Hz and a spatial resolution of 1,080 x 1,200 pixels. A fisheye lens for each eye magnifies and distorts the views of each eye to achieve a field of view of 110 degrees. Using inverse lens distortion, the visualization software transforms the rendered images such that both distortions cancel out and the final image looks normal again. Because of this mechanism, the spatial resolution of the HTC Vive cannot be directly compared to the CRT. The special active-matrix OLED screen has an individual light source for every pixel. To minimize motion blur in virtual reality applications, pixels are illuminated for only a very short time period during one time frame and stay dark for the rest of the frame. This low-persistence display technique of the HTC Vive is therefore similar to the display technique of CRT screens, where a sweeping cathode-ray excites individual pixels for a short period of time (Elze, 2010).

The software for presenting the stimuli on either the CRT or the HTC Vive was written in C++, using the OpenGL library (version 4.1.0, Khronos Group, Beaverton, OR, USA) for visualization, the GLFW library (version 3.2.1, <u>http://www.glfw.org/license.html</u>) for user interface handling, and the OpenVR library (version 1.0.6, Valve, Bellevue, WA, USA) for controlling the HTC Vive. Steam (Valve, Bellevue, WA, USA) was used to operate the HTC Vive. Color and luminance of all stimuli presented on the CRT were measured in CIE Lxy coordinates using an X-Rite i1 Pro spectrophotometer (Munich, Germany).

----- Figure 1 about here ------

# 2.3 Stimulus presentation

All stimuli were displayed on a black background (L = 0.3 cd/m<sup>2</sup>, x = .3, y = .3). A white plus (L = 120.8 cd/m<sup>2</sup>, x = .3, y = .3; bold Courier New, font size 18, corresponding to about 14 pixels = .4°v.a.) served as central fixation cross. A set of 20 red and blue capital letters (ABDEFGHJKLMNOPRSTUVX; bold Arial, font size 68, corresponding to about 51 pixels =  $1.4^{\circ}v.a.$  in height; red with L = 36.0 cd/m<sup>2</sup>, x = .6, y = .3, and blue with L = 18.4, cd/m<sup>2</sup>, x = .2, y = .1) was used. Two or six letters were presented per trial at a center-to-center distance of 7.3°v.a. (= 256 pixels) at 45°, 90°, 135°, 225°, 270°, and 315° of an imaginary circle around the central fixation cross. One of eight different red-blue pattern masks (100 x 100 pixels = 2.8 x 2.8°v.a., L = 38.6, cd/m<sup>2</sup>, x = .6, y = .3 red, and L = 23.1, cd/m<sup>2</sup>, x = .2, y = .1 blue) followed on each position.

Due to the 3D stereoscopic visualization and the fisheye lens distortion in the HTC Vive, the effective sizes, distances, and viewing angles of the presented visual stimuli after lens distortion cannot be determined exactly from the stimulus' pixel sizes and positions before lens distortion. We manually adjusted the respective sizes, screen positions, and viewing parameters of the stimuli presented in the HTC Vive such that they matched as closely as possible the stimuli presented on the CRT.

#### 2.4 Procedure

Participants completed two TVA-based assessment days with each device (CRT and HTC Vive) separated by 7 days. On each day, participants completed one 45 minutes lasting TVA session with each device separated by a 30 minutes break, in which participants left the

experimental room. Each assessment day thus lasted for about two hours. The four possible assessment orders were counterbalanced across participants. However, after excluding two participants based on implausible fitting parameters, two orders were represented by 10 datasets, while the other two orders were represented by 9 datasets (CRT followed by Vive on both days as well as Vive followed by CRT on the first and CRT followed by Vive on the second day) in the analyzed sample. An individual participant was tested by the same of four experimenters at about the same time of day.

A modified version of the combiTVA assessment by Vangkilde, Bundesen, and Coull (2011) was used (see Figure 2). The combiTVA assessment consisted of 9 blocks each containing 36 trials. A block of 36 practice trials that was not included in the analysis preceded the nine experimental blocks. Participants were instructed to keep central fixation throughout the presentation. Each trial started with a red central fixation cross displayed for 1,000 ms. Next, two red letters (9 x per block), six red letters (18 x per block) or a combination of four blue and two red letters (9 x per block) appeared for a brief duration. While two red letters were always shown for 78 ms, the six-letter whole report display was presented for 11, 22, 56, 78, 144 or 200 ms (each duration 3 x per block). Each position on the imaginary circle around the fixation cross was occupied equally often by a target letter throughout the experiment. On each trial, the identity of the letters was randomly chosen with each letter appearing only once per trial. One pattern mask followed on each of the six possible letter positions for 500 ms. Each mask was randomly chosen from the set of the eight pattern masks with each mask appearing only once per trial.

----- Figure 2 about here ------

At the end of each trial, the screen went black and participants had to give an unspeeded verbal report of all letters they remembered in any order. The experimenter typed in the reported letters. The typed-in letters appeared on the display in white (bold letters of Verdana font with 32 pixel height =  $.9^{\circ}v.a.$ ; L =  $120.8 \text{ cd/m}^2$ , x = .3, y = .3), so that participants could ask for a correction if necessary. The experimenter started a new trial by pressing the space bar after the participant verified the reported letters and indicated to be ready for the next trial. Participants were instructed to report the letters they were fairly certain of having seen but to refrain from guessing (Vangkilde et al., 2011). Specifically, they were to aim at an accuracy of reported (typed-in) letters between 80 and 90 % (i.e. at error rates between 20 and 10 %). After each block, participants were informed about their report accuracy and reminded of the accuracy range (see Table 1 for the effectiveness of this instruction).

## 2.5 Implementation

While presenting the stimuli on the CRT is straightforward, visualizing them on the HTC Vive is slightly more involved and requires a three-step process: First, the scene is rendered as seen from the left eye and stored in a texture image. In the second step, another texture is filled with the right eye's view. In the third step, these two images are distorted in order to compensate for the effect of the fisheye lenses and are then shown on the two displays of the HTC Vive. Besides the different eye positions, the first two render passes are equivalent to the CRT rendering procedure. The third step is conveniently performed by the compositor functionality of the OpenVR library.

In order to present a stimulus for a specific duration, the required number of display refreshes is computed from the duration in milliseconds and the display refresh rate, here 90 Hz for both devices. We enable V-sync (vertical synchronization) in order to guarantee that the

individual frames produced by the graphics card are in synchronization with the refresh rate of the device. We made sure that our hardware setup is powerful enough for rendering the stimuli in less than one display refresh duration, such that we can guarantee the absence of dropped frames.

#### 2.6 Data analysis

We obtained for each participant and each assessment the five components of visual processing as defined within the TVA framework: The visual capacity parameters threshold of conscious perception  $t_0$  (in ms), the capacity of visual working memory *K* (in number of letters), visual processing speed *C* (letters per second), and the visual attention parameters top-down controlled selectivity  $\alpha$ , and the lateral attentional bias  $w_{index}$ . Top-down selectivity ranges from 0, indicating perfect selection, to 1, indicating unselective processing. The lateral attentional bias ranges from 0 as complete rightward bias to 1 as complete leftward bias.

The five TVA components are based on the number of correctly reported letters and obtained by analyzing the data with the LIBTVA toolbox (Dyrholm, Kyllingsbæk, Espeseth, & Bundesen, 2011), which implements an extension of the classic TVA model (Bundesen, 1990). The LIBTVA is a toolbox for MATLAB (R2013b, The Mathworks, Natick, MA, USA) and provides maximum-likelihood fitting routines for estimating each of the five components (see also, Kyllingsbæk, 2006). Fitting details were the same as in Foerster et al. (2016).

Besides the four components of visual processing, participants' error rates (rates of erroneously reported letters) were assessed to check whether participants were in the required accuracy range between 80 and 90%. Test-retest reliabilities of the five visual processing components were computed as Pearson's product-moment correlations between participants' components in the first and the second session for each device. Pearson's correlation was chosen over the intraclass correlation to quantify test-retest reliability of the TVA parameters because it

is well known that the TVA parameters are susceptible to learning effects. Thus, a linear relationship in the form of y = m \* x + b was expected across testing days rather than identical values in the form of y = x. We calculated significance values and confidence intervals for each correlation. Moreover, we calculated Steiger's (Steiger, 1980) *Z*-test for independent correlations in order to compare the test-retest reliabilities of the HTC Vive and the CRT. All data analyses and plotting procedures were performed using R3.4.0 (R Development Core Team, 2016) and the packages ggplot2 (Wickham, Chang, & RStudio, 2016), gridExtra (Auguie & Antonov, 2016), plyr (Wickham, 2016), and psych (Revelle, 2017).

# **3** Results

Each participant completed the TVA-based assessment (Figure 2) with a standard CRT screen and with the HTC Vive (Figure 1) on a first day and repeated the two assessment types seven days later. Descriptive statistics of the resulting five visual processing components as well as of the error rates are provided in Table 1. The processing components were significantly correlated between the HTC Vive and the CRT on the first as well as on the second day (Table 2). Learning effects from the first to the second day for each device as well as their differences across devices are reported in Table S1 in the Supplementary Material.

----- Table 1 about here ------

----- Table 2 about here ------

Figure 3 visualizes the test-retest reliabilities of all five visual processing components per assessment device as linear regression lines along with the individual participants' data in the two sessions.

----- Figure 3 about here ------

----- Table 3 about here ------

As can be seen in Table 3, the test-retest reliabilities of the five visual processing components were significant for both, the HTC Vive and the CRT. Moreover, none of the five test-retest reliabilities differed significantly between the HTC Vive and the CRT (Table 3). As has been reported before, the strength of the test-retest reliabilities of the five components differ with top-down controlled selectivity being at the lower end (Habekost et al., 2014).

# 4 Discussion

A prerequisite for differential assessment and fine-grained specification of deficits of visual attention and processing capacity in neuropsychological disorders is a highly standardized and reliable assessment across clinical institutions. A high standardization across varying testing environments can be achieved by utilizing HMDs, the increasingly available virtual reality devices, for neuropsychological assessment. HMDs offer the crucial advantage of an inherently standardized visual environment, so that the visual stimulation is exclusively dependent on the programmed stimulation. In this way, a constant visual stimulation (luminance, color, viewing distance etc.) can be guaranteed across different testing conditions and institutions. At the same time, assessment is facilitated and bedside testing is enabled by using these portable HMDs.

A second prerequisite for a suitable assessment instrument is, however, that it provides sufficient reliabilities. Here, we compared the test-retest reliabilities of five components of visual apprehension in terms of visual selective attention and processing capabilities measured with the HTC Vive to the respective reliabilities measured with a standard CRT screen which is the current gold-standard for assessing visual processing components. Specifically, we applied the combiTVA assessment (Vangkilde, Coull, & Bundesen, 2012) on both devices to estimate the visual attention parameters top-down controlled selective attention and lateral attentional bias, as well as the visual capacity parameters threshold of conscious perception, visual processing speed, and capacity of visual working memory. Results revealed that the test-retest reliabilities of all five TVA components of visual apprehension were in a comparable range across the HTC Vive and the standard CRT assessment and did not differ significantly. In addition, most test-retest reliabilities were in an acceptable range for neuropsychological assessment. As found before (Habekost et al., 2014), top-down controlled selectivity provided the lowest reliability of .42 measured with the CRT and .64 measured with the HTC Vive..

This finding extends the results of our previous study (Foerster et al., 2016) substantially, in which we have demonstrated that the developer version of the HMD Oculus Rift can be used to reliably assess three parameters of non-selective visual processing capacity. Specifically, in this precursor study, we used the TVA-based whole report paradigm (Bundesen, 1990; Duncan et al., 1999; Habekost, 2015) to measure the threshold of conscious perception, visual processing speed, and visual working memory. The test-retest reliabilities of these three components of nonselective visual processing capacity were comparable across the developer version of the Oculus Rift and a standard CRT.

The present investigation extended this previous finding by revealing that the second key component of visual apprehension, namely visual selective attention in terms of top-down

controlled selectivity and lateral attentional bias can also be measured with an HMD that is commercially available as reliable as with a standard CRT. This ability to process relevant information with higher priority than non-relevant information in crowded natural environments is of key importance for all types of intelligent behavior (e.g., Desimone & Duncan, 1995). The finding of mostly reliable components of visual apprehension assessed by an HMD is of high importance for clinical purposes. Firstly, reliable and highly standardized neuropsychological diagnostic is possible with a commercially available HMD at a relatively low cost. Secondly, diagnosing patients with portable HMDs allows highly standardized bedside testing, especially as the assessment we tested here does not require a permanent internet connection.

The reliable measurement of visual selective attention is of special interest in neuropsychological purposes as several neurological and psychiatric disorders show various types of attention impairments (for a review see, Habekost, 2015). Specifically, top-down controlled selectivity has been found to be affected in patients with frontal and parietal lesions after stroke, spina bifida myelomeningocele, mild cognitive impairment, and Alzheimer's disease (Bublak et al., 2005; Caspersen & Habekost, 2013; Peers et al., 2005; Redel et al., 2012). Moreover, unusual lateral attentional biases have been observed in patients with neglect, stroke, ventral simultanagnosia, dyslexia, Huntington's disease, mild cognitive impairment, and Alzheimer's disease (Bublak et al., 2005; Duncan et al., 1999; Finke et al., 2011; Habekost, 2015; Habekost & Rostrup, 2007; Kraft et al., 2015; Peers et al., 2005; Redel et al., 2012; Stenneken et al., 2011).

The second key ability of visual apprehension refers to visual processing capacity. Within the TVA framework, it is subdivided into the components of threshold of conscious perception, visual processing speed, and working memory capacity. Especially the latter two are differentially affected in neuropsychological disorders. As an example, simultanagnosia is

characterized by deficits in processing speed, but not working memory capacity (Duncan et al., 2003; Finke et al., 2007; Neitzel et al., 2016). The same differentiation was observed in children with attention deficit hyperactivity disorder (ADHD, McAvinue et al., 2015), while adults with ADHD were impaired in working memory capacity rather than processing speed (Finke et al., 2011). Processing speed as well as working memory capacity are impaired in neglect patients, and patients with parietal stroke, Alzheimer's disease, Huntington's disease, alexia, and dyslexia, and processing speed declines with advanced age (Habekost, 2015; Habekost & Starrfelt, 2008). The threshold of conscious perception is also affected in mild cognitive impairment and Alzheimer's disease (Bublak et al., 2011). These clinical examples illustrate the importance of the assessment of differential aspects of visual selective attention and processing capacities for the purpose of clinical research and diagnostics. With the help of an even higher standardized and reliable assessment procedure achieved by using HMDs the partly contrasting results such as in the case of ADHD (Finke et al., 2011; McAvinue et al., 2015) might be resolved in the future. Note that the TVA assessment often needed to be adapted to the specific patient group, e.g. by splitting up partial and whole report trials in separate sessions and adapting the presentation durations (Bublak et al., 2005; Duncan et al., 1999, 2003; Finke et al., 2005; Finke, Bublak, Dose, Müller, & Schneider, 2006; Redel et al., 2012). Shortening test sessions prevent mental fatigue and adapting presentation durations ensures reliable fitting (enough data points within the highly curved region of the function that predicts performance by presentation duration). Testretest reliabilities of several different TVA versions assessed with an HMD thus need to be obtained in future studies.

When using HMDs for clinical purposes, patients who cannot be transported or sit upright can be assessed with exactly the same setup in bedside testing. In this way, assessment of visual selective attention and processing is rendered possible for patient groups and disease stages that

could not be tested heretofore. Moreover, the assessment of individual differences in visual apprehension (e.g., visual processing speed, selective attention, etc.) is also of interest for differential psychology as they depict important aspects of basic visual-cognitive abilities that are involved in many everyday tasks such as keyboarding, reading or handling cups (e.g., Foerster, Carbone, Koesling, & Schneider, 2011; Land & Tatler, 2009).

Therefore, we would like to encourage researchers and practitioners in neuropsychology and beyond to employ our software for standardized visual presentation within HMDs, which can be used for combiTVA, as well as for separate partial and whole report sessions and is available at <u>www.uni-bielefeld.de/psychologie/abteilung/arbeitseinheiten/01/Research/VR/</u>. The software works with the three commercially available HMDs HTC Vive, HTC Vive Pro, and Oculus Rift and a PC or a laptop with a graphics card that is at least as powerful as the NVIDIA GTX 1070 (check the system requirements provided on the websites by HTC Vive or Oculus Rift). All three HMDs can be operated without an internet connection which allows high flexibility. This is an important advantage for clinical applications in which a stable internet connection might not always be available.

Besides these numerous important advantages of using HMDs for neuropsychological purposes, there are also some limitations of course. Numerous studies reported that HMDs can induce cybersickness or motion sickness, especially in highly immersive VR environments (LaViola, 2000; Merhi, Faugloire, Flanagan, & Stoffregen, 2007). Participants in our study did not report any motion sickness or cybersickness during the TVA assessment. Motion sickness or cybersickness was presumably not an issue in our study because of our sparse visual environment, i.e. presenting stationary red and blue letters relatively near to the center (7.3°) on a black background in combination with the fact that participants did not move their head or body. Future studies have to be conducted to investigate the experience during such sparse visual

stimulations in VR when it comes to different participant groups, especially neurological patients and older adults who seem to be more susceptible to cybersickness (Arns & Cerney, 2005). For older adults, additional limitations of HMD testing might arise. Firstly, as older adults are less acquainted to this kind of technology, they might have reservations against it, so that a familiarization might be needed. Secondly, adjusting the HMD so that it is comfortable is a bit more taxing when the person is wearing glasses. As the HMD is attached to the head, wearing might become uncomfortable over a longer period of time. However, as technology develops fast, this problem might be solved in the near future. Finally, it is known that with longer testing on an electronic display, eve fatigue or evestrain can arise (Jeong, 2012). Evestrain is higher for watching 3D vs. 2D content (Lee, Heo, & Park, 2010). However, here, we presented all stimuli as perceived on a flat plane. Nevertheless, the plane itself is not flat and eye fatigue might be higher when perceiving items on a virtually flat plane than on a real flat plane. Eye fatigue also increases with time on test, which was balanced for the two devices in the present study (balanced test orders). Eye fatigue is also increased with higher rates of flicker (LaViola, 2000). In the present study, the same frame rate and thus flicker rate was applied in both devices. The periphery of the eye is more sensitive to flicker than its center, so that flickering in HMDs with their larger field of view has a higher potential to introduce eye fatigue. In addition, the display is placed closer to the person's eye in an HMD. However, stimulus arrangement and distance in the HTC Vive were manually adjusted to match the CRT presentation. Thus, it is unlikely that eye fatigue was differently influencing the results. Importantly, even if absolute eye fatigue would have been different across devices, test-retest reliabilities were comparable arguing that eve fatigue can be neglected, at least with healthy participants. Future studies have to tell whether these limitations of HMD testing can be overcome and whether the advantages of extremely high standardization,

portability, and presentation possibilities (3D, far periphery, single eye, etc.) will outweigh the limitations of using HMDs for neuropsychological purposes.

In summary, we demonstrated for the first time that two key characteristics of visual apprehension, the visual selective attention parameters top-down controlled selectivity and lateral attentional bias as well as the processing capacity parameters threshold of conscious perception, visual processing speed, and working memory capacity can be measured by commercially available HMDs as reliable as with a CRT, the current gold standard. Crucially, the inherent standardized visual presentation of these HMDs allows to compare individual values of visual apprehension components of healthy individuals and patient groups at various disease stages (including bedside testing) and across clinical institutions. Based on reliable measurements and standardized visual presentation, statistical norm distributions can be established that are of key importance to neuropsychological diagnostics.

#### **5** Acknowledgements

This research was supported by the Cluster of Excellence Cognitive Interaction Technology 'CITEC' (EXC 277) at Bielefeld University, which is funded by the German Research Foundation (DFG). We thank Signe Vangkilde for providing us the pattern masks.

#### **6** Author Contributions

RMF, CHP, and WXS designed the experiment. CB developed the experimental software under supervision by MB. CHP and RMF analyzed the data. RMF and CHP interpreted the results. RMF, CHP, WXS, and MB discussed the study. RMF wrote the first draft of the manuscript and RMF, WXS, CHP, and MB revised it. RMF prepared the supplementary material.

#### 7 References

- Arns, L. L., & Cerney, M. M. (2005). The relationship between age and incidence of cybersickness among immersive environment users. In *IEEE Proceedings. VR 2005. Virtual Reality, 2005.* (Vol. 2005, pp. 267–268). IEEE. doi:10.1109/VR.2005.1492788
- Auguie, B., & Antonov, A. (2016). gridExtra: Miscellaneous Functions for "Grid" Graphics. R package version 2.2.1. Retrieved from https://cran.r-project.org/package=gridExtra
- Bowie, C. R., & Harvey, P. D. (2006). Administration and interpretation of the trail making test. *Nature Protocols*, *1*, 2277–2281. doi:10.1038/nprot.2006.390
- Bublak, P., Finke, K., Krummenacher, J., Preger, R., Kyllingsbaek, S., Müller, H. J., &
  Schneider, W. X. (2005). Usability of a theory of visual attention (TVA) for parameterbased measurement of attention II: Evidence from two patients with frontal or parietal damage. *Journal of the International Neuropsychological Society*, *11*, 843–854. doi:10.1017/S1355617705050988
- Bublak, P., Redel, P., Sorg, C., Kurz, A., Förstl, H., Müller, H. J., ... Finke, K. (2011). Staged decline of visual processing capacity in mild cognitive impairment and Alzheimer's disease. *Neurobiology of Aging*, 32, 1219–1230. doi:10.1016/j.neurobiolaging.2009.07.012
- Bundesen, C. (1990). A theory of visual attention. *Psychological Review*, 97, 523–547.
- Bundesen, C., & Habekost, T. (2008). Principles of visual attention: Linking mind and brain.

Oxford: Oxford University Press.

- Bundesen, C., & Habekost, T. (2014). Theory of Visual Attention (TVA). In S. Kastner & K. Nobre (Eds.), *The Oxford handbook of attention* (pp. 1095–1121). Oxford: Oxford University Press.
- Bundesen, C., Habekost, T., & Kyllingsbaek, S. (2005). A neural theory of visual attention:
  Bridging cognition and neurophysiology. *Psychological Review*, *112*, 291–328.
  doi:10.1037/0033-295X.112.2.291
- Buschman, T. J., Siegel, M., Roy, J. E., & Miller, E. K. (2011). Neural substrates of cognitive capacity limitations. *Proceedings of the National Academy of Sciences*, *108*, 11252–11255. doi:10.1073/pnas.1104666108
- Caspersen, I. D., & Habekost, T. (2013). Selective and sustained attention in children with spina bifida myelomeningocele. *Child Neuropsychology*, *19*, 55–77.
  doi:10.1080/09297049.2011.639753
- Cowan, N. (2000a). Processing limits of selective attention and working memory: Potential implications for interpreting. *Interpreting*, *5*, 117–146. doi:10.1075/intp.5.2.05cow
- Cowan, N. (2000b). The magical number 4 in short-term memory: A reconsideration of mental storage capacity Dep. *Behavioral and Brain Sciences*, *24*, 87–185. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11515286
- Desimone, R., & Duncan, J. S. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*, 193–222. doi:10.1146/annurev.ne.18.030195.001205
- Duncan, J., Bundesen, C., Olson, A., Humphreys, G., Chavda, S., & Shibuya, H. (1999).
   Systematic analysis of deficits in visual attention. *Journal of Experimental Psychology: General*, 128, 450–478.
- Duncan, J., Bundesen, C., Olson, A., Humphreys, G., Ward, R., Kyllingsbæk, S., ... Chavda, S.

(2003). Attentional functions in dorsal and ventral simultanagnosia. *Cognitive Neuropsychology*, *20*, 675–701. doi:10.1080/02643290342000041

- Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, *96*, 433–458. doi:10.1037/0033-295X.96.3.433
- Dyrholm, M., Kyllingsbæk, S., Espeseth, T., & Bundesen, C. (2011). Generalizing parametric models by introducing trial-by-trial parameter variability: The case of TVA. *Journal of Mathematical Psychology*, 55, 416–429. doi:10.1016/j.jmp.2011.08.005
- Elze, T. (2010). Misspecifications of stimulus presentation durations in experimental psychology:
  A systematic review of the psychophysics literature. *PLoS ONE*, *5*, e12792.
  doi:10.1371/journal.pone.0012792
- Estes, W. K. (1972). Interactions of signal and background variables in visual processing. *Perception and Psychophysics*, *12*, 278–286. doi:10.3758/BF03207204
- Finke, K., Bublak, P., Dose, M., Müller, H. J., & Schneider, W. X. (2006). Parameter-based assessment of spatial and non-spatial attentional deficits in Huntington's disease. *Brain*, 129, 1137–1151. doi:10.1093/brain/awl040
- Finke, K., Bublak, P., Krummenacher, J., Kyllingsbæk, S., Müller, H. J., & Schneider, W. X. (2005). Usability of a theory of visual attention (TVA) for parameter-based measurement of attention I: Evidence from normal subjects. *Journal of the International Neuropsychological Society*, 11, 832–842. doi:10.1017/S1355617705050976

Finke, K., Schneider, W. X., Redel, P., Dose, M., Kerkhoff, G., Müller, H. J., & Bublak, P.
(2007). The capacity of attention and simultaneous perception of objects: A group study of Huntington's disease patients. *Neuropsychologia*, 45, 3272–3284. doi:10.1016/j.neuropsychologia.2007.06.006

Finke, K., Schwarzkopf, W., Müller, U., Frodl, T., Müller, H. J., Schneider, W. X., ... Hennig-

Fast, K. (2011). Disentangling the adult attention-deficit hyperactivity disorder
endophenotype: Parametric measurement of attention. *Journal of Abnormal Psychology*, *120*, 890–901. doi:10.1037/a0024944

- Finkel, D., Reynolds, C. A., McArdle, J. J., & Pedersen, N. L. (2007). Age changes in processing speed as a leading indicator of cognitive aging. *Psychology and Aging*, 22, 558–568. doi:10.1037/0882-7974.22.3.558
- Foerster, R. M., Carbone, E., Koesling, H., & Schneider, W. X. (2011). Saccadic eye movements in a high-speed bimanual stacking task: Changes of attentional control during learning and automatization. *Journal of Vision*, *11(7):9*, 1–16. doi:10.1167/11.7.9
- Foerster, R. M., Poth, C. H., Behler, C., Botsch, M., & Schneider, W. X. (2016). Using the virtual reality device Oculus Rift for neuropsychological assessment of visual processing capabilities. *Scientific Reports*, 6, 37016. doi:10.1038/srep37016
- Foerster, R. M., & Schneider, W. X. (2013). FuncSim Toolbox for Matlab: Computation of eyetracking scanpath similarity. Bielefeld University. doi:10.4119/unibi/citec.2013.7
- Foerster, R. M., & Schneider, W. X. (2015). Expectation violations in sensorimotor sequences: Shifting from LTM-based attentional selection to visual search. *Annals of the New York Academy of Sciences*, 1339, 45–59. doi:10.1111/nyas.12729
- Habekost, T. (2015). Clinical TVA-based studies: A general review. *Frontiers in Psychology*, *6*, 1–18. doi:10.3389/fpsyg.2015.00290

Habekost, T., Petersen, A., & Vangkilde, S. (2014). Testing attention: Comparing the ANT with TVA-based assessment. *Behavior Research Methods*, *46*, 81–94. doi:10.3758/s13428-013-0341-2

Habekost, T., & Rostrup, E. (2007). Visual attention capacity after right hemisphere lesions. *Neuropsychologia*, 45, 1474–1488. doi:10.1016/j.neuropsychologia.2006.11.006

- Habekost, T., & Starrfelt, R. (2008). Visual attention capacity: A review of TVA-based patient studies. *Scandinavian Journal of Psychology*, *50*, 23–32. doi:10.1111/j.1467-9450.2008.00681.x
- Habekost, T., & Starrfelt, R. (2009). Visual attention capacity: A review of TVA based patient studies. *Scandinavian Journal of Psychology*, *50*, 97–109.
- Jeong, H. (2012). A comparison of the influence of electronic books and paper books on reading comprehension, eye fatigue, and perception. *The Electronic Library*, *30*, 390–408. doi:10.1108/02640471211241663
- Kraft, A., Dyrholm, M., Bundesen, C., Kyllingsbæk, S., Kathmann, N., & Brandt, S. A. (2013).
  Visual attention capacity parameters covary with hemifield alignment. *Neuropsychologia*, *51*, 876–885. doi:10.1016/j.neuropsychologia.2013.01.021
- Kraft, A., Irlbacher, K., Finke, K., Kaufmann, C., Kehrer, S., Liebermann, D., ... Brandt, S. A. (2015). Dissociable spatial and non-spatial attentional deficits after circumscribed thalamic stroke. *Cortex*, 64, 327–342. doi:10.1016/j.cortex.2014.12.005
- Kyllingsbæk, S. (2006). Modeling visual attention. *Behavior Research Methods*, *38*, 123–133. doi:10.3758/BF03192757
- Land, M. F., & Tatler, B. W. (2009). Looking and acting. New York: Oxford University Press.
- LaViola, J. J. (2000). A discussion of cybersickness in virtual environments. *ACM SIGCHI Bulletin*, *32*, 47–56. doi:10.1145/333329.333344

Lee, E., Heo, H., & Park, K. (2010). The comparative measurements of eyestrain caused by 2D and 3D displays. *IEEE Transactions on Consumer Electronics*, 56, 1677–1683. doi:10.1109/TCE.2010.5606312

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, 390, 279–281. doi:10.1038/36846

- Ma, W. J., Husain, M., & Bays, P. M. (2014). Changing concepts of working memory. *Nature Neuroscience*, 17, 347–356. doi:10.1038/nn.3655
- McAvinue, L. P., Vangkilde, S., Johnson, K. A., Habekost, T., Kyllingsbaek, S., Bundesen, C., & Robertson, I. H. (2015). A componential analysis of visual attention in children with ADHD. *Journal of Attention Disorders*, *19*, 882–894. doi:10.1177/1087054712461935
- Merhi, O., Faugloire, E., Flanagan, M., & Stoffregen, T. A. (2007). Motion sickness, console video games, and head-mounted displays. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 49, 920–934. doi:10.1518/001872007X230262
- Neitzel, J., Ortner, M., Haupt, M., Redel, P., Grimmer, T., Yakushev, I., ... Finke, K. (2016).
   Neuro-cognitive mechanisms of simultanagnosia in patients with posterior cortical atrophy.
   *Brain*, 139, 3267–3280. doi:10.1093/brain/aww235
- Oberauer, K. (2009). Design for a working memory. In B. H. Ross (Ed.), *Advances in research and theory* (Vol. 51, pp. 45–100). San Diego: Elsevier Academic. doi:10.1016/S0079-7421(09)51002-X
- Pashler, H. (1987). Target-distractor discriminability in visual search. *Perception and Psychophysics*, *41*, 285–292. doi:10.3758/BF03208228
- Peers, P. V., Ludwig, C. J. H., Rorden, C., Cusack, R., Bonfiglioli, C., Bundesen, C., ... Duncan, J. (2005). Attentional functions of parietal and frontal cortex. *Cerebral Cortex*, 15, 1469– 1484. doi:10.1093/cercor/bhi029
- Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual Review* of Neuroscience, 13, 25–42.
- Poth, C. H., Foerster, R. M., Behler, C., Schwanecke, U., Schneider, W. X., & Botsch, M. (2018).
   Ultrahigh temporal resolution of visual presentation using gaming monitors and G-Sync.
   Behavior Research Methods, 50, 26–38. doi:10.3758/s13428-017-1003-6

- Poth, C. H., Petersen, A., Bundesen, C., & Schneider, W. X. (2014). Effects of monitoring for visual events on distinct components of attention. *Frontiers in Psychology*, *5*, 1–11. doi:10.3389/fpsyg.2014.00930
- Poth, C. H., & Schneider, W. X. (2018). Attentional competition across saccadic eye movements. *Acta Psychologica*, *190*, 27–37. doi:10.1016/j.actpsy.2018.06.011
- Redel, P., Bublak, P., Sorg, C., Kurz, A., Förstl, H., Müller, H. J., ... Finke, K. (2012). Deficits of spatial and task-related attentional selection in mild cognitive impairment and Alzheimer's disease. *Neurobiology of Aging*, *33*, 195.e27-195.e42. doi:10.1016/j.neurobiolaging.2010.05.014
- Reitan, R. M., & Wolfson, D. (1994). A selective and critical review of neuropsychological deficits and the frontal lobes. *Neuropsychology Review*, 4, 161–198. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7881456
- Revelle, W. (2017). psych: Procedures for psychological, psychometric, and personality research. R package version 1.7.5. Retrieved from https://cran.r-project.org/package=psych
- Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin*, 87, 245–251.
- Stenneken, P., Egetemeir, J., Schulte-Körne, G., Müller, H. J., Schneider, W. X., & Finke, K. (2011). Slow perceptual processing at the core of developmental dyslexia: A parameterbased assessment of visual attention. *Neuropsychologia*, 49, 3454–3465. doi:10.1016/j.neuropsychologia.2011.08.021
- Unsworth, N., & Engle, R. W. (2007). The nature of individual differences in working memory capacity: Active maintenance in primary memory and controlled search from secondary memory. *Psychological Review*, *114*, 104–132. doi:10.1037/0033-295X.114.1.104

Vangkilde, S., Bundesen, C., & Coull, J. T. (2011). Prompt but inefficient: Nicotine differentially

modulates discrete components of attention. *Psychopharmacology*, *218*, 667–680. doi:10.1007/s00213-011-2361-x

- Vangkilde, S., Coull, J. T., & Bundesen, C. (2012). Great expectations: Temporal expectation modulates perceptual processing speed. *Journal of Experimental Psychology: Human Perception and Performance*, 38, 1183–1191. doi:10.1037/a0026343
- Verghese, P., & Nakayama, K. (1994). Stimulus discriminability in visual search. Vision Research, 34, 2453–2461.
- Wickham, H. (2016). plyr: Tools for splitting, applying and combining data. R package version 1.8.4. Retrieved from https://cran.r-project.org/package=plyr
- Wickham, H., Chang, W., & RStudio. (2016). ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics. R package version 2.2.1. Retrieved from https://cran.rproject.org/package=ggplot2

*Table 1.* Descriptive statistics. Means (*M*) and standard deviations (*SD*) of the five visual processing components (threshold of conscious perception, processing speed, working memory capacity, top-down controlled selectivity, and lateral attentional bias) for the two sessions performed using the HTC Vive and the two sessions performed using the CRT.

÷	HTC Vive		CRT		
	Session 1 M(SD)	Session 2 M(SD)	Session 1 M(SD)	Session 2 M(SD)	
Threshold of conscious perception	30.12 (12.48)	29.30 (12.10)	23.06 (10.98)	22.53 (8.88)	
Visual processing speed	73.32 (27.44)	90.53 (34.03)	61.75 (25.9)	78.84 (34.72)	
Visual working memory capacity	3.14 (.89)	3.29 (.73)	2.94 (.86)	3.23 (.82)	
Top-down controlled selectivity	.41 (.19)	.35 (.22)	.37 (.21)	.24 (.12)	
Lateral attentional bias	.49 (.11)	.48 (.10)	.48 (.11)	.48 (.12)	
Error rate	.10 (.06)	.08 (.06)	.10 (.06)	.07 (.05)	

*Table 2.* Pearson correlations r and their corresponding p-values in parentheses between the measurements of the HTC Vive and the CRT for each of the five processing components (threshold of conscious perception, processing speed, working memory capacity, top-down controlled selectivity, and lateral attentional bias) on both days.

Session 1	Session 2
$r\left(p ight)$	$r\left(p ight)$
.73 (< .001)	.62 (< .001)
.60 (< .001)	.62 (< .001)
.75 (< .001)	.85 (< .001)
.67 (< .001)	.54 (< .001)
.75 (< .001)	.83 (< .001)
	Session 1 r (p) .73 (< .001) .60 (< .001) .75 (< .001) .67 (< .001) .75 (< .001)

*Table 3*. Test-retest reliabilities of the five visual processing components (threshold of conscious perception, processing speed, working memory capacity, top-down selectivity, and lateral attentional bias) for the HTC Vive and the CRT. Test-retest reliabilities are provided as Pearson's r along with their corresponding p-values and confidence intervals (CI). The last column provides Steiger's Z-test values and the corresponding p-values for the comparisons between the HTC Vive and the CRT regarding the test-retest reliabilities of the five visual processing components.

	HTC Vive		CRT		
	$r\left(p ight)$	CI of <i>r</i>	$r\left(p ight)$	CI of <i>r</i>	Steiger's Z (p)
Threshold of conscious perception	.81 (< .001)	[.66; .90]	.63 (< .001)	[.39; .79]	1.59 (.11)
Visual processing speed	.74 (< .001)	[.55; .86]	.54 (< .001)	[.27; .74]	1.44 (.15)
Visual working memory capacity	.75 (< .001)	[.56; .86]	.84 (< .001)	[.71; .92]	1.10 (.27)
Top-down controlled selectivity	.64 (< .001)	[.41; .80]	.42 (.008)	[.12; .66]	1.30 (.19)
Lateral attentional bias	.81 (< .001)	[.67; .90]	.80 (< .001)	[.65; .89]	.14 (.89)



*Figure 1*. Assessment set-ups. A participant sitting in front of a standard CRT with head position fixed by a chin rest (left) and wearing the HTC Vive virtual reality device (right).



*Figure 2*. Trial sequence of the TVA-based assessment (based on the combiTVA paradigm, Vangkilde et al., 2011.



*Figure 3.* Test-retest reliabilities of the five visual processing components as linear regression lines along with the individual participants' data. Threshold of conscious perception (top-left panel), working memory capacity (top-right panel), processing speed (middle-left panel), top-down controlled selectivity (middle-right panel), and lateral attentional bias (bottom-left panel) for both devices, i.e. the HTC Vive (red [grey] points and lines) and the CRT (black points and lines). The values of session 1 are depicted on the x-axis and the values of session 2 on the y-axis. The straight lines represent the linear regressions.

# **Supplementary Material**

# Neuropsychological assessment of visual selective attention and processing capacity with head-mounted displays

Rebecca M. Foerster<sup>1,3</sup>, Christian H. Poth<sup>1,3</sup>, Christian Behler<sup>2,3</sup>, Mario Botsch<sup>2,3</sup>, & Werner X. Schneider<sup>1,3</sup>

<sup>1</sup>Neuro-cognitive Psychology, <sup>2</sup>Computer Graphics and Geometry Processing, <sup>3</sup>Cluster of Excellence Cognitive Interaction Technology, Bielefeld University, P. O. Box 100131, D-33501 Bielefeld, Germany

Correspondence to Rebecca M. Foerster (rebecca.foerster@uni-bielefeld.de)

*Table S1*. Paired *t*-test statistics (*t*-value, *p*-value, and *Cohen's d<sub>z</sub>*) for the comparison of the five visual processing components (threshold of conscious perception, processing speed, working memory capacity, top-down controlled selectivity, and lateral attentional bias) between the first and the second day for the HTC Vive and the CRT as well for the comparison of the component changes (day 2 - day 1) of the two devices.

<u> </u>	HTC Vive		CRT		HTC Vive vs. CRT	
	<i>t</i> ( <i>p</i> )	Cohen's $d_z$	t(p)	Cohen's $d_z$	t (p)	Cohen's $d_z$
Threshold of conscious perception	.67 (.51)	.11	.37 (.71)	.06	.16 (.87)	.03
Visual processing speed	4.63 (< .001)	.75	3.52 (< .01)	.57	.02 (.98)	.003
Visual working memory capacity	1.53 (.14)	.25	3.74 (< .001)	.61	1.02 (.31)	.17
Top-down controlled selectivity	1.82 (.08)	.29	4.23 (< .001)	.69	2.22 (<.05)	.36
Lateral attentional bias	.87 (.39)	.14	.73 (.47)	.12	1.28 (.21)	.21