Multidimensional assessment of Virtual Reality applications in clinical neuropsychology: the "VR-Check" protocol

Stephan Krohn
Department of Neurology
Charité-Universitätsmedizin Berlin
Berlin, Germany
stephan.krohn@charite.de

Johanne Tromp
Department of Neurology
Max-Planck Institute for Human Cognitive and Brain Sciences
Leipzig, Germany
tromp@cbs.mpg.de

Eva M Quinque
Clinic for Cognitive Neurology
University Hospital Leipzig
Leipzig, Germany
eva.quinque@medizin.uni-leipzig.de

Julia Belger
Clinic for Cognitive Neurology
University Hospital Leipzig
Leipzig, Germany
julia.belger@medizin.uni-leipzig.de

Felix Klotzsche
Department of Neurology
Max-Planck Institute for Human Cognitive and Brain Sciences
Leipzig, Germany
klotzsche@cbs.mpg.de

Michael Gaebler
Department of Neurology
Max-Planck Institute for Human Cognitive and Brain Sciences
Leipzig, Germany
gaebler@cbs.mpg.de

Angelika Thöne-Otto
Clinic for Cognitive Neurology
University Hospital Leipzig
Leipzig, Germany
angelika.thoene@medizin.uni-leipzig.de

Carsten Finke
Department of Neurology
Charité-Universitätsmedizin Berlin
Berlin, Germany
carsten.finke@charite.de

***Abstract*— The experimental features of Virtual Reality (VR) applications in clinical neuropsychology are only inadequately captured by classical psychometric quality criteria, warranting an extended paradigm evaluation framework. Here we propose such a framework by means of a multidimensional checklist (VR-Check). Evaluation dimensions encompass cognitive domain specificity, ecological relevance, technical feasibility, user feasibility, user motivation, task adaptability, performance quantification, immersive capacities, training feasibility, and predictable pitfalls. Assessment along these VR-Check dimensions enables a systematic and comparative evaluation of neuropsychological VR applications, as we illustrate here for the exemplary case of assessing spatial cognition with immersive VR.**

Keywords— Virtual Reality, evaluation, neuropsychology, paradigm design

I. Introduction

The rapidly growing literature on Virtual Reality (VR) applications for clinical purposes has recently led Rizzo and König to suggest that clinical VR may be "ready for primetime" [1]. Similarly, a consortium of international VR experts has just put forward a catalogue of best-practice recommendations for running clinical trials with VR healthcare tools [2]. In clinical neuropsychology, the increasing relevance of VR applications presents researchers with the challenge of evaluating these applications systematically in order to optimize paradigm design for their specific research question. Moreover, the extensive design possibilities of VR yield qualitatively new task features that are not adequately captured by classical psychometric quality criteria (objectivity, reliability, validity), highlighting the need of an extended paradigm evaluation framework. Here we propose such a framework for VR applications in form of a multidimensional checklist: VR-Check.

II. VR-Check Dimensions

Fig. 1 summarizes the proposed evaluation criteria. Domain specificity is judged in light of existing evidence regarding the domain of interest (e.g. relation to other assessments with well-established specificity) and foreseeable domain confounds. Next, we assess how relevant the virtual world, the experimental stimuli, and the activities performed to solve the task are to the user's everyday life. Technical feasibility concerns whether implementation of the task is feasible in a virtual environment in general and specifically in the desired head-mounted display (HMD), 2D display, or both. User interaction and navigation in the virtual environment are judged with respect to the feasibility and necessity of input devices such as VR controllers or a mouse. Next, the paradigm is assessed with respect to its feasibility in healthy subjects, the target patient population and different age groups. Here, potential caveats due to navigation and interaction complexity must be addressed. Similarly, long task duration or high attentional demands may restrict feasibility in some user groups and must thus inform paradigm design. Likelihood and severity of adverse effects (e.g. cybersickness) in the user groups of interest are considered. Potential ethical concerns are also made explicit here. Regarding user motivation, we appraise the likely user-subjective expected benefit, the entertainment factor of the paradigm, the feasibility of a within- or across-session reward system, and the feasibility of a direct performance feedback to the user. Task adaptability is assessed with respect to the creation of parallel versions, the (automated) grading of task difficulty, and the possibility to modify the task to induce sufficient performance variance in the target population. Next, we address if outcome variables to quantify user performance exist or can be derived, if these are validated, and if they can be measured and evaluated independent of the experimenter administering the task. Moreover, we examine task and system factors that influence the likelihood of creating the sense of

The present work emanates from the research consortium VReha – virtual worlds for digital diagnostics and cognitive rehabilitation, funded by the German Federal Ministry for Education and Research (BMBF grant 13GW0206D).

being present in the virtual environment. Importantly, we address the paradigm‘s feasibility for cognitive training with regard to factors that may hinder repeated application and minimum requirements of task adaptability. The possibility of supporting cognitive strategies is assessed, and the likelihood and quality of transfer effects (especially near vs. far) is estimated, ideally based on prior evidence. Finally, we address predictable pitfalls to examine adherence to task requirements. Time, know-how, and monetary costs of implementing or modifying a paradigm are weighed against potential scientific and patient benefit in order to optimize the allocation of study resources.

VR-Check	
Evaluation dimension	*Application features*
Domain specificity	Evidence from existing literature Potential domain confounds
Ecological relevance	Relevance of virtual environment Relevance of experimental stimuli Relevance of user response to task
Technical feasibility	Implementable in VR Compatible with HMD Compatible with 2D screen Navigation and interaction devices
User feasibility	Feasible in healthy population Feasible in patient target groups Navigation and interaction complexity Duration and attentional demands Adverse effects (VR exposure) Ethical concerns
User motivation	Expected benefit Entertainment factor Reward system Performance feedback
Task adaptability	Parallel versions Level of difficulty Induction of performance variance
Performance quantification	Existing or derivable outcome variables Experimenter-independent evaluation
Immersive capacities	Immersive task and system factors Facilitates a sense of being in VE
Training feasibility	Possibility of repeated application Necessary adaptability conditions met Possibility of strategy training Likelihood and quality of transfer
Predictable pitfalls	Adherence to task requirements Allocation of study resources

Fig. 1. VR-Check evaluation dimensions.

III. Example case: spatial cognition

Fig. 2 summarizes the proposed framework (panel A) and an exemplary application in a specific research project (panels B and C). We sought to define an immersive VR paradigm for the neuropsychological assessment of spatial cognition, with maximum feasibility in a wide range of neurological patient groups, high potential for cognitive training, and where interaction with the virtual environment could be implemented with gesture recognition and body-tracking devices. The literature was screened for existing computerized paradigms assessing spatial memory and navigation. All candidate tasks were evaluated along the VR-Check dimensions by an interdisciplinary research consortium including cognitive neuroscientists, physicians, and clinical neuropsychologists. We here present the four most promising candidate tasks for our task specifications: the Starmaze (STM, navigation through a point-symmetric star-shaped labyrinth) [3], the Virtual

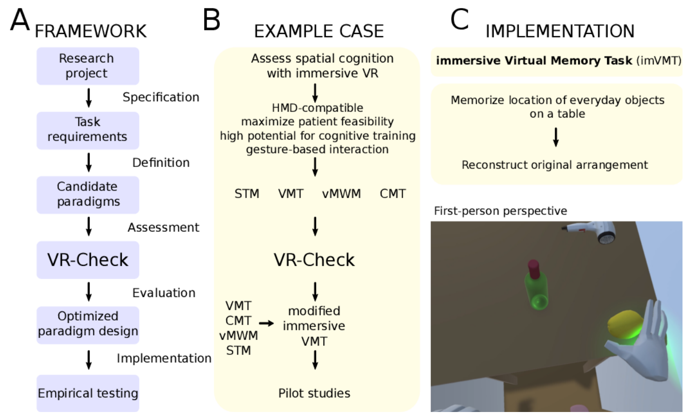

Fig. 2. Application of the VR-Check framework.

Memory Task (VMT, spatial memory task involving the location of everyday objects on a table) [4], the virtual Morris Water Maze (vMWM, subjects are required to learn the location of a concealed platform) [5], and the Cognitive Map Task (CMT, subjects construe and retrieve a cognitive map of a virtual town) [6]. All tasks exist in computerized 2D versions. While a detailed report of the comparative evaluation will be reported elsewhere, fig. 2B shows the relative paradigm ranking. The VMT was evaluated to be the most favorable basis for development as it fulfilled all task requirements and showed many positive properties along the VR-Check dimensions, e.g. high ecological relevance and low interaction complexity (no locomotion required), increasing patient feasibility. As a result, we defined a modified and immersive version of the VMT (imVMT), the implementation of which is shown in fig. 2C.

IV. Conclusion

VR-Check represents a systematic and general evaluation framework for VR applications in clinical neuropsychology, enabling researchers to optimize paradigm design in a flexible and project-specific manner.

References

[1] A. S. Rizzo and S. T. Koenig, "Is clinical virtual reality ready for primetime?," *Neuropsychology,* vol. 31, no. 8, pp. 877-899, Nov 2017.

[2] B. Birckhead *et al.*, "Recommendations for Methodology of Virtual Reality Clinical Trials in Health Care by an International Working Group: Iterative Study," *JMIR Ment Health, 6(1):e11973, Jan 2019.*

[3] K. Igloi, M. Zaoui, A. Berthoz, and L. Rondi-Reig, "Sequential egocentric strategy is acquired as early as allocentric strategy: Parallel acquisition of these two navigation strategies," *Hippocampus*, vol. 19, no. 12, pp. 1199-211, Dec 2009.

[4] S. Koenig, G. P. Crucian, A. Dünser, C. Bartneck, and J. C. Dalrymple-Alford, "Validity evaluation of a spatial memory task in virtual environments," *Int J Des Innov Res*, *6*, 1-13, 2011.

[5] Z. Nedelska *et al.*, "Spatial navigation impairment is proportional to right hippocampal volume," *Proc Natl Acad Sci U S A*, vol. 109, no. 7, pp. 2590-4, Feb 14 2012.

[6] G. Iaria, J. K. Chen, C. Guariglia, A. Ptito, and M. Petrides, "Retrosplenial and hippocampal brain regions in human navigation: complementary functional contributions to the formation and use of cognitive maps," *Eur J Neurosci*, vol. 25, no. 3, pp. 890-9, Feb 2007.