Article:

Wade, A. R. orcid.org/0000-0003-4871-2747, McCall, C. orcid.org/0000-0003-0746-8899, Karapanagiotidis, T. orcid.org/0000-0002-0813-1019 et al. (10 more authors) (2018) A neuroscientific approach to exploring fundamental questions in VR. Journal of Imaging and Science. 435. ISSN 1943-3522

https://doi.org/10.2352/ISSN.2470-1173.2018.03.ERVR-435 https://doi.org/10.2352/ISSN.2470-1173.2018.03.ERVR-435

A neuroscientific approach to exploring fundamental questions in VR

Wade, A.R.[1,2], McCall, C.[1], Karapanagiotidis, T.[1], Schofield, G.[3], Preston, C.[1], Hartley T.[1], Kaestner, M.[1], Horner, A.[1], Maloney, R.T.[1], Smallwood, J.[1], Jefferies, E.[1], Bloj, M.[4], Harris J.[5]

1: Dept Psychology University of York, UK
2: York Neuroimaging Centre, University of York, UK
3: Dept Theatre, Film and TV, University of York, UK
4: Department of Optometry and Vision Science, Bradford University, UK
5: Department of Psychology, University of St Andrews, UK

Abstract
Virtual reality presents a new set of challenges and opportunities for both engineers and neuroscientists. Here we provide an overview of a programme designed by a group of psychologists, neuroscientists and VR specialists to address some of the most outstanding issues in the field ranging from the very low-level (for example, how the brain processed motion-in-depth signals generated by stereoscopic display devices) to the very high level (how virtual environments can lead to a sense of immersion and emotional engagement).

We present data from psychophysical, electrophysiological and neuroimaging experiments and explain how different research methodologies can be applied to different problems in the field of VR/AR. We end by describing an open-source, extensible software package for studying issues in VR that can interface to common laboratory measurement equipment and discussing future directions and challenges facing the neuroscience and VR engineering communities.

Overview
At the sensory level, the failure of even the most sophisticated VR systems to simulate all aspects of a virtual world (for example, ocular accommodation and vergence mismatches for objects moving in depth or a decoupling of vestibular input from visual experience) leads to a degradation in the immersive experience which can range from mild annoyance to profound nausea.

At a more cognitive level, virtual worlds allow subjects to experience environments in which social, emotional and physical norms are violated. Examples might be games which present highly threatening or aversive stimuli, non-Newtonian motion trajectories, teleportation, scaling of size or time dimensions and changes in local spatial connectivity – for example a door from A to B may also be a door from B to C.

Here we describe recent work to build an integrated systems neuroscience environment to examine human cortical activity driven by VR stimuli. A variety of technologies to measure human brain function exist and each has its own strengths and limitations. For example, functional magnetic resonance imaging (fMRI) can interrogate activity within the brain at millimeter resolution but MR scanners cannot tolerate subject motion. In comparison, it is possible to measure bulk neuronal activity from the scalp in freely-moving subjects using electroencephalography (EEG) or functional near-infra-red spectroscopy (fNIRS) but the spatial resolution of such techniques is limited. An optimal approach combines data from multiple imaging methodologies to answer questions about how the brain processes novel stimuli in VR environments and how it adapts to deviations from normal real-world experience.

We present examples of neuroimaging and neurophysiological data from both low-level (full-color visual motion in depth) and high-level (emotional response) paradigms and we demonstrate a novel, open-source stimulus generation framework ('The Underwood Project') that allows neuroscience researchers to study key questions relating to VR (including wayfinding, emotionality, immersion and memory) in an environment that integrates at a millisecond resolution with common neuroimaging hardware including MRI scanners, MEG, EEG and TMS systems and eye trackers.

Finally, we outline what we consider to be the key challenges facing VR from the point of view of neuroscience and also the remarkable opportunities that VR affords the neuroscience community. We focus in particular on navigation, body representation and immersion – research areas that, we believe, have significant potential to benefit from VR and which have the potential to lead to clinical applications in the near future.

Low-level: visual neuroscience: 3D motion

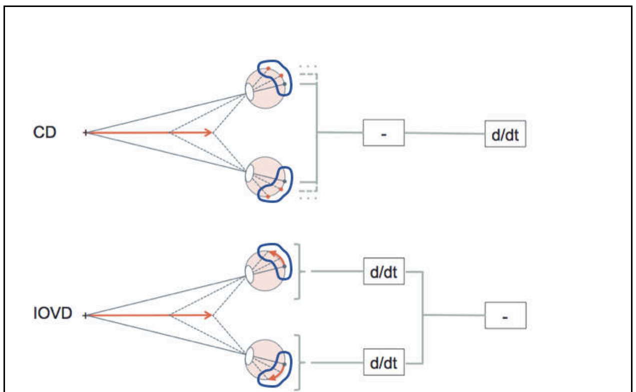

***Figure 1**: Two complimentary cues to motion in depth can be computed from the retinal signals generated by an object moving in 3D. Computing the rate of change of the disparity signal generates a 'changing disparity' cue (CD) while computing the difference between local retinal velocities provides an 'Interocular Velocity Difference' cue (IOVD). The brain appears to use both cues over different velocity ranges.*

The problem of presenting realistic 3D stimuli using a head-mounted display is not yet solved. It is well known that physical objects generate static multiple depth cues, only some of which (typically interocular disparity and occlusion) are replicated using VR goggles [1]. While technologies to incorporate other cues (most notably accommodation cues) are under investigation [2], they are not present in the current generation of commercial VR devices.

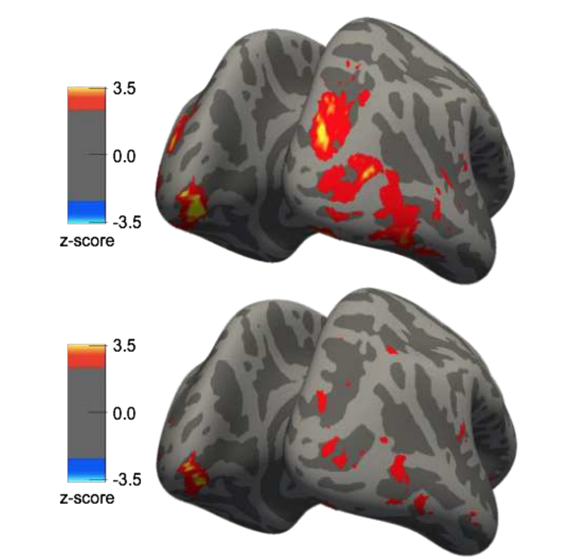

Figure 2: fMRI responses to two different motion cues. (Top) CD, (Bottom) IOVD. Shared activations over posterior (visual) cortex suggest largely-shared cortical networks. IOVD responses are significantly weaker than CD responses. Attentional modulation data (not shown) indicates an additional network of areas specific to individual cue types.

It is less well known that objects *moving* in depth can also generate two independent motion cues depending on the order in which interocular retinal positions are differenced and differentiated with respect to time. These two cues are often referred to as 'changing disparity' (CD) and 'interocular disparity differences' (IOVD) (See Figure 1) [3]–[5].

We have used a combination of psychophysics and fMRI to examine the way that these two cues to motion in depth are used in the human cortex. Psychophysically, we find that these cues provide information over complementary velocity ranges that span a wide range of potential natural stimulus configurations. Using fMRI and both stimulus and attentional manipulations, we find that while they share significant amounts of neuronal machinery, there are subtle differences in their cortical representation that indicate that two different neuronal pathways are involved at certain points (Figure 2).

This information is potentially important from an engineering perspective: the contributions of the different cues will depend on the ability of image display hardware to reproduce high and low temporal frequencies as well as fine-grained spatial detail. Specifically, the IOVD cue appears to be relatively robust to small losses of spatial resolution but requires a display device to be able to display objects moving at relatively high speeds. The CD cue has an almost inverse dependence on TF and resolution.

Moreover, while both cues contribute to a sense of motion in depth, only the CD cue is able to support a sense of moving 'form' – the percept elicited by the IOVD cue appears to alert the observer to the presence and polarity of 3D motion via a dedicate motion pathway but cannot provide information as to the identity of the moving object.

Mid-level: Wayfinding

Navigation in virtual environments (as in real environments) can be accomplished with a combination of strategies based on absolute directions and waypoints. Much of the work on the neuroscience of navigation in humans has been informed by relatively recent findings showing that mammals (including humans and rats) have cells in the hippocampal formation whose firing reflects their location within the environment (see [6] for review). These include 'place cells' that fire when the organism is in a particular location [7], 'grid-cells' that fire in a grid-like pattern that spans the environment and thus encodes the distance and direction of movements within it [8], 'border cells' that respond to nearby boundaries [9] and 'head direction cells' that fire according to the animal's heading (acting as a neural compass) [8].

Neural pathways that support memory and language presumably also contribute to more semantic aspects of navigation (for example, 'turn left at the red mailbox'). We have begun to use fMRI, psychophysics and MEG to examine the neural basis of wayfinding in humans. fMRI can be used to interrogate populations of cells in entorhinal cortex as subjects navigate through a maze or arena with well-defined landmarks that provide cues to absolute direction (Figure 2). Subjects' performance on these navigation tasks is used as a regressor in a general linear model fit of the fMRI signal timecourse to identify locations that represent location or target acquisition.

Most remarkably, the act of *imagining* trajectories through a virtual environment triggers responses in the entorhinal cortex grid cells that resemble those driven by actual navigation. fMRI can therefore be used to probe ways in which subjects prepare to navigate a virtual world as well as the ways that they perform that navigation and the cues (both semantic and directional) that they use to guide them [10].

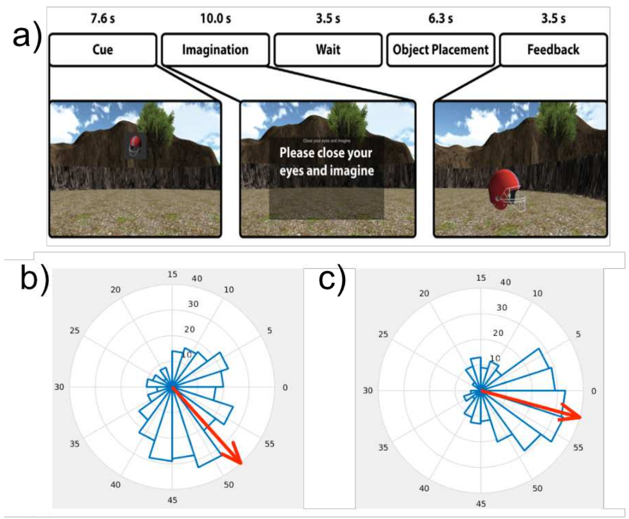

Figure 3. Imaging grid cells driven by imagined navigation. a) Subjects imagine trajectories through an arena to reach pre-learned targets. The act of imagining the trajectory stimulates 'grid' cells in entorhinal cortex that have a six-fold rotational symmetry. b) Biases in entorhinal cortex voxels plotted as a function of direction in a 6-fold symmetry pattern. Imagined navigation drives voxels that have clear direction preferences and statistically significant 6-fold symmetry indicating that they are reflecting the activity of grid cell populations. c) These activation patterns are far more tightly tuned when stimuli are presented using stereoscopic 3D. Panel (a) adapted with permission from [10].

From an engineering perspective, an important observation is that the tuning of grid-cell responses under these navigation tasks appears to become tighter when stimuli are presented in full stereo (See Figure 3). Our preliminary data therefore indicate that stereoscopic display hardware not only generates a more complete sense of immersion but that may facilitate wayfinding (and even route planning) in a 3D environment.

High-level: Emotional responses and body image

Figure 4. An aversive stimulus presented within a VR environment. Subjects' skin conductance and aversion ratings are measured in real-time while navigating the room. Physiological measures of stress correlate with the aversiveness of the stimulus at almost zero temporal lag.

Monitoring emotional response and immersion

Virtual environments elicit powerful emotional responses. At a basic level, the immersive nature of VR makes stimuli inherently self-relevant in ways that 2D presentations are not (Fgure 4). One can turn away from the screen to avoid seeing the monster in a horror film, but turning away in VR means turning your back on a predator. Furthermore, VR experiences are not singular events but unfold over time, just as emotional experiences do in everyday life. The psychology and neuroscience of human emotion demonstrate that human affect is time-dependent (e.g. [11]). Accordingly, our approach treats emotions within virtual environments not as momentary responses to evocative stimuli, but as responses that can evolve from anticipation, to reactivity, to recovery and adaptation.

This approach to the study of human affect requires continuous measurement of user responses. Peripheral physiological measures e.g., electrocardiograms (ECG), and galvanic skin responses (GSR), allow us to monitor changes in autonomic nervous system responses [12]. Neuroimaging techniques further allow us to measure central nervous system processes. While some imaging environments impose constraints on movement (i.e., fMRI or MEG), others are more forgiving (EEG or fNIRS). To determine the subjective correlates of these physiological processes, we use novel techniques for "playing back" virtual experiences on a desktop computer while eliciting continuous self-reports of the user's unfolding experience [13]. Finally, VR allows us to examine the effects of physiology and subjective experience on actual real-world behaviours, such as freezing or tonic immobility, which are difficult to elicit and measure in traditional laboratory contexts [14]. Together these methods allow us to study affect and affect regulation in the midst of virtual experiences.

Probing subject’s sense of body identity and ownership

Subjects in VR environments usually have an associated body (even if it its presence is only ever implied by the viewpoint and locomotion). How does the brain represent the virtual body and what (if any) are the limits on changing the body from the one it is accustomed to in real life?

Remarkably, neuroscientists have discovered that immersive environments created with VR and AR in combination with multisensory feedback can cause individuals to feel as though fake bodies or avatars are their own body (Maselli & Slater, 2013; Petkova & Ehrsson, 2008; Preston et al., 2015). Such virtual body illusions in combination with fMRI have enabled the identification of multisensory regions within the human posterior parietal, and premotor cortices (Brozzoli et al., 2012; Petkova et al., 2011; Preston & Ehrsson, 2016) analogous to bimodal cells recorded in non-human primates (Graziano et al., 1997; Rizzolatti et al., 1988).

Recent studies have expanded these immersive techniques to induce virtual modulation of the body appearance during fMRI, which is found to influence high-level emotional responses. Furthermore, psychophysical interaction analysis of the BOLD response during these body appearance manipulations reveals direct neural connections between multisensory (body perception) regions and networks associated with emotional processing (Preston & Ehrsson, 2016) – see Figure 5.

VR engineers can use this information both to test display hardware and to design new types of immersive devices. The

observation that multisensory integration is critical for generating a full immersive experience is important: VR systems already incorporate haptic feedback to some degree but it is possible that increased levels of embodiment can be achieved using haptic feedback devices that cover more or the body and/or by incorporating short, automated multisensory ‘body calibration’ sequences prior to beginning a VR session.

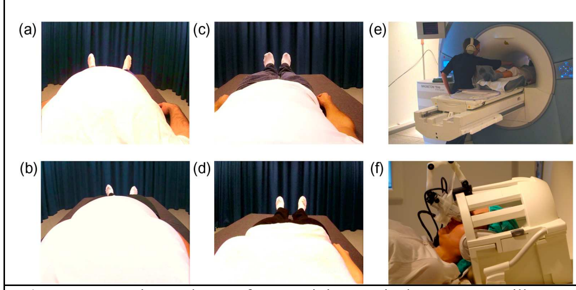

Figure 5. Experimental setup for examining cortical responses to illusory changes in body mass index. Subjects view a VR body located in the same position as their own while lying in a scanner. The studies reveal direct neural connections between regions coding body representations and emotion. Adapted with permission from [15]

Software: The Underwood Project

Our work in this domain to date has been implemented using a variety of different display devices and stimulus generation software. Although there is unlikely ever to be a ‘one size fits all’ solution to neuroscience VR stimulus generation, some common issues have arisen in a wide range of different studies. In particular, for many projects we require a highly configurable but relatively standardised dynamic environment in which stimulus events and subject actions are logged at high temporal resolution.

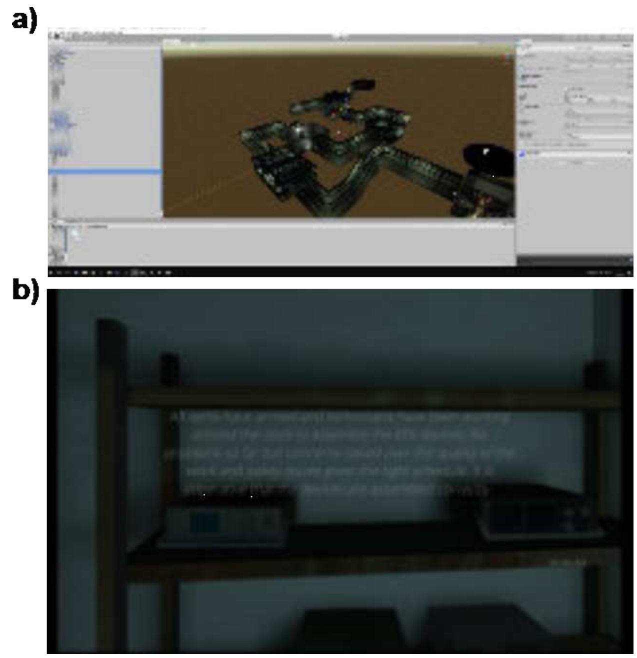

Figure 6 The Underwood Project a) Screenshot of design page in Unity. A standardised set of assets allows researchers to design dynamic environments to study navigation, emotion, immersion and memory. Additional routines provide interfaces to standard laboratory equipment and neuroimaging systems with millisecond lag and precision. b) Screenshot of a location in the game – narrative text can be presented after goals are acquired to direct future goals or provide feedback on performance.

In addition, the stimulus display system must interface at low latency with a variety of data acquisition systems - in our case specifically fMRI and MEG scanners and EEG amplifiers to provide triggers to- and receive control signals from- these device.

To address these issues, we have generated an environment (‘The Underwood Project’) based on the Unity 3D game development platform (Unity Technologies, SF). We maintain a standard drag-and-drop interface for stimulus / environment layout (see Figure 6) but provide relatively complex event handling and timing code that can generate timestamped log files at millisecond resolution and communicate with both serial and parallel ports on the host machines (with the option to add ethernet connections to remote servers). Both semantic and directional navigation clues can be embedded within the environment, subject performance can be monitored dynamically and it is trivial to implement both ‘open world’ and ‘closed world’ (for example, mazes) as well as more reduced environments such as stereoscopic dot fields used for examining low-level binocular cues.

The codebase for The Underwood Project is open source and hosted on our git repository at the York Neuroimaging Center (https://www.ynic.york.ac.uk/repo/).

Conclusion

Challenges and opportunities

Virtual and augmented reality technologies clearly face significant challenges to widespread adoption. At the lowest level, many of these challenges are related to the optical properties of the display devices and the efficiency and size of associated image processing systems. These are not domains that neuroscientists can contribute to directly. However, we can, perhaps, provide some guidance on which technologies should be prioritized and which are already ‘good enough’ to replicate real-world sensory experience. In addition, neuroscientists may be able to provide guidance on addressing outstanding issues in low-level composition of visual stimuli. For example, the observation that VR sickness can be reduced by restricting the field of view during periods of gaze change [16] has clear links to the historical literature on saccadic masking [17], [18] and the general understanding that the visual system is robust to relatively large changes in the visual scene providing they are masked by full-field transients [19] may yet prove useful to the VR/AR engineering community.

Higher-level issues in VR/AR have far clearer relevance to the neuroscience community and it is in these areas that we expect some of the most fruitful dialogues to develop in the future. In particular, the ability of VR to present novel environments, locomotion mechanisms and bodies is one of its primary strengths (because these alterations are interesting, fun or industrially important) but it is also a major challenge because it requires us to understand how human brain operates (or sometimes fails to operate) in the face of sensory inputs and feedback loops that it has not experienced in its evolutionary history. The observation (central to our work on the emotional effects of body dysmorphia) that we are able to ‘bind’ virtual bodies very different to our own, especially when a small amount of proprioceptive feedback is provided [20], is a remarkable illustration of the flexibility of the human brain to cope with novel situations but there may be equally non-intuitive counterexamples lurking in the space of possible VR stimuli that it would be useful to discover. A systematic neuroscientific approach has much to offer the VR community in this respect and we expect that as the technology for presenting VR stimuli improves, ‘higher order’ issues such as these will be an important research focus.

The conversation between the VR engineering community and neuroscientists is a two-way dialogue. VR and AR are powerful enabling technologies that allow the neuroscience community to rapidly and accurately conduct experiments on human subjects under conditions that would be impossible or unethical in the real world. As one example, VR allows us to conduct high-resolution neuroimaging experiments on subjects who are navigating extended environments [10] thereby providing insights into the fundamental neural mechanisms of wayfinding. VR also allows us to measure the physiological responses of humans embedded in highly dangerous or aversive scenes which would be prohibitively expensive to simulate in the real world [12].

To summarize, we believe that collaboration between neuroscientists and VR/AR engineers has huge potential for the advancement of both fields at many levels. Our group is currently addressing issues ranging from low-level stereo image processing to high level problems relating to body ownership, wayfinding and emotional regulation and we anticipate that advances in VR technology will lead to significant breakthroughs in all these domains in the near future.

References

[1] J. P. Wann, S. Rushton, and M. Mon-Williams, ‘Natural problems for stereoscopic depth perception in virtual environments’, *Vision Res.*, vol. 35, no. 19, pp. 2731–2736, Oct. 1995.

[2] G. Kramida, ‘Resolving the Vergence-Accommodation Conflict in Head-Mounted Displays’, *IEEE Trans Vis Comput Graph*, vol. 22, no. 7, pp. 1912–1931, 2016.

[3] C. Rashbass and G. Westheimer, ‘Disjunctive eye movements’, *J Physiol*, vol. 159, pp. 339–360, 1961.

[4] B. Rokers, L. K. Cormack, and A. C. Huk, ‘Strong percepts of motion through depth without strong percepts of position in depth’, *Journal of Vision*, vol. 8, no. 4, 2008.

[5] L. T. Likova and C. W. Tyler, ‘Stereomotion processing in the human occipital cortex’, *Neuroimage*, vol. 38, no. 2, pp. 293–305, Nov. 2007.

[6] T. Hartley, C. Lever, N. Burgess, and J. O’Keefe, ‘Space in the brain: how the hippocampal formation supports spatial cognition’, *Philos. Trans. R. Soc. Lond., B, Biol. Sci.*, vol. 369, no. 1635, p. 20120510, Feb. 2014.

[7] J. O’Keefe and J. Dostrovsky, ‘The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat’, *Brain Res.*, vol. 34, no. 1, pp. 171–175, Nov. 1971.

[8] T. Hafting, M. Fyhn, S. Molden, M.-B. Moser, and E. I. Moser, ‘Microstructure of a spatial map in the entorhinal cortex’, *Nature*, vol. 436, no. 7052, pp. 801–806, Aug. 2005.

[9] T. Solstad, C. N. Boccara, E. Kropff, M.-B. Moser, and E. I. Moser, ‘Representation of geometric borders in the entorhinal cortex’, *Science*, vol. 322, no. 5909, pp. 1865–1868, Dec. 2008.

[10] A. J. Horner, J. A. Bisby, E. Zotow, D. Bush, and N. Burgess, ‘Grid-like Processing of Imagined Navigation’, *Curr. Biol.*, vol. 26, no. 6, pp. 842–847, Mar. 2016.

[11] M. Résibois *et al.*, ‘The neural basis of emotions varies over time: Different regions go with onset- and offset-bound processes underlying emotion intensity’, *Soc Cogn Affect Neurosci*, Apr. 2017.

[12] L. K. Hildebrandt, C. McCall, H. G. Engen, and T. Singer, ‘Cognitive flexibility, heart rate variability, and resilience predict fine-grained regulation of arousal during prolonged threat’, *Psychophysiology*, vol. 53, no. 6, pp. 880–890, Jun. 2016.

[13] C. McCall, L. K. Hildebrandt, B. Bornemann, and T. Singer, ‘Physiophenomenology in retrospect: Memory reliably reflects physiological arousal during a prior threatening experience’, *Conscious Cogn*, vol. 38, pp. 60–70, Dec. 2015.

[14] C. McCall, L. K. Hildebrandt, R. Hartmann, B. M. Baczkowski, and T. Singer, ‘Introducing the Wunderkammer as a tool for emotion research: Unconstrained gaze and movement patterns in three emotionally evocative virtual worlds’, *Computers in Human Behavior*, vol. 59, pp. 93–107, 2016.

[15] C. Preston and H. H. Ehrsson, ‘Illusory Obesity Triggers Body Dissatisfaction Responses in the Insula and Anterior Cingulate Cortex’, *Cereb. Cortex*, vol. 26, no. 12, pp. 4450–4460, Dec. 2016.

[16] A. S. Fernandes and S. K. Feiner, ‘Combating VR sickness through subtle dynamic field-of-view modification’, in *2016 IEEE Symposium on 3D User Interfaces (3DUI)*, 2016, pp. 201–210.
[17] B. Erdmann and R. Dodge, *Psychologische Untersuchungen über das Lesen auf experimenteller Grundlage*. Niemeyer, 1898.
[18] B. Bridgeman, D. Hendry, and L. Stark, ‘Failure to detect displacement of the visual world during saccadic eye movements’, *Vision Res.*, vol. 15, no. 6, pp. 719–722, Jun. 1975.
[19] D. J. Simons and D. T. Levin, ‘Change blindness’, *Trends Cogn. Sci. (Regul. Ed.)*, vol. 1, no. 7, pp. 261–267, Oct. 1997.
[20] V. I. Petkova and H. H. Ehrsson, ‘If I Were You: Perceptual Illusion of Body Swapping’, *PLoS ONE*, vol. 3, no. 12, Dec. 2008.

Author Biography

(Note – due to the extended author list, we present a biography only for the first two authors)

Alex Wade received his BA in Natural Sciences from the University of Cambridge (1994) and his PhD in Neuroscience from UCL (1999). After a postdoctoral fellowship at Stanford University, he was a Principle Investigator at the Smith-Kettlewell Eye Research Institute in San Francisco until 2011. He is currently a Professor at the Psychology Department of the University of York where he directs the Centre for Future Health and is Deputy Director of the York Neuroimaging Centre.

Cade McCall received an M.A.in Psychology from the New School for Social Research (2003) and a Ph.D. in Social Psychology from the University of California, Santa Barbara (2009). He did his post doctoral work at the Max Planck for Human Cognitive and Brain Sciences and has been a lecturer in the Psychology Department at the University of York since 2016.