Supplementary of
Impact of correct and simulated focus cues on perceived realism

JOSEPH MARCH, University of Cambridge, United Kingdom
ANANTHA KRISHNAN, Bangor University, United Kingdom
SIMON J. WATT, Bangor University, United Kingdom
MAREK WERNIKOWSKI, University of Cambridge, United Kingdom
HONGYUN GAO, University of Cambridge, United Kingdom
ALI ÖZGÜR YÖNTEM, University of Cambridge, United Kingdom
RAFAŁ K. MANTIUK, University of Cambridge, United Kingdom

ACM Reference Format:
Joseph March, Anantha Krishnan, Simon J. Watt, Marek Wernikowski, Hongyun Gao, Ali Özgür Yöntem, and Rafał K. Mantiuk. 2022. Supplementary of Impact of correct and simulated focus cues on perceived realism. 1, 1 (November 2022), 4 pages. https://doi.org/10.1145/8888888.7777777

1 LIGHT FIELD CAPTURE RIG

To capture light fields with correct view-dependent shading, we used a motorized camera rig, which let us capture a full 4D light field. Two linear actuator were combined to allow for movement of the camera in the horizontal and vertical directions, as shown in Figure 1. The range of movement was ≈160 mm in each direction. Each actuator consisted of a stepper motor (NEMA23 2303HS200AW 175OZ, 2.00A) and individual control board (UStepper S), all supplied by Ooznest. A custom controller was programmed using Arduino IDE and controlled from MATLAB.

The image were captured captured with a Sony a7R3 mirrorless camera. For each image, we captured two RAW exposures at the aperture f5.6 and merged them using publicly available software [1]. The camera pose was estimated using AprilTag markers located on both sides of the object. Similarly as in [Zhong et al. 2021], we used a differentiable renderer to register 3D mesh of the object with the captured images.

[1]HDRUtils software for merging HDR images https://github.com/gfxdisp/HDRutils

Authors' addresses: Joseph March, Dept. of Computer Science and Technology, University of Cambridge, United Kingdom; Anantha Krishnan, School of Human and Behavioural Sciences, Bangor University, United Kingdom; Simon J. Watt, School of Human and Behavioural Sciences, Bangor University, United Kingdom; Marek Wernikowski, Dept. of Computer Science and Technology, University of Cambridge, United Kingdom; Hongyun Gao, Dept. of Computer Science and Technology, University of Cambridge, United Kingdom; Ali Özgür Yöntem, Dept. of Computer Science and Technology, University of Cambridge, United Kingdom; Rafał K. Mantiuk, Dept. of Computer Science and Technology, University of Cambridge, United Kingdom.

XXXX-XXXX/2022/11-ART $15.00
https://doi.org/10.1145/8888888.7777777

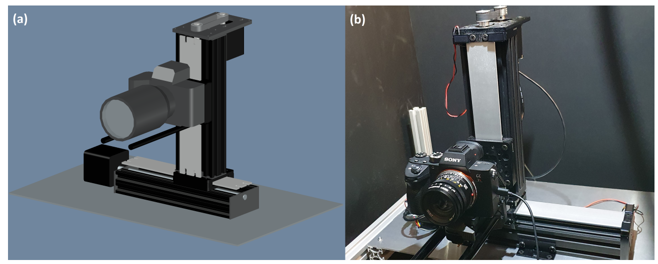

Fig. 1. CAD model of the light field capture rig (a) and its photograph (b).

2 PHOTOGRAPHS OF THE STIMULI

The examples of stimuli used in the experiment are shown in Figure 2. The images were captured with a mirrorless camera (Sony a7R3) and the aperture was set to 4 mm in diameter. This diameter should be close to the expected pupil size in the experiment. The top row was captured for camera focus on the near object (and near display) and the bottom row for the camera focus on the far object (and far display).

NearCorrect condition shows camera defocus blur on the far (top) or near (bottom) object, which should roughly correspond to to natural blur due to eye accommodating to different distances.

Stereo condition shows both objects in focus in the top row and out of focus in the bottom row. This is because only the near display plane was used in Stereo (and all Blur) conditions. Note that both objects appear sharp in the top row, which demonstrates the lack of natural defocus blur.

ChromaBlur, RerinalBlur and FakeBlur conditions were also rendered on the near plane, as the Stereo condition. However, the far (top) or near (bottom) object has defocus blur synthetically generated. Refer to the main paper for the details of the depth-of-field (DoF) rendering.

3 RESULTS FOR ALL OBSERVERS

In Figures 3 and 4 we provide plots for the data collected from each participant. Note that Participant 10 was known Deuteranope. The outlier analysis [Perez-Ortiz and Mantiuk 2017] revealed that participants 1 and 7 responses were the most different from the mean responses. We decided to include those participants as the general trend of their data did not reveal any systematic problems.

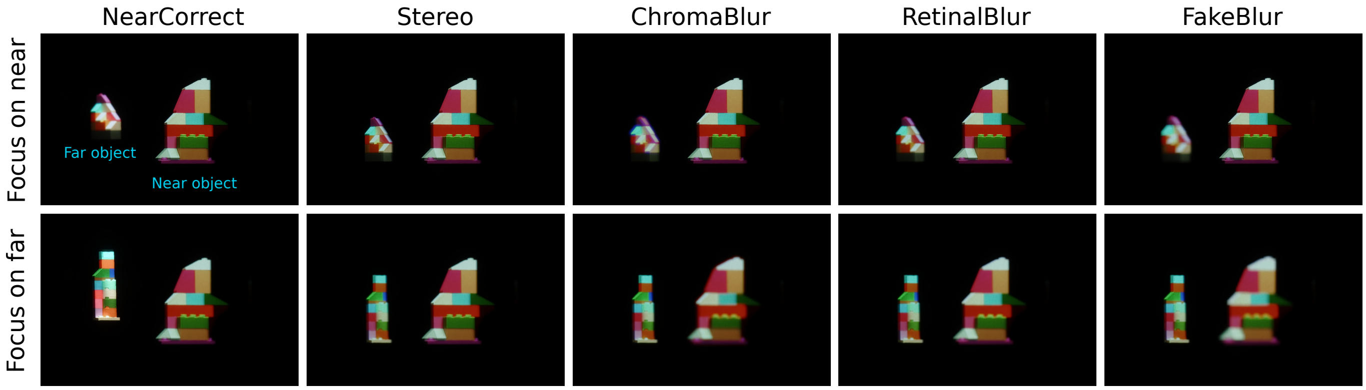

Fig. 2. Photographs of the stimuli used in the experiment. The top row shows photographs for camera focus and gaze point (for DoF rendering) set to the near object. The bottom row shows the same for the far object. See text for further details. Because we could not put the camera at the exact position of the eye, the LCD and DLP of the HDR displays were slightly misaligned, causing a small shadow on the right and a double image on the left of each object. Those artifacts were not present in the experiment.

Only a subset of participants could participate in the second (*fixed-on-near*) and third (*fixed-on-near-and-blank*) session of the experiment.

REFERENCES

Perez-Ortiz, M. and Mantiuk, R. K. (2017). A practical guide and software for analysing pairwise comparison experiments. *arXiv preprint*.

Zhong, F., Jindal, A., Yöntem, A. O., Hanji, P., Watt, S. J., and Mantiuk, R. K. (2021). Reproducing reality with a high-dynamic-range multi-focal stereo display. 40(6).

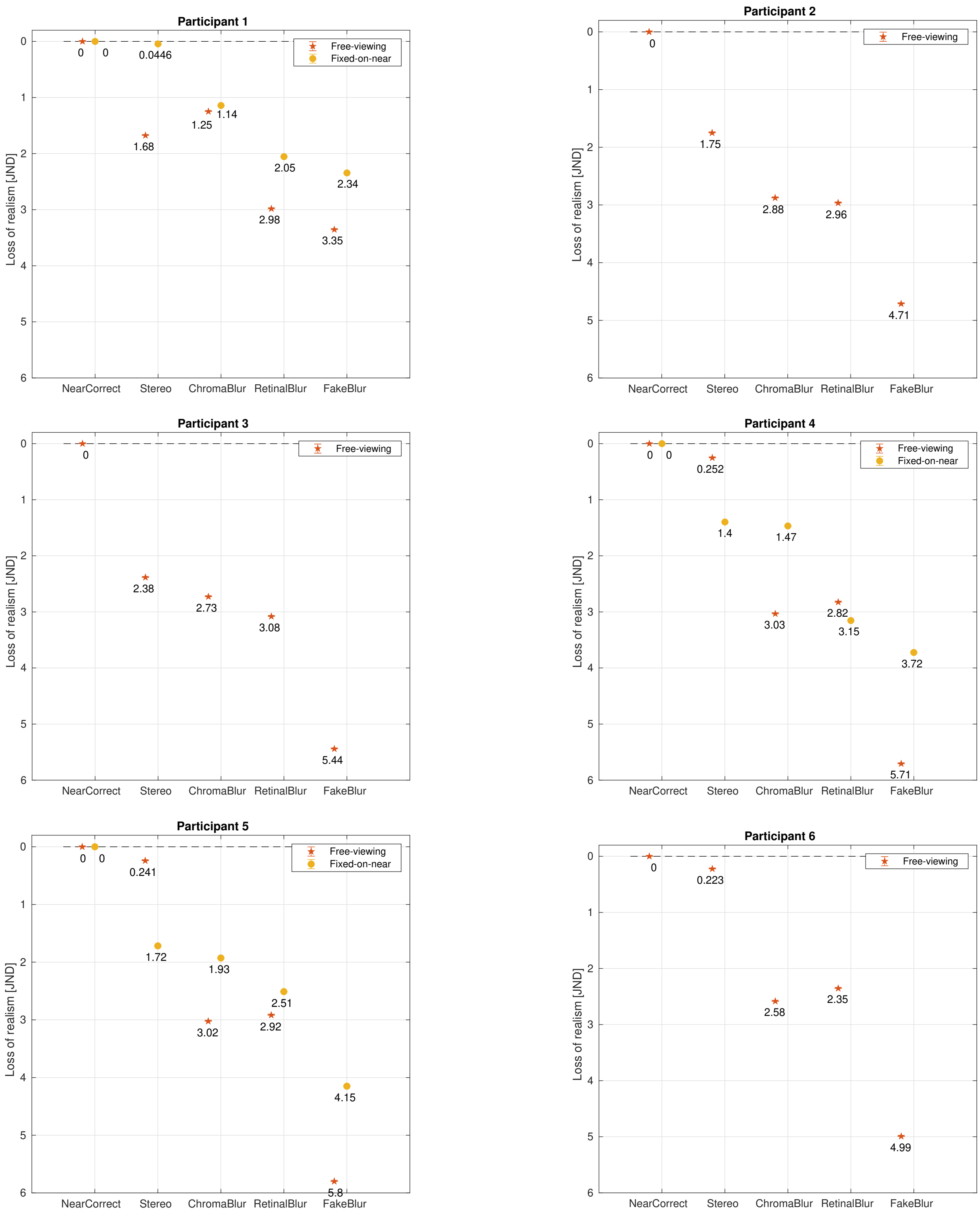

Fig. 3. Result plots for participants 1–6. Results for participants who took part in multiple sessions are all shown on a single plot.

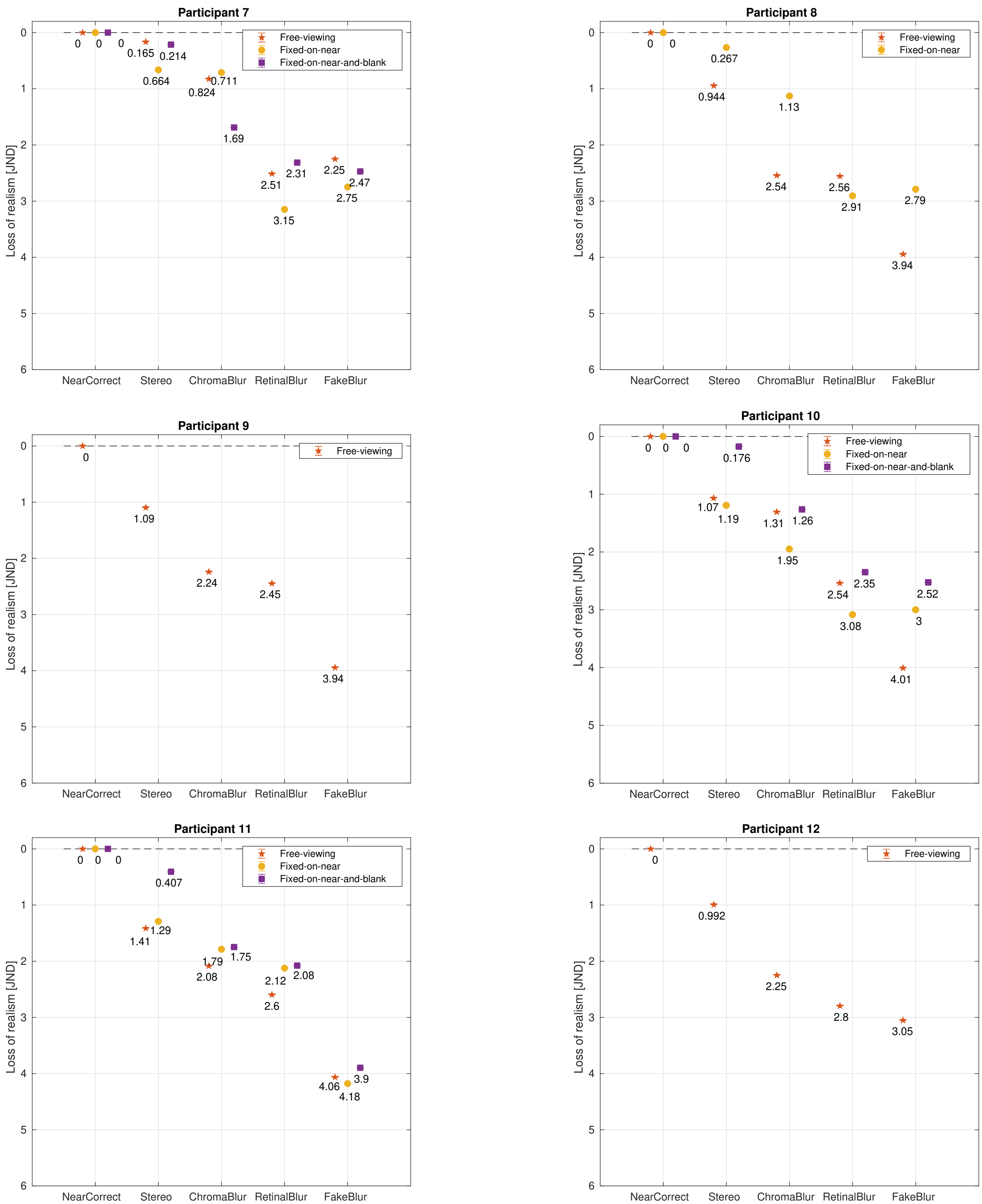

Fig. 4. Result plots for participant 7–12. Results for participants who took part in multiple sessions are all shown on a single plot.