AUDITORY STIMULI DEGRADE VISUAL PERFORMANCE IN VIRTUAL REALITY

SANDRA MALPICA, ANA SERRANO, JULIA GUERRERO-VIU, DANIEL MARTIN,
EDURNE BERNAL, DIEGO GUTIERREZ AND BELEN MASIA
Universidad de Zaragoza, I3A

SIGGRAPH 2022 VANCOUVER+ 8-11 AUG
SCIENTIFIC REPORTS
Graphics and Imaging Lab

ABSTRACT

We report an auditory effect of visual performance degradation in a virtual reality (VR) free-viewing setting for the first time. With the presentation of temporally congruent but spatially incongruent sound in the virtual world, we can degrade visual performance significantly at **detection** and **recognition**.

DETECTION: the visual target was seen or not (binary response).

RECOGNITION: after detected, the shape of the visual target was correctly identified (categorical response).

UNDERLYING CAUSES

Previous work suggests a neurological-level interplay between different modalities, which allows for crossmodal effects. This interplay, often facilitatory (specially when both modalities are synchronized) can trigger suppressive effects under certain conditions. For example, Hidaka and Ide [2015] suggest that white noise bursts via headphones can reduce performance when discriminating the orientation of Gabor patches in conventional displays. We believe that an involuntary shift of attention towards the sudden sound source may result in the degradation of visual performance due to a deactivation of the visual input [Mozolic et al. 2008].

APPLICATIONS

Suppressive effects are useful in applications where certain subtle changes have to be made without the users' awareness. For example, saccadic suppression is leveraged to rotate the virtual environment around the user in redirected walking techniques [Sun et al. 2018].

REFERENCES

[Hidaka and Ide 2015] Hidaka, S., & Ide, M. (2015). Sound can suppress visual perception. Scientific reports, 5(1), 1-9.

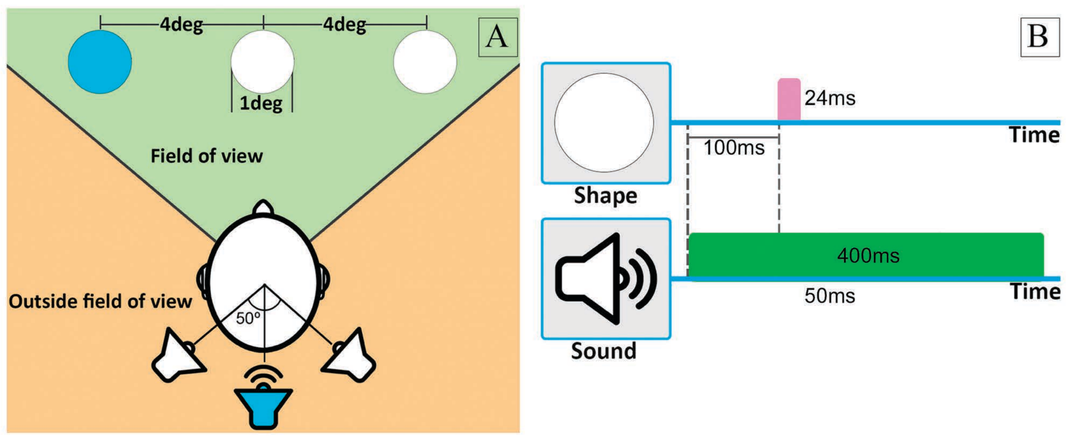
[Sun et al. 2018] Sun, Q., Patney, A., Wei, L. Y., Shapira, O., Lu, J., Asente, P., ... & Kaufman, A. (2018). Towards virtual reality infinite walking: dynamic saccadic redirection. ACM Transactions on Graphics (TOG), 37(4), 1-13.

This work was published in Scientific Reports. For more information, we refer the reader to our complete paper: Malpica, S., Serrano, A., Gutierrez, D., & Masia, B. (2020). Auditory stimuli degrade visual performance in virtual reality.

Corresponding author: Sandra Malpica (smalpica@unizar.es). This work was funded by the Europen Research Coundil (Grant no. 682080, project CHAMELEON), by the Spanish Ministry of Economy and Competitiveness (projects TIN2016-78753-P and TIN2016-79710-P) and by the Government of Aragon's Departamento de Ciencia, Universidad y Sociedad del Conocimiento through the Reference Research Group "Graphics and Imaging Lab" (ref T34_20R).

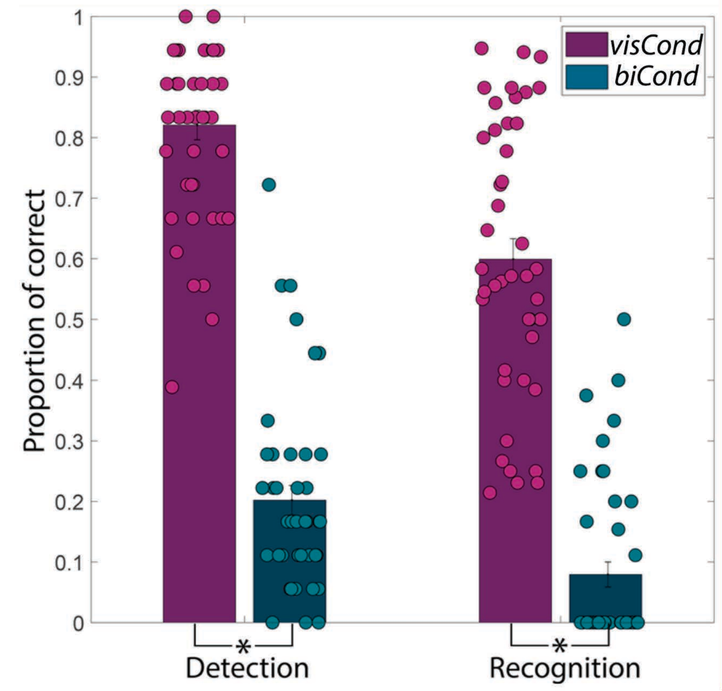
EXPERIMENTAL PROCEDURE

Top: 360º panorama of the virtual environment used in the experiments. The inset shows the five different shapes to be detected and recognized (not in scale). Participants could move freely in a physical space of 4x1.5m. Bottom: Spatial and temporal profile of the presented audiovisual stimuli. A) The visual targets appear at three possible locations inside the field of view (FoV), while auditory stimuli are located outside the FoV. B) Temporal layout. Visual targets are shown 100ms after the sound starts, for a duration of 24ms.



RESULTS

Top: Mean detection and recognition for visCond (no sound) and biCond (audiovisual) conditions. Both detection and recognition are significantly lower for the biCond condition where sound is present (marked with an asterisk). Individual performance is shown as scattered points over the bars. Note that recognition is calculated with respect to the detection percentage and not to the total number of stimuli presented (e.g., a 0% detection rate implies a 0% recognition rate). Error bars show 2xSEM.



Bottom: We discarded saccades towards the sound source as the underlying cause by means of an eye tracker. The plot depicts an example of gaze behavior during the presentation of a visual target that was not detected, corresponding to an audiovisual stimulus. The colored points represent gaze position over time. Target duration is marked with a gray column, its position with a red dotted line.