Evidence of Fear Conditioning in Virtual Reality Revealed by Eye Movements

Catherine Kim[1], Veronica Lennon[1,2], Dhwani Shah[1],
Temidayo Orederu[1,2], Daniela Schiller[1,2], Angela Radulescu[1]

[1]Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, NY, USA
[2]Department of Neuroscience, Icahn School of Medicine at Mount Sinai, New York, NY, USA
catherine.kim@mssm.edu, veronica.lennon@icahn.mssm.edu, shahdhwani0706@gmail.com,
temidayo.orederu@mssm.edu, daniela.schiller@mssm.edu, angela.radulescu@mssm.edu,

Evidence of fear conditioning in virtual reality revealed by eye movements

Abstract

Understanding how humans form associations between outcomes and environmental cues is well studied, but it is not clear how these associations are learned in real-world settings. Virtual reality (VR) provides a robust and ecologically valid platform for investigating contextual influences on learning, memory, and emotion. In this study, we use VR to study contextual fear conditioning. We conducted an integrated VR and eye-tracking study with 11 adults. To understand the effect of fear conditioning on behavior and emotional state, we analyzed gaze fixation data and self-report ratings of valence and arousal. We trained a long short-term memory (LSTM) classifier to differentiate responses to conditioned and neutral stimuli. We found that participants exhibited significantly higher gaze concentration on conditioned stimuli compared to neutral stimuli, demonstrating an enhanced emotional engagement in the fear-inducing virtual environments. Furthermore, the LSTM model not only successfully differentiated between fear and neutral conditions with high accuracy, but also revealed different patterns of fear generalization between pre- and post- fear session.

Keywords: contextual conditioning; virtual reality; eye-tracking

Introduction

From an ethological perspective, contextual conditioning serves an adaptive function, facilitating flexible responses to dynamic environments (Nairne & Pandeirada, 2008; Eichenbaum, 2017). Extensive research on contextual processes has been conducted in laboratory animals, but systematic investigations of context conditioning in humans remain relatively limited (Kroes, Dunsmoor, Mackey, McClay, & Phelps, 2017). This scarcity can be largely attributed to inherent difficulty of creating and experimentally manipulating spatio-temporal contexts for humans. Additionally, it is difficult to acquire richly annotated quantitative contextual information while simultaneously instilling a sense of immersion in laboratory experiments.

Virtual reality (VR) offers a powerful platform to address these challenges, because it provides the opportunity to render carefully manipulated environments with rich contextual stimuli. This naturalistic approach increases the likelihood of engaging the contextual learning systems that are at work in real world settings. VR is particularly effective because it creates a strong sense of "presence," such that individuals think, act, and feel as if they are in the virtual space rather than their physical surroundings. Leveraging eye-tracking and a multi-day VR task, we studied naturalistic fear conditioning by investigating participants' gaze in response to fear stimuli in various virtual contexts.

Methods

Experimental Setup

11 participants were recruited based on having (1) normal or corrected-to- normal vision without color-blindness (2) no neurological or psychiatric conditions, and (3) no prior history of epilepsy. We received consent from all participants. VR simulations were rendered using Unity game engine and presented through an HTC Vive Pro Eye head-mounted display, which includes dual AMOLED screens (resolution of 2880 x 1600 pixels, and 110° field of view), equipped with a Tobii XR eye-tracker. We designed three distinct apartment contexts (Fig. 1A), depicting a realistic bedroom, bathroom, and kitchen containing standard furniture and props.

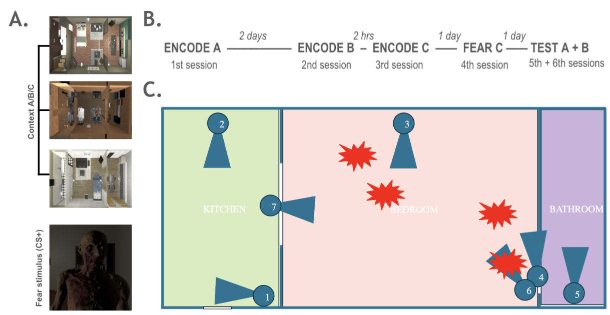

Figure 1: **A.** Three neutral contexts (i.e., apartments) and fear contextual stimulus (i.e., zombie). **B.** Timeline of the study. **C.** Floorplan of the apartment depicting 7 pre-defined camera locations. The sequence of zombie presentation, indicated by the red spots, always happens on camera location 6.

Study timeline

The study consisted of 6 sessions, where each participant initially explored three apartments that were separated by 2 days or 2 hours. Each session featured 7 pre-determined viewpoints ("cameras") each lasting one minute (Fig. 1C). Fear conditioning always took place on the 4th session (the day following the third context exploration): participants explored the third apartment, which included a 30-second sequence of a zombie approaching (CS+). The CS+ was only visibile during Camera 6. The fear conditioning sequence consists of the zombie suddenly appearing and quickly approaching the participant's location. For the context test sessions, participants returned to the designated first and second contexts, without the presence of the CS, and were again instructed to freely explore (Fig. 1B). After each VR session, participants completed three behavioral rating tasks performed on a laptop computer: an attention check, a valence-arousal scale (Russell, 1980), and an object categorization task.

Eye-tracking data analysis

In pre-processing eye-tracking data, normalized world-space gaze coordinates (x,y,z) were first converted to user-centric Cartesian coordinates by adjusting for head rotation angles. Data were further processed to compute polar coordinates (radial, azimuth) and pixel coordinates (x,y). To define fixations, a dispersion-threshold identification method was used, which assesses gaze stability by calculating the dispersion in sliding windows of gaze points across spherical coordinates.

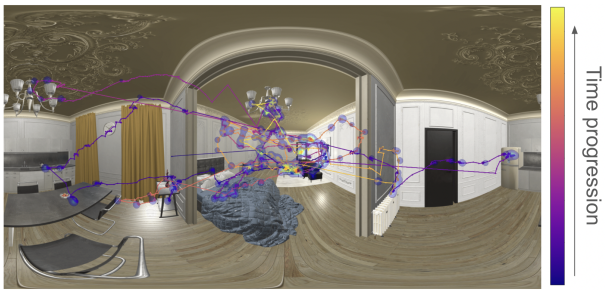

Figure 2: Example temporal trajectory of gaze fixations.

Results and Conclusions

Gaze entropy results

We used gaze entropy to quantify eye movements during and after fear conditioning to determine whether gaze became more concentrated, particularly during CS+ presentation. Gaze entropy was computed using Shannon's entropy: $H(X) = -\sum_{i=1}^{n} p(x_i)\log_2 p(x_i)$. It measures the randomness or concentration of gaze distribution across different gaze points, where higher entropy indicates more exploration and lower entropy suggests more focus. A paired t-test comparing camera location 6 and mean of other camera locations showed significantly lower entropy at location 6 (i.e., zombie appearance) compared to other locations ($t(10)$ = -4.64, $p < 0.001$), indicating concentrated gaze on the CS+, consistent with tracking of the CS+ stimulus in VR.

Valence-Arousal ratings

We assessed changes in participants' perceptions towards apartments following zombie exposure using the valence-arousal scale. A high valence rating indicates a more positive perception, whereas a high arousal rating reflects heightened alertness and excitability. Specifically, a negative mood would correspond to high arousal but low valence. We found that participants in general report high valence and moderate alertness. However, compared to other sessions, the fear conditioning (4th) session exhibited a notable decrease in valence ($t(10)$ = -11.50, $p < 0.001$) and an increase in arousal ($t(10)$ = 4.82, $p < 0.008$), suggesting that the zombie exposure shifted subjects' overall emotional response to the context (Table 1).

Long short-term classifier

We constructed an LSTM classifier (see Fig. 3A), trained on sequences of eye-tracking data (length = 1500; stride = 40)

Table 1: Mean arousal and valence scores across sessions.

Ratings	S1	S2	S3	S4	S5	S6
Valence	4.9	5.1	4.5	3.2	4.4	4.6
Arousal	3.1	3.3	4.1	5.4	4.6	4.5

S1, S2, S3 are neutral, pre-conditioning sessions. S4 is the fear conditioning session. S5 and S6 are post-conditioning.

from contexts with CS- and CS+ (sessions 3 and 4, respectively). We employed the LSTM classifier to take advantage of the sequential time-series eye-tracking data, reaching validation accuracy of 0.90.

To test for fear conditioning generalization, we assessed the percentage of categorization of pre- (sessions 1 and 2) and post- (sessions 5 and 6) fear sessions. We observed that the classifier identified sessions following zombie exposure as significantly more akin to the fear conditioning scenario within the apartment than to the same environment without the zombie presence. The difference in classification percentage (CS+ and CS-) from 0.58 to 0.79 post-fear conditioning suggests that participants' eye movements closely resembled those in the fear context, suggesting not only efficacy of fear conditioning in a VR setting, but also generalization of fear conditioning to other contexts.

In conclusion, the findings underscore VR's potential to closely mimic real-life conditions and enhance the ecological validity of fear conditioning across different contexts.

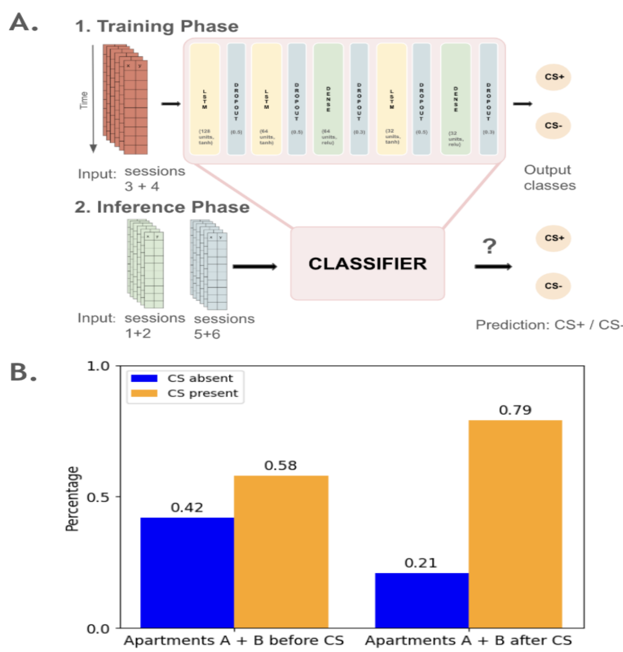

Figure 3: **A.** Schematic of LSTM classifier. The network is trained on pixel coordinates (input's shape: 1500 x 2) from session 3 (CS-) and 4 (CS+). **B.** Before CS includes session 1 and 2, and after CS includes session 3 and 4. There is an increase of percentage of fear classification from 0.58 to 0.79.

Acknowledgments

This research is supported by Icahn School of Medicine and National Institute of Mental Health (R21MH120789, R01MH122611, R01MH123069) and SFARI Award 877761. We thank members of the Radulescu and Schiller Labs at the Mount Sinai Center for Computational Psychiatry for insightful discussion and feedback.

References

Eichenbaum, H. (2017). Prefrontal–hippocampal interactions in episodic memory. *Nature Reviews Neuroscience*, *18*(9), 547–558.

Kroes, M. C., Dunsmoor, J. E., Mackey, W. E., McClay, M., & Phelps, E. A. (2017). Context conditioning in humans using commercially available immersive virtual reality. *Scientific Reports*, *7*(1), 8640.

Nairne, J. S., & Pandeirada, J. N. (2008). Adaptive memory: Is survival processing special? *Journal of memory and language*, *59*(3), 377–385.

Russell, J. A. (1980). A circumplex model of affect. *Journal of personality and social psychology*, *39*(6), 1161.