2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

# Affective response to volitional input perturbations in obstacle avoidance and target tracking games

Aashish N. Patel[1,3], Geeling Chau[1,2], Cheng Chang[2], Allan Sun[1], Jingya Huang[1],
Tzyy-Ping Jung[3], Vikash Gilja[*,1]

*Abstract*— **We present the use of two game-like tasks, Catnip and Dinorun, to explore affective responses to volitional control perturbations. We analyze behavioral and physiological measures with the self-assessment manikin (SAM), pupillometry, and electroencephalography (EEG) responses to provide intra-trial emotional state as well as inter-trial correlates with self-reported survey responses. We find that subject gameplay characteristics significantly correlate with valence and dominance scores for both games, and that perturbations to the games produce a measurable decrease in response scores for Dinorun. During perturbation events, pupillometry analysis reveals considerable SAM-agnostic dilation, with stronger responses in more rigid trialized event structures. Furthermore, analyses of neural activity from central and parietal regions demonstrate significant measurable evoked responses to perturbed events across the majority of subjects for both games. By introducing perturbations, this set of experiments and analyses inform and enable further studies of affective responses to the loss of volitional control during engaging, game-like tasks.**

## I. INTRODUCTION

Brain-computer interfaces enable direct access to neural signals that can be used to control systems that interact with and augment human capabilities. Current high-performance, non-invasive interfaces map evoked signals recorded from the brain to control signals through task and controller design. Such systems have yielded compelling demonstrations in which individuals are able to control spellers and interact with computers [1]. In addition to direct control, neural interfaces have also been used to extract neural correlates of cognitive and emotional state to evaluate responses to novel stimuli [2]. To date, both classes of systems work well under constrained, laboratory-based environments but generalization to dynamic interactions more similar to naturalistic human behavior remains a challenge. Thus, it remains unclear how well the current knowledge of human brain function translates into the highly dynamic interactions.

The use of games in task design aims to enable more complex interactions without modifying existing, physically constrained recording setups. Consideration of both the game dynamics and the neural response signals has generated a series of effective pairings: time-locked continuous control using event-related responses (ERPs) including P300 and N400 [3], visual-stimuli driven choice using steady-state visual evoked potentials (SSVEP), and for motor imagery-based control using spectral power features [4].

Evaluating goal-oriented control schemes, passive-BCIs [5] present an alternative where ERP error signals are used to evaluate task correctness. This approach releases users from encoding intermediate objectives, but relies on more rigid goal definitions to detect the neural response. Affective, or more broadly non-volitional, neural responses provide a complementary signal that may more generally measure analogues to reward. A few such studies using games (e.g. Pacman [6]) have demonstrated the utility of frustration conditions including noisy input to influence subject's valence, arousal, and dominance. Other game studies have utilized stress-inducing scenarios to elicit emotional responses [7].

Beyond neural responses, many studies have demonstrated the effectiveness of extracting physiological responses from other bio-sensors. Notably, pupillometry has rapid response times to stimuli and serves as an autonomic benchmark. Though susceptible to visual confounds in dynamic scenes including luminance, pupil size can provide measures of surprise and excitement. Gaze patterns can enable scene saliency identification, and pupil fixation can provide measures of interaction and attention [8]. Other physiological measures include heart-rate variability, galvanic skin response, and facial response; these also provide affective correlates, but generally have longer timescales or weaker responses.

Here we present two engaging, variable complexity games, Catnip and Dinorun. Catnip is a simple optimal-pursuit cursor task that presents visual stimuli perturbed to effectively modify closed loop control reliability. Dinorun is an obstacle avoidance task that increases in difficulty over time, and presents added difficulty through noisy controller input. Leveraging scalp EEG, pupillometry, subject gameplay, and self-assessment survey responses, we demonstrate changes in affective responses tied to tunable game parameters. We use these results to motivate further work utilizing game-like tasks to explore more informative affective dimensions and physiological measures sensitive to dynamic interactions.

## II. METHODS

### A. Subjects

Five healthy, college-aged subjects (Table I) participated in a dual-task game study in a purely voluntary manner, after providing informed written consents, under experimental

The authors declare the following financial interests which may be considered as potential competing interests with the work reported in this paper: VG holds shares in Neuralink Corp. and is an advisor/options holder at Paradromics, Inc.

[1]Department of Electrical and Computer Engineering, University of California, San Diego, La Jolla, CA 92092.

[2]Department of Cognitive Science, University of California, San Diego, La Jolla, CA 92092.

[3]Institute for Neural Computation, University of California, San Diego, La Jolla, CA 92092.

[*]Correspondence: vgilja@ucsd.edu

protocols approved by the Institutional Review Board of the University of California, San Diego (#140053).

| Subject | Age | Gender | Handedness | Catnip Sessions | Dinorun Sessions |
|---|---|---|---|---|---|
| 620 | 28 | Male | Right | 1 | 1 |
| 850 | 28 | Male | Right | 2 | 2 |
| 1003 | 18 | Male | Right | 2 | 2 |
| 1005 | 26 | Male | Left | 2 | 2 |
| 1008 | 22 | Male | Right | 2 | 2 |

TABLE I. Subject details

### B. Data Collection

Participants were oriented in a chair facing an experiment table providing direct visibility to a monitor and easy access to a keyboard and mouse. The experiment room was dimly lit, soundproof, and electrically isolated. Each subject was monitored by video from an observation room and directly instrumented with an EEG capture system.

The EEG data were acquired using a battery-powered Biosemi ActiveTwo system recording at a down-sampled rate of 512 Hz. A 64-channel electrode cap in standard 10-20 configuration was used with active wet Ag/AgCl electrodes. CMS and DRL channels were used as references. An optical interconnect from the amplifier to the recording system was used to minimize transmission noise.

A GazePoint GP3 eye-tracking device recording 1280 x 1024 pixels at 60 Hz was used to record pupil responses. The camera was positioned at the base of the display and oriented to provide a full view of each participant's face.

The task stimuli were presented on a 20.1-inch display screen positioned $\sim$ 70 $cm$ away from the subject. Two tasks were developed using PsychoPy and PyGame. Self-paced instructions were provided to the subjects prior to each session on the monitor. Event markers generated synchronously to stimuli presentation were sent to the EEG and camera recording systems. Data streams were timestamped and synchronized using Lab Streaming Layer (LSL).

### C. Tasks

The games in this study were designed to engage participants while presenting task stimuli. A cursor tracking game, Catnip (Fig. 1a), provides an engaging optimal-pursuit task. The subject controls a paw sprite using mouse movement and attempts to capture the catnip by clicking. As the player moves, the target attempts to maximize its distance from the cursor. After each successful capture, the target spawns at a random location. Visual perturbations in the form of user movement feedback loss are introduced at a fixed probability during specific, randomly ordered trials. Each trial lasts 30 seconds and performance feedback is provided following the game through a subject-specific high-score page.

The second game, Dinorun [9] (Fig. 1b), provides an obstacle avoidance task. The subject uses the spacebar and down arrow key on a standard keyboard to control a dinosaur sprite. As the environment scrolls across the screen, the subject must avoid different obstacles (cacti, pterodactyl). The

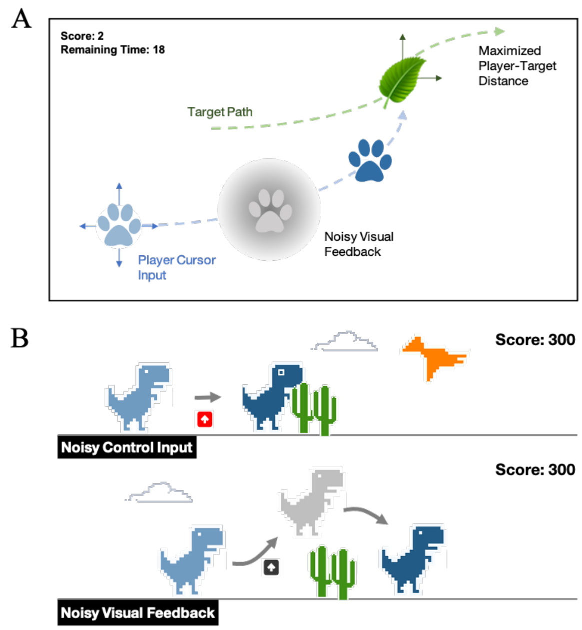


Fig. 1: a) Catnip game during a visual perturbation trial depicting sample player and target avoidance paths. b) Dinorun game during control and visual perturbation scenarios. Game backgrounds are removed for visualization.

game increases in difficulty over time through faster movement speeds. Input perturbations in the form of randomly dropped key-presses are introduced at a fixed probability throughout gameplay during specific trials. Each trial lasts approximately 60 seconds during which the subject plays through multiple games under the same conditions.

After each trial of both games, the participants answered the self-assessment manikin survey (SAM) [10] and indicated whether the game included any frustration conditions. Five randomly ordered blocks of standard and perturbed gameplay comprised a single session. Each subject completed up to two sessions for each game.

### D. Signal Conditioning

EEG data for each subject are loaded into EEGLAB toolbox [11]. A standard BioSemi 64-channel head model provided by the manufacturer is used for electrode positioning. Signals are bandpass filtered 0.5-55 Hz and `clean_rawdata` is used to remove bad channels. Independent component analysis is then applied, independent components (ICs) characterized with ICLabel [12], brain ICs projected back to the electrode space, and lastly z-scored. In the analysis, primarily the central and parietal electrodes [13] are examined against ERP signals referenced from -0.3 to 0 seconds from condition onset.

The 60 Hz Gazepoint video streams were analyzed to extract the pupil size. A semi-automated pipeline was followed: (1) identify task event time windows, (2) manually select a region of interest in image frame, (3) run the PyGaze [14] algorithm to fit an ellipse to the pupil at 30 Hz. Each frame was then manually inspected for correctness and re-annotated where necessary. Prior to analysis, subject blinking frames were removed, pupil size was normalized per session across annotated data, and linear interpolation was used to resample the pupil data to the recording frequency.

## III. RESULTS

To evaluate the elicited perceived-affect of Catnip, trends in the SAM responses across subjects are measured against game variables. Pooling the responses for all the trials across subjects, Fig 2A depicts a clear positive trend between player-target distance, valence and dominance measures. A Spearman rank-order test was used to measure the correlation between player-target distance and perceived emotion. All reported results are Benjamini-Hochberg false-discovery rate (FDR) corrected between conditions across subjects ($\alpha = 0.05$). Valence ($v$) and dominance ($d$) both had significant correlations, $r_v = 0.44$ and $r_d = 0.61$ ($p < 0.05$). Player performance measured by game score at the end of each trial similarly presented significant correlations, $r_v = 0.47$ and $r_d = 0.49$ ($p < 0.05$). Comparing the perturbation and standard gameplay (Fig 2B), subject survey responses were lower on average, indicating weak influence on affective responses during Catnip gameplay.

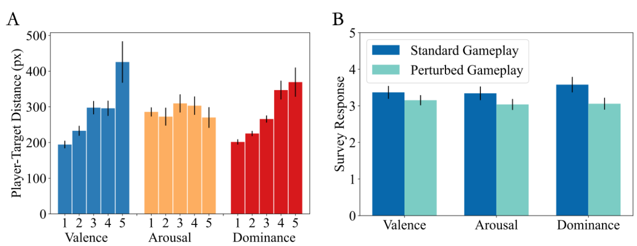


Fig. 2: Catnip task performance and SAM responses across subjects. (A) Average player-target distance, and (B) average survey response for standard and perturbed gameplay.

Evaluating the game trends during Dinorun sessions, Fig 3A depicts a similar positive trend for valence and dominance and a neutral response for arousal with respect to game score. Fig 3A can also be related to game duration due to the direct relationship with the score. The FDR corrected Spearman correlation between conditions and across subjects for valence and dominance both had significant correlations, $r_v = 0.53$ and $r_d = 0.56$ ($p < 0.05$). Examining the perturbation conditions independently from standard gameplay (Fig 3B), subject survey responses were significantly lower for valence and dominance ($p < 0.05$).

The pupillometry feature extracted for Catnip and Dinorun trials provides an accessible physiological marker of decision outcomes. The difference between the median pupil diameter of 250 ms windows, 300 ms preceding and 50 ms following perturbation onset, are used. Fig 4 shows significant dilation between each condition for both games ($p < 0.05$).

Neural responses to the Catnip perturbation events are shown in Fig 5A. Three electrodes from the central-parietal region, previously shown to be modulated by surprise [13], were selected that maximized evoked responses in hold-out trials. Standard events are neural activity with an onset 1-2

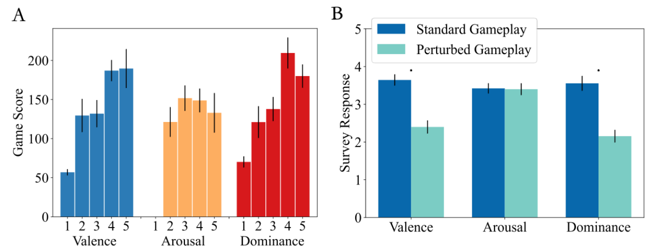


Fig. 3: Dinorun task performance and SAM responses across subjects. (A) Average game score, and (B) average survey response for standard and perturbed conditions.

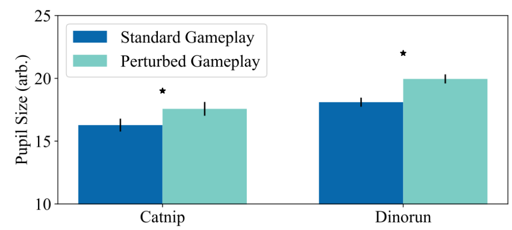


Fig. 4: Pupil responses to perturbation and standard events during Catnip and Dinorun, summarized across subjects. The difference in pupil sizes between the two games are likely due to variations in game scene illumination.

seconds prior to perturbation onset. A 300ms window prior to condition onset for each epoch was used as a baseline. To evaluate the significance, a bootstrap analysis applied to a window between +50 ms and +150 ms aligned to minimize P300 overlap per subject was used. The bootstrap was evaluated over 20,000 random samplings using the difference of the median normalized amplitude in the window between perturbed and standard conditions. FDR corrected p-values for all subjects resulted in statistically significant ($p < 0.05$) differences between the two conditions.

Dinorun perturbation event response potentials are shown in Fig 5B. Electrodes in the central and parietal regions were examined and similarly down selected to three electrodes that maximized evoked responses in hold-out trials. Standard and perturbed conditions are extracted from game-over conditions caused by collisions with obstacles. A 300ms window prior to condition onset is used as a baseline. A bootstrap analysis applied to a window from 100 ms to 200 ms post-onset was selected to maximize the condition response difference per subject. The bootstrap was evaluated over 20,000 samplings using the difference of the median normalized amplitude in the window between perturbed and standard conditions. FDR corrected p-values for all except subject 620 resulted in statistically significant ($p < 0.05$) differences between the two conditions.

## IV. DISCUSSION

Affective decoding has benefited from studies leveraging psychometric insights to neural correlates. Linking these two domains requires exploring task constructs that allow the analysis of physiological responses to intentional control, perturbations to volitional commands, and affective

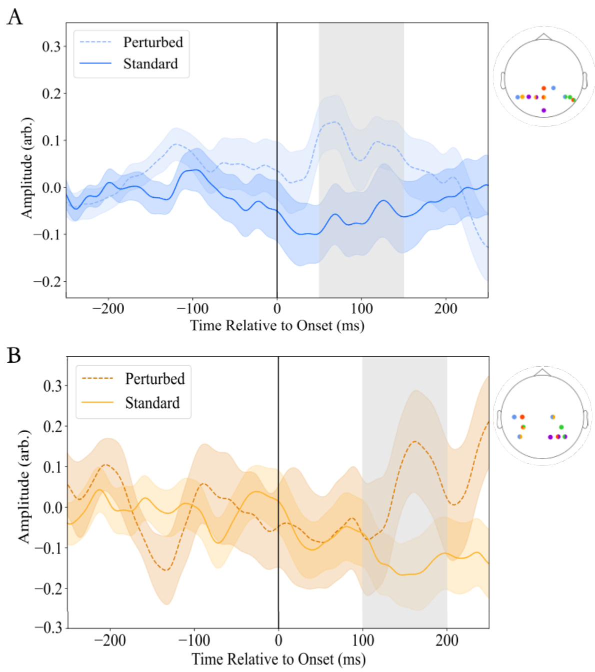


Fig. 5: Average evoked responses for a single exemplar subject between perturbed and standard event conditions for (A) Catnip and (B) Dinorun. The grey region highlights the 100ms used for bootstrap evaluation. The electrode map depicts the electrodes, individual colors per subject, evaluated. The blue and orange electrodes identify the depicted neural response channels for Catnip and Dinorun, respectively.

responses. This study proposes utilizing engaging, game-like tasks to explore their utility in eliciting such responses.

For two such games, we demonstrated that across subjects, there is a clear positive correlation between controllable game conditions including active target avoidance and game duration to reported SAM metrics of valence and dominance. This can be attributed to both satisfaction and direct influence in being able to complete the game tasks successfully. The lack of a strong correlation in Catnip or Dinorun for arousal indicates this measure is not continuously modulated by the task, but may instead be transient or measure task engagement. Moreover, the strong correlation between valence and dominance suggests an interaction between the two measures and a need for additional affective dimensions.

Physiological markers more outwardly accessible confirm the separability of events at a finer timescale than the trial surveys in both tasks. While task perturbations resulted in significant pupil responses, no significant trend was observed across subjects in relation to the SAM responses. This lack of sensitivity, however, may be due to the small evaluation window (<500ms): the weaker event onset and subject observed outcome alignment for both games might increase the time before a response is effectively measurable. Furthermore, specific task and game design considerations should be made to ensure isoluminant conditions and minimize subject vision fatigue due to active infrared emission to better ensure clean pupillometry measures.

The neural responses were also analyzed at the event scale and demonstrated a unique response to the conditions. While the evoked response in Catnip was less pronounced, there is a measurable response prior to the P300 window. Dinorun similarly demonstrated a separable evoked response, and with a stronger response across subjects indicating a more stereotyped response from the task. The difference between the two games is likely due to the stronger event alignment and feedback of Dinorun compared to Catnip.

Overall, we have presented a unique pair of games with different interaction schemes, perturbation conditions, and trial-definitions for which we have explored relationships between self-reported, behavioral, and pupillometry responses. We have also identified areas that require further analysis, specifically around the sensitivity of physiological responses to graded measures, which will further enable game-driven studies that allow the assessment of more complex affective responses and interactions.

## References


[1] Reza Abiri, Soheil Borhani, Eric W Sellers, Yang Jiang, and Xiaopeng Zhao. A comprehensive review of eeg-based brain–computer interface paradigms. *Journal of neural engineering*, 16(1):011001, 2019.

[2] Soraia M Alarcao and Manuel J Fonseca. Emotions recognition using eeg signals: A survey. *IEEE Transactions on Affective Computing*, 10(3):374–393, 2017.

[3] Jing Jin, Brendan Z Allison, Yu Zhang, Xingyu Wang, and Andrzej Cichocki. An erp-based bci using an oddball paradigm with different faces and reduced errors in critical functions. *International journal of neural systems*, 24(08):1450027, 2014.

[4] Petar Horki, Teodoro Solis-Escalante, Christa Neuper, and Gernot Müller-Putz. Combined motor imagery and ssvep based bci control of a 2 dof artificial upper limb. *Medical & biological engineering & computing*, 49(5):567–577, 2011.

[5] Thorsten O Zander and Christian Kothe. Towards passive brain–computer interfaces: applying brain–computer interface technology to human–machine systems in general. *Journal of neural engineering*, 8(2):025005, 2011.

[6] Boris Reuderink, Anton Nijholt, and Mannes Poel. Affective pacman: A frustrating game for brain-computer interface experiments. In *International conference on intelligent technologies for interactive entertainment*, pages 221–227. Springer, 2009.

[7] Bram van de Laar, Hayrettin Gürkök, Danny Plass-Oude Bos, Mannes Poel, and Anton Nijholt. Experiencing bci control in a popular computer game. *IEEE Transactions on Computational Intelligence and AI in Games*, 5(2):176–184, 2013.

[8] Wolfgang Einhäuser. The pupil as marker of cognitive processes. In *Computational and cognitive neuroscience of vision*, pages 141–169. Springer, 2017.

[9] S Shekhar. T-rex-rush. https://github.com/shivamshekhar/Chrome-T-Rex-Rush, 2020.

[10] MM Bradley and PJ Lang. Measuring emotion: the Self-Assessment Manikin and the Semantic Differential. *J Behav Ther Exp Psychiatry*, 25(1):49–59, 1994.

[11] Arnaud Delorme and Scott Makeig. Eeglab: an open source toolbox for analysis of single-trial eeg dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134(1):9–21, 2004.

[12] Luca Pion-Tonachini, Ken Kreutz-Delgado, and Scott Makeig. Iclabel: An automated electroencephalographic independent component classifier, dataset, and website. *NeuroImage*, 198:181–197, 2019.

[13] Rogier B Mars, Stefan Debener, Thomas E Gladwin, Lee M Harrison, Patrick Haggard, John C Rothwell, and Sven Bestmann. Trial-by-trial fluctuations in the event-related electroencephalogram reflect dynamic changes in the degree of surprise. *Journal of Neuroscience*, 28(47):12539–12545, 2008.

[14] ES Dalmaijer, S Mathot, and S Van der Stigchel. PyGaze: An open-source, cross-platform toolbox for minimal-effort programming of eyetracking experiments. *Behav Res*, 46(1):913–912, 2014.