Auditory Cues Enhance Recall of Emotional Transitions in the Offline Temporal Dominance of Emotions Method

Haoke Yu
Tokyo Metropolitan University
Tokyo, Japan

Shogo Okamoto
Tokyo Metropolitan University
Tokyo, Japan

***Abstract*—The Temporal Dominance of Emotions (TDE) method tracks dynamic emotional changes but requires continuous visual and manual input, limiting its use in interactive contexts. The offline TDE method addresses this by allowing retrospective reporting during video playback. To enhance emotional recall, we incorporated auditory cues from the original experience into the playback. Using a card game, Blackjack, eight participants completed both online and offline TDE sessions, with and without auditory cues. A similarity index with variable temporal resolution was used to compare the online and offline emotional profiles. The results showed that under high-resolution settings, the similarity scores for the silent condition were more dispersed, indicating reduced consistency in emotional recall. These findings suggest that auditory cues improve the temporal coherence of offline TDE, particularly when fine-grained alignment with online data is required.**

***Index Terms*—Video game, user experience, psychological assessment**

I. Introduction

The Temporal Dominance of Emotions (TDE) method [1] assesses the transient evolution of consumer emotions over time. Although TDE has been primarily applied in food sciences [2]–[4], it has also been extended to non-food-related stimuli [5]–[8]. For example, Peltier et al. [6] used TDE to assess the emotional impact of coffee advertisements. Despite nearly a decade of development, TDE remains difficult to apply in experiences that require continuous visual and manual engagement, due to its reliance on real-time button-press reporting.

Yu et al. [9] proposed an offline version of the TDE method to enable it for interactive experiences such as video games, driving, and other physically engaging activities. In this study, we investigate the role of auditory cues in offline TDE by assessing the similarity between online and offline TDE tasks across two conditions: one with auditory cues and one without. Each condition involved both online and offline TDE sessions, and the similarities between them were analyzed. This study aims to refine experimental guidelines for future applications of the offline TDE methods.

This study was in part supported by MEXT Kakenhi (24K22310).

Fig. 1. Graphical interface used in the TDE tasks. Each button corresponds to an emotional attribute.

II. Methods

A. *Online and Offline Temporal Dominance of Emotions (TDE) Methods*

A graphical user interface, as shown in Fig. 1, was used to conduct the TDE tasks. At the beginning of the game, participants initiated the evaluation by pressing the start button. They then selected the button representing the emotion that best captured their most dominant emotional state at each moment. Once an emotion was selected, the corresponding button remained active until another was selected. The system continuously recorded the timestamps of each button press throughout the task. Participants concluded the evaluation by pressing the stop button at the end of the game. The offline TDE method [9] follows the same protocol as the original (online) version but allows participants to retrospectively evaluate emotional changes by viewing a recording shortly after their experience.

B. *Gamification as Stimuli*

A computer-based version of the card game Blackjack was used as the experimental task. In this game, players aim to achieve a card total closer to 21 than the dealer, without exceeding it. The game involves strategic choices such as hitting, standing, or doubling down, which often induce dynamic emotional responses. Its simple control scheme makes

979-8-3315-2415-9/25/$31.00 ©2025 IEEE

it suitable for emotion assessment within the online TDE framework.

C. Participants

Eight university students (mean age: 23.0) participated after providing informed consent.

D. Auditory Cues

Regarding the auditory cues, adjectives corresponding to the attributes used in experiments were strictly excluded from the speech content. Participants' utterances during gameplay mainly consisted of three types: meaningless reaction voices, verbalized thoughts during decision-making, and statements about their next actions.

E. Procedures

Participants first confirmed their understanding of the ten emotional attributes investigated: "frustrated," "excited," "anxious," "tense," "joyful," "angry," "dominant," "disappointed," "confident," and "hopeful," as determined through a preliminary experiment. A brief explanation of the game rules was also provided. A practice session was conducted to ensure that participants could accurately reflect their emotional states using the TDE interface and perform the game operations smoothly.

Each participant then played five rounds of Blackjack while performing real-time emotional assessment (online TDE), lasting approximately 150 s in total. Immediately after gameplay, they conducted an offline TDE task by watching a video recording of their session and recalling the emotions experienced during the online TDE task, rather than those felt during the viewing.

One online–offline pair was considered a single set. Two sets were conducted: one with silent video playback and the other with auditory cues generated during the original experience.

F. Data Analysis

To evaluate the consistency between the online and offline TDE tasks, we computed a similarity index [10] based on time-series emotional data under each condition (with and without auditory cues). This method assesses whether the same emotional attributes were selected at approximately the same times in both TDE sessions.

The similarity between online and offline TDE results was calculated using binarized time-series emotion data. Each emotional attribute i at time t was represented as $x_i(t) \in \{0, 1\}$, where 1 indicates that the emotion was selected at that moment. The entire evaluation period was divided into R equal-length intervals. Within each interval $k \in \{1, ..., R\}$, the average dominance value of each emotion was computed as:

$$X_i[k] = R \int_{(k-1)/R}^{k/R} x_i(t)\, dt. \tag{1}$$

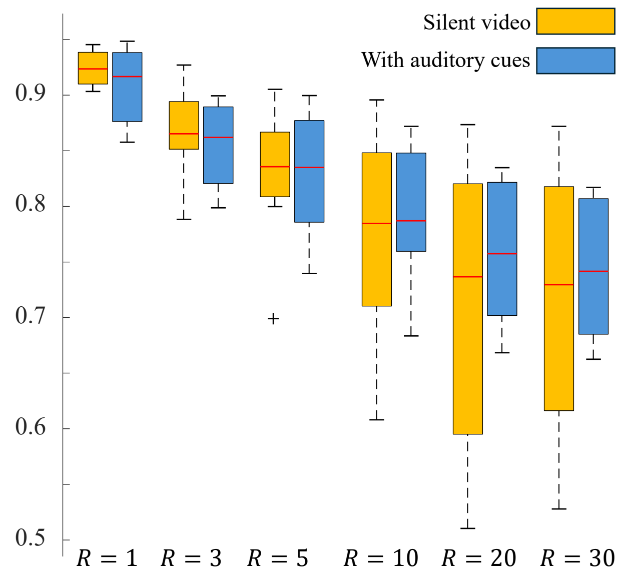

Fig. 2. Boxplot of similarity for each R.

The similarity S_R was then defined as follows:

$$S_R\left(X^{(\text{on})}, X^{(\text{off})}\right) = 1 - \frac{1}{\sqrt{2}R} \sum_{k=1}^{R} \left(\sqrt{\sum_{i=1}^{q} \left(X_i^{(\text{on})}[k] - X_i^{(\text{off})}[k] \right)^2} \right) \tag{2}$$

where $X_i^{(\text{on})}[k]$ and $X_i^{(\text{off})}[k]$ represent the dominance values for the online and offline conditions, respectively, and q is the total number of emotional attributes. The similarity index ranges from 0 to 1, with higher values indicating greater similarity between the two emotional time series.

The strictness of temporal alignment is controlled by the parameter R. The total duration of the task is divided into R equal segments, and within each segment, only the duration during which an attribute was selected is considered. Minor temporal misalignments within the same segment are therefore tolerated.

For example, when $R = 30$ and the gameplay lasts 150 s, a temporal misalignment of less than 5 s is considered acceptable. When $R = 1$, the entire task is treated as a single segment, and similarity reflects only the overall proportion of each reported emotional attribute, regardless of timing. As R increases, the analysis becomes more sensitive to the timing of emotional transitions.

This index enables flexible adjustment of temporal alignment strictness by varying R, depending on the purpose of the analysis. Researchers can choose an appropriate R value based on whether they aim to evaluate general emotional trends or precise temporal consistency.

III. Results

The similarity between online and offline TDE under the auditory and silent conditions varied between 0.908–0.816 and 0.924–0.689, respectively, across $R = \{1, ..., 30\}$.

The distribution of similarity across participants for $R = 1, 3, 5, 10, 20, 30$ is shown in Fig. 2. As observed in the figure, under the silent condition, the distribution of similarity values becomes noticeably wider when $R = 10, 20, 30$, compared to when $R = 1, 3, 5$, suggesting increased variation in temporal alignment at higher resolution levels.

To further assess whether the degree of variability differed significantly between conditions, an F-test was conducted for each value of R.

Among the R values shown in Fig. 2, significant differences in variation between the auditory and silent conditions were observed at $R = 1$ $(F(8, 8) = 0.20, p = 0.046)$ and $R = 30$ $(F(8, 8) = 5.60, p = 0.037)$. In addition, although not shown in the figure, significant differences were also detected at several other larger R values, including $R = 21, 24, 26$, and 27. Hence, with R greater than 20, the variance of similarity scores tends to be higher in the condition without auditory cues. No significant differences were found at smaller R values such as $R = 3, 5, 10$, and 20.

IV. Discussion

Based on the distribution trend of similarity across different R values, the effectiveness of auditory cues was confirmed for large R values. As R increased, the criterion for evaluating timing misalignment become stricter, and the condition with auditory cues exhibited a smaller deviation between the online and offline conditions. This indicates that the timing alignment in evaluations was better achieved with the support of auditory cues.

At $R = 1$, the variation in similarity was significantly smaller under the silent condition than under the auditory condition. However, the similarity values in both conditions remained within a consistently high range (0.891–0.948 with auditory cues and 0.902–0.945 without), suggesting that the observed difference may stem from limited sample size and individual differences across participants.

This observation highlights an important methodological point: as the value of R increases, the interpretation of the similarity index fundamentally changes. In this study, the average task duration was 119.64 s. When $R > 20$, we observed a noticeable broadening of the similarity distribution under the silent condition. This suggests that, without auditory cues, more participants struggled to recall the timing of emotional transitions with sufficient precision to meet the 6-second temporal resolution associated with $R = 20$.

In contrast, when $R = 1$, the similarity index reflects only the total duration of selected emotional attributes across the entire task, without sensitivity to temporal alignment. As such, similarity values tend to be high but do not capture the potential benefit of auditory cues in improving the timing consistency of evaluations. These findings suggest that when assessing the impact of the auditory cues on temporal consistency using the similarity index, one should be cautious about interpreting similarity values across different R settings using a single evaluation criterion.

According to participants' feedback, the offline condition often enabled greater awareness of their emotional experiences, allowing them to report emotions that might have been overlooked during the online task. Notably, real-time reporting in the online TDE can disrupt the natural flow of gameplay, despite efforts to select the type of games that elicits emotional fluctuations while permitting simultaneous gameplay and emotion reporting.

This observation raises an important point: online and offline TDE results may complement each other rather than compete for superiority. The offline TDE method was originally proposed to accommodate situations where online TDE is not feasible. However, during the experimental phase, we also relied on the online results as a reference to evaluate the validity of the offline data. This reliance highlights a key limitation of the current study design—without concurrent online measurements, assessing the validation of offline reports becomes difficult.

To address this issue in future work, we plan to explore ways to optimize the emotion-reporting interface in online settings to reduce interference, as well as to apply offline TDE to content types beyond video games.

Aside from these exceptions and limitations, our results support the conclusion that auditory cues have a positive effect on offline emotional recall. In future work, we plan to explore in detail how auditory cues influence the temporal alignment of offline emotional recall, particularly with respect to their impact on specific emotions, their stability over time, and their effectiveness across different task types.

References

[1] G. Jager, P. Schlich, I. Tijssen, J. Yao, M. Visalli, C. de Graaf, and M. Stieger, "Temporal dominance of emotions: Measuring dynamics of food-related emotions during consumption," *Food Quality and Preference*, vol. 37, pp. 87–99, 2014.

[2] S. Okamoto, "Structural modeling of temporal dominance responses using covariances of contemporary changes in subjective qualities," *International Journal of Affective Enginering*, vol. 20, pp. 127–130, 2021.

[3] M. Galmarini, R. Silva Paz, D. Enciso Choquehuanca, M. C. Zamora, and B. Meszd, "Impact of music on the dynamic perception of coffee and evoked emotions evaluated by temporal dominance of sensations (TDS) and emotions (TDE)," *Food Research International*, vol. 150, 2021.

[4] T. Okada, S. Okamoto, and Y. Yamada, "Affective dynamics: Causality modeling of temporally evolving perceptual and affective responses," vol. 13, no. 2, pp. 628–639, 2022.

[5] K. Kantono, N. Hamid, D. Shepherd, Y. Lin, S. Skiredj, and B. Carr, "Emotional and electrophysiological measures correlate to flavour perception in the presence of music," *Physiology & Behavior*, vol. 204, pp. 118–129, 2019.

[6] C. Peltier, M. Visalli, and A. Thomas, "Using temporal dominance of emotions at home. impact of coffee advertisements on consumers' behavior and methodological perspectives," *Food Quality and Preference*, vol. 71, pp. 311–319, 2019.

[7] H. Nagano, N. Saito, K. Matsumori, T. Kazama, M. Konyo, and Y. Yokokohji, "On the analysis of tactile sensation based on time measurement: An experimental case study on the interaction between skin and lotion," *IEEE Transactions on Haptics*, vol. 16, no. 2, pp. 339–344, 2023.

[8] Y. Kosuge and S. Okamoto, "Emohance: Real-time emotional amplification in gaming via physiological vibratory feedback," in *IEEE International Conference on Systems, Man and Cybernetics*, 2024.

[9] H. Yu, Y. Kosuge, and S. Okamoto, "Offline temporal dominance of emotions method using recorded videos," in *IEEE 13th Global Conference on Consumer Electronics*, 2024, pp. 879–881.

[10] H. Natsume, S. Okamoto, and H. Nagano, "TDS similarity: Outlier analysis using a similarity index to compare time-series responses of temporal dominance of sensations tasks," *Foods*, vol. 12, no. 10, p. 2025, 2023.