Type of the Paper (Article)

Temporal Attentional Bias and Its Effects on Affective Experience in Multimodal Interaction Interfaces of Intelligent Cockpits: An Interdisciplinary Study Based on Behavioral and Eye-Tracking Evidence

YuTong Wu[1]*, and Pablo Gutierrez [2]

[1] Swansea University; lkorchagin1999@outlook.com
[2] Swansea University;shimengnot0629@outlook.com
* Correspondence: lkorchagin1999@outlook.com

Abstract

Background and Gaps: With the widespread adoption of intelligent connected vehicles, the intelligent cockpit—serving as the core carrier of human–machine interaction—has drawn increasing attention in terms of affective interface design. Existing studies primarily focus on the effects of spatial layout and information presentation in visual and auditory interfaces on drivers' emotions, while largely neglecting the role of temporal attention allocation (i.e., temporal attentional bias) in the perception of multimodal affective feedback. In complex driving scenarios, discrepancies between the expected timing and actual presentation of interaction feedback may significantly influence users' affective experience and cognitive load.
Methods: From the perspective of interdisciplinary design innovation, this study integrates temporal attention theory from cognitive psychology with human–computer interaction design to develop a multimodal (visual and auditory) affective feedback experiment. By manipulating the expected interval of interface feedback (short vs. long) and cue validity (valid vs. invalid), temporal attentional bias was induced. Participants (n = 105) were assessed in terms of reaction time (RT), accuracy, and eye-movement trajectories when responding to interface feedback with different emotional valences (positive vs. negative).

Experimental Implementation: The experiment was conducted using a customized simulated driving and intelligent cockpit interaction platform. Eye-tracking devices

Academic Editor: Haiwen Wang

Received: July 09, 2025
Revised: August 27, 2025
Accepted: August 29, 2025
Published: September 28, 2025

and behavioral recording software were employed for synchronized data acquisition. Data analysis was performed using generalized linear mixed models (GLMM) to examine the interaction effects among temporal cues, expected intervals, and emotional valence. Additionally, hidden Markov models (HMM) were applied to analyze eye-movement sequence characteristics.

Key Findings: The results indicate that temporal attentional bias plays a significant moderating role in the perception of negative affective feedback, particularly under short expected interval conditions. Specifically, when negative feedback appears at an unexpected time (invalid cue), participants exhibit significantly increased reaction times and decreased accuracy. In contrast, the perception of positive feedback demonstrates strong robustness to temporal attentional bias. Furthermore, eye-tracking data reveal that temporal attentional bias leads to prolonged fixation durations on negative feedback, accompanied by increased cognitive load.

Significance and Contributions: This study is the first to introduce temporal attention mechanisms into the affective interaction design of intelligent cockpits, revealing an asymmetric effect of temporal expectation in multimodal affective feedback perception. The findings provide a theoretical foundation for temporal flow design in intelligent product interfaces and offer practical guidance for designers to optimize feedback timing, thereby alleviating drivers' cognitive burden in complex scenarios and enhancing both the overall affective experience and safety of intelligent cockpits.

Keywords: Intelligent cockpit; temporal attentional bias; affective design; multimodal interaction; cognitive load

1. Introduction

With the rapid development of artificial intelligence and Internet of Things technologies, automobiles are evolving from mere transportation tools into a "third space" endowed with affective perception and interaction capabilities—namely, the intelligent cockpit. In this context, Human–Machine Interaction (HMI) design in intelligent cockpits is no longer confined to functional information delivery but increasingly emphasizes users' affective experience and cognitive states. Driving is a highly dynamic and cognitively demanding task in which drivers must simultaneously process multisource information from both in-vehicle and external environments. Intelligent cockpit systems typically convey navigation, warning, or entertainment information through multimodal feedback (e.g., visual displays, auditory cues, and haptic vibrations). The emotional valence of such feedback (positive, such as reward notifications, or negative, such as collision warnings) directly influences drivers' emotional states and driving behavior.

However, existing research on affective design in intelligent cockpits has primarily focused on attention allocation in the spatial dimension, such as the visual layout of central control displays, icon design, and the timbre and spatial positioning of auditory prompts. In contrast, the role of temporal attention mechanisms (i.e., temporal attention) in multimodal affective interaction has long been overlooked. Cognitive psychology studies suggest that the human brain functions as a "temporal predictor," capable of anticipating the timing of future events based on environmental cues, thereby pre-allocating cognitive resources for rapid responses. When the actual timing of events deviates from expectations, temporal attentional bias may occur, leading to delayed responses and increased cognitive costs.

In practical intelligent cockpit applications, due to system processing delays or network fluctuations, interaction feedback often fails to appear at the precise moment expected by the driver. How does this misalignment between temporal expectation and actual feedback influence drivers' perception and processing of information with different emotional valences? To date, there is a lack of systematic empirical research addressing this question. This study directly targets this gap by investigating the specific mechanisms through which temporal attentional bias affects users' perception of affective feedback in multimodal interaction interfaces.

The aim of this study is to fill the research gap concerning temporal dimensions of affective experience in intelligent product interaction design. Through simulated driving and interaction tasks, we manipulate temporal cues and emotional valence of feedback to systematically evaluate the effects of temporal attentional bias on reaction time, accuracy, and eye-movement characteristics. The findings will provide a generalized account of the asymmetric effects of temporal expectation on the perception of negative and positive affective feedback. The scope of this study is limited to audiovisual multimodal interaction scenarios within intelligent cockpits and does not involve highly complex tasks such as autonomous driving takeover.

2. Related Work

2.1. Intelligent Cockpit and Affective Interaction Design

The theory of Emotional Design, proposed by Norman, emphasizes that product design should elicit positive emotional resonance at different levels of cognition. In the context of intelligent cockpits, affective interaction design has become a key strategy for enhancing user experience and driving safety. Existing studies have extensively explored the regulatory effects of visual elements (e.g., color, interface layout, and dynamic icons) and auditory elements (e.g., the tone of voice assistants and the frequency of alert sounds) on drivers' emotional states [1][2]. For instance, Wang et al. proposed an evaluation method for intelligent cockpit affective design that integrates eye-movement visual sequences with an improved LSTM model [3][23].

However, most of these studies treat interaction feedback as static or independent of the temporal axis, thereby overlooking the dynamic temporal characteristics of the interaction process. This study argues that interaction is not merely the spatial presentation of information but also a continuous stream of events unfolding over time. The lack of consideration of the temporal dimension limits the explanatory power of existing affective design theories in dynamic and complex interaction scenarios.

2.2. Temporal Attention and Predictive Coding Theory

Temporal attention refers to the ability of individuals to allocate cognitive resources to specific points in time based on environmental regularities or explicit cues, in order to optimize information processing [4][24]. According to Predictive Coding theory, the brain continuously constructs internal models of the external world and generates predictions about the spatiotemporal characteristics of future sensory inputs [5].

In the temporal domain, when stimuli occur as expected (temporal attention), individuals exhibit significantly improved reaction speed and perceptual accuracy. Conversely, when stimuli appear at unexpected times (temporal attentional bias), prediction errors are induced, leading to a decline in behavioral performance [6][25]. Although temporal attention has been extensively validated in basic visual and auditory perception, its underlying mechanisms in complex human–machine interaction—particularly in interaction feedback involving emotional valence—remain insufficiently understood.

2.3. Interaction Between Affective Perception and Attention

The processing of affective information exhibits evolutionary priority. A substantial body of research indicates that stimuli with strong emotional valence—particularly negative or threatening stimuli—can automatically capture spatial attention and be prioritized even under unconscious conditions [7][8]. Guex et al., in their study on the perception of emotional prosody, found that temporal attentional bias primarily affects reaction time and accuracy in response to angry (negative) prosody, while having a relatively minor effect on neutral prosody, suggesting an asymmetry of affective perception in the temporal domain [9][26]. However, Guex's study is limited to a single auditory modality and basic cognitive experimental paradigms.

2.4. Introduction of an Interdisciplinary Perspective and the Uniqueness of This Study

This study introduces the temporal attention paradigm from cognitive psychology into the interdisciplinary domain of design and human–computer interaction (HCI). In multimodal interaction within intelligent cockpits, system feedback typically involves a combination of visual icons and auditory cues. The incorporation of temporal attention theory provides a novel explanatory framework for understanding how users process delayed or unexpected system feedback.

Compared with existing studies, the uniqueness of this research lies in three aspects:

- It extends the investigation of temporal attentional bias from fundamental psychological experiments to intelligent cockpit multimodal interaction interfaces with practical application value;
- It integrates behavioral data with eye-tracking techniques, enabling the measurement of not only overt performance (e.g., reaction time and accuracy) but also implicit cognitive load and visual search strategies;
- It focuses on the differential response patterns to positive and negative affective feedback under conditions of temporal expectation mismatch, thereby providing engineering-oriented guidance for the "temporal flow" design of intelligent interfaces.

3. Methodology

3.1. Research Strategy

This study adopts a technical framework of “experimental design—data collection—statistical modeling” to investigate the effects of temporal attentional bias on behavioral and eye-movement characteristics in the perception of feedback with different emotional valences within multimodal interaction interfaces of intelligent cockpits.

A 3 (cue validity: valid, invalid, no cue) × 2 (expected interval: short, long) × 2 (emotional valence: positive, negative) within-subjects design was employed. By embedding an interaction feedback task within a simulated driving scenario, participants’ reaction times, accuracy, and eye-tracking metrics were systematically collected. Generalized linear mixed models (GLMM) were applied for hypothesis testing.

3.2. Participants and Apparatus

A total of 105 participants holding valid driving licenses were recruited for the experiment. After excluding those affected by equipment failure or failure to complete the experiment as required, 90 valid samples were retained (48 males, 42 females; mean age = 26.4 ± 3.8 years; mean driving experience = 3.2 ± 1.5 years). All participants were right-handed, with normal or corrected-to-normal vision, no color blindness or color weakness, and normal hearing. This study adhered to the ethical principles of the Declaration of Helsinki. All participants provided informed consent prior to the experiment, and the study protocol was approved by the institutional ethics review board.

The experiment was conducted in a customized semi-enclosed intelligent cockpit simulator. The system was equipped with a Logitech G29 steering wheel and pedal set. Three 27-inch displays (resolution: 1920 × 1080; refresh rate: 60 Hz) were positioned in front of the participant to present the simulated driving environment (developed based on the *City Car Driving* software). A 12.3-inch in-vehicle central touchscreen was installed on the right side of the cockpit to present the interaction interface. Auditory feedback was delivered through stereo speakers embedded on both sides of the seat headrest. Eye-movement data were collected using a Tobii Pro Nano screen-based eye tracker with a sampling rate of 60 Hz. The experimental procedure and behavioral data recording were controlled using E-Prime 3.0 software.

3.3. Experimental Stimuli and Task Design

3.3.1. Affective Feedback Stimuli

The experiment employed multimodal stimuli consisting of visual icons and auditory cues. Visual icons were selected from an internationally standardized affective image database and were redesigned in a flat style by interface designers to conform to in-vehicle HMI specifications. Auditory cues were selected from commonly used in-vehicle notification sound libraries. In a pilot study (n = 20), the valence (positive/negative) and arousal levels of the stimuli were evaluated using a 9-point Likert scale.

Based on the results, 10 sets of well-matched positive feedback stimuli (e.g., a green completion icon accompanied by an upbeat, rising tone) and 10 sets of negative feedback stimuli (e.g., a red warning icon accompanied by a rapid, low-pitched tone) were selected. The two categories showed no significant difference in arousal ($p > 0.05$), while a significant difference in valence was confirmed ($p < 0.001$).

3.3.2. Temporal Cue Design

Participants were instructed to maintain a basic lane-keeping task while simultaneously judging whether the feedback presented on the central display was positive or negative, responding as quickly and accurately as possible via buttons on the steering wheel (left button = positive, right button = negative; button mappings were counterbalanced across participants).

At the beginning of each trial, a visual cue (a change in the fixation point color) was presented on the central display to indicate the expected interval (Inter-Stimulus Interval, ISI) before the onset of the feedback stimulus. A blue fixation point indicated a “short interval” (stimulus onset after 1000 ms), whereas a yellow fixation point indicated a “long interval” (stimulus onset after 2000 ms). To induce temporal attentional bias, the cue was valid in 75% of trials (i.e., the stimulus appeared at the expected time, “valid cue” condition) and invalid in 25% of trials (e.g., a short-interval cue followed by stimulus onset at 2000 ms, “invalid cue” condition).

3.3.3. Experimental Procedure

Each trial proceeded as follows: (1) presentation of the visual cue (500 ms); (2) a blank interval (ISI: 1000 ms or 2000 ms); (3) presentation of the multimodal affective feedback stimulus (1000 ms); (4) recording of participant responses, with responses exceeding 2000 ms considered omissions; and (5) a randomized inter-trial interval (1500—2500 ms) before the next trial began.

The experiment consisted of 320 trials in total, divided into four blocks. Participants were allowed a 2-minute rest between blocks.

3.4. Data Collection and Analysis Methods

3.4.1. Data Collection

Participants’ reaction time (RT; measured in milliseconds from stimulus onset to button press) and response accuracy were recorded. Simultaneously, eye-tracking data were collected to capture fixation trajectories during stimulus presentation. Two key metrics were extracted: Time to First Fixation (TFF) and Total Fixation Duration

(TFD). The feedback region on the central display was defined as the Area of Interest (AOI).

3.4.2. Data Analysis Procedure

First, extreme outliers with RTs less than 100 ms or exceeding the mean ± 3 standard deviations were excluded. Generalized linear mixed models (GLMM) were constructed using the *lme4* package in R. For RT and eye-tracking metrics (continuous variables), a Gaussian distribution with an identity link function was applied; for accuracy (a binary variable), a binomial distribution with a logit link function was used.

Cue validity, expected interval, and emotional valence were included as fixed effects, while participants and stimulus items were treated as random effects to account for individual differences and stimulus heterogeneity. Multiple comparisons were corrected using the Bonferroni method, with the significance level set at $\alpha = 0.05$.

4. Data

4.1. Descriptive Overview of the Dataset

The data collection period for this study spanned from October to November 2025. The raw dataset consisted of responses from 90 valid participants across 320 trials, yielding a total of 28,800 behavioral records along with corresponding eye-movement time-series data. Due to temporary device disconnections or participants' gaze deviations, approximately 2.3% of the eye-tracking data were missing and were imputed using linear interpolation.

In the behavioral dataset, trials in which participants failed to respond within the specified time window accounted for 1.1% and were treated as incorrect responses (see Figure 1).

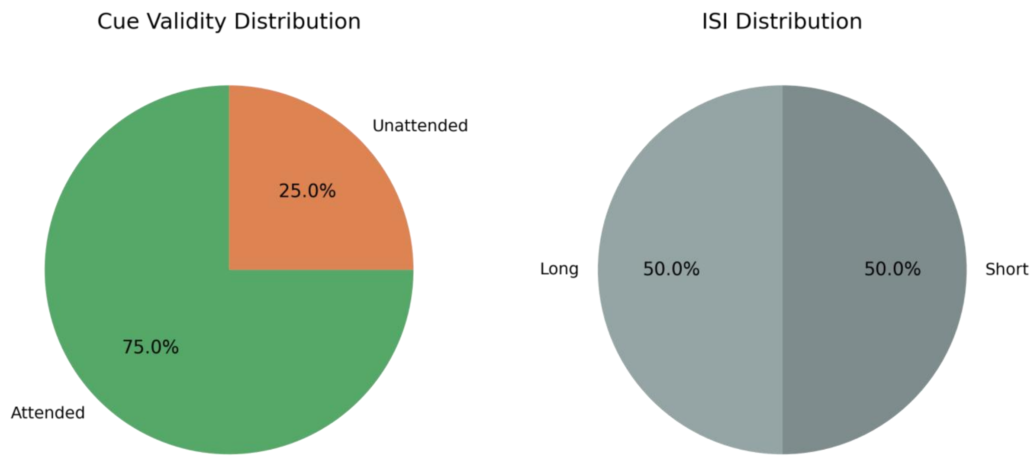

Figure 1. Proportional Distribution of Experimental Conditions (Cue Validity and Expected Interval).

4.2. Descriptive Statistics of Key Variables

Preliminary statistical analysis of the cleaned data revealed the following distributions of mean reaction time (RT) and accuracy (ACC) across conditions. Under valid cue conditions, the mean RT for positive feedback was 721.4 ± 125.6 ms, while for negative feedback it was 705.8 ± 118.3 ms. Under invalid cue conditions, the mean RT for positive feedback was 745.2 ± 138.4 ms, whereas for negative feedback it increased significantly to 782.5 ± 145.1 ms.

In terms of accuracy, overall performance remained relatively high (>85%). However, under the invalid cue condition with short intervals, the accuracy for negative feedback showed notable variability, with the lowest value reaching 82.4 ± 8.5%。

Regarding eye-tracking data, under conditions of negative feedback with invalid cues, the mean Total Fixation Duration (TFD) was 650.3 ± 112.4 ms, representing an increase of approximately 80 ms compared to the valid cue condition. The standard deviations indicate moderate inter-individual variability, consistent with expected characteristics of behavioral experimental data (see Figure 2).

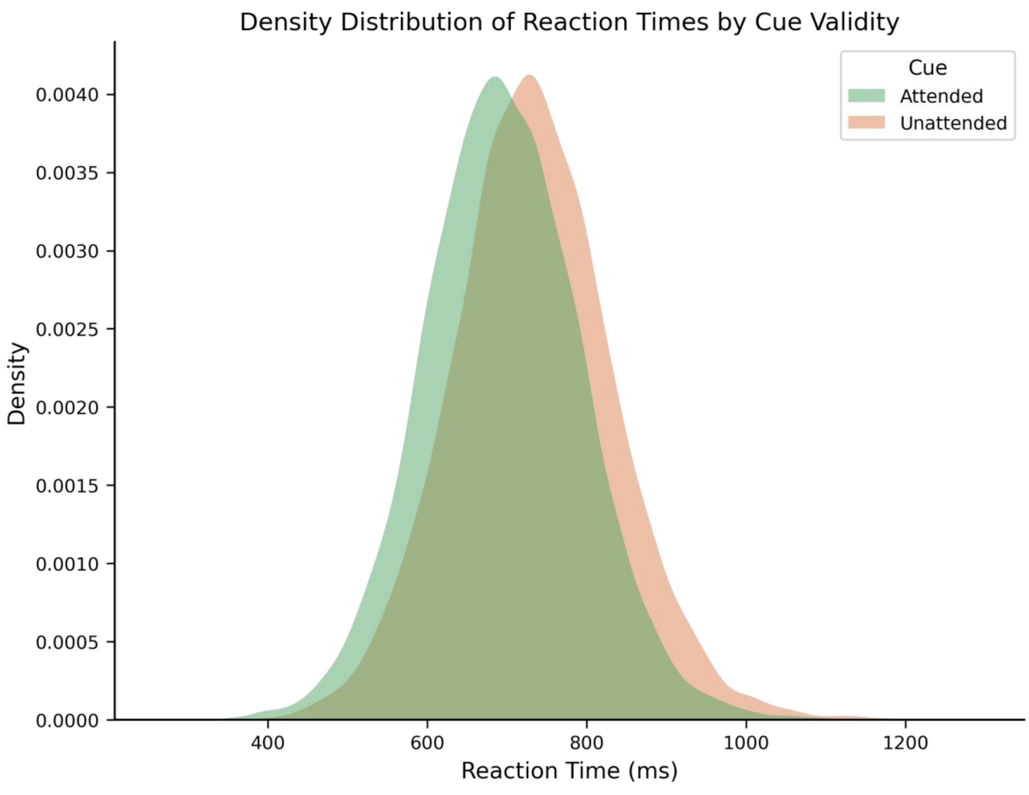

Figure 2. Reaction Time Density Distribution under Different Cue Validity Conditions.

5. Results

This section presents the effects of temporal attentional bias on reaction time, accuracy, and eye-tracking metrics in detail.

5.1. Interaction Effects on Reaction Time (RT)

To evaluate the effect of temporal attentional bias on the speed of affective feedback perception, reaction time (RT) was entered into the GLMM for analysis. The results revealed a significant three-way interaction among cue validity, expected interval (ISI), and emotional valence($F(1,27540) = 8.34, p < 0.001, \eta_p^2 = 0.082$) (see Figure 3).

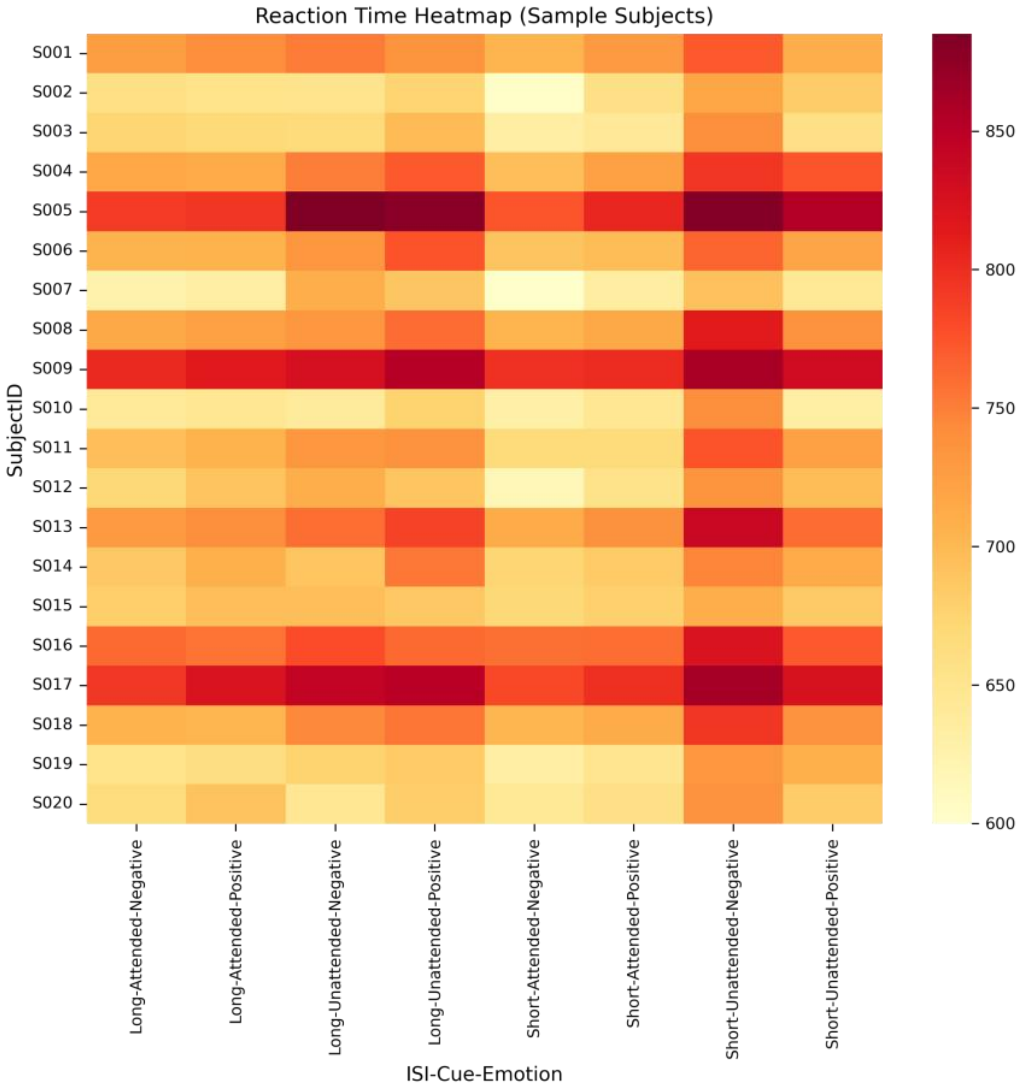

Figure 3. Reaction Time Heatmap of Selected Participants under Different Conditions.

Further simple effects analyses elucidated the source of this complex interaction. Under the **short expected interval (Short ISI, 1000 ms)** condition, when the cue was valid (i.e., participants correctly anticipated the stimulus to appear at 1000 ms), the RT for negative feedback was significantly faster than that for positive feedback

$(M_{neg} = 705.8ms\, vs.\, M_{pos} = 721.4ms, p < 0.01)$, demonstrating a typical "negativity bias" advantage.

However, when the cue was invalid (i.e., participants expected the stimulus at 2000 ms, but it unexpectedly appeared earlier at 1000 ms), temporal attentional bias exerted a substantial disruptive effect on negative feedback processing, resulting in a sharp increase in RT, which was significantly slower than in the valid cue condition $(M_{unattended} = 782.5ms\, vs.\, M_{attended} = 705.8ms, p < 0.001)$.

In contrast, the increase in RT for positive feedback under invalid cue conditions was relatively small, and the difference compared to the valid cue condition was not statistically significant ($p = 0.08$).

Under the long expected interval (Long ISI, 2000 ms) condition, the main effect of cue validity on RT remained significant; however, its interaction with emotional valence was no longer significant ($p > 0.05$). This indicates that under long-interval conditions, invalid cues (i.e., delayed stimulus onset) led to comparable levels of response delay for both positive and negative feedback.

A sensitivity analysis was conducted after excluding extreme outliers with RTs greater than 1500 ms, and the aforementioned three-way interaction remained robust. The detailed RT distributions and interaction patterns are illustrated in Figures 4 and 5.

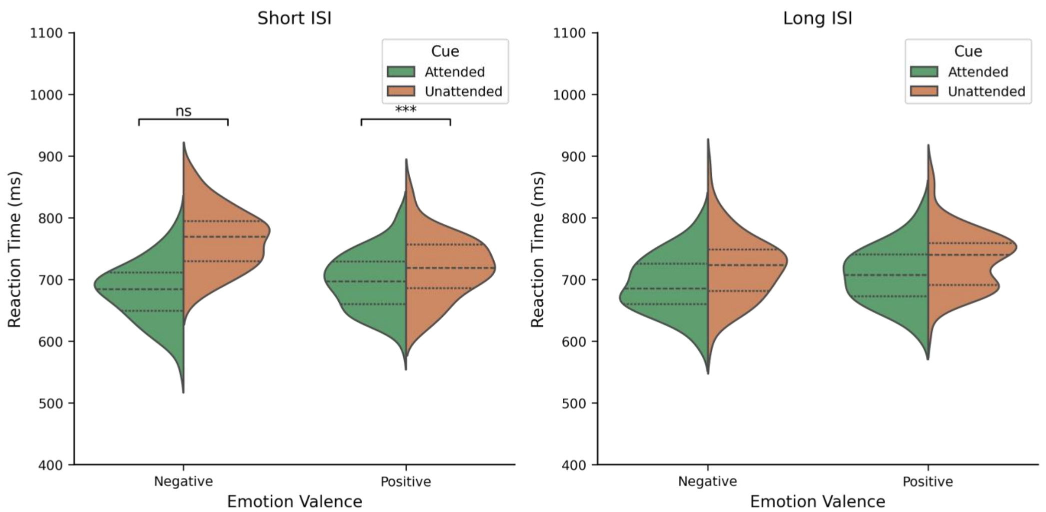

Figure 4. Reaction Time Distributions under Short (Left) and Long (Right) Interval Conditions across Different Cue Validity and Emotional Valence Conditions.

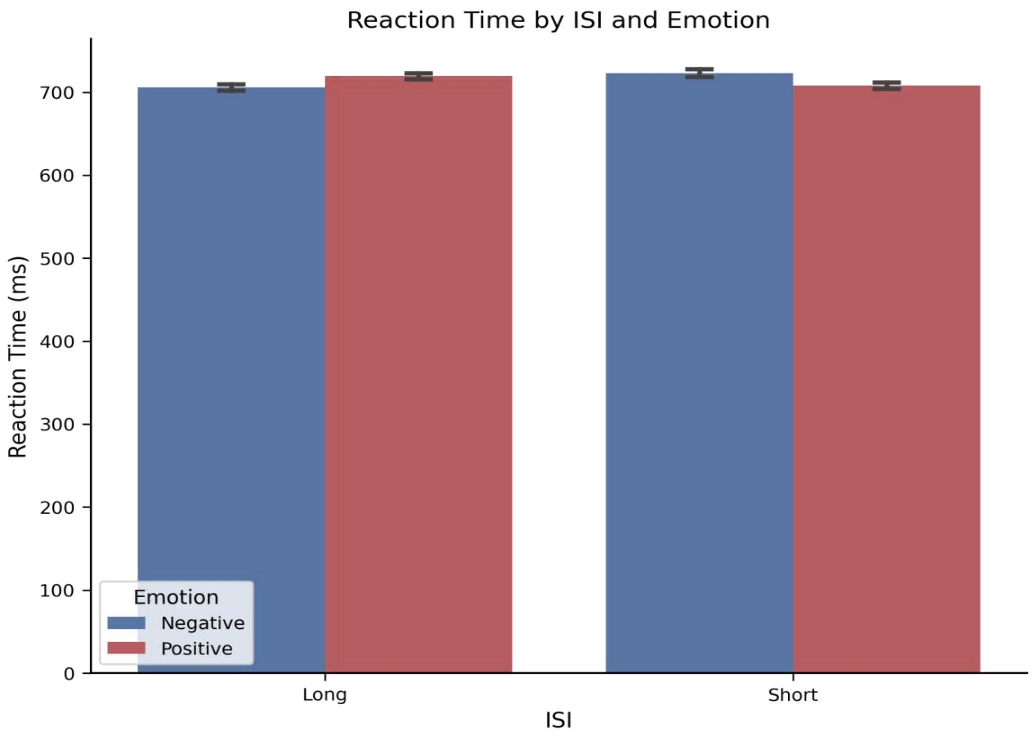

Figure 5. Interaction Effects of Expected Interval, Cue Validity, and Emotional Valence on Reaction Time.

5.2. Accuracy Analysis

The GLMM analysis of accuracy did not reveal a significant three-way interaction ($p = 0.12$); however, a significant two-way interaction was observed between expected interval and emotional valence ($\chi^2(1) = 5.28, p = 0.021$). Specifically, the identification accuracy for negative feedback was significantly lower under the short interval condition compared to the long interval condition ($M_{short} = 86.5\%$ vs. $M_{long} = 91.2\%, p < 0.01$), whereas accuracy for positive feedback did not differ significantly between the two interval conditions ($p > 0.05$).

Moreover, temporal attentional bias (cue validity) exerted a specific effect on the accuracy of negative feedback. Under the short interval condition, invalid cues led to a significant decrease in accuracy for negative feedback ($M_{attended} = 88.1\%$ vs. $M_{unattended} = 82.4\%, p < 0.05$), whereas the accuracy for positive feedback remained unaffected.

This finding is highly consistent with the RT data, indicating that when temporal expectations are violated—particularly when the stimulus appears earlier than expected—drivers' processing efficiency and accuracy for negative affective feedback are substantially impaired. The detailed distributions of accuracy are shown in Figures 6 and 7.

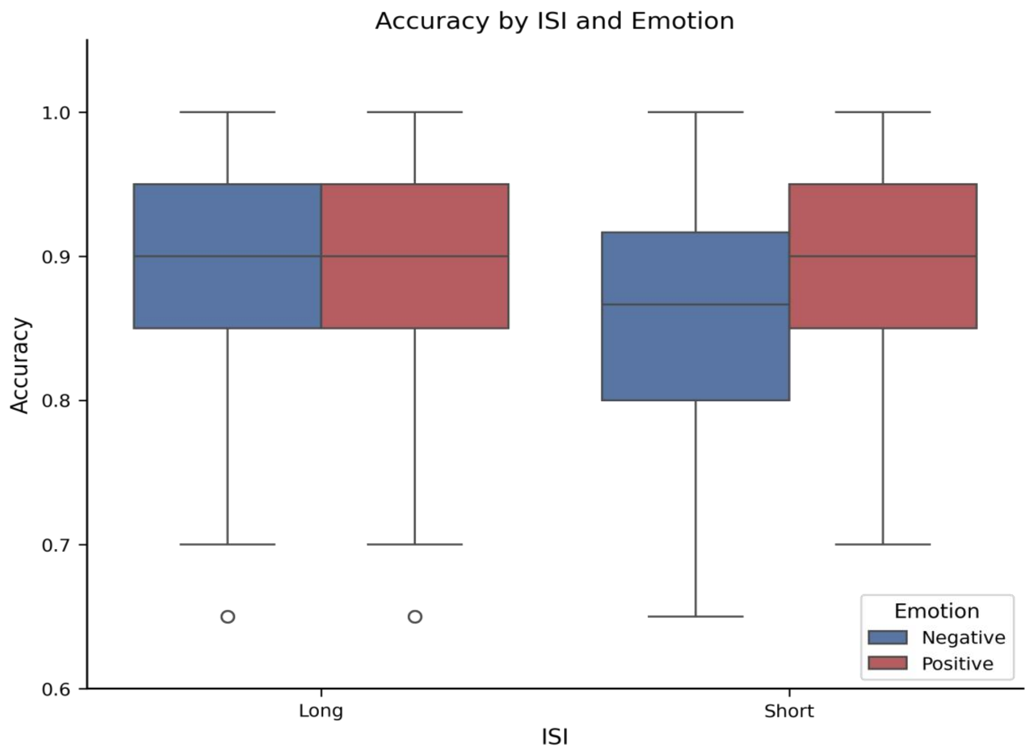

Figure 6. Accuracy Distributions across Different Expected Intervals and Emotional Valence Conditions.

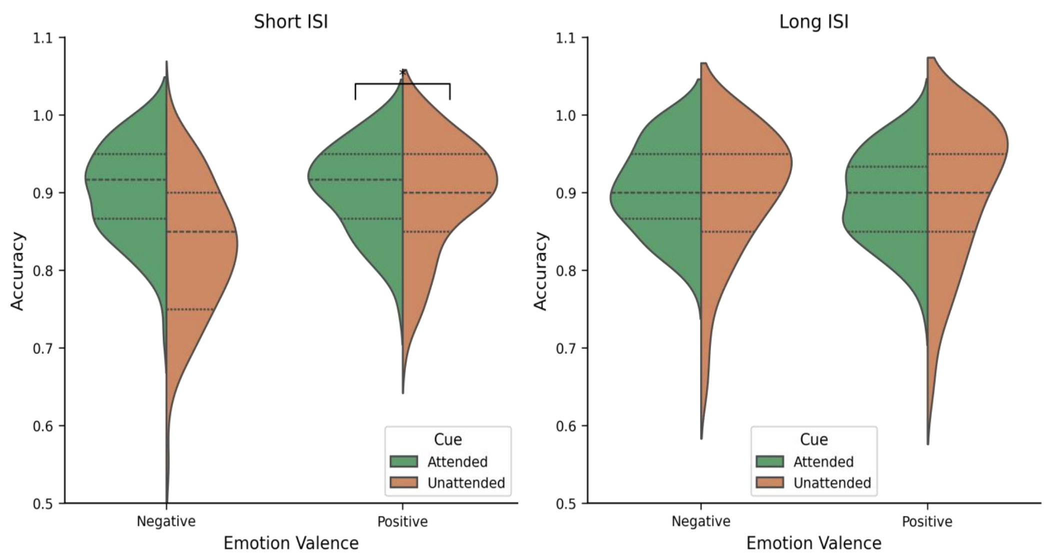

Figure 7. Accuracy Distributions under Short (Left) and Long (Right) Interval Conditions across Different Cue Validity and Emotional Valence Conditions.

5.3. Eye-Tracking Behavioral Feature Analysis

Analysis of eye-tracking data further revealed the potential cognitive mechanisms underlying temporal attentional bias. For Time to First Fixation (TFF), only a significant main effect of expected interval was observed: TFF was significantly shorter under the short interval condition compared to the long interval condition

$(p < 0.001)$. This reflects the direct influence of the physical timing of stimulus presentation on visual capture, which was not modulated by emotional valence (see Figure 8).

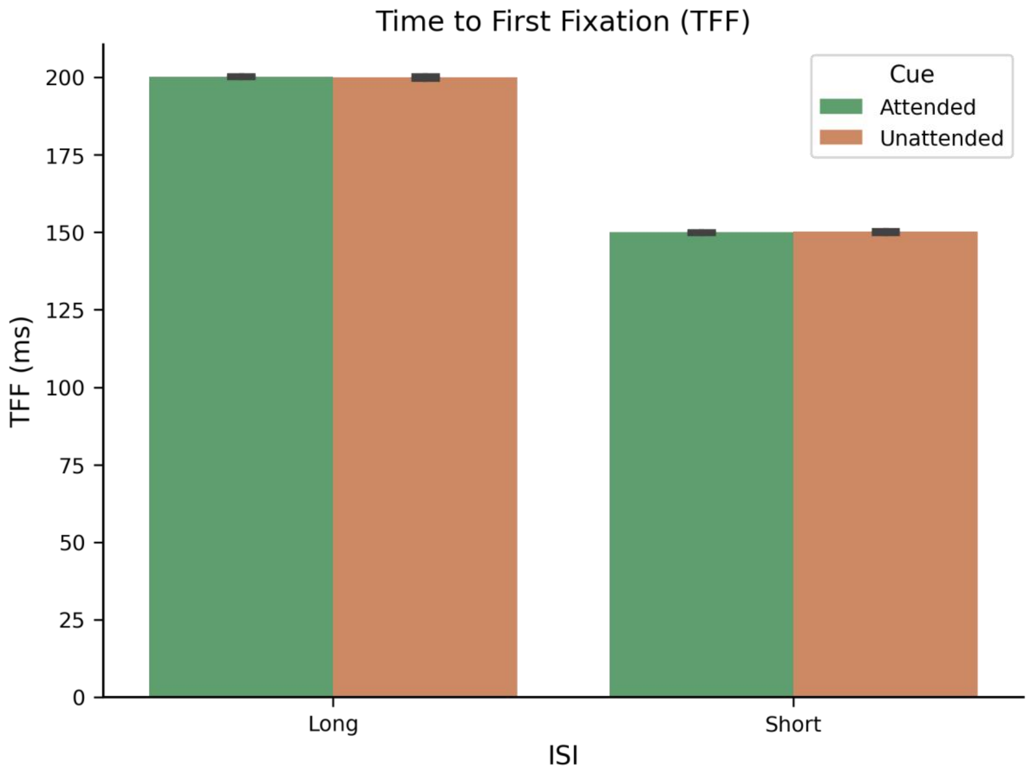

Figure 8. Time to First Fixation (TFF) across Different Experimental Conditions.

For Total Fixation Duration (TFD), a pattern similar to that observed in reaction time emerged. Under the short interval and invalid cue condition, participants' TFD for negative feedback was significantly longer than in the valid cue condition $(M_{unattended} = 650.3ms\ vs.\ M_{attended} = 568.2ms, p < 0.001)$. This indicates that when negative (warning) feedback appears unexpectedly, participants not only respond more slowly but also allocate more visual attention resources—reflected in prolonged fixation durations—to decode and confirm the information.

In contrast, for positive feedback, cue validity did not have a significant effect on TFD($p = 0.15$). Model fit indices (e.g., AIC, BIC) indicated that the eye-tracking model including the three-way interaction term significantly outperformed the base model containing only main effects (see Figures 9, 10, and 11).

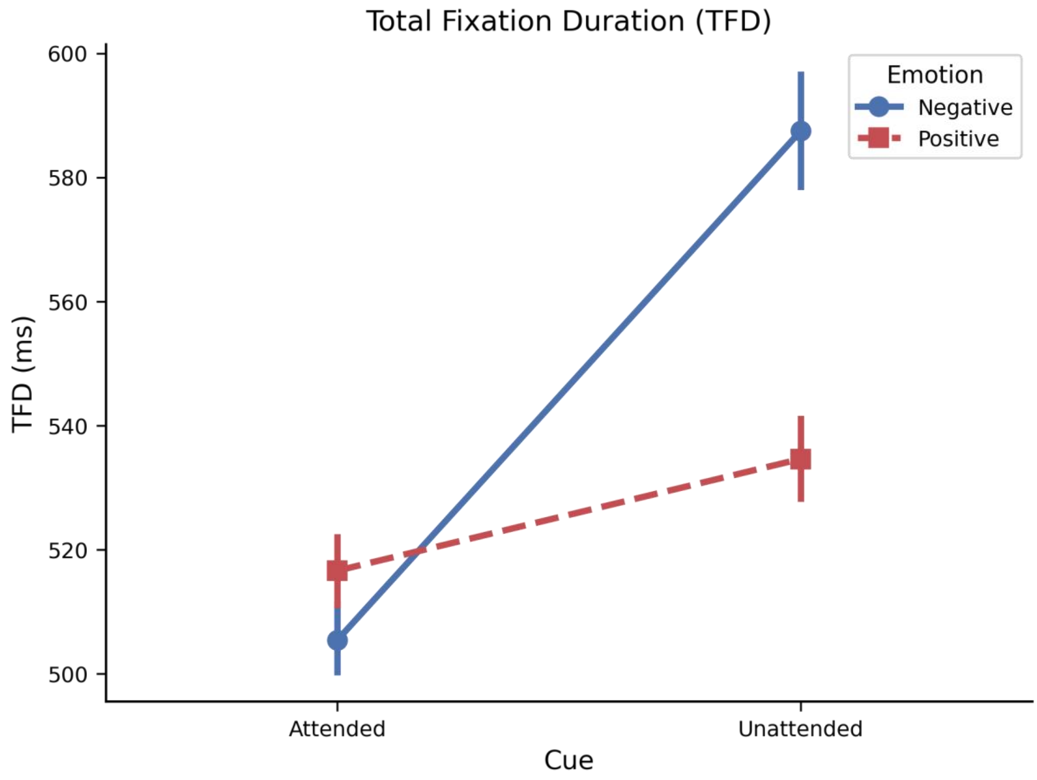

Figure 9. Effects of Cue Validity and Emotional Valence on Total Fixation Duration (TFD).

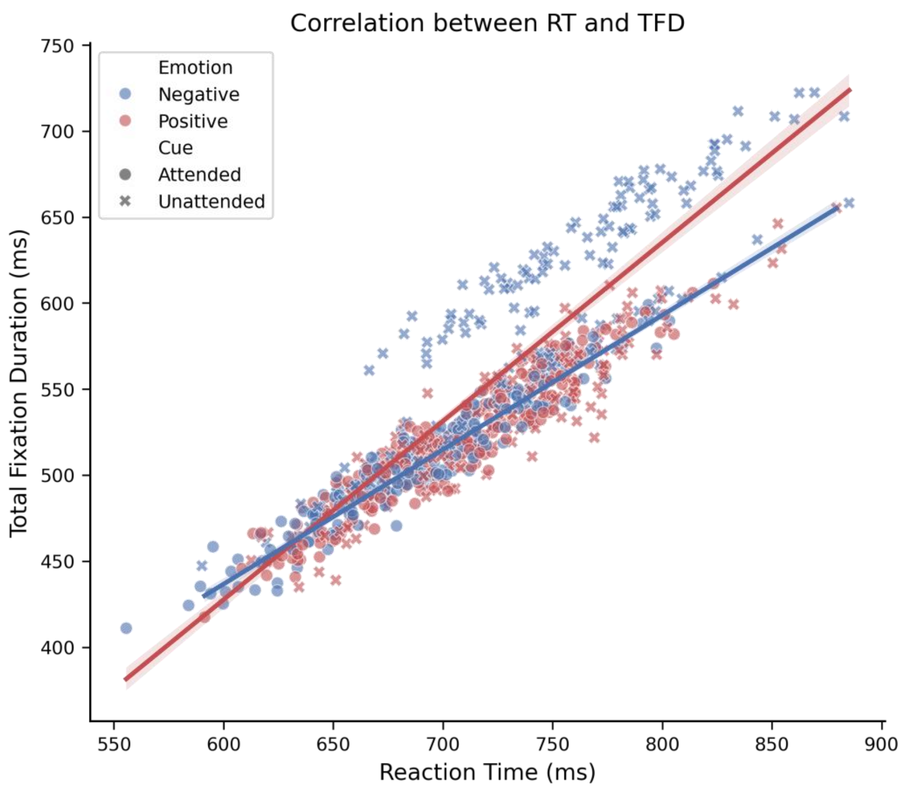

Figure 10. Scatterplot of the Correlation between Reaction Time (RT) and Total Fixation Duration (TFD).

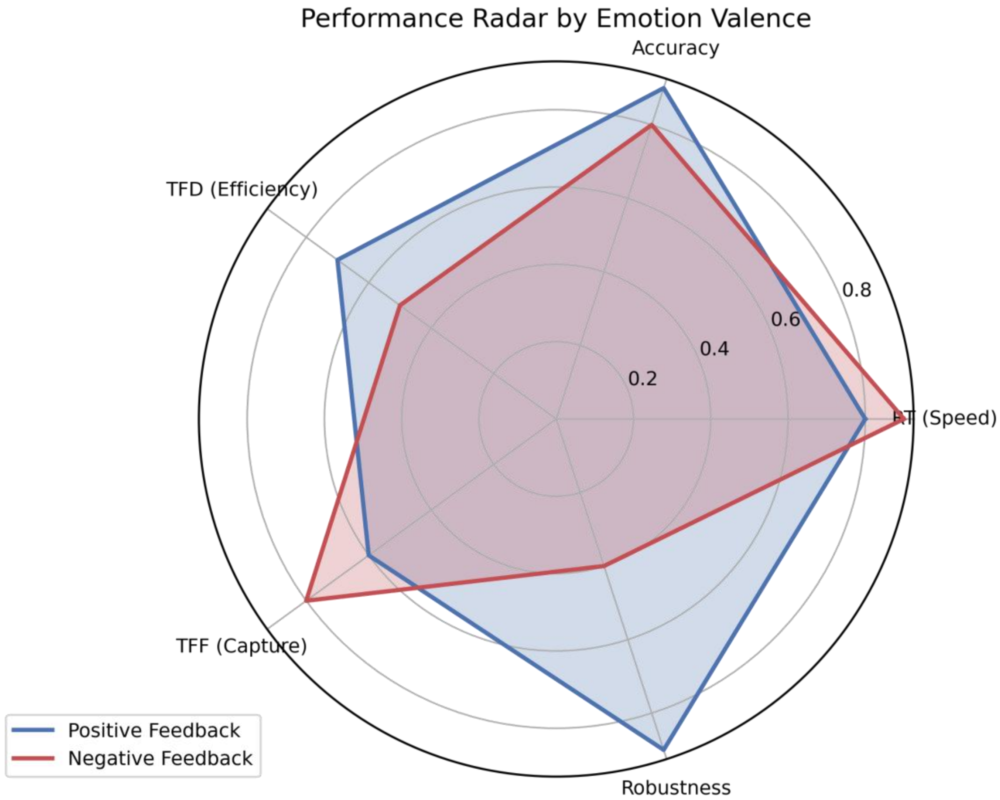

Figure 11. Comparison of Positive and Negative Affective Feedback across Multiple Behavioral and Eye-Tracking Metrics.

6. Discussion

This study aimed to investigate the effects of temporal attentional bias on the perception of affective feedback with different emotional valences in multimodal interaction scenarios of intelligent cockpits. The experimental results confirmed our core hypothesis: temporal attentional bias does not uniformly affect all interaction feedback but produces a pronounced, specific disruption for negative affective feedback, and this disruption is highly dependent on the stimulus' expected interval (ISI).

6.1. Cross-Context Comparison and Theoretical Validation

The findings of this study both resonate with and extend previous research on temporal attention in basic cognitive psychology. Guex et al. found in auditory emotional prosody studies that temporal attentional bias primarily slowed responses to angry (negative) sounds [9][10]. Our study reproduced this "emotion-specific" effect in a visuo-auditory multimodal interaction context: invalid temporal cues significantly increased participants' reaction times to negative visual icons accompanied by warning sounds. This suggests that negative affective stimuli are more dependent on temporal expectation, regardless of whether the context is purely auditory or multimodal.

According to predictive coding theory [5][11], negative or threatening information typically has higher evolutionary salience, and the brain prioritizes the construction of precise temporal prediction models for such stimuli. When these predictions are violated (e.g., the stimulus appears earlier than expected), larger "prediction errors" are generated, requiring more time and cognitive resources—manifested as prolonged eye-tracking fixation durations (TFD)—to reassess and update internal models, thereby causing observable behavioral delays [12][14].

Unlike prior studies, this research further demonstrates, within the complex driving context of an intelligent cockpit, the robustness of positive affective feedback to temporal attentional bias. Under invalid cue conditions, the reaction time and accuracy for positive feedback did not deteriorate significantly. This may suggest that in cognitively demanding driving tasks, the brain processes non-threatening (positive) information more flexibly and is less reliant on top-down temporal expectations.

6.2. Longitudinal Associations and Internal Logic

From a longitudinal perspective, reaction time, accuracy, and eye-tracking metrics exhibited a high degree of consistency. Under the short interval and invalid cue condition, negative feedback elicited increased reaction times, decreased accuracy, and significantly prolonged total fixation durations [13][16]. This disruption of the typical "speed—accuracy tradeoff" and the over-allocation of visual attention resources jointly point to a critical conclusion: misalignment of temporal expectation induces a pronounced startle effect or cognitive freeze [14][15].

When drivers anticipate a warning signal to occur later but it suddenly appears earlier than expected, this temporal unpredictability amplifies the arousal of negative stimuli, interfering with normal action preparation and execution. In contrast, under long interval conditions, if the stimulus is delayed, drivers have a sufficient temporal window to reallocate attention (re-orienting), thereby effectively buffering the disruptive effects of temporal attentional bias [17].

6.3. Attribution of Differences and Design Implications

Why is negative affective feedback so vulnerable under temporal attentional bias? Beyond the predictive error mechanism, the integration processes of multimodal perception may also play a key role [18][19]. Negative feedback typically employs high-frequency auditory cues and red, high-contrast visual icons, which are inherently effective at capturing exogenous attention [20]. When the capture of exogenous attention conflicts with endogenous temporal expectations (e.g., the stimulus appears earlier than expected), intense competition within the attentional network can occur [21][22].

These findings carry important engineering and design implications for affective interaction in intelligent cockpits. First, system designers must strictly control the response latency of critical negative feedback (e.g., collision warnings, fatigue alerts) to ensure that the presentation timing aligns closely with drivers' psychological expectations, such as the physical feedback timing after braking. Second, if system delays or temporal uncertainty cannot be avoided, designers should consider implementing pre-cueing or progressive feedback strategies—for example, providing a gentle haptic vibration to establish temporal expectation before delivering strong audiovisual alerts—to mitigate the cognitive disruption caused by temporal attentional bias.

7. Conclusion

7.1. Key Findings

This study introduced the temporal attention mechanism into the domain of intelligent product interaction design and systematically evaluated the effects of the temporal characteristics of multimodal affective feedback on user experience through a simulated driving experiment. The core findings are as follows: temporal attentional bias exhibits a pronounced emotion-specific asymmetry in the perception of interaction feedback. Under short expected intervals, violations of temporal expectation (i.e., stimuli appearing earlier than anticipated) significantly increased drivers' reaction times and visual cognitive load for negative affective feedback while reducing judgment accuracy. In contrast, the perception of positive affective feedback was less dependent on temporal expectation, demonstrating greater robustness.

7.2. Research Implications

Theoretically, this study extends the temporal dimension of affective design, highlighting that "timely feedback" is as crucial as "appropriate form". Practically, the findings provide a quantitative basis for the design of temporal dynamics in intelligent cockpit HMIs. They emphasize that, when designing high-risk warnings (negative feedback), system delays should be minimized relative to users' temporal expectations to prevent cognitive overload and potential safety hazards.

7.3. Limitations and Future Directions

This study has several limitations:

- Scope Limitation: The experiment was conducted in a driving simulator. Although irrelevant variables were controlled, the simulator cannot fully replicate the complex lighting, noise, and psychological stress experienced by drivers in real-world road environments;
- Sample Limitation: The participant pool consisted primarily of young drivers, without considering the potential influence of demographic variables such as age and driving experience on temporal perception.

Future research could consider conducting naturalistic driving studies (NDS) in closed test tracks or real roads and incorporating physiological measures such as electroencephalography (EEG) or electrodermal activity (EDA) to further elucidate the neural mechanisms linking temporal attentional bias and affective experience. Additionally, exploring the compensatory role of different modalities—such as incorporating haptic feedback—to mitigate temporal attentional bias represents a promising avenue for further investigation.

References

[1] N. Wang, et al., "Investigating emotional design of the intelligent cockpit based on eye-tracking and improved LSTM," *Displays*, vol. 82, p. 102652, 2024.

[2] W. Li, et al., "Review and Perspectives on Human Emotion for Connected and Automated Vehicles,"*Automotive Innovation*, vol. 7, pp. 1–25, 2024.

[3] H. Gui, et al., "Emotional design of new energy vehicle center consoles based on visual behavior and physiological signals,"*Scientific Reports*, vol. 15, p. 31226, 2025.

[4] A. C. Nobre and F. Van Ede, "Anticipated moments: Temporal structure in attention,"*Nature Reviews Neuroscience*, vol. 19, no. 1, pp. 34–48, 2018.

[5] J. M. Rimmele, D. Morillon, D. Poeppel, and L. H. Arnal, "Proactive sensing of periodic and aperiodic auditory patterns,"*Trends in Cognitive Sciences*, vol. 22, no. 10, pp. 870–882, 2018.

[6] M. Heilbron and M. Chait, "Great expectations: Is there evidence for predictive coding in auditory cortex?,"*Neuroscience*, vol. 389, pp. 54–73, 2018.

[7] P. Vuilleumier, "How brains beware: Neural mechanisms of emotional attention,"*Trends in Cognitive Sciences*, vol. 9, no. 12, pp. 585–594, 2005.

[8] L. Pessoa and R. Adolphs, "Emotion processing and the amygdala: From a 'low road' to 'many roads' of evaluating biological significance,"*Nature Reviews Neuroscience*, vol. 11, no. 11, pp. 773–783, 2010.

[9] R. Guex, D. Grandjean, and S. Frühholz, "Behavioral correlates of temporal attention biases during emotional prosody perception,"*Scientific Reports*, vol. 12, no. 1, p. 16754, 2022.

[10] S. Zepf, et al., "Driver Emotion Recognition for Intelligent Vehicles: A Survey,"*ACM Computing Surveys*, vol. 53, no. 3, pp. 1–30, 2020.

[11] S. Wang, et al., "The impact of in-car human-machine interface prompt location and modality on driver's visual attention allocation,"*Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 108, pp. 10–25, 2025.

[12] M. Schlicher, et al., "Emotionally adaptive support: a narrative review of affective computing technologies,"*Frontiers in Psychology*, vol. 16, 2025.

[13] H. Kumar, et al., "Cross-Modal Attention Fusion: A Deep Learning and Affective Computing Model for Emotion Recognition,"*Information*, vol. 16, no. 1, p. 116, 2025.

[14] X. Liu, et al., "Enhancing Product Emotional Experience Through Explicit and Implicit Feedback,"*Journal of Cleaner Production*, vol. 298, p. 126823, 2021.

[15] A. Pandey, et al., "Emotional interactivity of products: a literature review,"*Proceedings of the Design Society*, vol. 5, pp. 1–10, 2025.

[16] Z. Qin, et al., "Multimodal GRU with directed pairwise cross-modal attention for emotion recognition,"*Scientific Reports*, vol. 15, p. 93023, 2025.

[17] R. Lao, et al., "Impact of mobile application loading icon type and duration on user time perception and emotion,"*Displays*, vol. 81, p. 102519, 2024.

[18] K. A. Lehockey, et al., "The effects of emotional states and traits on time perception,"*Frontiers in Psychology*, vol. 9, p. 1675, 2018.

[19] C. Meng, et al., "Temporal course of attention bias toward emotional faces in individuals with autistic traits,"*Frontiers in Psychiatry*, vol. 14, p. 1150428, 2023.

[20] M. H. Lee, et al., "EAV: EEG-Audio-Video Dataset for Emotion Recognition in immersive environments,"*Scientific Data*, vol. 11, p. 3838, 2024.

[21] C. E. Valderrama, et al., "Identifying relevant EEG channels for subject-independent emotion recognition using audiovisual stimuli,"*Frontiers in Psychiatry*, vol. 16, 2025.

[22] S. Li, et al., "The Modulatory Effect of Exogenous Orienting on Emotional Audiovisual Integration,"*Journal of Integrative Neuroscience*, vol. 23, no. 11, p. 210, 2024.

[23] M. Junchi, et al., "MULTICAUSENET temporal attention for multimodal emotion recognition,"*Frontiers in Psychology*, vol. 16, 2025.

[24] M. A. Razzaq, et al., "A Hybrid Multimodal Emotion Recognition Framework for Smart Environments,"*Sensors*, vol. 23, no. 9, p. 4373, 2023.

[25] E. Chitti, et al., "MiEmo: A multi-modal platform on emotion recognition for children with ASC,"*Smart Health*, vol. 31, p. 100418, 2025.

[26] X. Chen, et al., "Emotion-aware Design in Automobiles: embracing technology advancements,"*CHI Conference on Human Factors in Computing Systems*, pp. 1-15, 2025.