

Physiological Precursors That Precede the Awareness of Cognitive Stress

Takuya Endo, Toshikazu Kato, and Toru Nakata

Faculty of Science and Engineering, Chuo University, Tokyo, 112-855, Japan

ABSTRACT

In collaboration, early detection of partner's dissatisfaction is important to keep communication in good shape. Such early detection is also essential for AI systems and agents that interact with humans. This paper examined whether physiological changes associated with cognitive stress precede the conscious experience of dissatisfaction during consensus building dialogue. Twenty male undergraduate participants engaged in controlled one-on-one prioritization tasks based on a snow-mountain rescue scenario. Participants continuously reported dissatisfaction in real time using a visual analogue scale (VAS). Physiological data were temporally aligned to peaks of subjective dissatisfaction. The results showed that subtle but sustained decreases in nasal skin temperature frequently occurred prior to subjective dissatisfaction peaks. Nasal skin temperature may serve as an early autonomic marker of emerging cognitive stress, whereas heart rate may reflect later or more intense stress responses. By elucidating the temporal dynamics between physiological responses and subjective stress, this study provides foundational insights for cognitive computing in sociotechnical systems and supports the design of adaptive dialogue support frameworks.

Keywords: Cognitive stress, Consensus-building dialogue, Nasal skin temperature, Thermal imaging, Human-centered system design

INTRODUCTION

Consensus building constitutes a fundamental mechanism in sociotechnical systems, including organizational decision-making, collaborative problem solving, and safety-critical operations. Although consensus is often interpreted as successful agreement, the process through which it is achieved may involve cognitive strain, negotiation pressure, and asymmetric conversational influence.

Agreements reached through compromise rather than mutual acceptance may conceal latent “cognitive stress”, which refers to stress arising from cognitively appraised negative interpretations of the situation. Behavior under such cognitive stress may degrade individual decision quality and collective system performance.

Within the field of Human Systems Integration (HSI), optimizing system performance requires explicit consideration of human cognitive and affective capacities throughout system design and operation (Booher, 2003; Chapanis, 1996).

Stress assessment in collaborative settings, however, has traditionally relied on post-hoc questionnaires or retrospective evaluations. Such methods capture consciously recognized states but fail to identify when cognitive strain begins to emerge during interaction.

Physiological sensing offers a complementary approach because autonomic responses can emerge prior to consciously recognized dissatisfaction.

Previous psychophysiological studies have shown that nasal skin temperature decreases under acute stress and mental load due to sympathetic vasoconstriction. Because the nasal region is highly sensitive to autonomic modulation, sustained temperature reductions measured by infrared thermography have been proposed as indicators of stress-related peripheral blood flow changes.

These findings provide physiological grounding for examining nasal temperature as a potential early marker of emerging cognitive stress. Facial thermography enables non-contact measurement of peripheral vasoconstriction, particularly in the nasal region, which is sensitive to sympathetic activation.

Heart rate provides additional information on cardiovascular autonomic modulation and may reflect later-stage or more intense stress responses.

Multimodal integration of these signals aligns with systems engineering approaches that emphasize comprehensive modeling of human-system interaction (Folds et al., 2008). The present study examines whether physiological changes temporally precede subjective dissatisfaction during consensus-building dialogue.

Stress is conceptualized broadly as cognitive load arising from judgment, negotiation, and decision-making demands. By focusing on onset timing rather than average signal magnitude, this study seeks to clarify the temporal structure of cognitive stress emergence and provide a proof-of-concept framework for adaptive intervention design.

We specifically hypothesize that event-to-event dispersion in response timing can mask anticipatory patterns in grand averages (i.e. averages over all subjects and all experiments), making onset-based analysis more suitable for detecting early precursors.

METHODS

Participants

Twenty male undergraduate students participated in the study. Participants reported no cardiovascular, neurological, or psychiatric conditions and refrained from caffeine intake prior to participation. All procedures were approved under institutional ethical guidelines and written informed consent was obtained.

The restriction to a relatively homogeneous sample was intended to reduce inter-individual physiological variance during this proof-of-concept phase, allowing clearer examination of temporal structure rather than demographic modulation.

For event-level analysis (which is observation on one event without averaging to other similar events), we used only data of participants and events with sufficient pre-peak baseline.

As a result, 18 candidate dissatisfaction peaks were detected, and 15 events from 12 participants remained after baseline adequacy filtering.

Dialogue Task

The experimental paradigm was designed to induce cognitive stress within a controlled yet ecologically interpretable interaction.

Participants engaged in a one-on-one consensus-building dialogue with an experimenter using a snow-mountain rescue prioritization scenario. The task required ranking limited survival resources and reaching agreement under time constraint.

Such prioritization tasks are widely employed in negotiation and collaborative decision-making research because they reliably elicit cognitive load, perspective conflict, and negotiation-driven adjustment while preserving procedural consistency.

Experiment Setup

The session followed a fixed temporal structure: a 10-minute acclimation period to stabilize peripheral temperature and cardiovascular responses to ambient conditions; a 3-minute resting baseline during seated silence; a 10-minute dialogue phase; and a 3-minute post-task rest. Room temperature, lighting conditions, seating distance, and camera positioning were standardized across participants.

Measurement of Facial Skin Temperature

Facial skin temperature was continuously recorded using an infrared thermal imaging camera (FLIR Boson 320; spatial resolution 320×256; thermal sensitivity 50 mK).

The device was positioned approximately 50 cm in front of participants to allow non-contact acquisition throughout dialogue. Automated facial landmark detection algorithms extracted regional temperature time series.

The nasal region was selected for primary analysis because it demonstrated the most stable visibility across head movements and is physiologically sensitive to sympathetic vasoconstriction.

Thermal signals were sampled at 1 Hz. Heart-rate samples were aligned to the 1-second grid using a sample-and-hold approach for synchronization; onset detection for heart rate respected its original sampling cadence.

Measurement of Heart Rate

Heart rate was simultaneously recorded using a wearable photoplethysmography-based device (Fitbit Sense 2) at approximately 0.2 Hz. While lower in sampling frequency, heart rate provides complementary

information reflecting cardiovascular autonomic modulation and may index stress intensity.

Measurement of Subjective Dissatisfaction

Participants continuously rated dissatisfaction toward the dialogue partner using a graphical slider interface ranging from 0 to 100. This produced a continuous subjective time series synchronized with physiological signals.

SIGNAL PROCESSING TO DETECT STRESS REACTIONS

Low-Pass Filtering

To reduce transient noise while preserving causal temporal structure, physiological time series were smoothed using a 5-second causal moving average, relying exclusively on past observations. This filtering choice reflects the long-term goal of real-time applicability.

Scaling, Elimination of Baseline Constants, and Event Spotting

All detection thresholds were treated as operational definitions for evaluating temporal precedence, not as predetermined and universal constants.

Because raw physiological values are strongly influenced by individual baseline levels and sensor placement variability, all analyses were conducted on relative deviations from baseline rather than absolute magnitudes.

Two baseline definitions were employed.

1. For descriptive visualization, the final 30 seconds of the pre-task resting period were used.
2. For event-based analysis, an event-specific baseline was defined as the interval from $t = -120$ s to $t = -60$ s relative to dissatisfaction peak ($t = 0$).

This time interval was selected to represent a comparatively stable pre-manifestation window while avoiding contamination from the immediate pre-peak escalation phase.

Events lacking at least 30 seconds of usable baseline data were excluded from onset-based evaluation.

Dissatisfaction events were defined as local maxima in the subjective time series during the dialogue phase. Each peak was treated as a temporal anchor ($t = 0$), and physiological signals were aligned accordingly.

Finding Onset Trends on Physiological Signals

Nasal skin temperature and heart-rate signals were resampled to a unified 1-second resolution using timestamp synchronization to enable cross-modal comparison.

The analysis focused on onset timing of sustained physiological deviations, rather than relying solely on mean-level amplitude comparisons. This approach was motivated by our hypothesis that early stress indicators may vary in timing across individuals. They would become obscured under simple averaging.

Detection of the onset requires enough deviation from baseline, otherwise we will mistakenly detect many false positives, which are mere transient fluctuations.

For nasal skin temperature, relative decreases were examined using operational criteria representing sustained drops maintained across consecutive seconds. We tested two values of magnitude to differentiate subtle early changes from more pronounced responses.

For heart rate, sustained increases relative to baseline across consecutive samples were evaluated, accounting for lower sampling frequency.

Threshold values were treated as analytical operationalizations intended to test temporal precedence rather than to define predetermined and universal constants. This distinction is critical to avoid overgeneralization.

Evaluation of Detection Reliability

For each event, physiological onset was classified as preceding ($t < 0$) or non-preceding ($t \geq 0$) relative to dissatisfaction peak.

A two-sided binomial test assessed whether the observed proportion of preceding onsets deviated from chance expectation ($p = 0.5$). Given the exploratory nature and modest sample size, inferential results are interpreted as indicative trends rather than definitive causal confirmation.

RESULTS

Detected Onsets

Fifteen dissatisfaction events met inclusion criteria for onset-based analysis after baseline adequacy filtering.

At the individual event-level, substantial heterogeneity was observed in both subjective trajectories and physiological responses. Subjective dissatisfaction peaks varied in morphology: some events displayed rapid, transient spikes, whereas others exhibited plateau-like persistence across several seconds.

This morphological variability suggests that peak amplitude alone does not fully characterize the temporal dynamics of cognitive stress.

Time Ahead of Subjective Recognition of Stress

In the majority of events, nasal skin temperature exhibited gradual decreases prior to subjective peak manifestation.

In several cases, decreases began between 20 and 40 seconds before dissatisfaction reached its maximum. These decreases were typically sustained rather than instantaneous, consistent with gradual sympathetic vasoconstriction.

In contrast, heart rate increases were observed in fewer events and showed greater variability in both magnitude and temporal alignment.

Figure 1 illustrates a representative example of an individual event-level time series aligned to Dissatisfaction peak ($t = 0$). (The plot displays nasal skin temperature, heart rate, and subjective dissatisfaction within the pre-manifestation window.) To evaluate whether anticipatory patterns were visible at the aggregate level, event-locked grand-average waveforms were

computed. When nasal temperature signals were averaged across events, no clear pre-peak decrease was evident in the mean waveform. Standard deviation bands indicated substantial dispersion in both onset timing and magnitude across events.

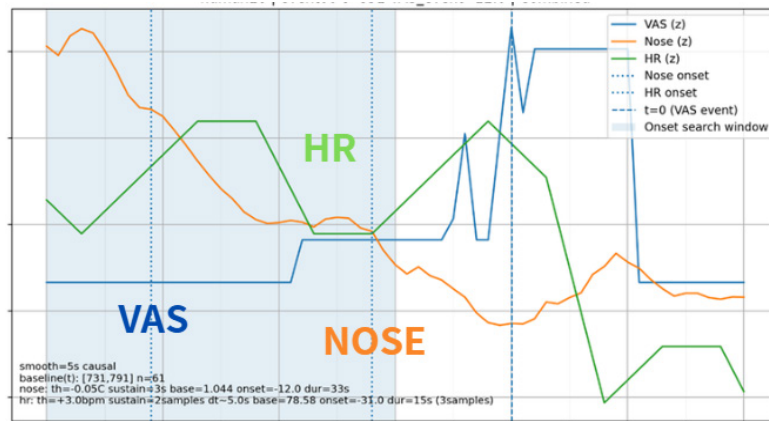


Figure 1: A representative example of physiological responses before and after subjective recognition of dissatisfaction. The horizontal axis is for time, as $t = 0$ is adjusted on the peak of subjective dissatisfaction. The vertical axis for the three data of nasal skin temperature, heart rate, and subjective dissatisfaction, which are shifted vertically for comparison convenience.

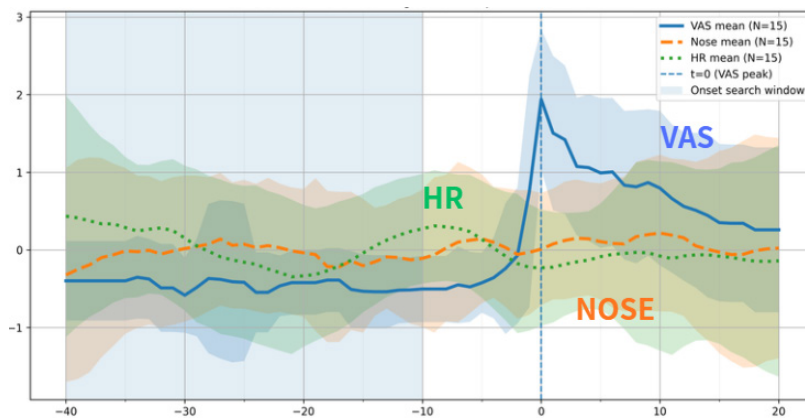


Figure 2: Grand-average event-locked physiological waveforms aligned to subjective dissatisfaction peak ($t = 0$). Shaded regions are standard-deviation bands of the data among all VAS spike events, and they indicate inter-event variability. Curves are mean of data among all VAS spike events.

A similar absence of consistent pre-peak elevation was observed in heart rate grand averages. (Nasal skin temperature and heart rate aligned to subjective dissatisfaction peak at $t = 0$. Shaded regions represent inter-event variability.)

The absence of clear anticipatory structure in averaged waveforms appears attributable to temporal dispersion. Onset timing varied considerably across events, causing phase misalignment that attenuated mean-level effects.

Reliability of Nasal Onset Detection

Fifteen peaks of subjective dissatisfaction were identified.

Onset-based classification revealed that there were 12 subtle but sustained decreases of the nasal temperature, which preceded subjective dissatisfaction.

Applying a two-sided binomial test assuming equal probability of preceding versus non-preceding onset, this distribution was found to show a significant bias toward temporal precedence ($p = 0.035$).

However, when we applied larger thresholds for the nasal temperature decreases to define onsets, correctness did not significantly deviate from chance expectation.

Figure 3 presents the histogram of onset timing for the subtle nasal temperature decreases. The distribution illustrates substantial dispersion in onset latency, with many events occurring in the -40 to -10 second window prior to subjective dissatisfaction peaks. This temporal spread supports the interpretation that early physiological changes are present but temporally heterogeneous across events.

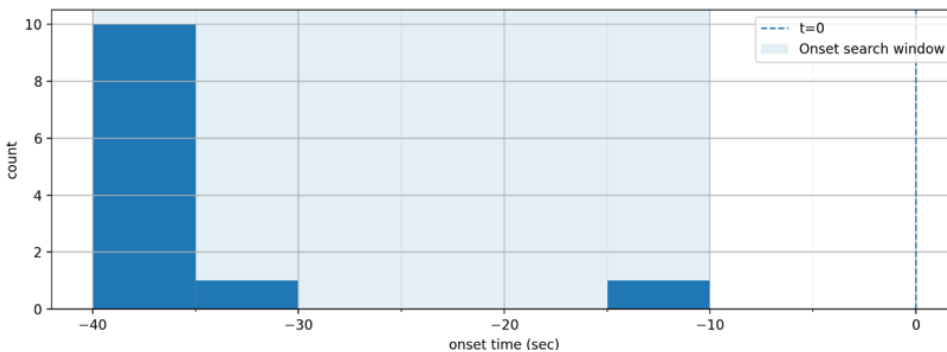


Figure 3: Histogram of onset timing for subtle nasal temperature decreases relative to subjective dissatisfaction peak ($t = 0$).

In contrast, Figure 4 illustrates that larger-magnitude nasal temperature decreases exhibited broader temporal dispersion, with onset timing distributed across the -40 to 0 second window. This variability likely contributed to the absence of statistically significant temporal bias under stricter detection criteria.

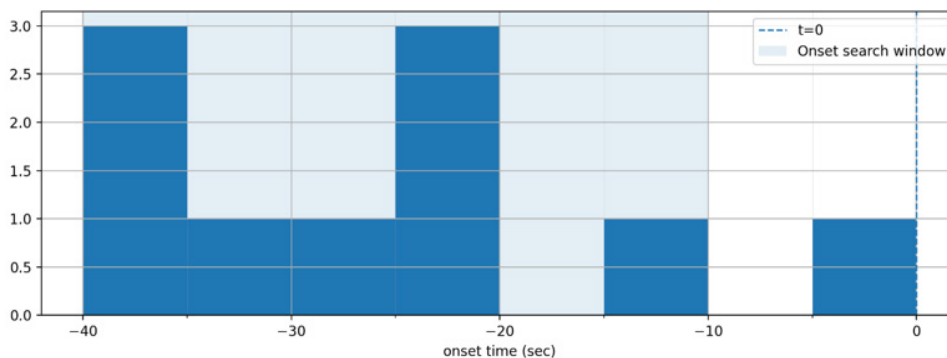


Figure 4: Histogram of onset timing for larger-magnitude nasal temperature decreases relative to subjective dissatisfaction peak ($t = 0$).

Absence of Concrete Pattern of Heart Rate Trend

For heart rate, preceding increases were observed in 5 of 15 events, and no significant temporal bias was detected.

These findings suggest that subtle peripheral (i.e. nasal) vasoconstriction responses may occur earlier in the stress emergence process than cardiovascular acceleration.

The distribution of nasal onset timing demonstrated broad dispersion. Some events showed early-phase responses more than 30 seconds prior to subjective peak, whereas others exhibited onset within 10 seconds of manifestation.

This dispersion explains why grand-average waveforms failed to reveal anticipatory trends and highlights temporal heterogeneity as a central characteristic of stress emergence during interactive negotiation.

Relationship Between Physiological Reactions and Verbal and Non-Verbal Content of Dialogue

Exploratory qualitative inspection of dialogue segments surrounding physiological onset windows revealed conversational markers such as hesitation fillers, mitigated agreement expressions, and brief acknowledgment responses.

Although not systematically quantified, these observations suggest potential coupling between autonomic changes and interactional micro-dynamics during emerging cognitive strain.

Collectively, the results indicate that nasal skin temperature demonstrates consistent event-level temporal precedence relative to subjective dissatisfaction, whereas heart rate does not exhibit comparable reliability under the present task conditions.

CONCLUSION

This study investigated whether physiological changes associated with cognitive stress precede the conscious recognition of dissatisfaction during consensus-building dialogue. While grand-average analyses did not reveal clear anticipatory trends, onset-based analysis showed that subtle decreases in nasal skin temperature frequently occurred prior to the subjective onset of dissatisfaction.

These findings suggest that peripheral facial thermography may function as an early autonomic marker of emerging cognitive stress, whereas heart rate may index later-stage or intensity-dependent responses. The substantial variability in onset timing across individuals and events further indicates that detection frameworks must accommodate temporal heterogeneity rather than rely exclusively on averaged patterns.

This work provides proof-of-concept evidence within a controlled dialogue paradigm. In practical applications, early nasal-temperature cues could be leveraged to trigger lightweight dialogue-support interventions (e.g., brief reflection prompts or facilitator signals) before dissatisfaction becomes overt. Future research should advance individualized detection models, incorporate multimodal linguistic features, and evaluate adaptive intervention strategies in real-world sociotechnical environments.

ACKNOWLEDGMENT

The authors also thank all participants for their cooperation in the experiment.

REFERENCES

- Booher, Harold, ed. (2003). *Handbook of human systems integration*. New Jersey: Wiley.
- Booher, H.R., Minninger, J. (2003). "Human systems integration in army systems acquisition", in: Booher, Harold (Ed.), *Handbook of human systems integration*, pp. 663–698.
- Chapanis, A. (1996). *Human factors in systems engineering*. Wiley Series in Systems Engineering and Management. Hoboken, NJ: Wiley.
- Folds, Dennis., Gardner, Douglas., Deal, Steve. (2008). Building Up to the Human Systems Integration Demonstration. *INCOSE INSIGHT*, 11(2).
- Friedenthal, S., Moore, A., Steiner, R. (2008). *A Practical Guide to SysML: The Systems Modeling Language*. Morgan Kaufmann; Elsevier Science.
- Honour, Eric C. (2006). *A Practical Program of Research to Measure Systems Engineering Return on Investment (SE-ROI)*. Proceedings of the Sixteenth Annual Symposium of the International Council on Systems Engineering, Orlando, FL.
- Meilich, Abe. (2008). *INCOSE MBSE Initiative Status of HSI/MBSE Activity Presentation*.
- Taubman, Philip. (2008). Top Engineers Shun Military; Concern Grow. *The New York Times*.