# Understanding Behind the Smile of Emotion Workers: Detecting After-Call Stress in Call Agents

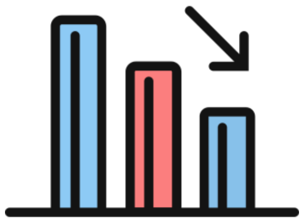
Duri Lee, Heejeong Lim, Vedant Das Swain*, Uichin Lee
KAIST & *New York University

# Risks of Repetitive Emotional Workload at Call Centers

Repetitive emotional workload of call agents cause high stress

Stress degrades task performance and prolonged exposure poses long-term health risks

Constant emotion regulation
following company's display rules



Performance
Drop

Mental health
risks

Zapf, D., Vogt, C., Seifert, C., Mertini, H., & Isic, A. (1999). Emotion work as a source of stress: The concept and development of an instrument. *European Journal of work and organizational psychology*, *8*(3), 371-400.

# Data-driven Stress Monitoring in HCI Contexts

- So far data-driven stress monitoring mainly focused on knowledge workers
- However, there’s a lack of exploration on emotional labor contexts
- For ecological validity, we need to consider real work contexts: i.e., work procedures and data types

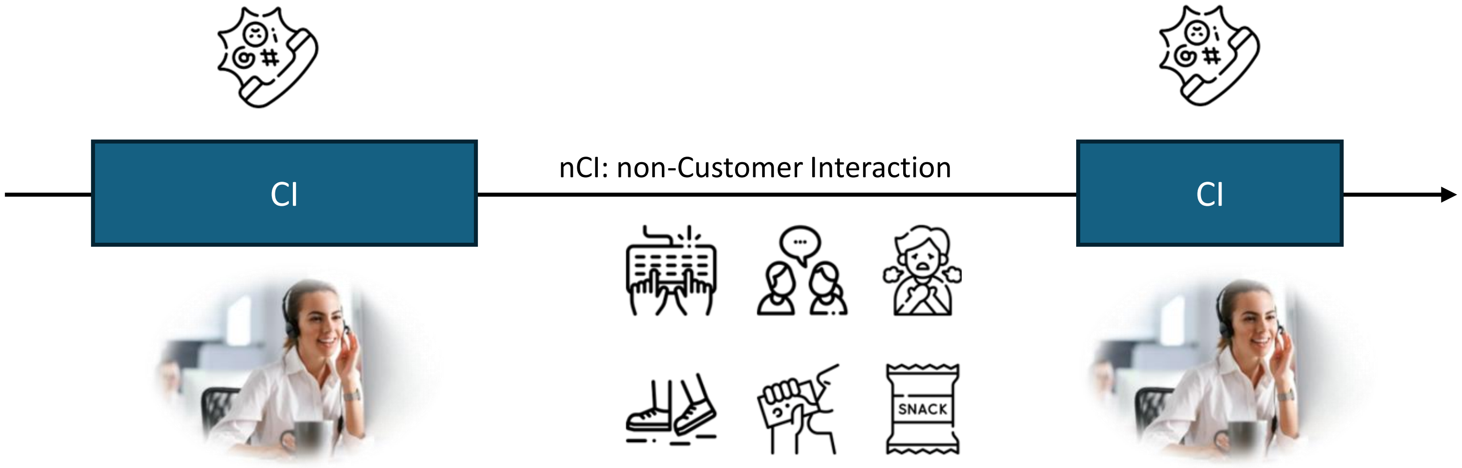

**Digital traces of knowledge work**: wearable sensing, email logs, calendar entries, desktop activities


**Emotional labor:** Work procedures & Data availability

Mark et al., Email duration, batching and self-interruption: Patterns of email use on productivity and stress , CHI 2016
Howe et al., Design of Digital Workplace Stress-Reduction Intervention Systems: Effects of Intervention Type and Timing, ACM CHI 2022
Hide-and-seek: Detecting Workers' Emotional Workload in Emotional Labor Contexts Using Multimodal Sensing, Park et al., PACM IMWUT 2024

# Repetitive Nature of Emotional Work

CI: Customer Interaction
nCI: non-Customer Interaction

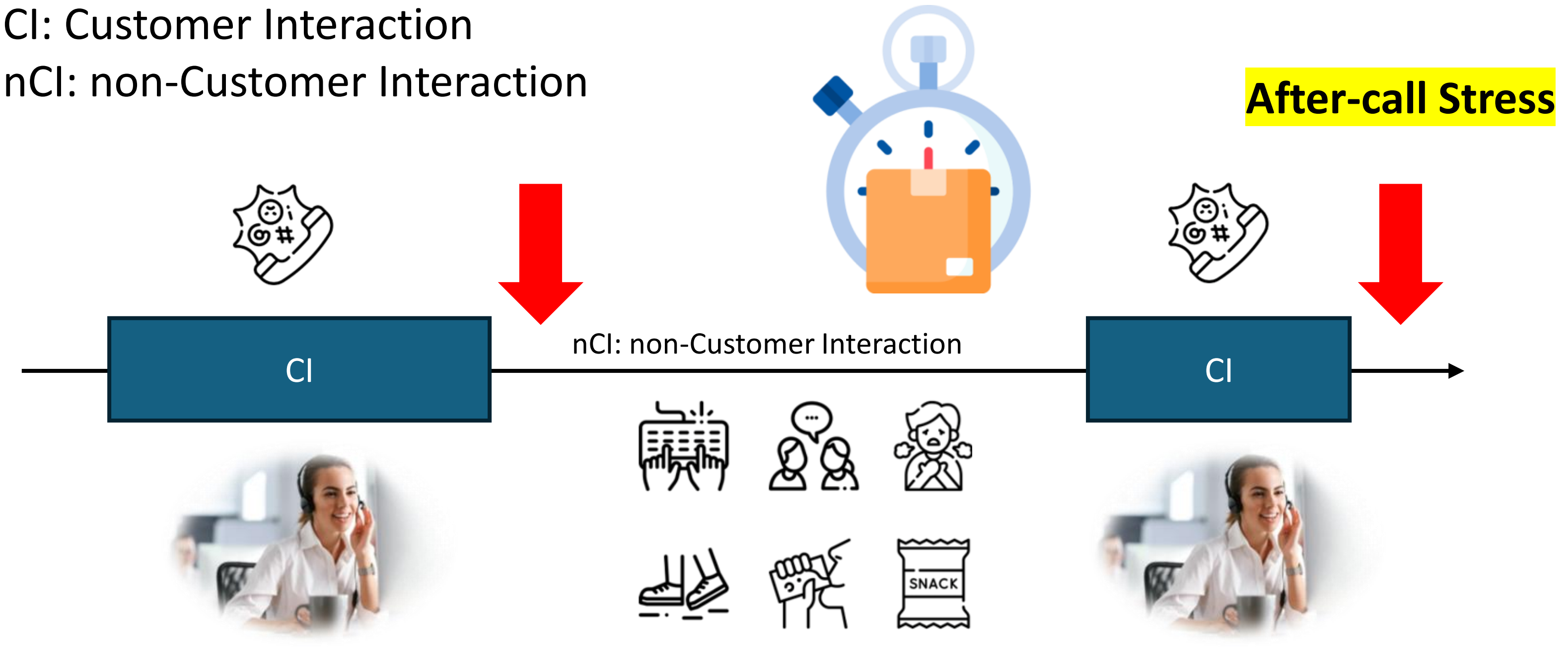




**Front stage**:
Emotion regulation
based on display rules

**Backstage**:
Prepare next CI
(e.g., summarizing, relaxing)

Insights: Cyclic work nature generates structured task logs that differ from knowledge-work traces

# Self-Tracking **After-call Stress** for Worker’s Stress Management


**Front stage**:
Emotion regulation
based on display rules

**Backstage**:
Prepare next CI
(e.g., summarizing, relaxing)

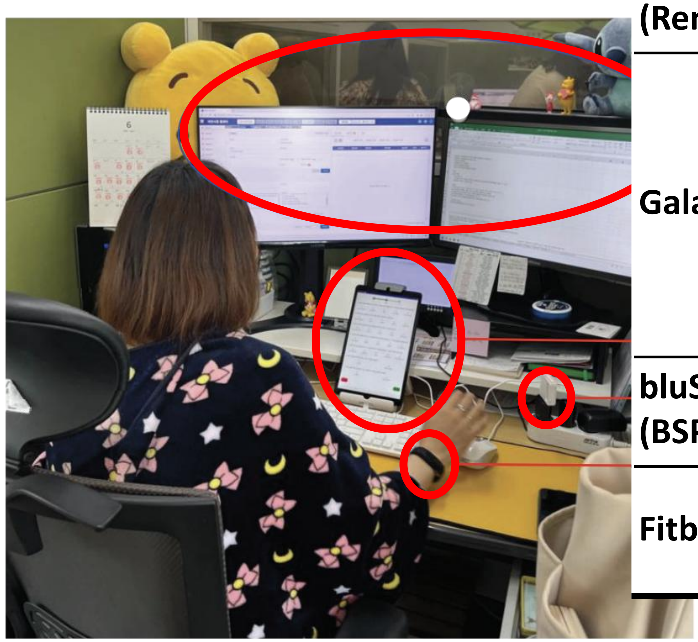
Task aligned stress tracking is an essential step toward designing effective strategies to protect workers from excessive stress and improve well-being (e.g., JIT interventions)

# Research Goal

**RQ1: Can we detect after-call stress by combining passive sensing data with call-center task logs, while respecting the task-aligned cycle of emotional labor?**

**RQ2: Beyond prediction performance, can interviews with call agents help us better understand the factors related to after-call stress?**

# Data Collection Methods



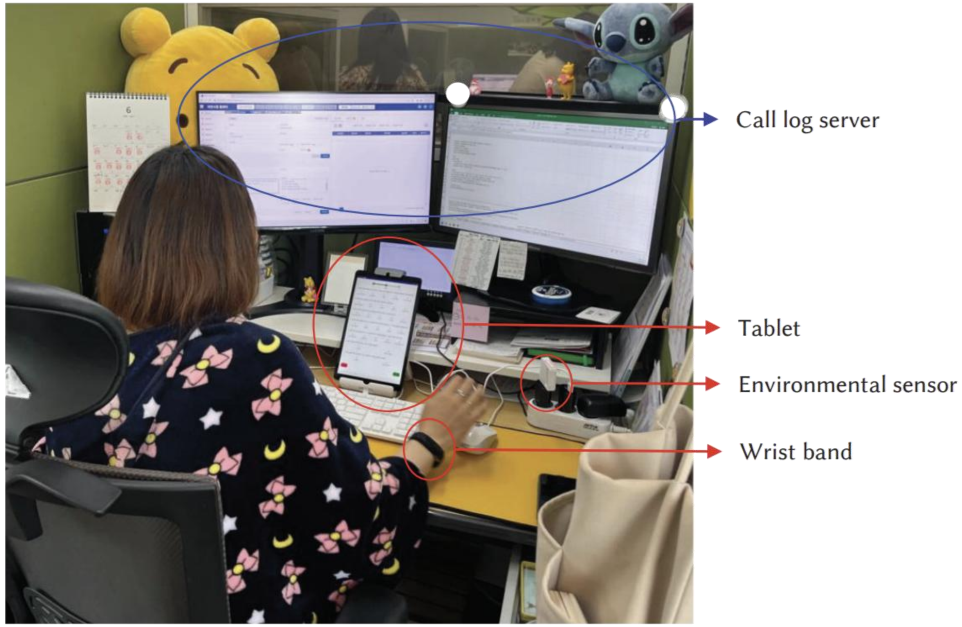
| Device | Freq | Sensing Type | Data type | Data |
|---|---|---|---|---|
| **Call Log Server (Remote)** | Per call | Passive sensing | Task | Call start time, call duration, question (text), answer (text), agreement, complaint |
| **Galaxy Tab** | Per call | Self reported | Behavioral | Eating, drinking |
| | Per call | Self reported | Psychological | Stress, arousal, valence, surface acting |
| | Per day | Self reported | Daily baseline | Daily health condition, stress, arousal, valence, bedtime, wake-up time |
| | 5.0 Hz | Passive sensing | Behavioral | Desktop activity (x, y, and z) |
| **bluSensor (BSP02AIR)** | 0.1 Hz | Passive sensing | Environmental | $CO_2$, humidity, temperature |
| **Fitbit Inspire 2** | 1 min | Passive sensing | Behavioral | Step counts |
| | 1 sec | Passive sensing | Physiological | Heart rate |

Collected sensor data and self-report data

from 18 call agents for 4 weeks

at a city hall call center in South Korea

# Data Collection Methods

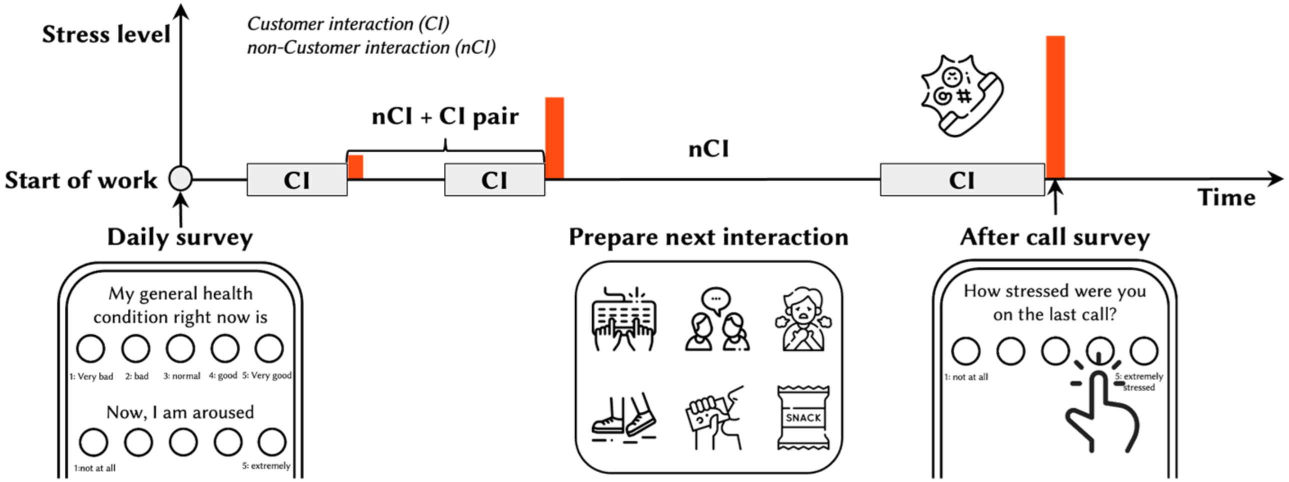



Collected sensor data and self-report data

from 18 call agents for 4 weeks

at a city hall call center in South Korea

Daily workflow for self-report data collection

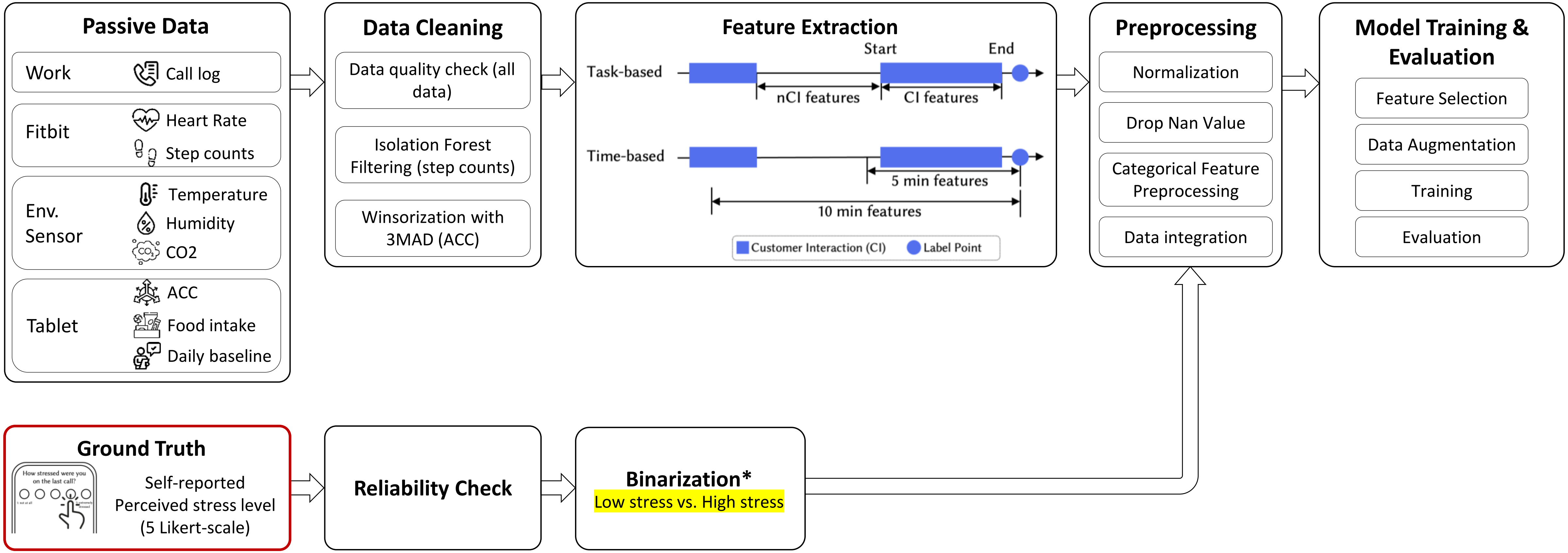



Daily baseline before work starts

After-call stress (self-report)

⇒ 7,442 self-reports

# Data Analysis Methods




* subject-wise mean threshold

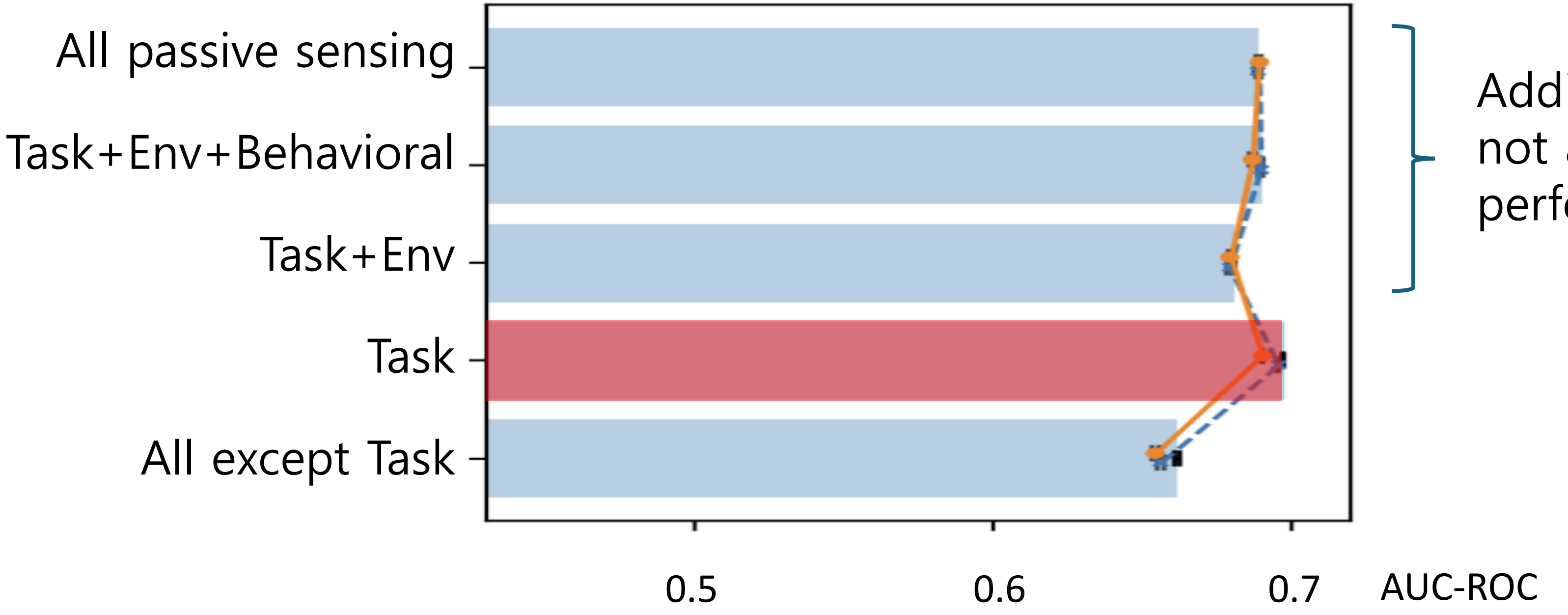
# Results – Task Alignment Matters for Stress Detection

**Exact alignment** **Close alignment** **Non-alignment (too large window size)**

| | Task-aligned | | Time-based | | | | |
|---|---|---|---|---|---|---|---|
| | | 5-min | 10-min | 15-min | 20-min | 25-min | 30-min |
| **ROC-AUC mean (STD)** | | | | | | | |
| RF | **0.685** (0.090) | **0.692** (0.082) | 0.671 (0.080) | 0.642 (0.074) | 0.616 (0.074) | 0.610 (0.067) | 0.605 (0.0638) |
| XGBoost | 0.679 (0.085) | 0.686 (0.088) | 0.664 (0.083) | 0.637 (0.082) | 0.625 (0.072) | 0.613 (0.078) | 0.599 (0.057) |
| CatBoost | 0.661 (0.094) | 0.673 (0.085) | 0.656 (0.085) | 0.627 (0.084) | 0.600 (0.086) | 0.591 (0.075) | 0.594 (0.072) |
| LDA | 0.661 (0.079) | 0.663 (0.081) | 0.643 (0.078) | 0.617 (0.071) | 0.599 (0.069) | 0.593 (0.066) | 0.589 (0.065) |
| SVM | 0.623 (0.104) | 0.624 (0.106) | 0.612 (0.099) | 0.584 (0.099) | 0.567 (0.093) | 0.560 (0.094) | 0.546 (0.092) |
| DT | 0.576 (0.071) | 0.601 (0.066) | 0.577 (0.068) | 0.583 (0.047) | 0.536 (0.031) | 0.523 (0.039) | 0.547 (0.044) |
| TabNet | 0.658 (0.039) | 0.676 (0.035) | 0.636 (0.030) | 0.612 (0.048) | 0.587 (0.042) | 0.595 (0.025) | 0.587 (0.015) |

More effective to segment data according to the work contexts, rather than using broad, generic time windows

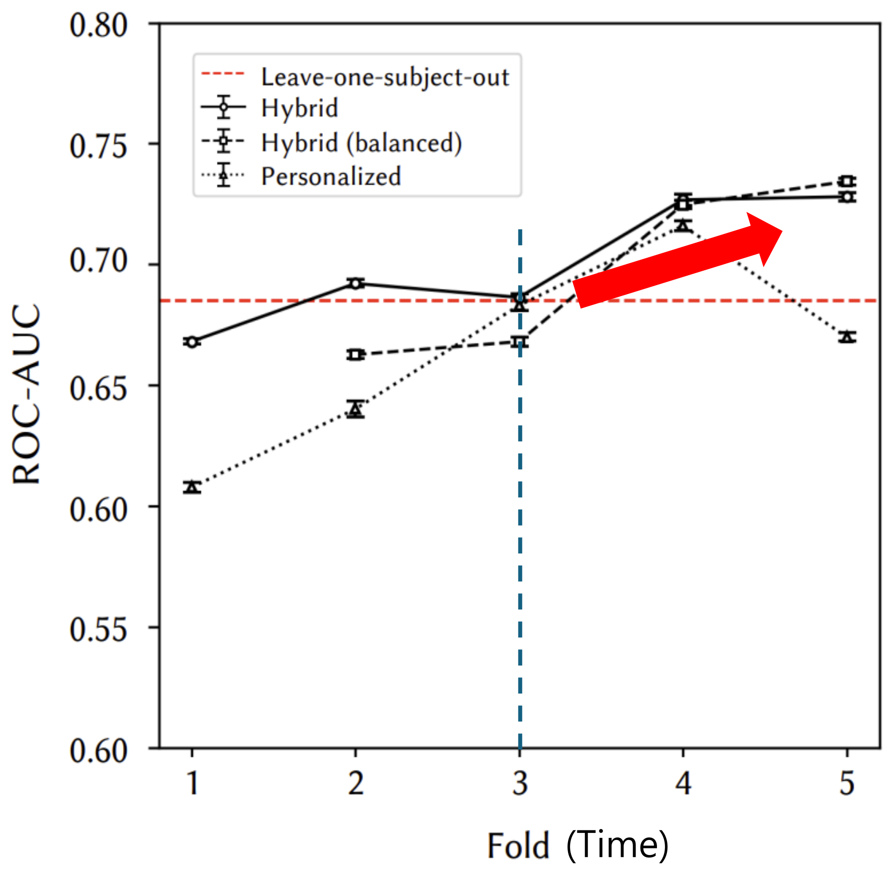
# Results – Task-log Features Considered Important




Adding more features does not always help improve performance

**Task logs** should not be treated as secondary context.
Instead, they should be considered **first-class sensors** in this work setting

# Results – Personalization Improves Detection Performance




After-call stress detection should not rely only on a one-size-fits-all population model, but it should adapt to individuals

** Metric : Mean (±SD) of AUC, # experiments per fold = 30*

* Time-serious cross-validation

# Understanding Factors Related to After-class Stress Dynamics

- **Semi-structured interview (N = 13) after data collection**
  - 1) **What triggers stress** during customer interaction (CI) ?
  - 2) How do you **respond to stressful conversations** during CI?Q
  - 3) **What do you do after a stressful CI** during non-customer interaction (nCI)?
  - 4) Besides stressful CI, **what other factors influence** your perceived stress level after a call?

# Results – Diversity of Stress Triggers

**Findings 1** Diverse stress triggers within similar customer interactions

- **Ambiguous problem description**

*"Even though it's stressful when the customer is unpleasant, it's even more challenging when they inquire without knowing their issue. The process of figuring it out is complicated, and it becomes harder if they are not cooperative."*

- **Expectation of emotional support**

*``After venting for about 11 minutes … they said,*

*`Talking about it like this has made me feel a bit better.' That really hurt me.''*

- **Rejection of proposed solutions** or **misunderstanding** of the caller on the suggested solution

→ ***Implication:***

Future modeling may need to go beyond simple call metadata and incorporate more content-level representations of customer interactions.

# Results – Different Approaches During Stressful Interactions

**Findings 2** Different approaches while engaging with stressful customer interactions

- **Disengagement driven by helplessness**

*“I **slump and zone out** because there is nothing I can do in the middle of the call.”*

- **Active coping strategies during the call**

*"Even though I know it's inappropriate, I make **rude gestures** or mouth **profanities directed** at the customer.” ”Just let him talk, and I take out a piece of paper and **write down song lyrics**.”*

*"I **drink water** to cool down.”*

→ ***Implication:***

Need for personalization: even when we observe similar or related behavioral signals, their meaning may differ across individuals

# Results – Variability in Recovery Patterns

**Findings 3 Variability in recovery patterns** during nCI intervals

- Agents use **different reset strategies** during non-customer interaction (nCI)

– talking with colleagues, sitting quietly, deep breathing, typing forcefully, having snacks/cold drinks

– leaving the seat or office when stress is overwhelming

*"Getting up from my seat usually means I'm really stressed, and I need to step out.."*

→ ***Implication:***

Can't treat non-customer interaction (nCI) segment as a uniform resting period.

Recovery is personal, and systems need to account for that variability.

# Results – Other Influential Personal and Environmental Factors

**Findings 4** Other influential factors related affecting after-call stress

- Agent's personal and environmental context matters

  - Personality, coping style, task preference, physical conditions (e.g., sore throat, menstrual cycle)

  - Environmental factors (e.g., noise, background chatter)

*“Too loud colleagues or background chatter influence my task”*

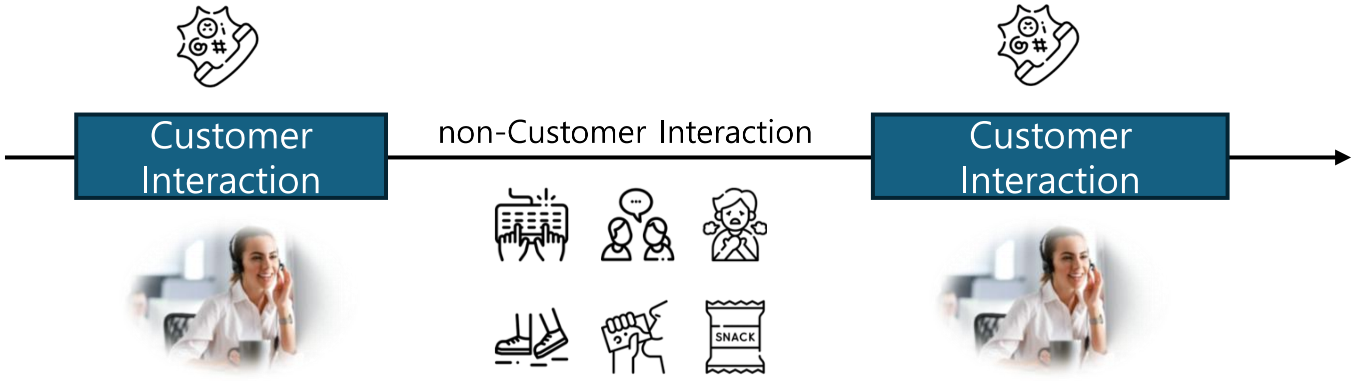
*“I feel burden when I response to the customer in overly quite environment,*

*because everyone can listen my voice”*

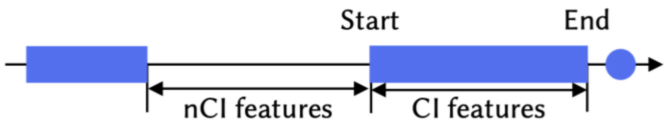
→ ***Implication:***

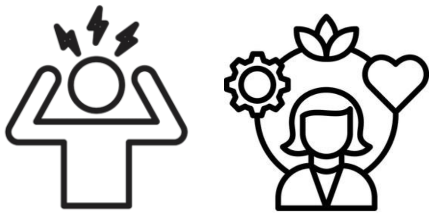
Even within the same person, stress signals can vary
depending on the day and the surrounding conditions

# Key Findings





**Stress Factors, Coping and Recovery Patterns**

**Task Logs as First-Class Sensors**

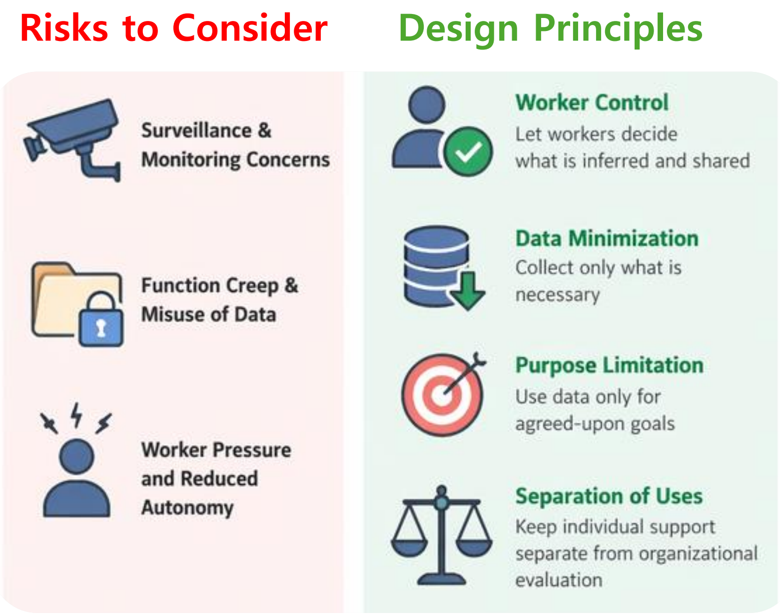
**Task-Aligned Segmentation for Feature Extraction**

**Personalization for Mitigating Inter-/Intra-Personal Variability**

# Implications for System Design


**Adopt Progressive Sensing with Minimal Intrusion**


**Designing Closed Loops: Sensing → Intervention**




**Ethics: Sensing Should Empower Workers, Not Monitor Them**

# Understanding Behind the Smile of Emotion Workers: Detecting After-Call Stress in Call Agents

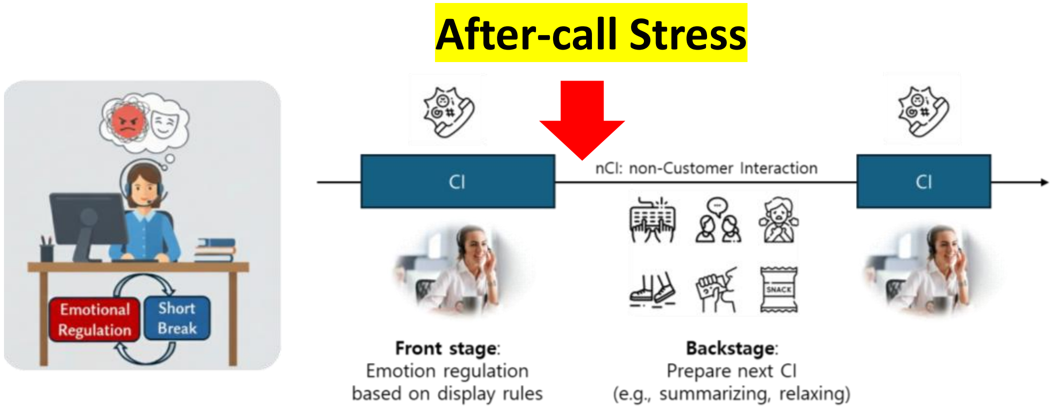

## Duri Lee, Heejeong Lim, Vedant Das Swain, Uichin Lee

KAIST School of Computing

NYU TANDON SCHOOL OF ENGINEERING


## Key Findings

- Task Logs as First-Class Sensors
- Task-Aligned Segmentation as Modeling Principle
- Personalization for Mitigating Inter-/Intra-Personal Variability




## Implications

- Adopt Progressive Sensing with Minimal Intrusion
- Designing Closed Loops: Sensing → Intervention
- Ethics: Sensing Should Empower Workers, Not Monitor Them