When Feeling and Physiology Diverge: Understanding Dual-Indicator Stress Sensemaking & Micro-Interventions in an Emotional-Labor Workplace

Jeonghyun Kim, Yugyeong Jung, **Junmo Lee**, Kwangyoung Lee,
Hwajung Hong, and Uichin Lee

Two Stress Indicators: Perceived & Physiological

Managing stress has become central to sustaining mental health in everyday life, but how these systems define and measure stress is different.

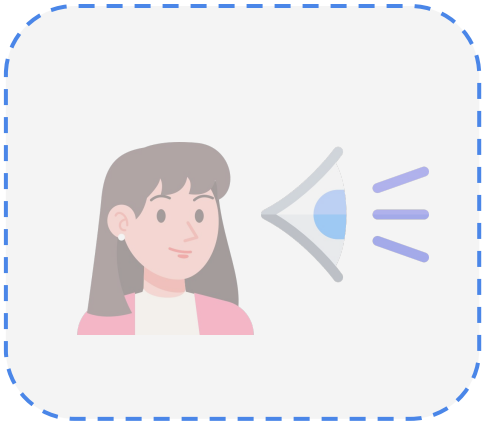
	Perceived Stress		Physiological Stress
		Not always the same ≠	
Reflection	Individuals' subjected experiences		Bodily arousals
Measurement	Self-reports, ESM		Physiological signals

Stress is inherently **multifaceted** and cannot be fully captured by a single indicator alone

Understanding How Users Interpret Stress

- Many prior works focus technical goals and reducing stress indicator metrics
- Limited attention to how users actually understand and interpret stress data

Less attention



Main Focus



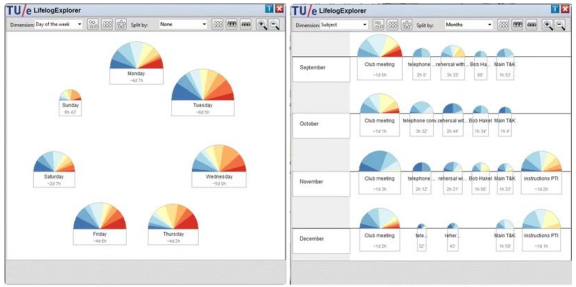
- Stress is not merely a physiological reaction, but also an experience constructed through individuals' appraisals, interpretations, and meaning-making processes[1]
- Therefore, how the users **reflect on the causes and effect of their stress** based on the data is important

[1] Richard S Lazarus. 1984. Stress, appraisal, and coping. Vol. 445. Springer.

Background and Motivation

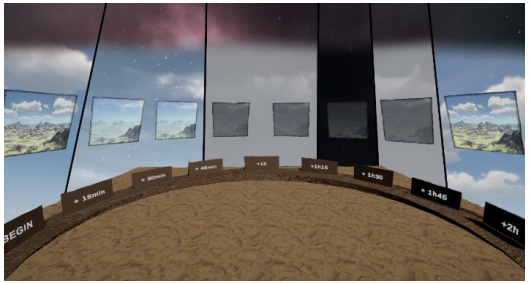
Data-Driven Stress Interpretation In Existing Works

Prior works have suggested systems that helps users understand and regulate stress by supporting data-driven reflections.

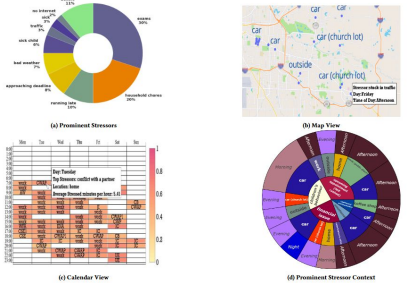


(a) Days of the week (b) Subject trends

LifelogExplorer[1]:
Visualizing physiological stress pattern over time with context



VeatherReflect[2]:
Representing physiological stress as weather



(a) Prominent Stressors (b) Map View (c) Calendar View (d) Prominent Stressor Context

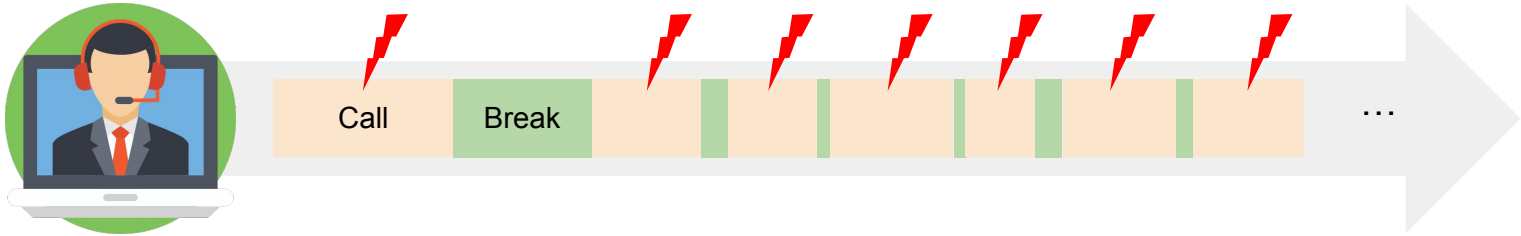
S. Neupane et al.[3]:
Visual reflection on perceived stress and related contexts

Organized around **single score** or aggregated score, lacks **comparative reasoning**

[1] R. Kocielnik and N. Sidorova, "Personalized stress management: enabling stress monitoring with lifelogexplorer," KI-K ünstliche Intelligenz, 2015.
[2] N. Wagener, et al., "Veatherreflect: Employing weather as qualitative representation of stress data in virtual reality," in Proceedings of the 2023 ACM Designing Interactive Systems Conference, 2023.
[3] S. Neupane, et al., "Momentary stressor logging and reflective visualizations: Implications for stress management with wearables," in Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems, 2024.

Emotional Labor Workplace

We focus on emotional labor workplace, particularly call-center environments

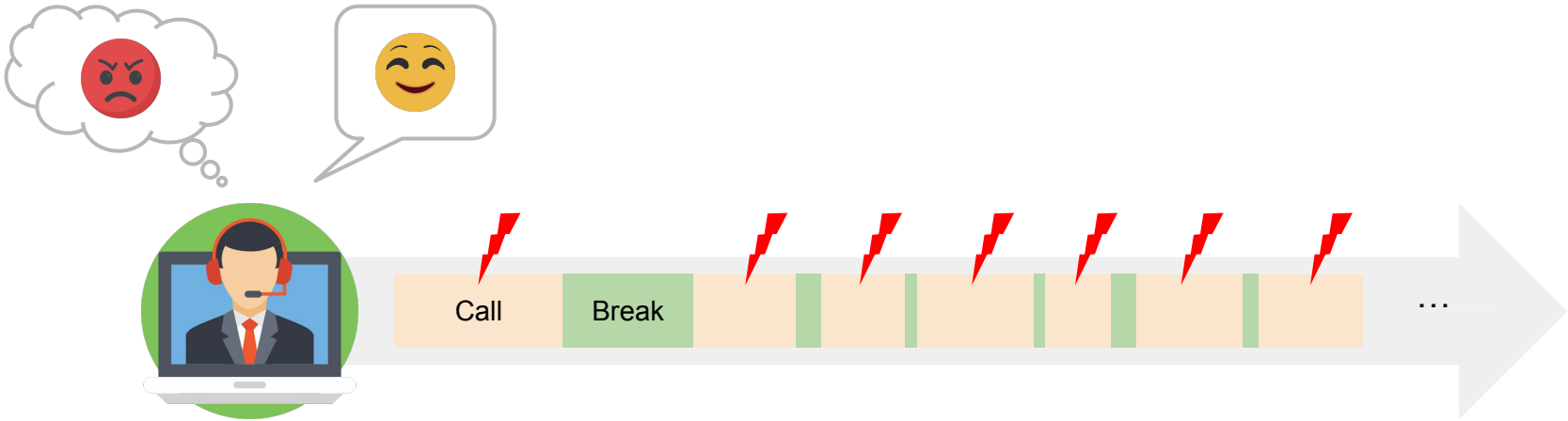


Intensive, repetitive, and emotionally demanding calls make

- Common perceived and physiological stress experience
- Providing understanding to stress for sustaining mental well-being and performance

Emotional Labor Workplace

We focus on emotional labor workplace, particularly call-center environments



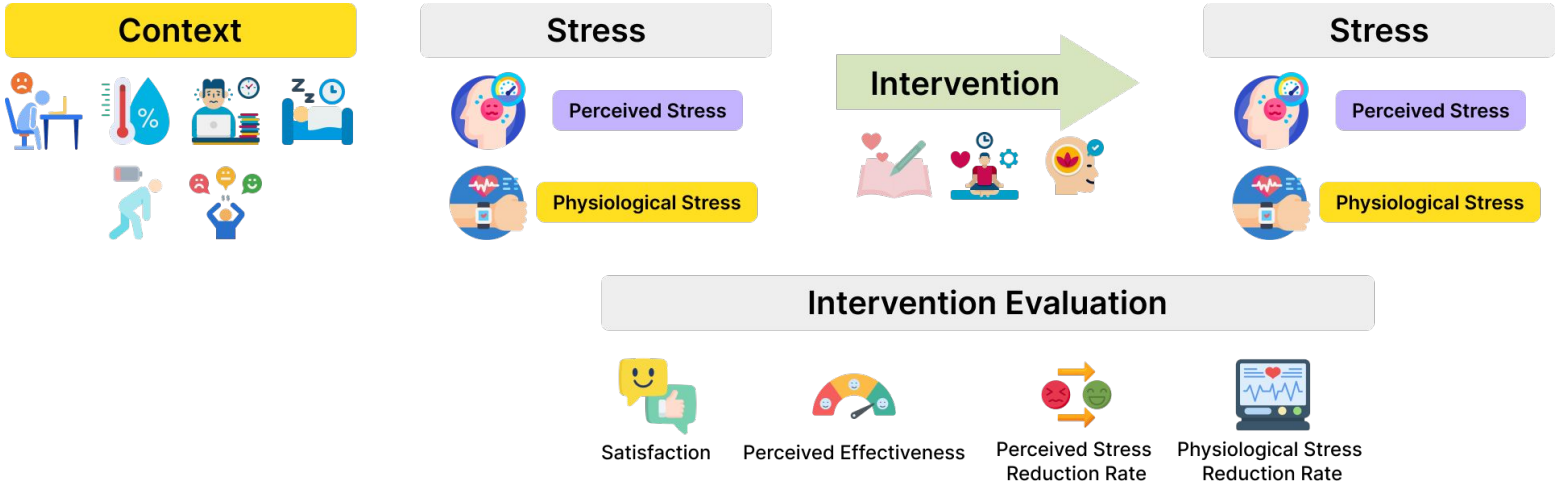
Intensive, repetitive, and emotionally demanding calls make

- Common perceived and physiological stress experience
- Providing understanding to stress for sustaining mental well-being and performance

Research Questions (RQs)

In emotional labor workplace settings,

- **RQ1:** How do participants integrate and interpret **perceived and physiological stress indicators with contextual data** to make sense of their stress?
- **RQ2:** How do participants evaluate the provided micro-interventions, and how do these interventions influence their sensemaking of subsequent **perceived and physiological stress changes**?



Study Design

Study Design

Study Overview

- **Location:** metropolitan call center in south Korea
- **Duration:** 3 days (pilot), 4 weeks (main study)
- **Demographics: N=19** (17 Female, 2 Male)
  - Age: Mean = 38.90 (SD = 5.81)
  - Years of Experience: Average tenure 3.58 years



Study Design

In-the-Wild Study Setup

Smartwatch (Galaxy Watch 7)

- To collect physiological sensor data (HR, PPG, skin temperature, ACC)

Sending step event data

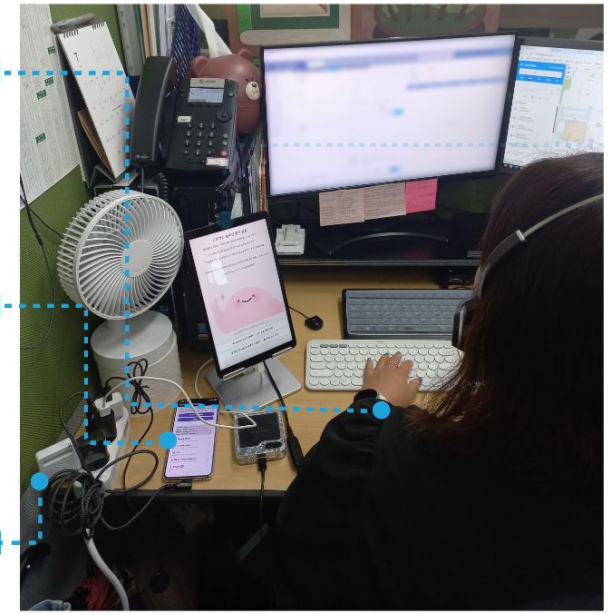
Smartphone (Galaxy S21/22)

- To receive step event data from smartwatch and calculating step count
- To store environmental sensor data

Sending environmental sensor data

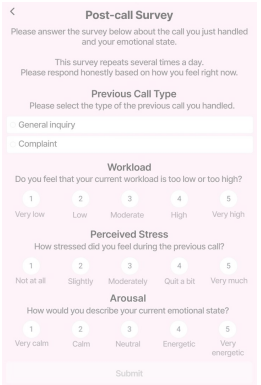
Environmental sensor (bluSensor)

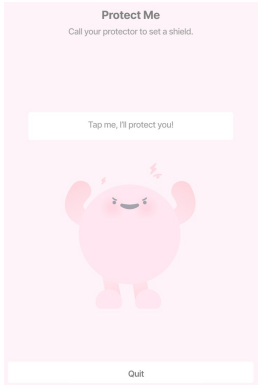
- To collect environmental sensor data (temperature, humidity, CO2, tVOC)



Tablet (Galaxy Tab A7 Lite)

- To provide diverse interventions via stress intervention system
- To collect self-report data:
  - 1) pre-shift survey,
  - 2) post-call survey,
  - 3) post-intervention survey
  - 4) post-intervention refusal survey





Study Design

In-the-Wild Study Setup

Smartwatch (Galaxy Watch 7)

- To collect physiological sensor data (HR, PPG, skin temperature, ACC)

Sending step event data

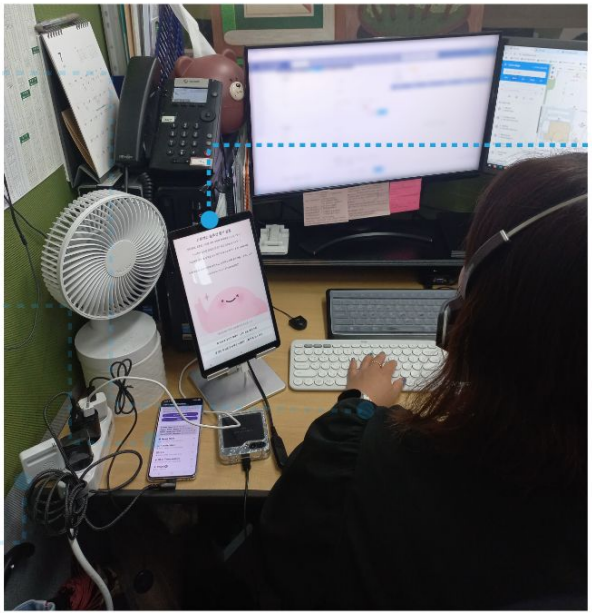
Smartphone (Galaxy S21/22)

- To receive step event data from smartwatch and calculating step count
- To store environmental sensor data

Sending environmental sensor data

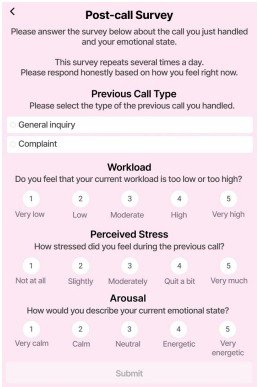
Environmental sensor (bluSensor)

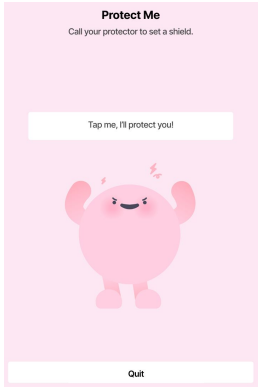
- To collect environmental sensor data (temperature, humidity, CO2, tVOC)



Tablet (Galaxy Tab A7 Lite)

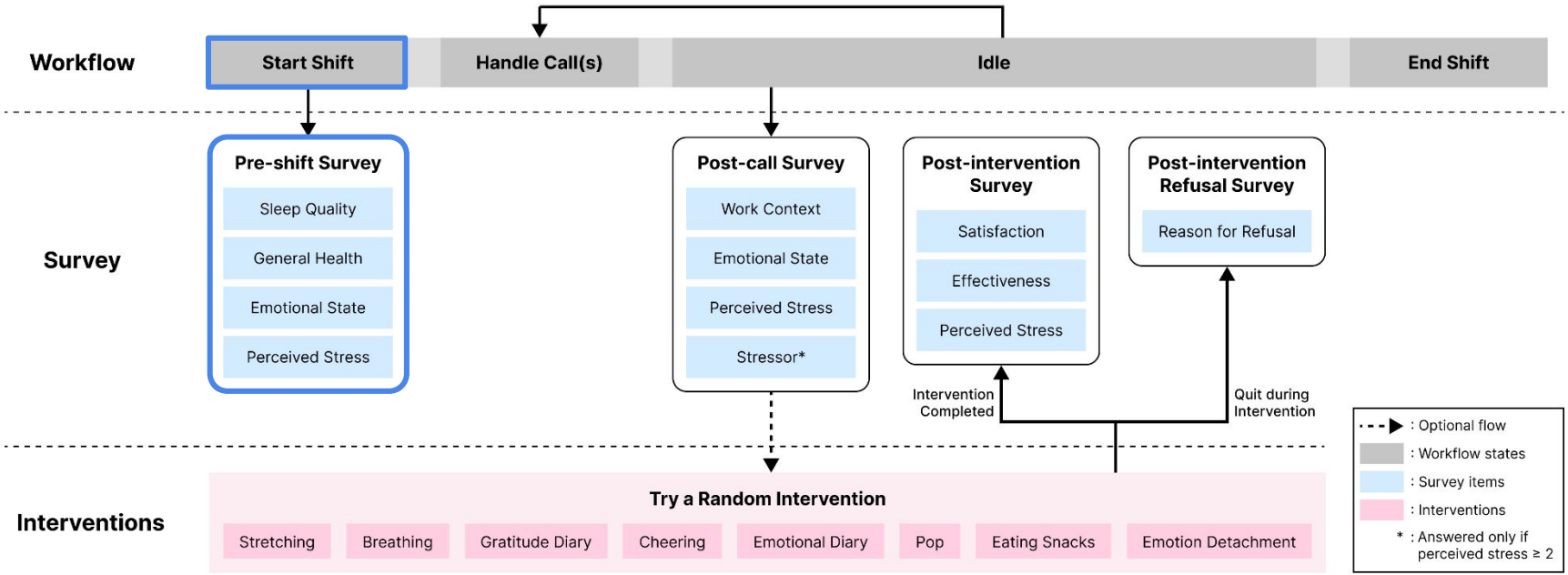
- To provide diverse interventions via stress intervention system
- To collect self-report data:
  - 1) pre-shift survey,
  - 2) post-call survey,
  - 3) post-intervention survey
  - 4) post-intervention refusal survey





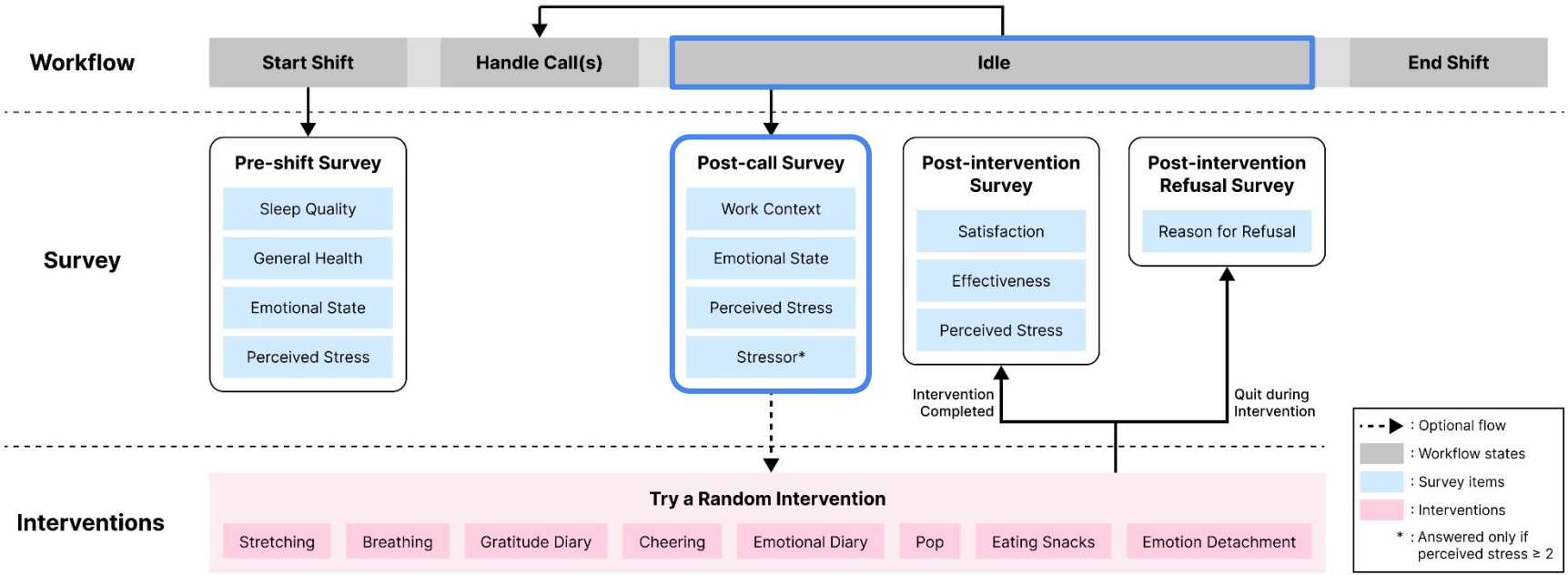
Study Design

Using the Intervention Application



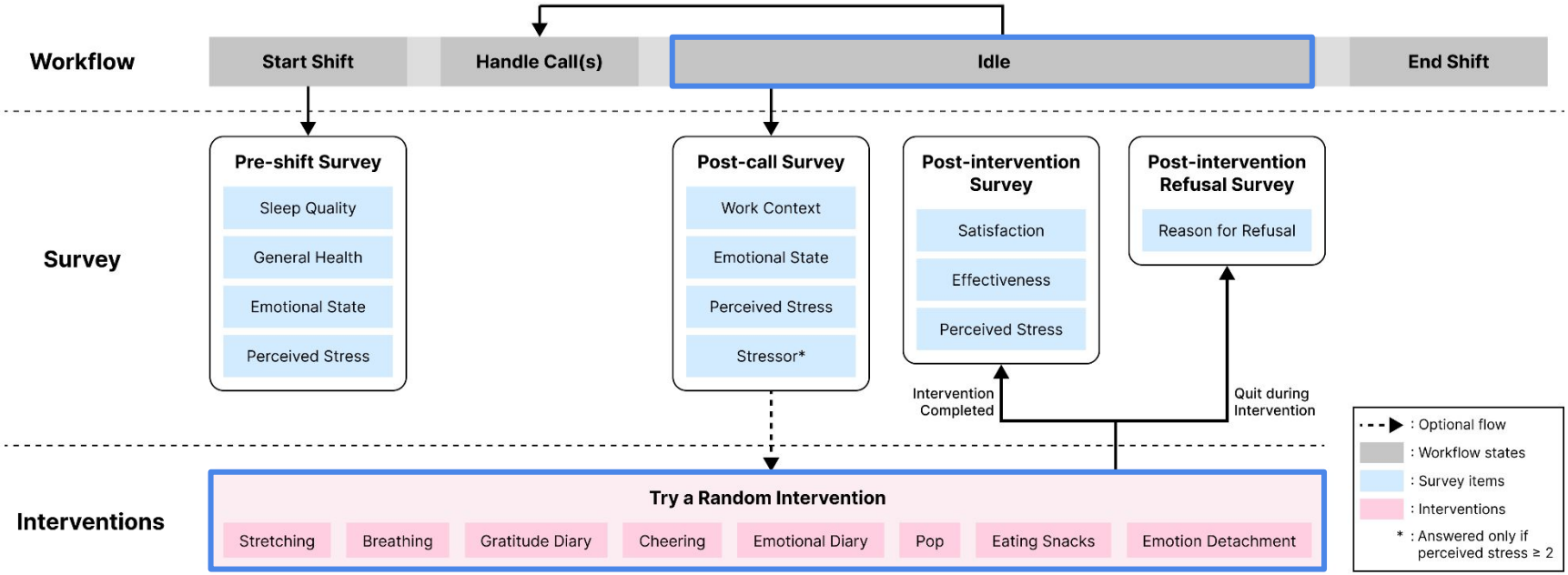
Study Design

Using the Intervention Application



Study Design

Using the Intervention Application

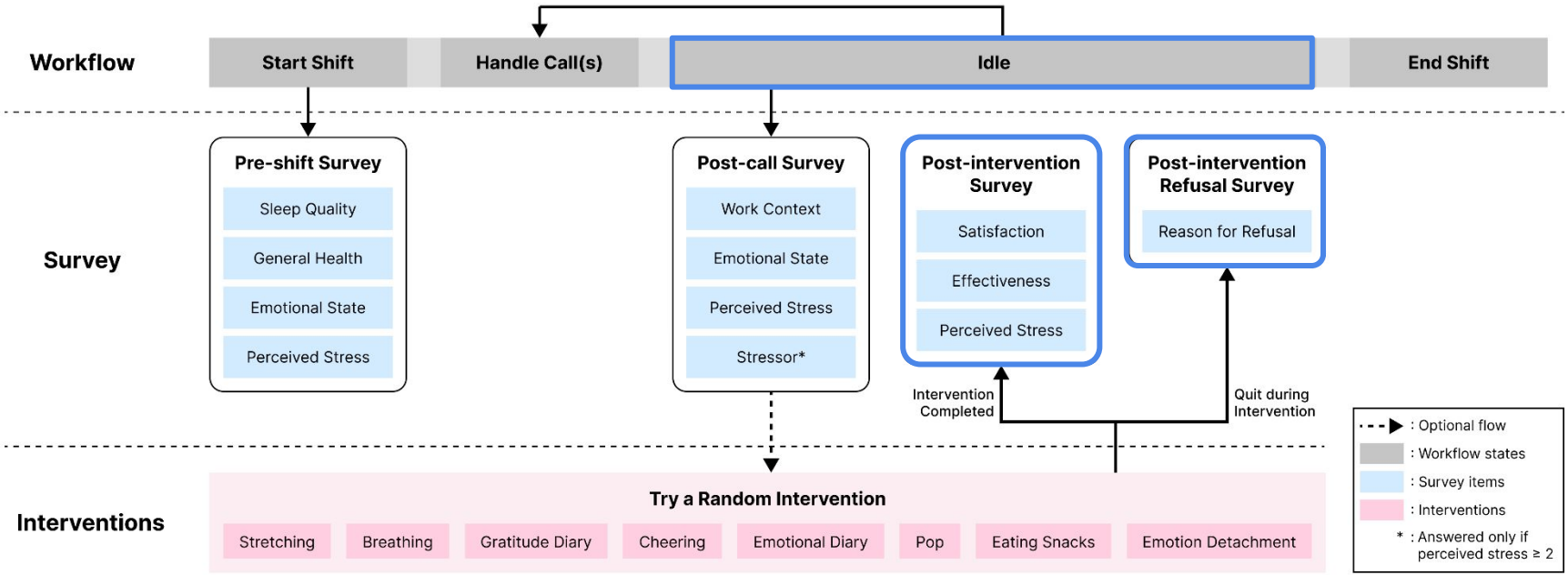




Stretching | Protect Me | Deep Breathing | Sugar Boost | Anger-Eating Fairy | I Did Well, Right? | Words I Need Right Now | Is It Because of Me?

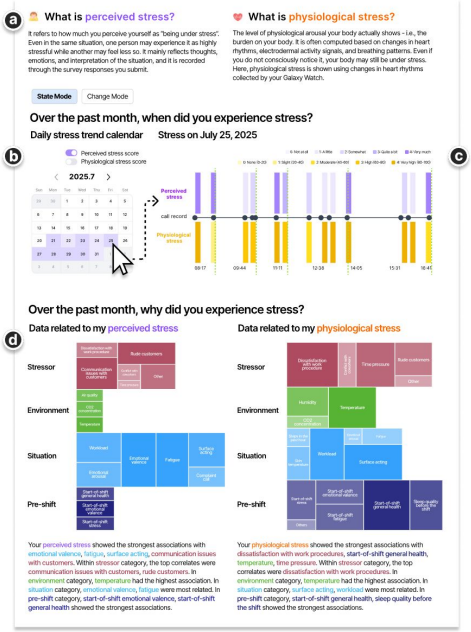
Study Design

Using the Intervention Application



Study Design

Dual-Indicator Reflection Dashboard





State Mode (RQ1)

Integrating stress states with context

Change Mode (RQ2)

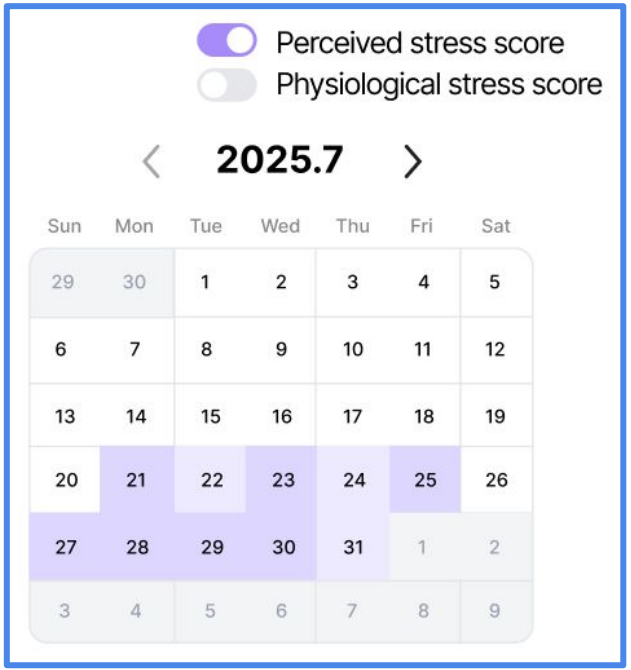
Visualizing changes after interventions

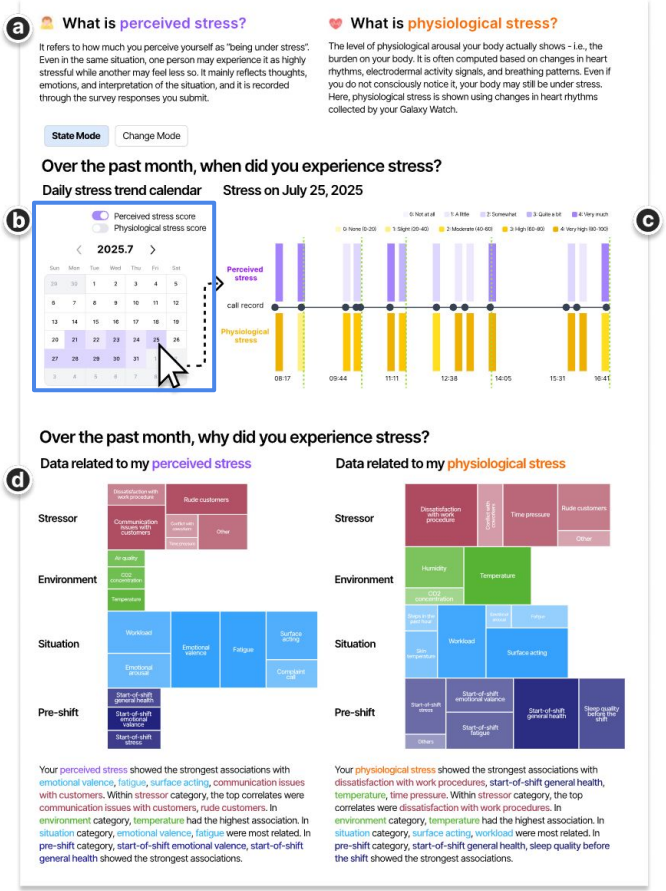
Study Design

Dashboard: State Mode

(b) Calendar

- Summarizes daily stress trends
- Allows identifying high/low stress days at a glance



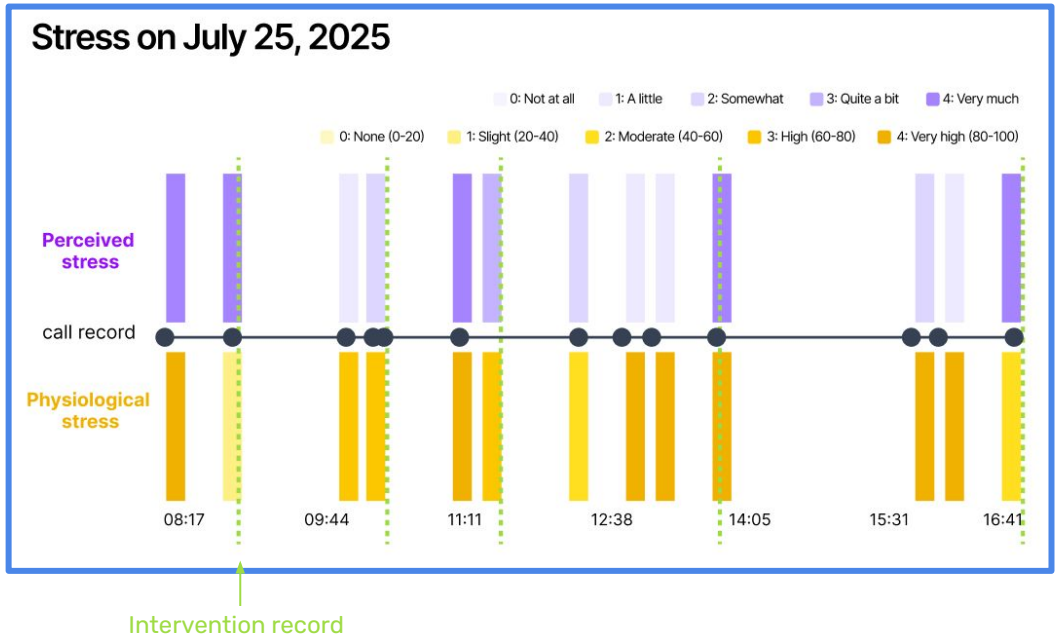


Study Design

Dashboard: State Mode

(c) Timeline

- Clicking on a specific day of a calendar reveals a timeline
- Visualizes alignment and divergence over time

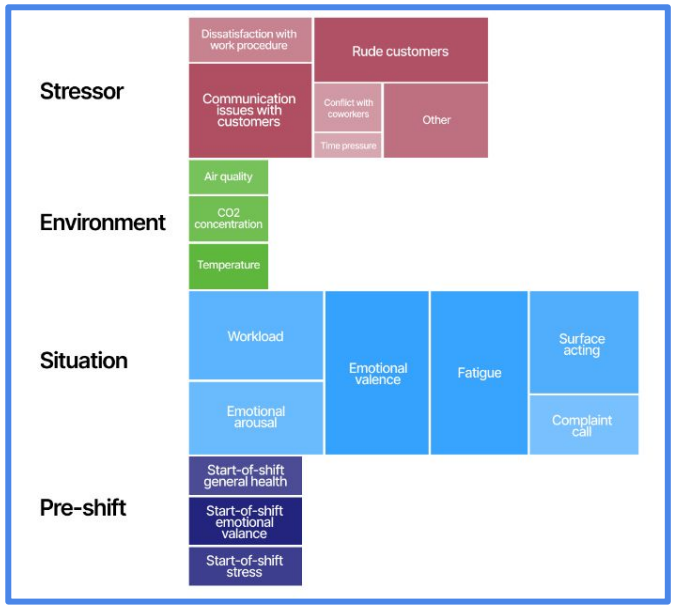




Dashboard: State Mode

(d) Treemap

- Correlates stress types with **<u>four categories of contextual factors</u>**
- Helps users explore “which” contextual factors are highly related to each stress indicator
    - Larger treemap boxes -> stronger associations to the context



Treemap Explanation



Study Design

Dashboard: Change Mode

Similar to state mode, but focus on stress **reduction effect** of the interventions

(b) Calendar

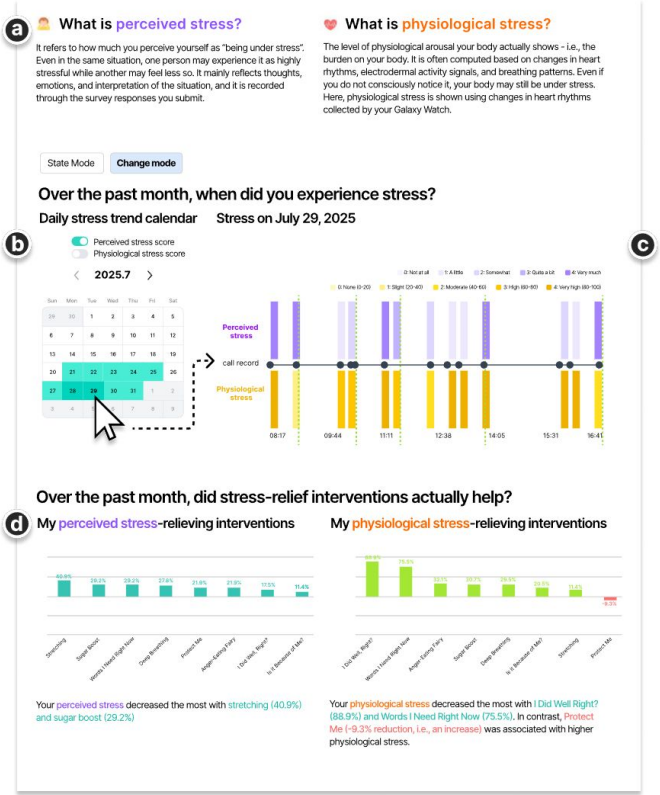
- Shows stress reduction trends (*increased, decreased, or no-change)*
- Darker colors = Greater reduction

(c) Timeline

- Participants can check whether each stress type has *increased or decreased*

(d) Bar Chart

- Ranks provided interventions by observed reduction effects
- Helps evaluate which strategy actually worked for each stress indicator

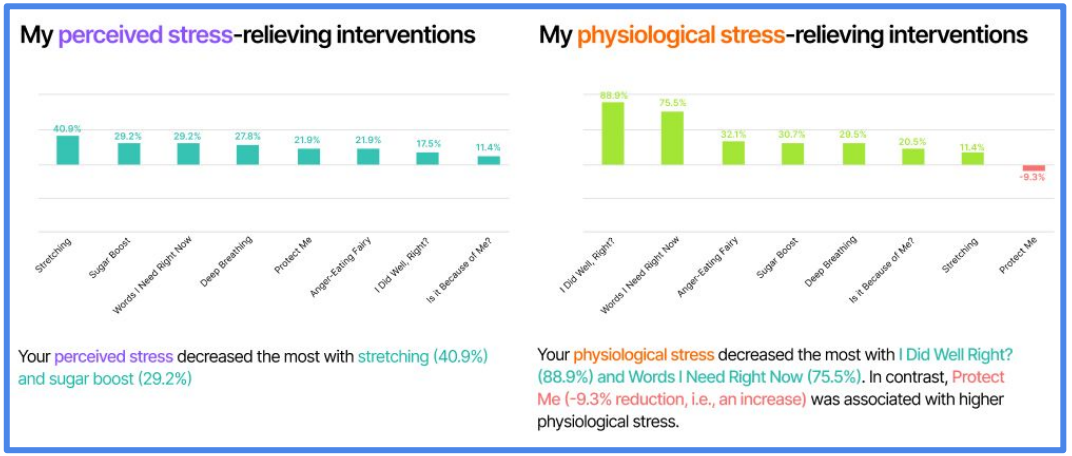


Study Design

Dashboard: Change Mode

(d) Bar Chart

- Ranks provided interventions by observed reduction effects
- Helps evaluate which strategy actually worked for each stress indicator





Study Design

Participant Interview



Warm-up

↓

State-mode dashboard exploration
(10-15 min.)

↓

Interview (RQ1)

↓

Change-mode dashboard exploration
(10-15 min.)

↓

Interview (RQ2) & Wrap-up

Results

Initial Perception of Perceived and Physiological Stress

Participants clearly distinguished between **perceived** and **physiological** stress

> *"Perceived stress is just the stress I feel in my mind, something I sense in my head" (P09)*

> *"When body reacts, like grabbing the back of my neck or feeling heat rise, that feels like physiological stress" (P01)*

- They described physiological stress through symptoms such as headaches (n=4), increased heart rate (n=4), muscle tension (n=2), chest tightness (n=2), indigestion (n=1), warmth or heat sensation (n=1), and shortness of breath (n=1)

However, most participants were <u>unaware how physiological stress could be measured</u>

Contextualizing Perceived and Physiological Stress

Participants interpreted their stress by reconciling the data with their **personal experiences and intuitions**, forming their own views

Perceived stress

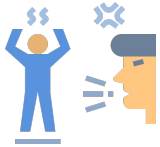


Communication issues, rude customers



Complaint calls, workload

Physiological stress



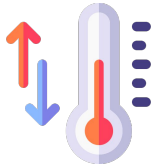
Communication issues, rude customers



Time Pressure



Fatigue



Environment

Contextualizing Perceived and Physiological Stress

Participants interpreted their stress by reconciling the data with their **personal experiences and intuitions**, forming their own views

Perceived stress

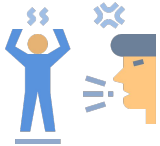


Communication issues, rude customers



Complaint calls, workload

Physiological stress



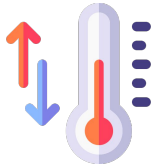
Communication issues, rude customers



Time Pressure

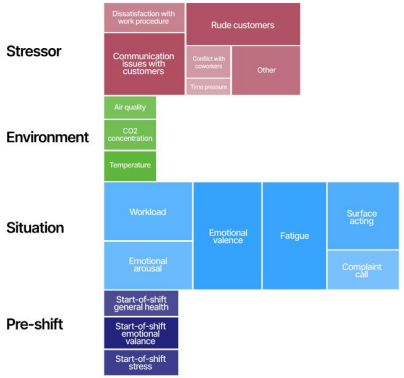


Fatigue



Environment

≠



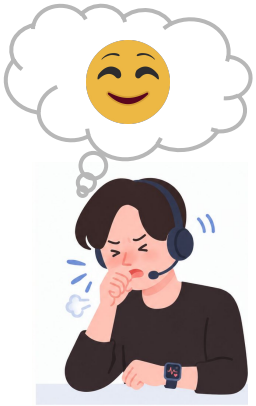
Dashboard Content

How Did Participants Understand the Discrepancies?

Participants viewed mismatches (feeling vs. physiology) as **informative signals**



Trusting objectively measured indicators even when they differed from their subjective judgment



Interpreting the discrepancy based on **situational factors,** such as bad physiological condition



Attributing higher perceived stress to **emotional suppression and accumulated tension**

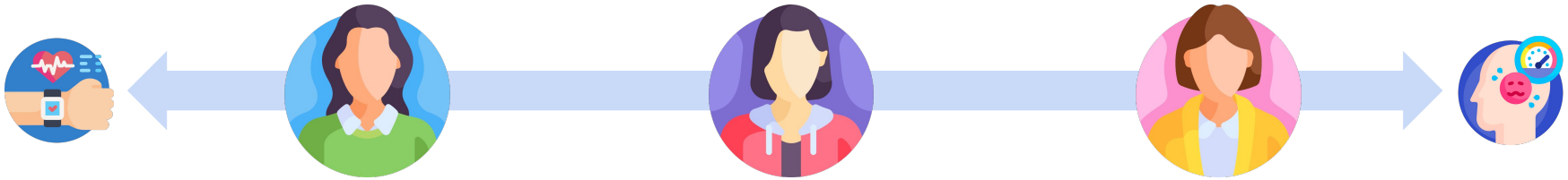
Results: Understanding of perceived and physiological stress with contextual data (RQ1)

Which Stress Indicator Better Reflects Their Stress?

Physiological stress as more trustworthy

Two indicators are not separate but **complementary**

Perceived stress better captured their actual condition



*"...it **even picks up moments when I wasn't consciously aware** that I was feeling stressed"*

*"When perceived stress is high, I believe the **body also reacts**, so I don't see them as unrelated"*

*"...since it was directly experienced and **consciously assessed**"*

Results: Preference of intervention and understanding of post-intervention changes (RQ2)

Intervention Preference: Which One and Why?



Stretching

Pros: Refreshing (14)
Cons: Lack of time & space



Protect Me

Pros: Relieve through fidgeting (14)
Cons: Cannot relate to the idea



Deep Breathing

Pros: Calm the mind (14)
Cons: Lack of time



Sugar Boost

Pros: Eating snacks relieves (13)
Cons: Health issue / not preferred



Anger-Eating Fairy

Pros: Emotional dumping ground (10)
Cons: Makes one reflect on stress



I Did Well, Right?

Pros: Can reflect one's own feelings (9)
Cons: Cannot think up a word as time passes



Words I Need Right Now

Pros: Enjoy the feeling of being talked to (9)
Cons: Vague feeling

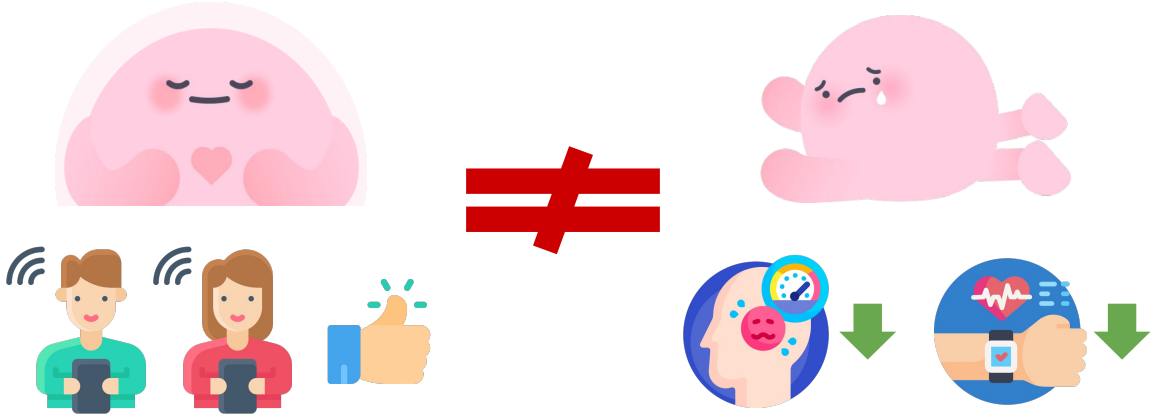


Is It Because of Me?

Pros: Can separate emotion from work (7)
Cons: Does not apply to one's situation

Intervention Preference: Was It Actually Effective?

Users “liked” interventions made them feel better, even if their actual stress didn't drop



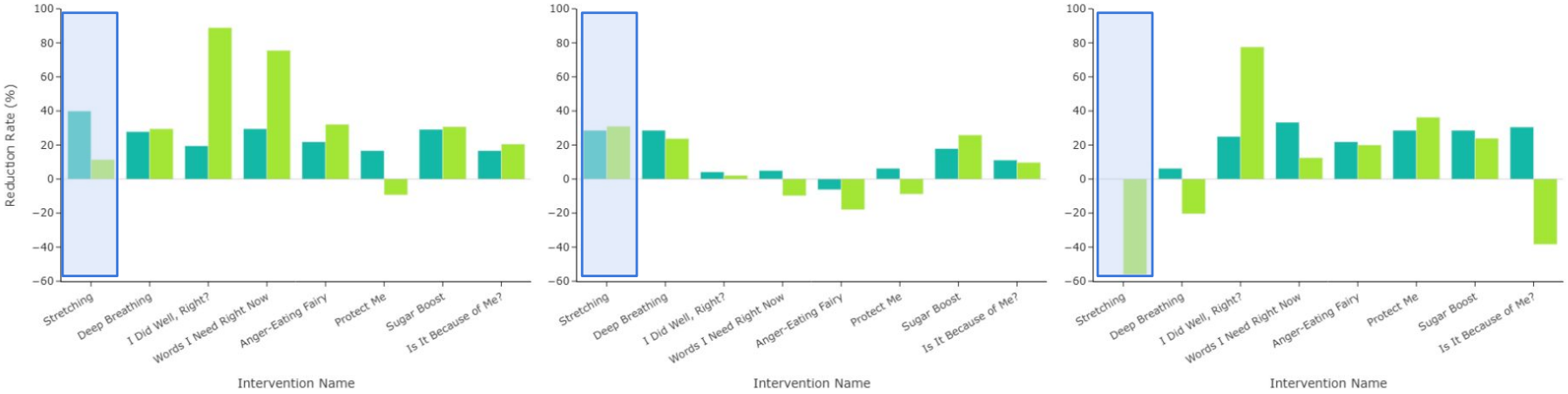
User satisfaction

Stress reduction

Results: Preference of intervention and understanding of post-intervention changes (RQ2)

Discrepancy on the Effect of Micro-interventions

Effect varied across participants & between perceived and physiological stress



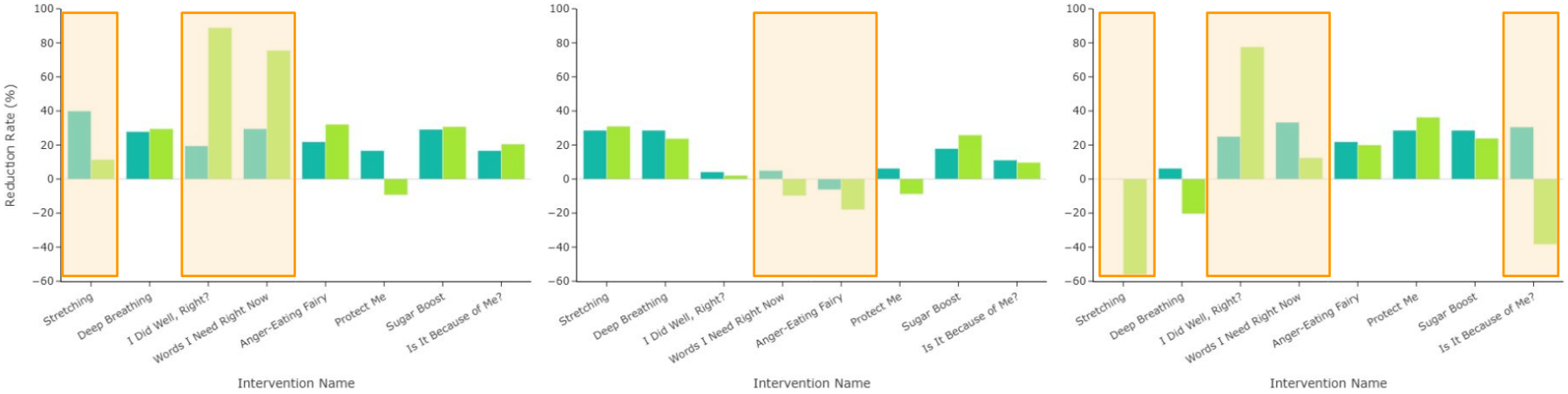
P04 P09 P14

Per-intervention **perceived** and **physiological** stress-reduction rates (%)

Results: Preference of intervention and understanding of post-intervention changes (RQ2)

Discrepancy on the Effect of Micro-interventions

Effect varied across participants & between perceived and physiological stress



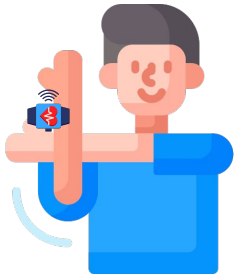
P04 P09 P14

Per-intervention **perceived** and **physiological** stress-reduction rates (%)

Interpreting Discrepancies Between Two Types of Stress Changes

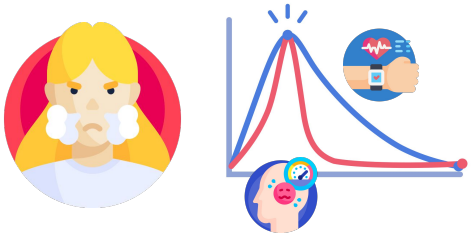
Participants did not regard this inconsistency as an error but instead contextualized these differences within their behaviors, conditions, and recovery pace

Attributing discrepancy to their **lifestyle or behavioral factors**



*"Maybe it's related to exercise intensity or heart rate; I **stretched quite actively**"*

Focusing on **temporal gap between cognitive and physiological recovery**



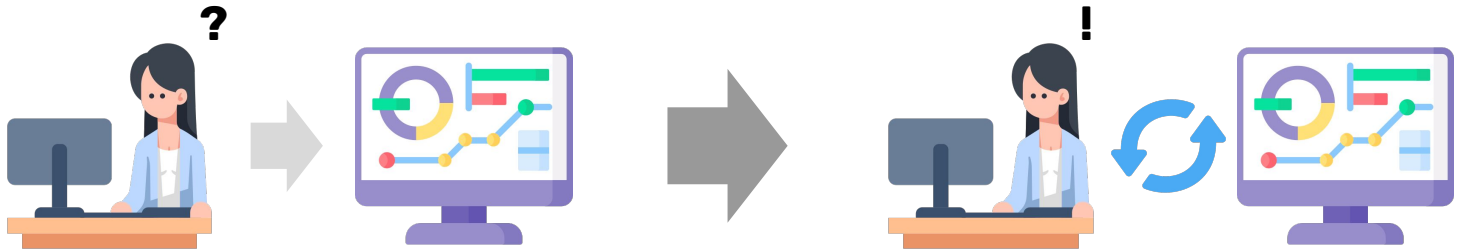
*"...when stress has been accumulating for a while, **physiological stress levels don't go down right away**"*

Discussion

Discussion

Supporting Stress Data Literacy in Call Center Context

- Users lacked accurate mental models for physiological stress
- To develop literacy, data should be repeatedly linked to lived experience.
- Design direction: Shift towards brief **reflection-in-action**



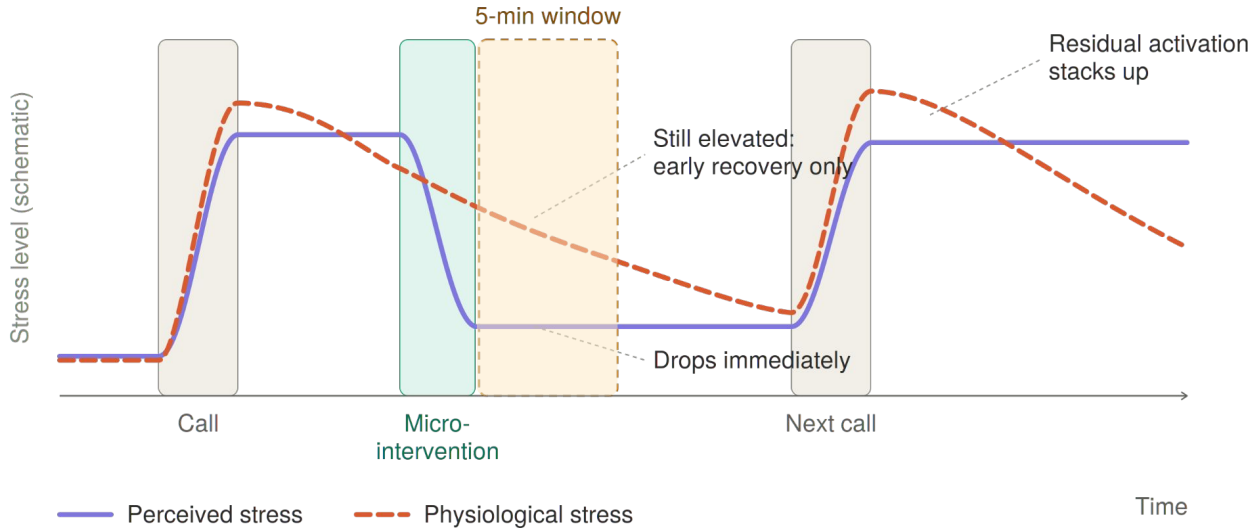
Reflection-**on-action** (post-hoc)

Reflection-**in-action** (in-situ)

Discussion

Understanding Discrepancy in Perceived and Physiological Stress

- Participants read mismatches through their condition, lifestyle, and recovery pace
- **Temporal mismatch**
  - Reactivity vs. recovery: body responds fast but recovers slowly, with **individual differences**[1]
  - Our 5-min window captures early recovery; back-to-back calls leave residual activation



[1] Melanie Bamert and Jennifer Inauen. 2022. Physiological stress reactivity and recovery: Some laboratory results transfer to daily life. Frontiers in Psychology 13 (2022), 943065.

Discussion

Understanding Discrepancy in Perceived and Physiological Stress

- Altogether, The discrepancy reflects the **temporal, individual, contextual complexity of stress regulation**
- Need to support within-person sensemaking **with complementary indicators**

Discussion

Discrepancy Between Objective Efficacy and User Value

Discrepancy between “What works” (Data) and “What They like” (Preference) to reduce stress

- Users valued immediate psychological relief more than data-driven efficacy
- This divergence presents a critical dilemma for system design



Actual stress reduction
(perceived/physiological)



User values
(preference/satisfaction)

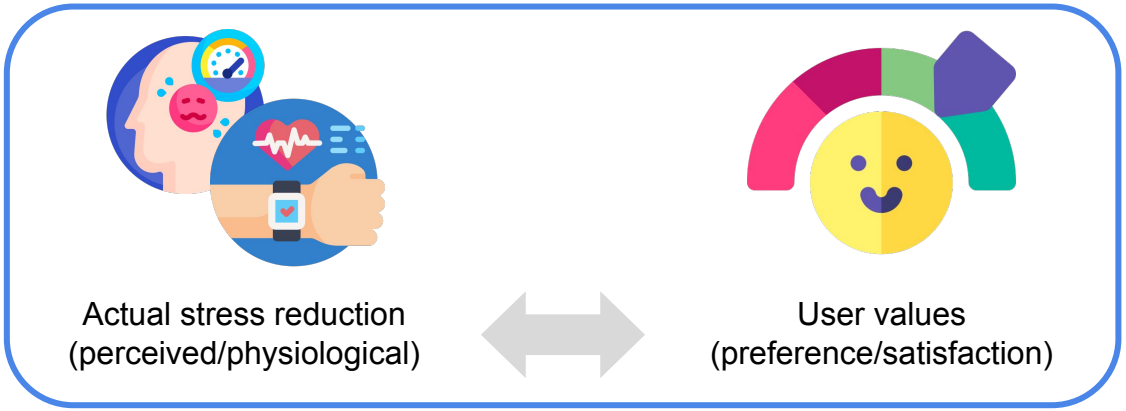
Discussion

Discrepancy Between Objective Efficacy and User Value

Discrepancy between “What works” (Data) and “What They like” (Preference) to reduce stress

- Users valued immediate psychological relief more than data-driven efficacy
- This divergence presents a critical dilemma for system design

Balance efficacy with user values (**Value-based interventions**)



Discussion

Discrepancy Between Objective Efficacy and User Value

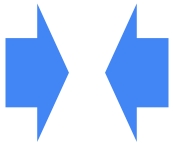
Discrepancy between “What works” (Data) and “What They like” (Preference) to reduce stress

- Users valued immediate psychological relief more than data-driven efficacy
- This divergence presents a critical dilemma for system design

Reducing the gap via reflection (**Value-based reflection**)



Actual stress reduction
(perceived/physiological)





User values
(preference/satisfaction)

Discussion

Limitations and Future Work

Limitations

- HRV as a noisy proxy for physiological stress in the wild (e.g., wrist motion)
- Retrospective reflection after 4 weeks → recall difficulty
- Intervention effects entangled with post-call recovery

Future Work

- Complementary signals with explicit uncertainty in the interface
- Reflection-in-action: brief, timely reflective support
- Extending to other high-stress occupations (e.g., soldiers, delivery workers)

When Feeling and Physiology Diverge: Understanding Dual-Indicator Stress Sensemaking & Micro-Interventions in an Emotional-Labor Workplace

Jeonghyun Kim, Yugyeong Jung, **Junmo Lee**, Kwangyoung Lee, Hwajung Hong, and Uichin Lee

Summary

- Investigated dual-indicator stress sensemaking in a call center
- Deployed a 4-week in-the-wild study with a custom dashboard
- Found that users anchor interpretation in context and value divergence as a signal

Contribution

- First systematic investigation of dual-indicator sensemaking in the wild
- Implications for surfacing divergence and scaffolding data literacy

Paper

Appendix

Presurvey Results

In the pre-survey, participants showed moderate levels of stress across most categories, with lower stress in job demands and less burden from impaired cognitive control

Survey type	Category	Score	Max score
KOSS-19	Physical environment	4.21±1.18	10
	Job demands	5.37±1.16	15
	Job autonomy	5.00±1.05	10
	Relationship conflict	6.47±1.12	10
	Job insecurity	5.05±1.08	10
	Organizational system	11.47±1.54	20
	Unfair compensation	5.84±0.83	10
	Work-life balance	5.52±0.70	10
K-BAT	Exhaustion	24.63±4.80	40
	Mental distance	11.11±2.82	20
	Impaired cognitive control	11.84±2.91	25
	Impaired emotional control	12.63±3.15	25

Participants

19 Call Center Workers (17 Female, 2 Male)

- Age: Mean = 38.90 (SD = 5.81)
- Years of Experience: Average tenure 3.58 years
- Shift Work: 8-hour shifts, continuous service coverage

PID	Gender	Age	Years of Experience	Pilot	PID	Gender	Age	Years of Experience	Pilot
P01	F	38	≥ 2 years	Y	P11	F	34	≥ 2 years	N
P02	F	43	≥ 5 years	Y	P12	F	50	≥ 5 years	N
P03	F	35	≥ 5 years	Y	P13	F	50	≥ 5 years	N
P04	M	37	≥ 3 years	N	P14	F	31	≥ 5 years	N
P05	F	38	≥ 3 years	N	P15	F	42	≥ 3 years	N
P06	F	38	≥ 1 year	N	P16	F	44	≥ 5 years	N
P07	M	42	≥ 3 years	N	P17	F	31	≥ 3 years	N
P08	F	34	≥ 5 years	N	P18	F	44	≥ 3 years	N
P09	F	42	≥ 5 years	N	P19	F	33	≥ 2 years	N
P10	F	33	≥ 3 years	N					

In-the-Wild Data Collection Setup

Method	Source	Device	Data	Frequency
Passive sensing	Call log server	Call center server	Call start time, Call end time, Shift schedule	Per call
	Wrist band	Galaxy Watch 7	HR	1 Hz
			IBI	-
			Skin temperature	$\frac{1}{60}$ Hz
			ACC	25 Hz
	Phone	Galaxy S21–22	Steps	0.2Hz
	Environmental sensor	bluSensor (BSP02AIR)	Temperature, Humidity, CO_2, tVOC	0.1 Hz
Self-report	Tablet	Galaxy Tab (A7 Lite)	*Pre-shift survey*: Sleep quality, General health, Perceived stress, Arousal, Valence, Fatigue	Per day
			Post-call survey: Previous call type, Workload, Perceived stress, Stressor, Arousal, Valence, Fatigue, Surface acting	≥ 15 per day
			Post-intervention survey: Intervention satisfaction, Stress relief effect, Perceived stress	≥ 9 per day
			Post-intervention refusal: Reason for refusal	-

Appendix - Study Design

Collected Survey Items

Survey	Category	Question	Answers
Pre-shift	Sleep quality	How would you rate the overall quality of your sleep last night?	Very poor (1) - Very good (5)
	General health	How would you rate your overall health at this moment?	Very poor (1) - Very good (5)
	Perceived stress	How stressed do you feel right now?	Not at all (1) - Very much (5)
	Arousal	How would you describe your current emotional state?	Very calm (1) - Very energetic (5)
	Valence	How would you describe your current emotional state?	Very negative (1) - Very positive (5)
	Fatigue	How fatigued do you feel right now?	Not at all (1) - Very much (5)
Post-call	Previous call type	Please select the type of the previous call you handled.	General inquiry / Complaint
	Workload	Do you feel that your current workload is too low or too high?	Very low (1) - Very high (5)
	Perceived stress	How stressed did you feel during the previous call?	Not at all (1) - Very much (5)
	Stressor	Please select all factors that caused you stress. (Multiple selections allowed)	Felt that my abilities were insufficient / Difficulty in understanding the tasks / Pressure from performance evaluation / Dissatisfaction with work procedures (e.g., having to follow fixed manuals) / Communication issues with customers (e.g., self-centered customers repeating the same thing) / Rude customers (e.g., verbal abuse, shouting) / Time pressure (e.g., backlog of tasks) / Noise from surrounding people / Conflicts or issues with colleagues / Other
	Arousal	During the previous call, how would you describe your emotional state?	Very calm (1) - Very energetic (5)
	Valence	During the previous call, how would you describe your emotional state?	Very negative (1) - Very positive (5)
	Fatigue	How fatigued did you feel during the previous call?	Not at all (1) - Very much (5)
	Surface acting	During the previous call, to what extent did you try to hide your true emotions?	Not at all (1) - Very much (5)
Post-intervention	Intervention satisfaction	How satisfied were you with the intervention you completed?	Not at all (1) - Very much (5)
	Perceived effectiveness	To what extent did you feel that this intervention helped relieve your stress?	Not at all (1) - Very much (5)
	Perceived stress	How stressed do you feel right now after completing the intervention?	Not at all (1) - Very much (5)

Data Reflection and Post-interview

Participants engaged in data reflection tasks by exploring their data via the dashboard and joined a semi-structured interview regarding what they observed.

State Mode (RQ1)

- Briefly reported their dashboard experience and newly recognized stress patterns
- Shared the data points they referenced when interpreting perceived and physiological mismatches (e.g., felt experience, call type/difficulty, environmental factors) along with the contextual factors judged most closely associated with each indicator

Change Mode (RQ2)

- Described their criteria for judging an intervention's effectiveness
  - e.g., the duration of the effect, and the (dis)agreement of pre- and post-intervention changes in perceived and physiological stress
- Reported the intervention they felt was most effective, its (mis)alignment with data-derived rankings, and any noticed bodily responses

Data Processing

Data Cleansing

- Removed missing values, modality-specific outliers, and invalid intervals before computing features
- From 5,699 post-call survey-anchored samples, 15.02% (n = 856) were excluded
  - Analyses used the remaining 4,843 valid samples, with an average of 254.89 responses per person (SD = 30.76)
  - The main reason for exclusion: the lack of validly timestamped samples in at least one of the data types (e.g., improper wristband wearing)

Windowing

- 5-min window anchored to post-call survey timestamp [t_s – 5min, t_s]
  - Exception: step counts (60-min window)
- Aggregated sensor features within each window using the mean (and standard deviation for the accelerometer)
  - Survey responses were aligned to sensor features by timestamp

Interpreting Discrepancies Between Two Types of Stress Changes

No consistent relationship between changes in perceived and physiological stress

- Per-participant Spearman's correlation between perceived and physiological stress reduction

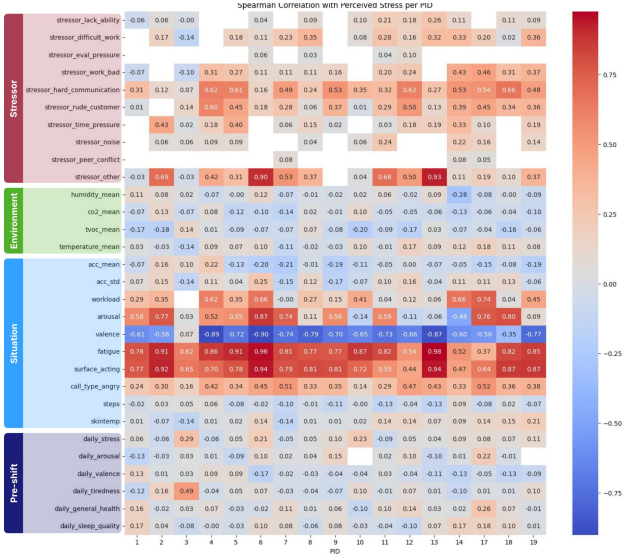
PID	Correlation	PID	Correlation	PID	Correlation
P01	0.038	P07	0.124	P13	0.058
P02	-0.092	P08	-0.111	P14	-0.148
P03	0.058	P09	0.152	P17	0.301
P04	-0.194	P10	0.081	P18	0.112
P05	0.066	P11	0.053	P19	0.084
P06	0.095	P12	-0.025		

- Perceived and Physiological stress reductions often moved in opposite directions
  - e.g., A breathing exercise might lower HRV (physio) but not annoyance (perceived)

Appendix - Results (RQ1)

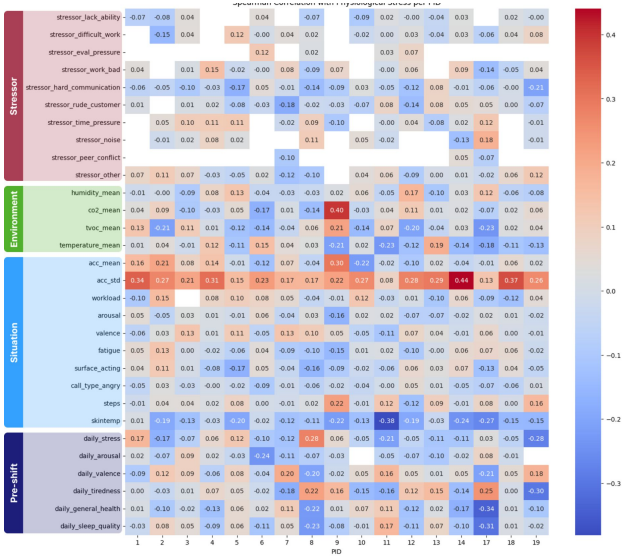
Correlation Between Contexts and Stress Indicators

Correlation with Perceived Stress per PID



Strongly modulated
by emotional/social stimuli

Correlation with Physiological Stress per PID



Correlations were smaller in
magnitude and more dispersed

Appendix - Results (RQ2)

Intervention Preference: Was It Actually Effective?

Per-participant Spearman's correlation between intervention satisfaction and other measures

PID	Correlation		
	Perceived effectiveness	Perceived reduction	Physiological reduction
P01	0.945***	0.525	0.195
P02	1***	0.279	-0.031
P03	0.989***	-0.03	0.028
P04	0.811*	0.611	-0.028
P05	1***	0.554	0.515
P06	1***	0.676	0.399
P07	0.991***	0.361	0.299
P08	0.989***	0.59	0.576
P09	0.977***	0.937**	0.856
P10	0.988***	0.476	0.648
P11	0.975***	0.494	0.101
P12	0.968***	0.668	-0.522
P13	0.86**	0.357	0.414
P14	0.787*	0.375	-0.233
P17	0.969***	0.874	0.745
P18	0.957***	0.687	0.153
P19	0.941***	0.463	0.034

Users "liked" interventions made them feel better, even if their actual stress didn't drop

Appendix - Results (RQ2)

Intervention Evaluation: Preference and Perceived Effectiveness

Preferences for each intervention and their reasons varied across individuals

Intervention (*N*)	Preference reasons	Dislike reasons
Stretching (14)	Refreshing, Following instructions is helpful	Lack of time, Lack of space, Feeling self-conscious (noticed by others)
Protect Me (14)	Relieve anger, Stress releases through fidgeting (like a game), Refreshing	Cannot relate to the idea of a character protecting me
Deep Breathing (14)	Refreshing, Calms the mind, Allows time to prepare for the next call	Lack of time, Difficult to follow instructions as directed
Sugar Boost (13)	Eating snacks relieves stress	Health issue or is not preferred, Effect is not immediate
Anger-Eating Fairy (10)	Emotional dumping ground, Can calmly reflect on and organize the stressful situation	Makes one reflect on stress, Cannot think up a word
I Did Well, Right? (9)	Can reflect on one's own feelings, Self-reassurance, Receives comfort	Don't want to recall stressful situation, Cannot think up a word as time passes, Lack of time
Words I Need Right Now (9)	Can reflect on one's own feelings, Enjoy the feeling of being talked to	Vague feeling, Cannot think up a word to say as time passes
Is It Because of Me? (7)	Separate emotion from work which helps to hold it in	Cannot relate as the content does not apply to their situation

Evaluation Results for Each Intervention

Participants' perceptions of each intervention and effects on stress

Intervention	Satisfaction	Perceived effectiveness	Stress reduction (%)	
			Perceived	Physiological
Stretching	3.62 ± 0.73	3.4 ± 0.67	16.82 ± 20.81	16.3 ± 36.32
Deep Breathing	3.87 ± 0.59	3.74 ± 0.51	14.04 ± 23.74	4.59 ± 17.32
I Did Well, Right?	3.22 ± 0.77	3.14 ± 0.67	14.28 ± 14.39	2.54 ± 42.18
Words I Need Right Now	3.37 ± 0.75	3.22 ± 0.73	14.6 ± 16.36	10.89 ± 32.95
Anger-Eating Fairy	3.47 ± 0.87	3.35 ± 0.79	16.06 ± 19.75	16.3 ± 23.67
Protect Me	3.7 ± 0.7	3.56 ± 0.68	17.49 ± 16.4	8.33 ± 21.61
Sugar Boost	3.5 ± 0.67	3.28 ± 0.64	15.65 ± 17.92	10.73 ± 21.13
Is it Because of Me?	2.83 ± 0.81	2.75 ± 0.77	8.33 ± 18.82	-0.21 ± 20.75

Intervention Evaluation: Reasons for refusing intervention

During the experiment, 2.56% of the recommended interventions (31 out of 1,211) were refused

Reason for refusal	Count
Too much work or other tasks to attend to	22
Did not feel like doing it at the moment	7
Felt that the activity would not be helpful	5
Felt okay as I am right now	4
Concerned about others' opinions	2
Total	**40**

Unused Slides

Research Limitations in HCI

#1: Sensing is multi-dimensional, but the feedback is often monolithic

- Dual-indicator interpretation is not supported

#2: Intervention evaluation rarely considers mismatched outcomes

- The evaluation measures for interventions are single-lens (e.g., satisfaction OR HRV)

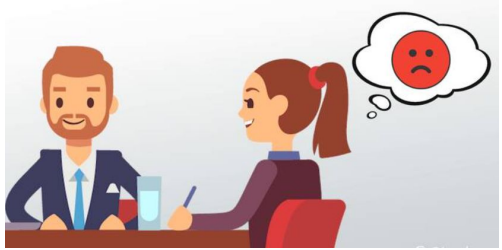
#3: Users' understanding in integrating both stress indicators is underexplored

- Users cannot check whether their bodily responses really match how they feel

Stress in Emotional Labor Workplaces

High emotional labor demands in call center context

- In call center setting, employees must regulate inner feelings to display organizationally required emotions (e.g., friendliness) [1]
  - This ***surface acting*** **creates** **emotional dissonance**, a gap between felt and displayed emotions, which leads to chronic stress, burnout, and blunted physiological responses [2]



- In such environments, **interoceptive awareness** is often compromised
  - Interoceptive awareness - how accurately people notice and interpret signals from within their own body (e.g., heartbeat, tension)
  - Heavy cognitive load can blunt a call center worker's ability to accurately perceive their own physiological arousal [3], increasing the likelihood of misalignment

[1] A. R. Hochschild, "The managed heart," in Working in America, 2022.
[2] A. A. Grandey, "Emotional regulation in the workplace: A new way to conceptualize emotional labor.," Journal of occupational health psychology, 2000.
[3] S. N. Garfinkel, et al., "Knowing your own heart: distinguishing interoceptive accuracy from interoceptive awareness," Biological psychology, 2015.

Limitations of Existing Stress Tracking Systems

Prior research in personal informatics has utilized a combination of psychological reports, physiological signals (e.g., HRV, EDA), and contextual cues to track stress [1-2]

While stress sensing is multi-dimensional, the feedback target is often monolithic:

- Current systems rely on single-indicator evaluation (as a single stress score), overlooking divergent outcomes
  - e.g., stress label using self-reported stress, automatically tracking biosignals only
- This aggregation approach oversimplifies their stress and obscures the nuance between mind (feeling) and body (physiology)
  - e.g., breathing exercises reducing heart rate but not anxiety

Users may distrust data when it doesn't match their feelings,
which leads to low trust & disengagement of the system [3]

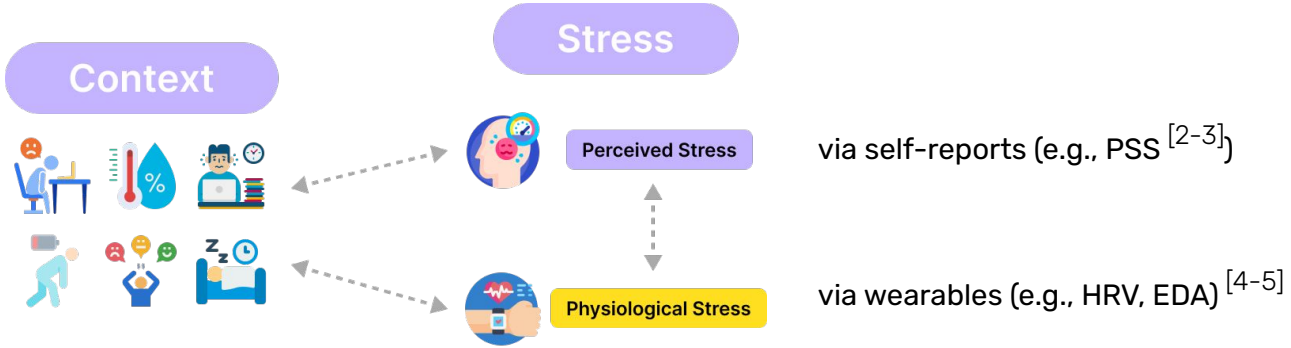
[1] Y. Jiang, et al., "Intimasea: exploring shared stress display in close relationships," in Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, 2023.
[2] V. Mishra, et al., "Investigating the role of context in perceived stress detection in the wild," in Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers, 2018.
[3] S. Kye, et al., "Multimodal data collection framework for mental stress monitoring," in Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers, 2017.

Defining, Measuring, and Contextualizing Stress

Stress is a complex phenomenon spanning subjective appraisal (perceived) and bodily response (physiological) [1]

- Since **perceived and physiological stress indicators represent distinct facets of the stress** experience that might move in opposite directions, neither alone provides a complete picture

To interpret stress signals, **contextual information** is also essential to identify triggers



Need to <u>view both stress indicators alongside contextual traces</u>

[1] A. M. Shahsavarani, et al., "Stress: Facts and theories through literature review," International Journal of Medical Reviews, 2015.
[2] G. Jung et al., "Counterstress: Enhancing stress coping planning through counterfactual explanations in personal informatics," in Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems, 2025.
[3] E. Jo, et al., "Understanding parenting stress through co-designed self-trackers," in Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, 2020.
[4] K. A. James, et al., "Understanding the relationships between physiological and psychosocial stress, cortisol and cognition," Frontiers in endocrinology, 2023.
[5] A. Schulz, et al., "Interoception, stress, and physical symptoms in stress associated diseases," European Journal of Health Psychology, 2021.

Results: Understanding of perceived and physiological stress with contextual data (RQ1)

How Did Participants Understand the Discrepancies?

Participants viewed mismatches (feeling vs. physiology) as **informative signals,** not errors

Trusting objectively measured indicators even when they differed from their subjective judgment



*"Perceived stress reflects my subjective judgment, while **physiological stress reflects bodily responses**. So I think physiological stress is accurate"*

Interpreting the discrepancy based on **situational factors**



*"When I had COVID-19, I was **coughing so much** that I couldn't handle calls properly…"*

Attributing higher perceived stress to **emotional suppression and accumulated tension**



*"As I keep doing work, stress builds up, and when emotions are suppressed like that, I **get upset over even small things**"*

Result Analysis

Qualitative Analysis

- Inductive thematic analysis of post-interviews

Quantitative Analysis

- For RQ1: Spearman correlations (stress indicators vs. contextual factors)
- For RQ2: Correlation between satisfaction vs. perceived effectiveness / actual stress reduction


Satisfaction


Perceived Effectiveness


Perceived Stress Reduction Rate


Physiological Stress Reduction Rate

* P15 and P16 were excluded from all analysis due to survey response quality issues
  ⇒ Results therefore report N=17 (P01-P14, P17-P19)

Study Design

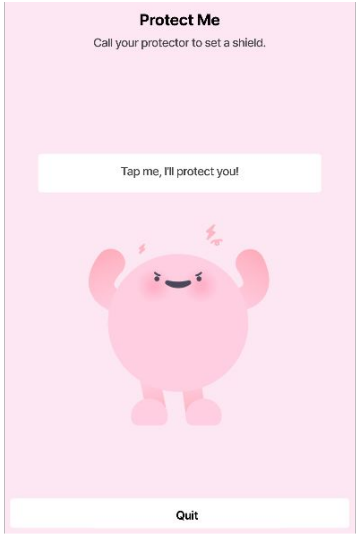
In-the-Wild Data Collection Setup



Pre-shift survey



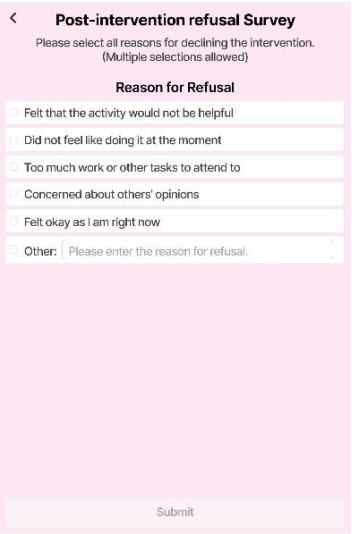
Post-call survey



Micro-intervention



Post-intervention survey



Post-intervention refusal survey

Discussion

Design Implications

#1: Expose Divergence

- Explicitly visualize the mismatch between feeling and physiology
- Do not hide it behind a single score
- Highlight “Recovery Lags” to help users understand their body’s delayed response

#2: Scaffold Literacy

- Provide Context-Linked Explanations
- Move beyond text definitions; use the user's own data to explain physiology
    - e.g., “Your heart rate is high because of the environmental noise, not just the call”

#3: Personalize Interventions

- Personalize based on Divergent Realized Effects
- Recommend interventions that target the specific deficit
    - e.g., “You need physiological recovery, try breathing” vs “You need mood lift, try stretching”

Conclusion

Key Takeaways

Future of Stress Management: We must move beyond monolithic stress tracking

- Surfacing complexity (Divergence)
- Fostering data literacy
- Balancing efficacy with user value