

AI for *Psy*ber-Physical Systems

Prof. Dr. Shoya Ishimaru (Osaka Metropolitan University)

ASPIRE "AI-Physical Systems" Kick-Off Meeting, 04.03.2026

Shoya Ishimaru



Master's student @OPU



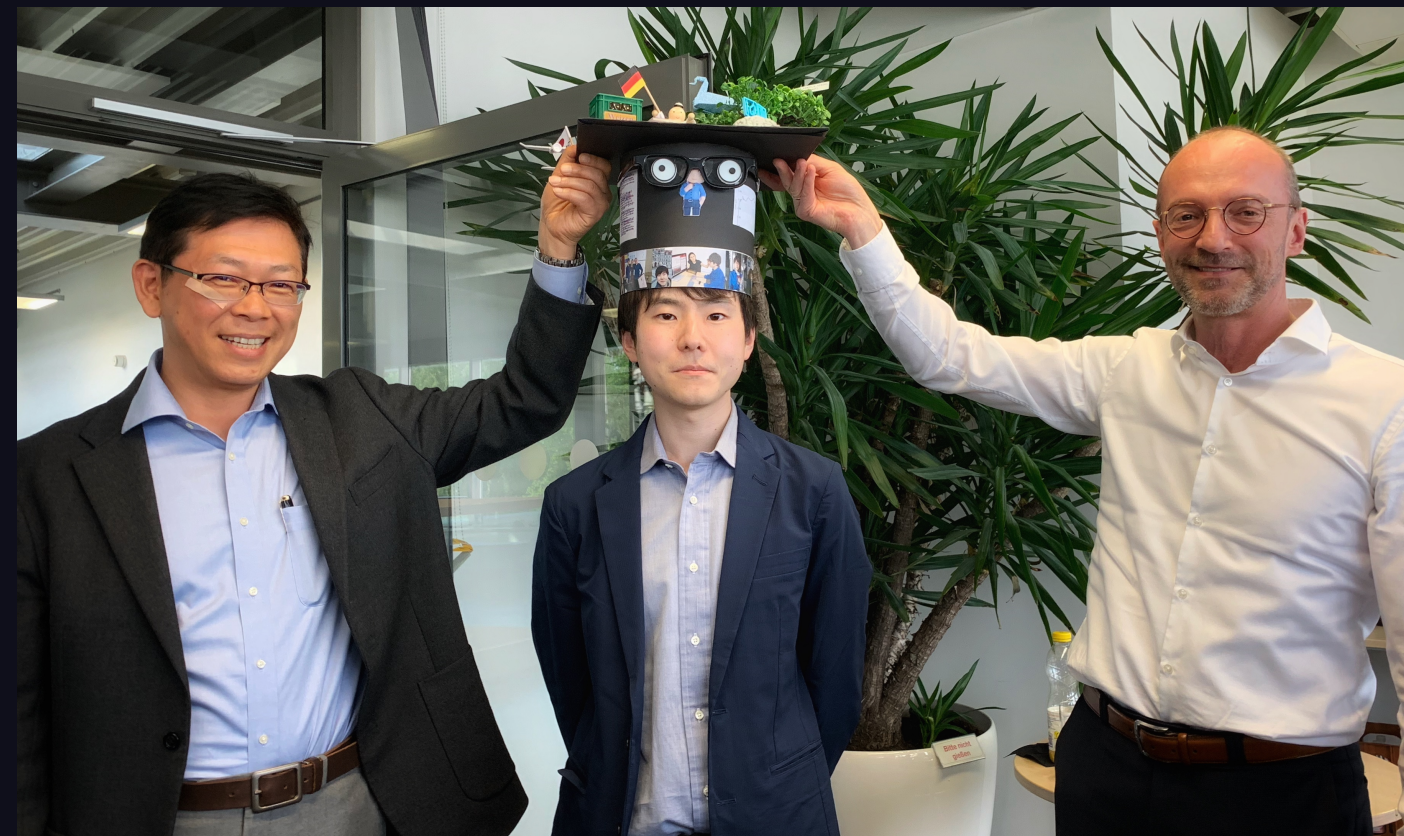
Super Creator @IPA Mitou



Professor @OMU



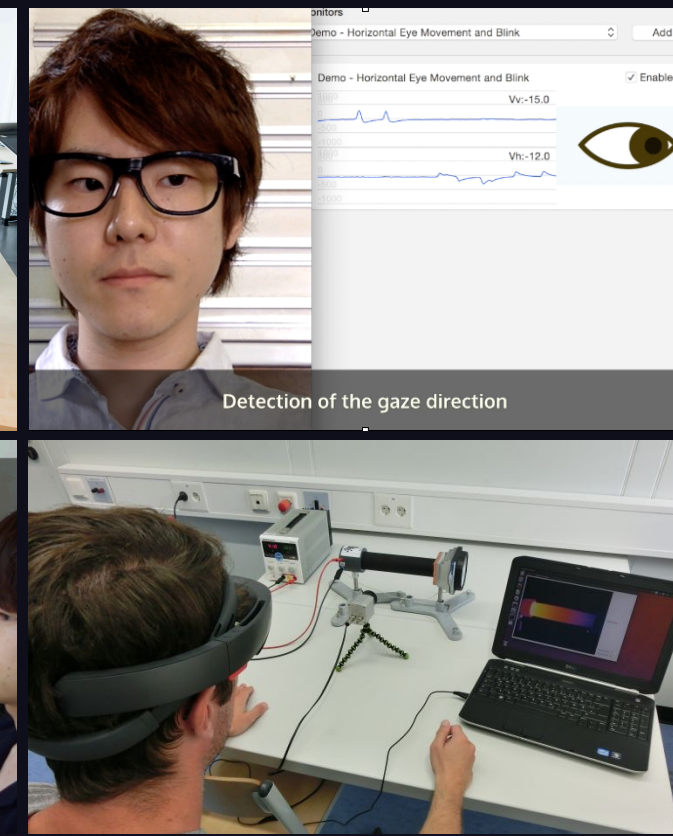
CEO @Affectify Inc.



Ph.D. student @TU Kaiserslautern



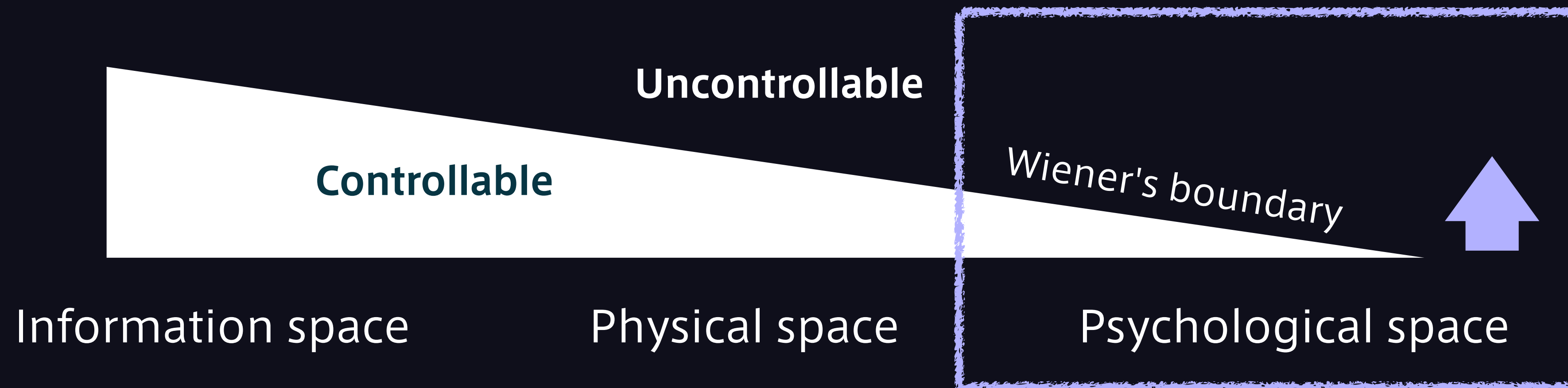
Co-Director @DFKI iQL-Lab
Junior Professor @RPTU



CRO @Alphaben GbR

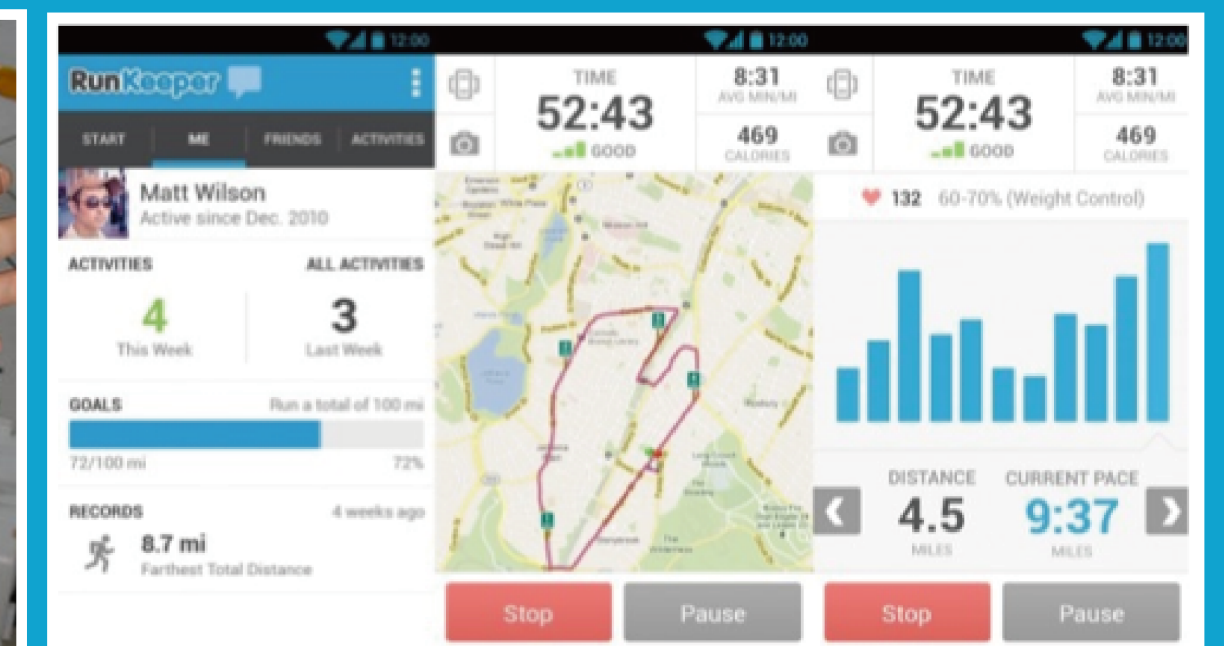
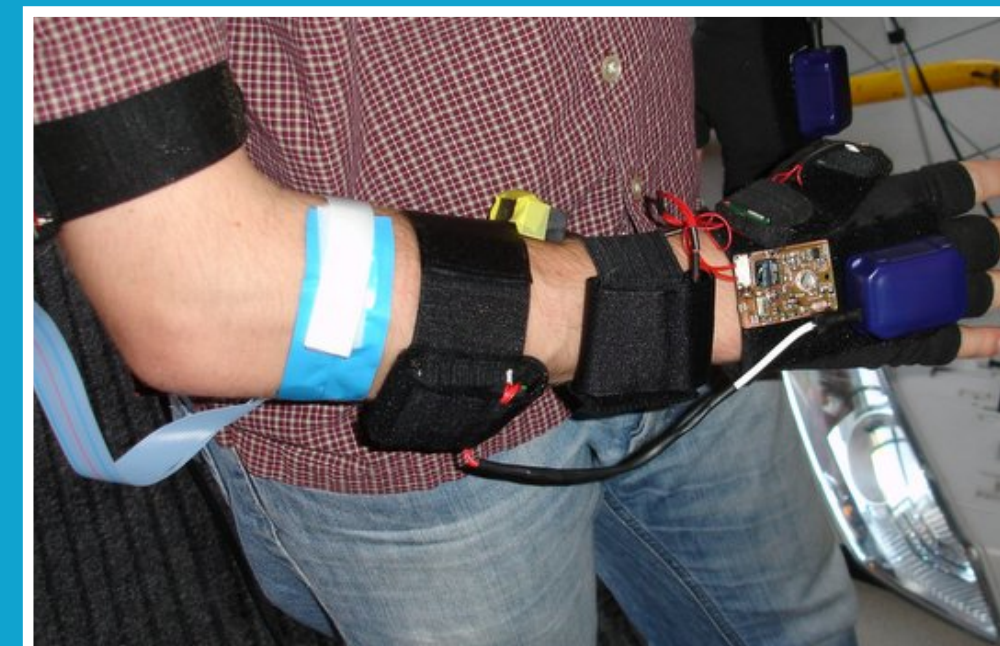
Psybernetics (PsyberLab)

- Psychological Cybernetics
- Expanding the controllable domain of Cybernetics, popularized by Norbert Wiener, not only to information space and physical space, but also to psychological space (human emotions, knowledge states, etc.)



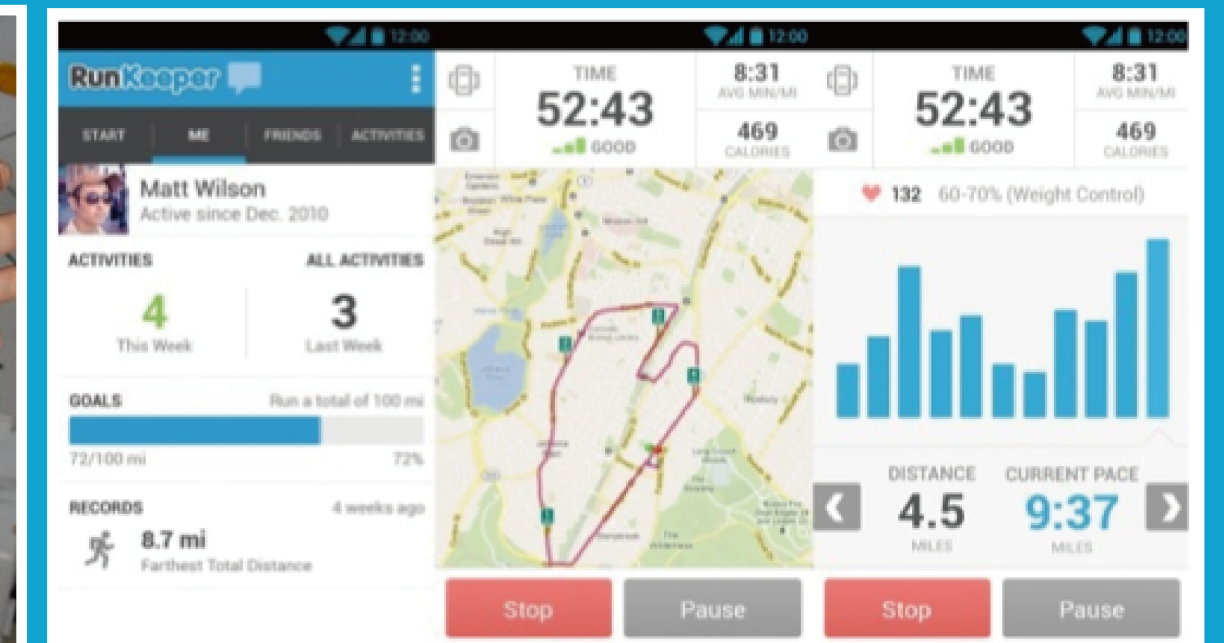
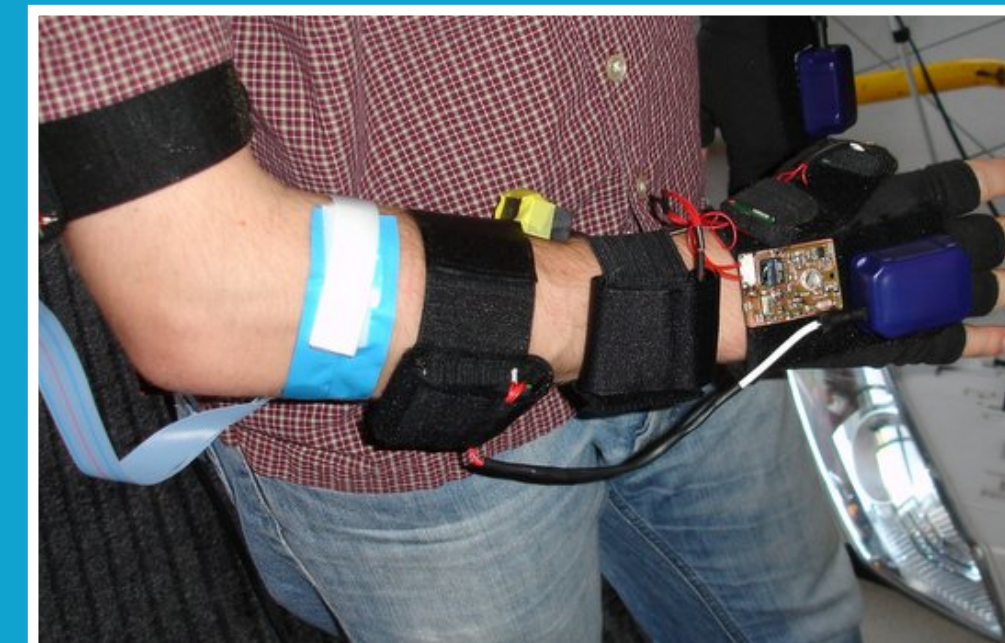
Iceberg Model of Activity Recognition

Physical Activity
walking, standing,
cycling, sleeping, ...

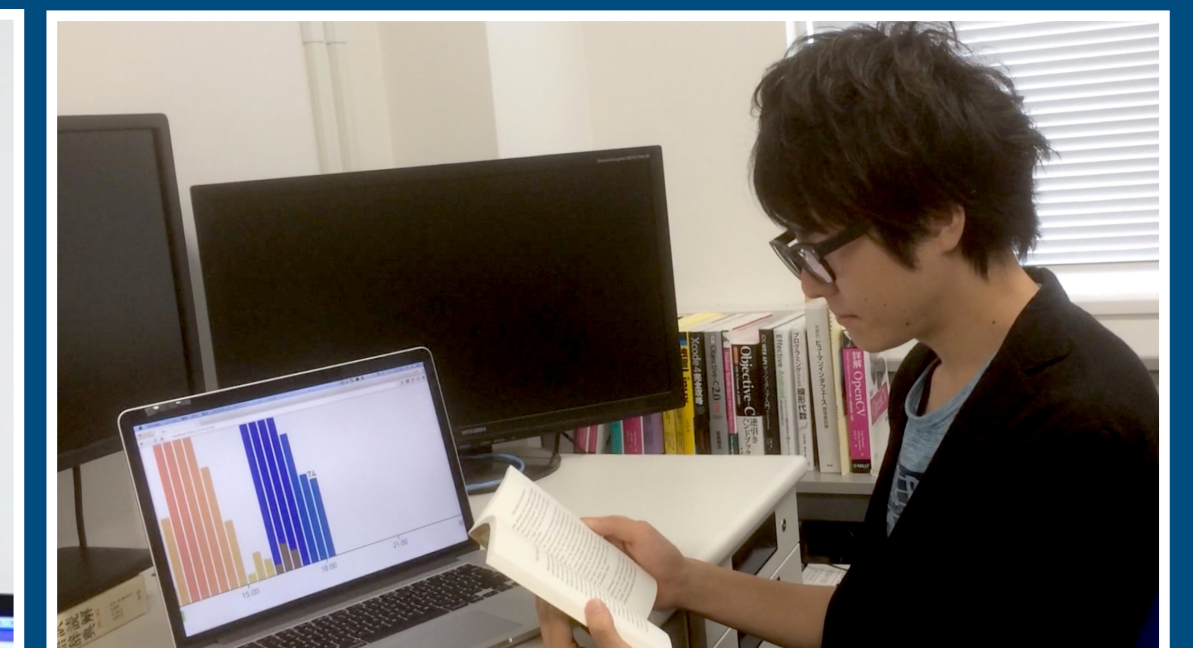


Iceberg Model of Activity Recognition

Physical Activity
walking, standing,
cycling, sleeping, ...



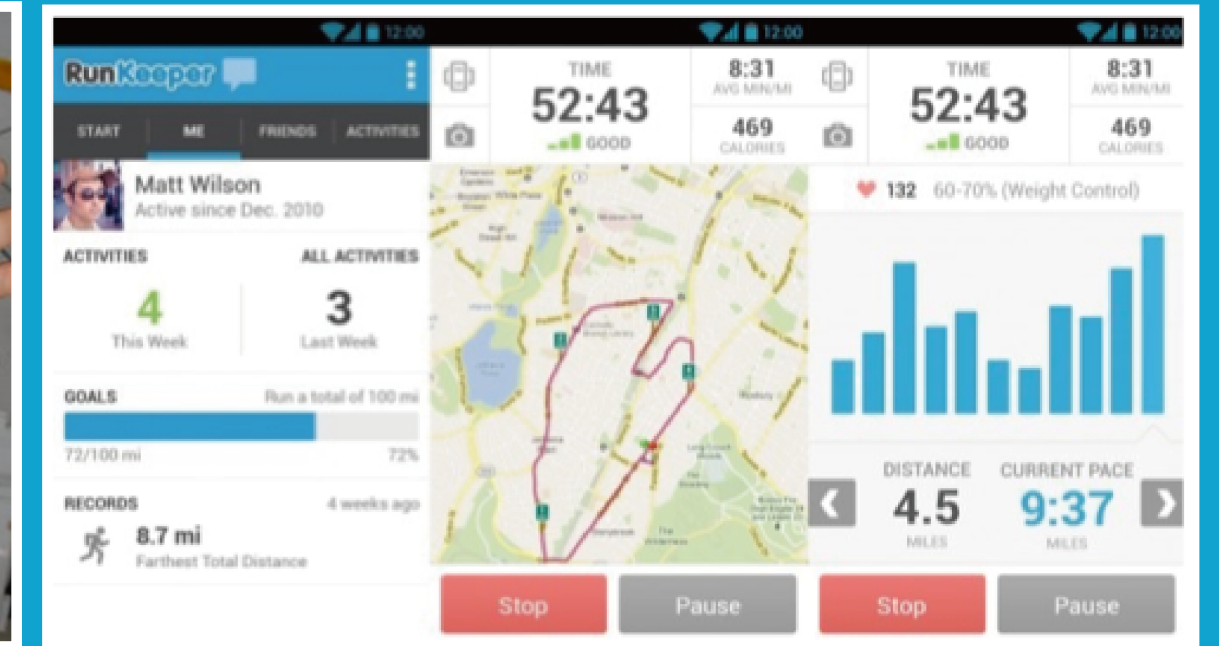
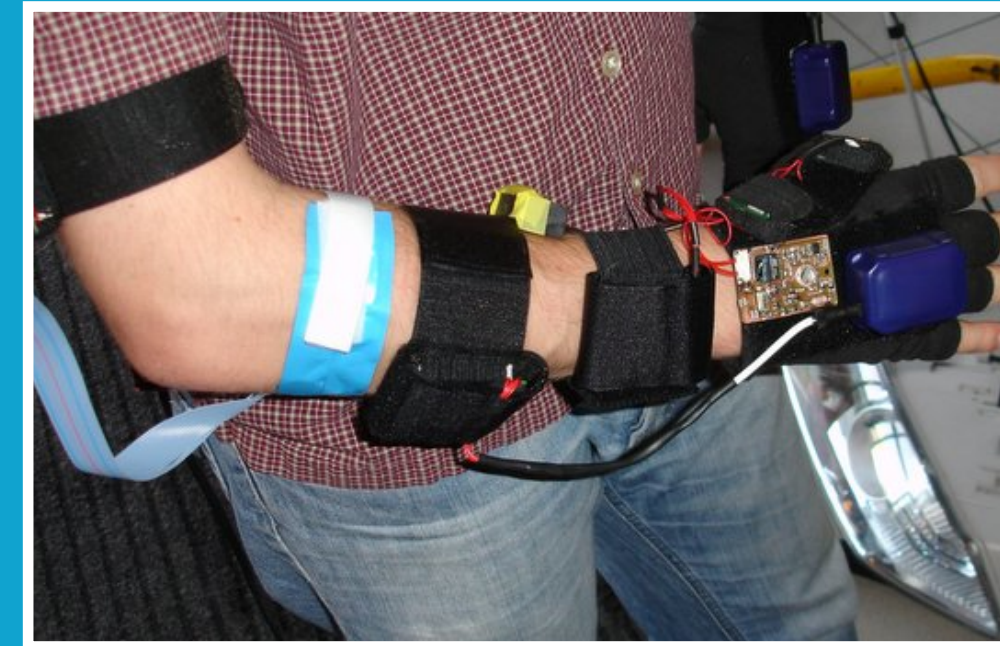
Cognitive Activity
reading, writing,
memorizing, talking, ...



Iceberg Model of Activity Recognition

Physical Activity

walking, standing,
cycling, sleeping, ...



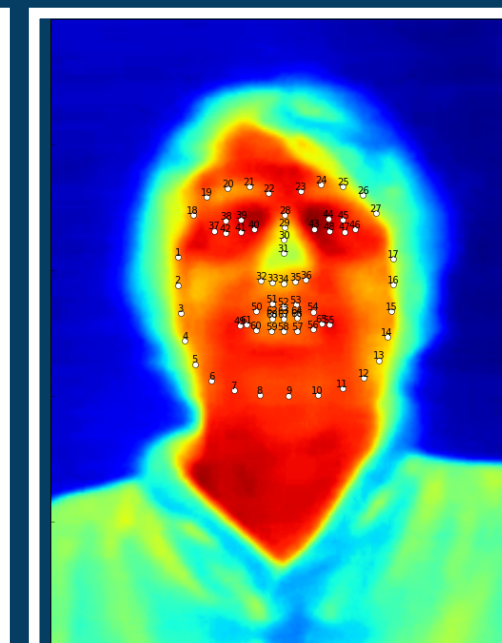
Cognitive Activity

reading, writing,
memorizing, talking, ...



Cognitive and Affective State

interest, workload,
confidence, fatigue, ...



Making people smarter... What does "smart" mean?

Learning

Acquiring knowledge in a usable form

We propose intelligent learning assistants that recognize cognitive/affective states (e.g., attention, interest, workload, self-confidence) and optimize learning materials for individuals to improve their learning performance.

Reasoning

Making a decision based on knowledge

Smart sensing technologies enable people to track their daily activities and make rational decisions on the basis of activity logs and future projections. We also investigate nudging techniques that encourage change in their behavior.

Communicating

Sharing knowledge with others

Communicating with others is an important activity for coming up with new ideas and transferring skills. We quantify social activities such as teaching, presenting, and discussing by sensors to support collaborative work.

PsyberLab defines intelligence as the ability of learning, reasoning, and communicating

Making people smarter... What does "smart" mean?

Learning

Acquiring knowledge in a usable form

We propose intelligent learning assistants that recognize cognitive/affective states (e.g., attention, interest, workload, self-confidence) and optimize learning materials for individuals to improve their learning performance.

Reasoning

Making a decision based on knowledge

Smart sensing technologies enable people to track their daily activities and make rational decisions on the basis of activity logs and future projections. We also investigate nudging techniques that encourage change in their behavior.

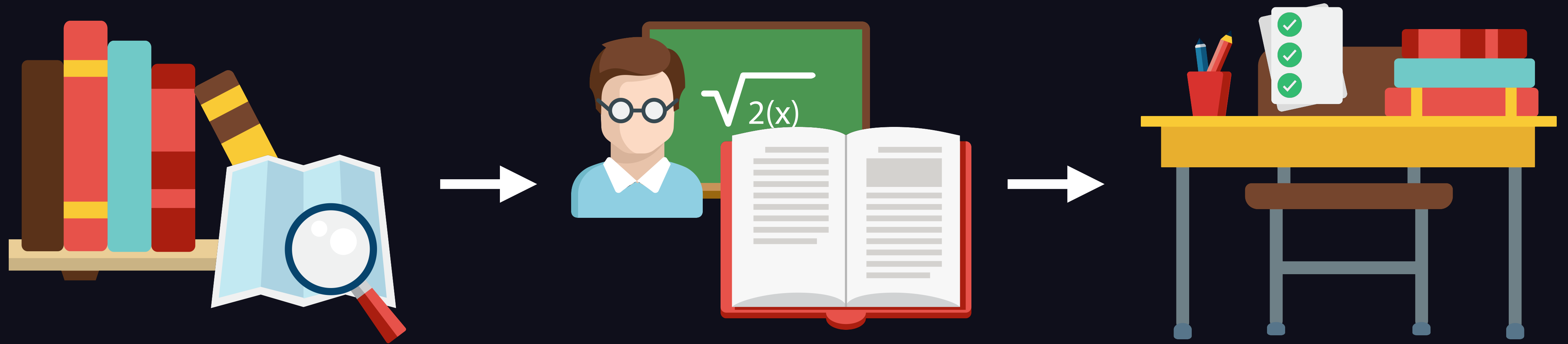
Communicating

Sharing knowledge with others

Communicating with others is an important activity for coming up with new ideas and transferring skills. We quantify social activities such as teaching, presenting, and discussing by sensors to support collaborative work.

PsyberLab defines intelligence as the ability of learning, reasoning, and communicating

Augmented Positive Learning Cycle



Daily habits and curiosity

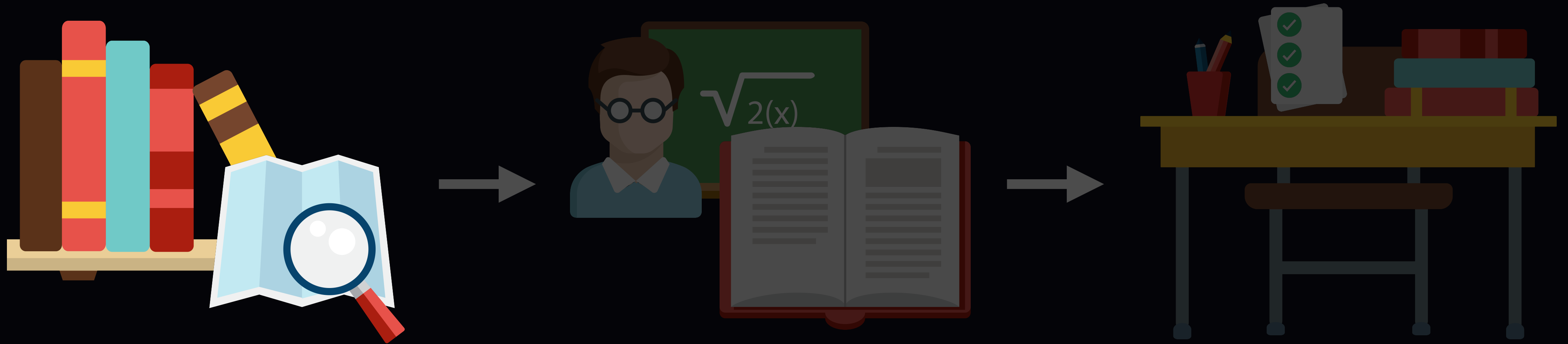
Lecture and textbook

Questions (test)

Can AI understand interest of students and nudge them to learn voluntarily?

Which question should be reviewed carefully? (considering priority & leakage)

Augmented Positive Learning Cycle



Daily habits and curiosity

Lecture and textbook

Questions (test)

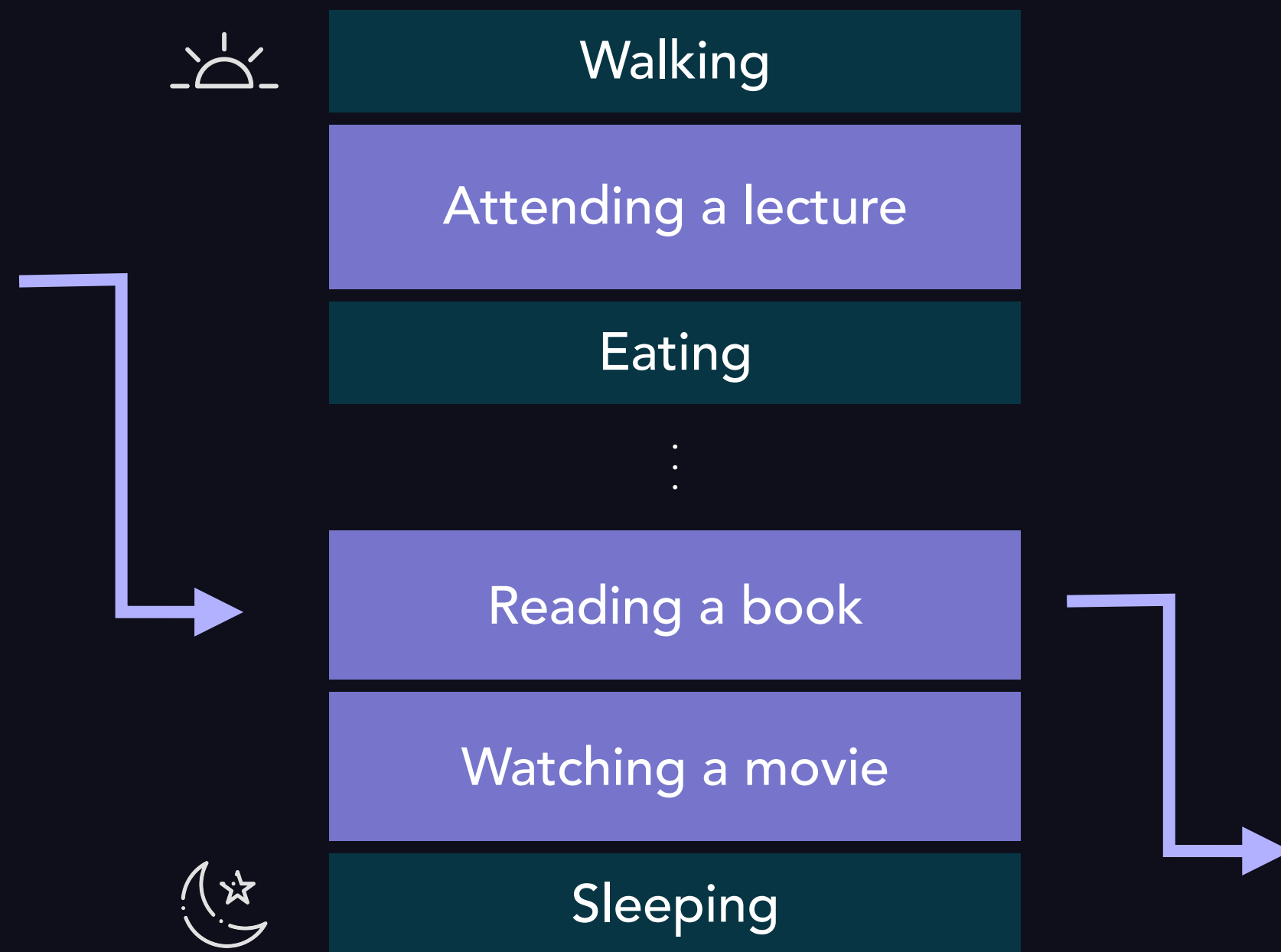
Can AI understand interest of students and nudge them to learn voluntarily?

Which question should be reviewed carefully? (considering priority & leakage)

Understanding Daily Activities, from Physical to Cognitive



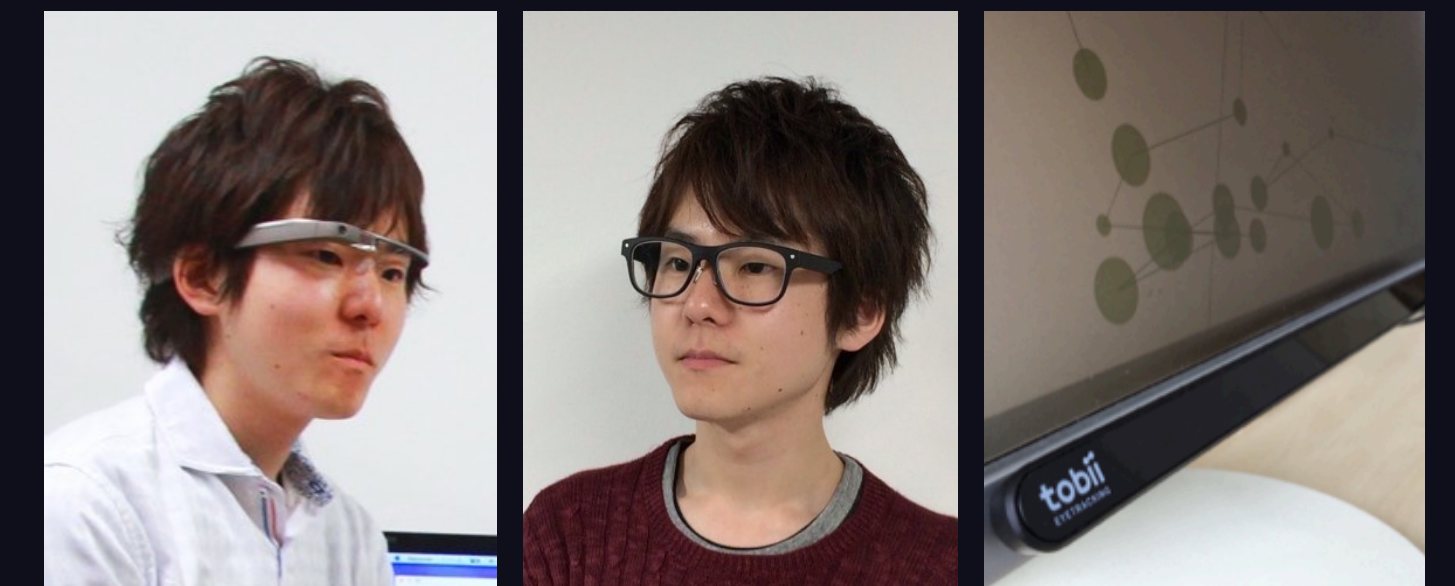
Commercially available activity trackers



Students' daily activities



Wearable device [Bulling, et al.]



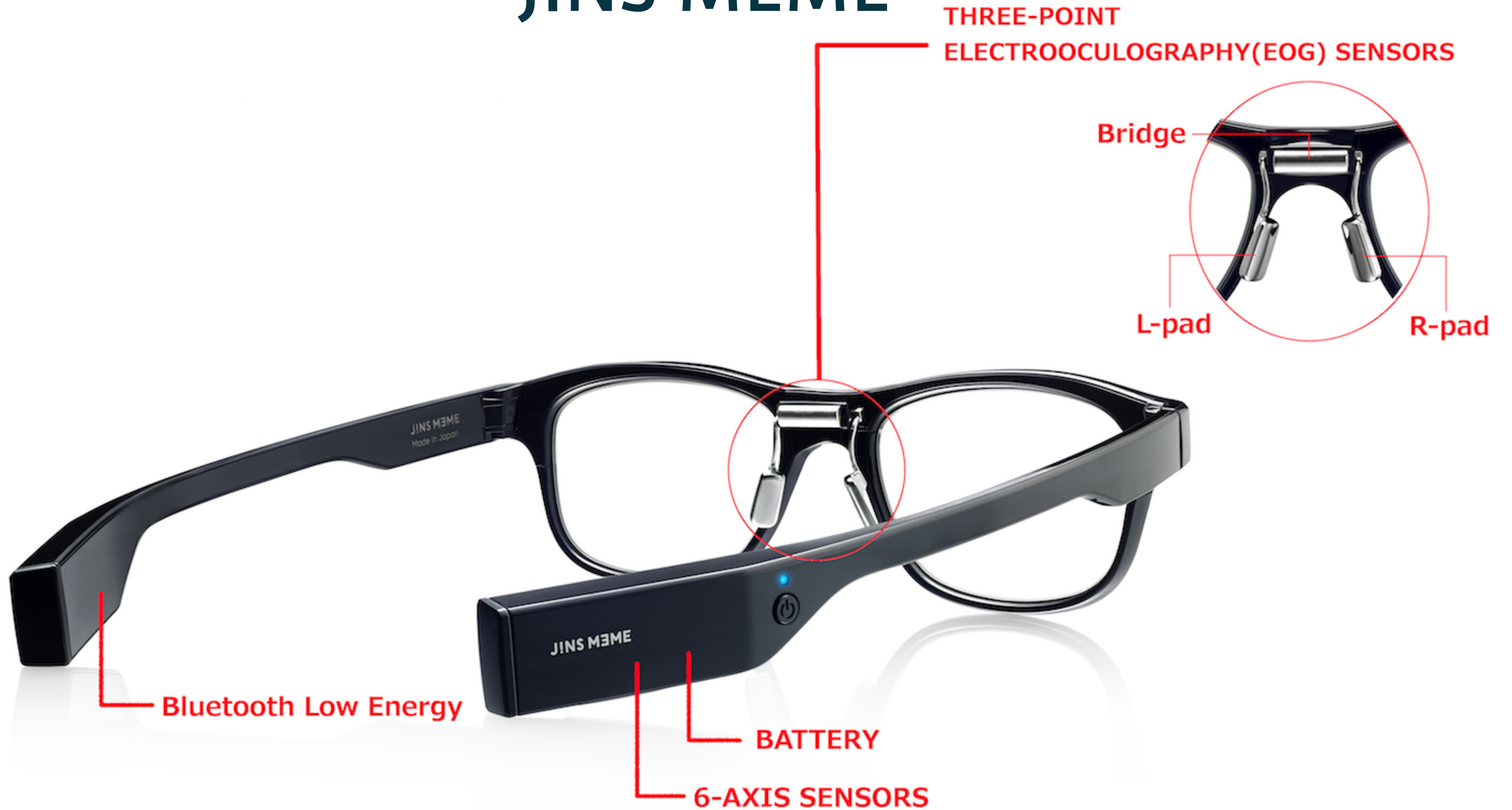
Smart eyewear

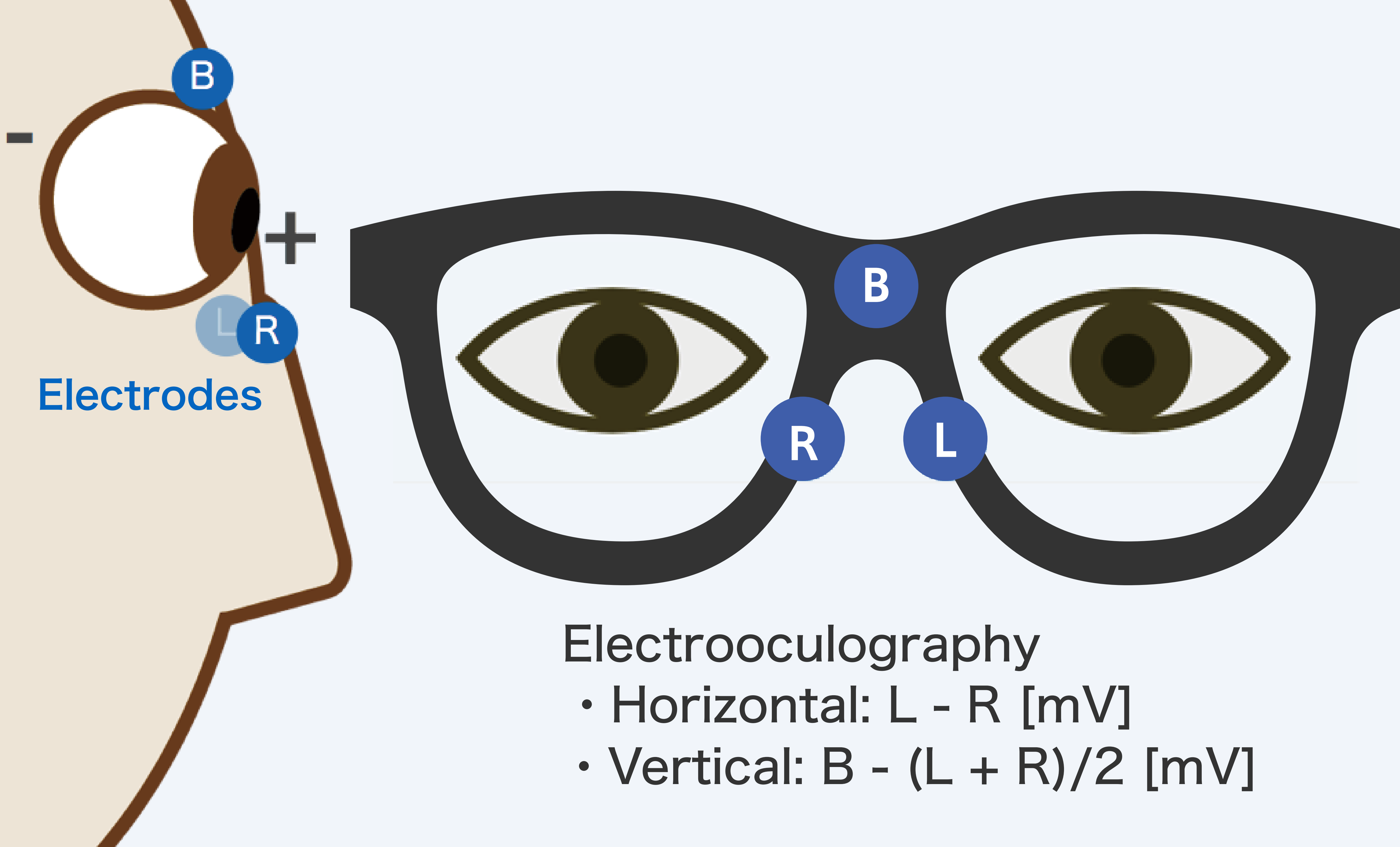
A. Bulling, et al. Eye movement analysis for activity recognition using electrooculography. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33.4 (2010): 741-753.

J. Steil & A. Bulling. Discovery of everyday human activities from long-term visual behaviour using topic models. *Proc. UbiComp 2015*, pp. 75-85, 2015.

S. Ishimaru, et al. Smart Eyewear for Interaction and Activity Recognition. *CHI 2015 Extended Abstracts*, pp. 307-310, 2015.

JINS MEME

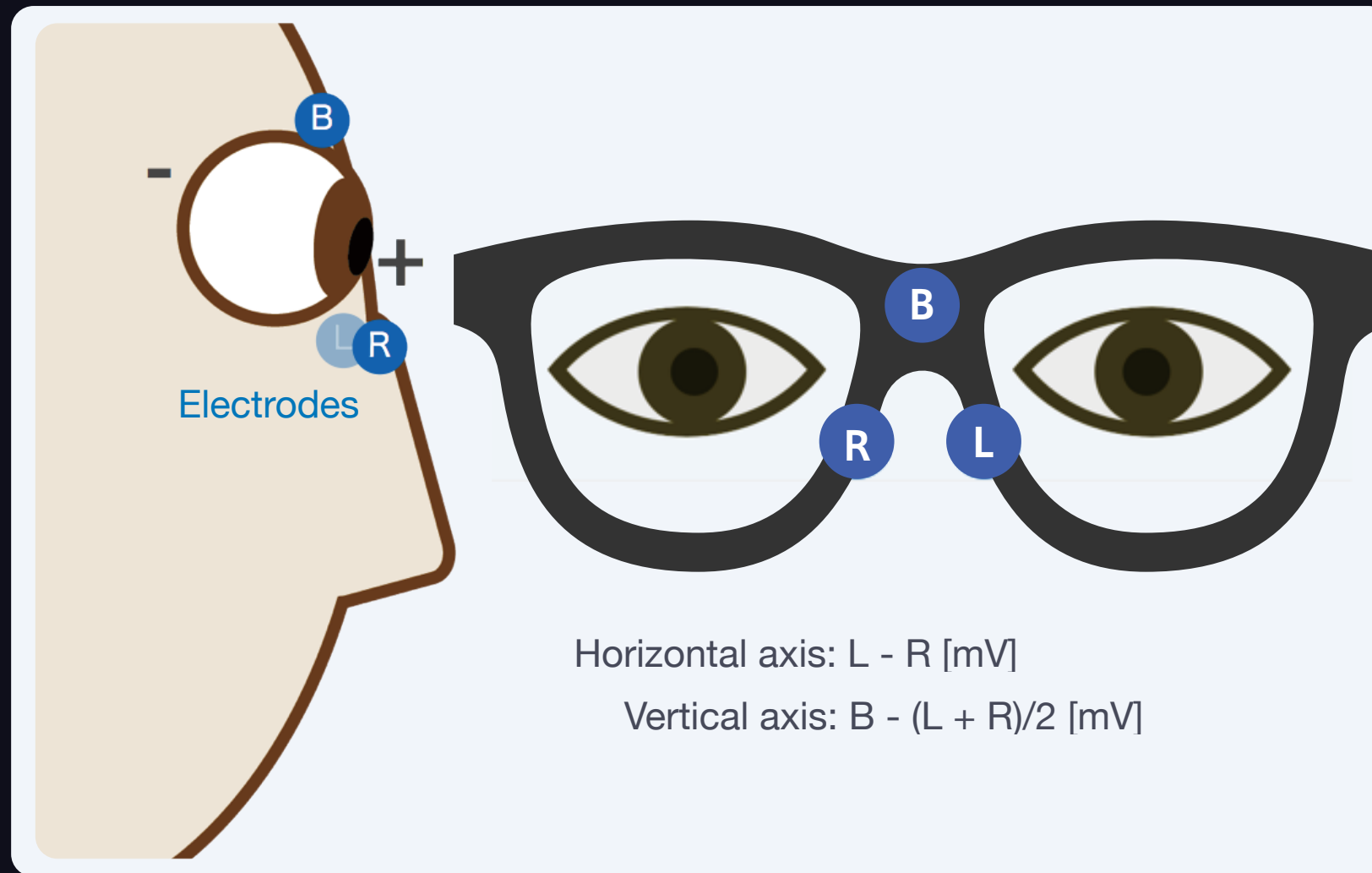
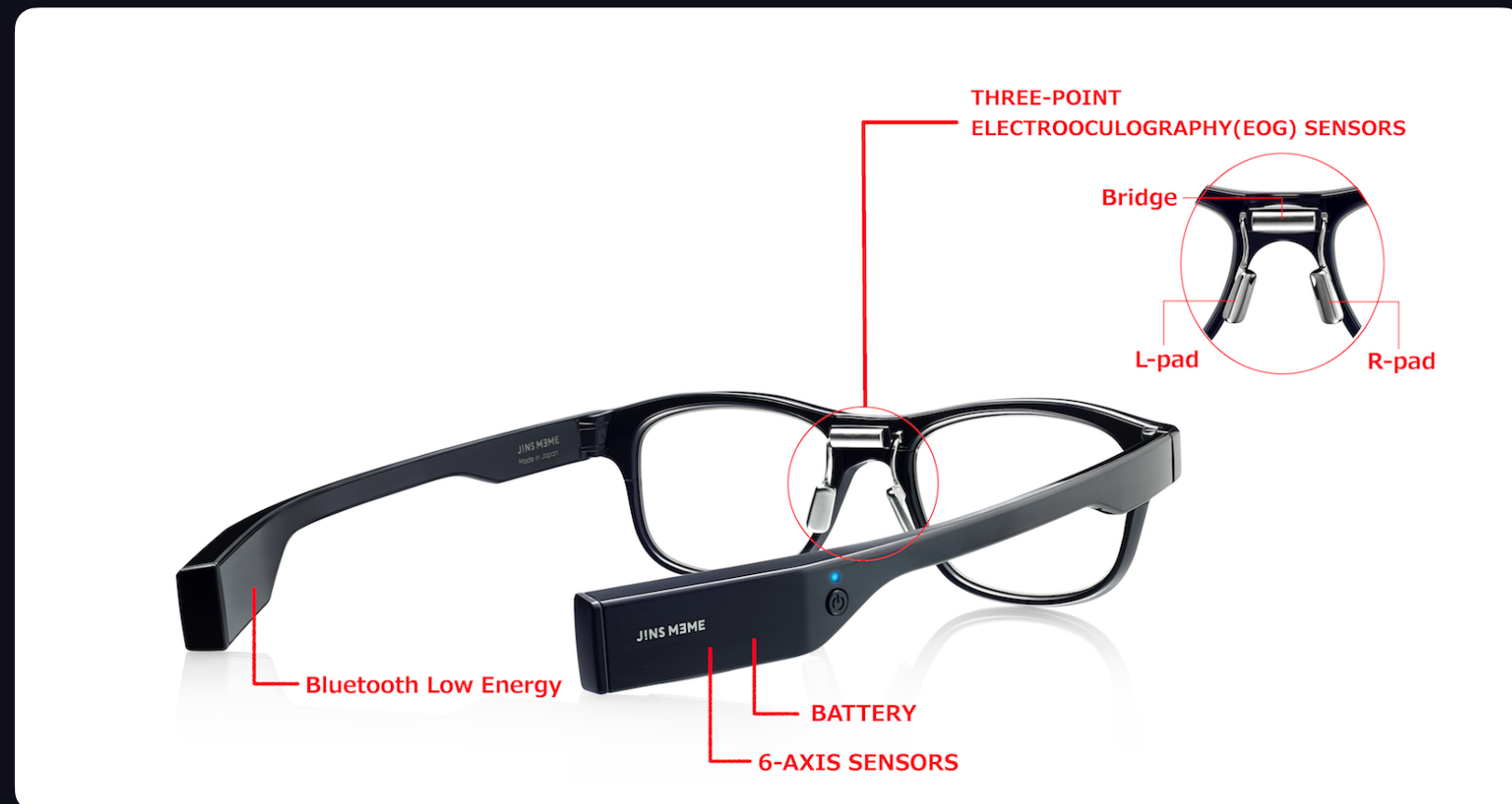




Electrooculography

- Horizontal: $L - R$ [mV]
- Vertical: $B - (L + R)/2$ [mV]

JINS MEME Electrooculography (EOG) Glasses



Software interface showing eye tracking data and recording controls:

- Windows: eye 0, eye 1, World, Recorder
- Eye 0: CPU 101.3, 50 FPS
- Eye 1: CPU 52.1, 55 FPS
- World: CPU 44.1, 31 FPS, Confidence: 0.91
- Recorder: Pupil recordings are saved like this: "path_to_recordings/recording_session_name" where "nnn" is an increasing number to avoid overwrites. You can use "/" in your session name to create subdirectories. Recordings are saved to "~/pupil_recordings". You can change the path here but note that invalid input will be ignored. Path to recordings: /Users/shoya/record Recording session name: 2016_01_26 Request additional user info: Compression: bigger file, less CPU Recording the raw eye video is optional. We use it for debugging.
- Graph - Display Sensor Values (Enabled): Value: Vv1 -108.5
- Graph - Display Sensor Values (Enabled): Value: Vh1 -455.0

Electrooculography

EOG + Head motion based Activity Classification

		Predicted class			
		reading	typing	eating	talking
Actual class	reading	78%	8%	0%	12%
	typing	36%	19%	1%	42%
	eating	0%	6%	93%	0%
	talking	12%	37%	7%	41%

EOG (Acc.: 54%)
majority vote -> 72 %

		Predicted class			
		reading	typing	eating	talking
Actual class	reading	69%	18%	0%	12%
	typing	23%	66%	0%	9%
	eating	3%	1%	77%	18%
	talking	10%	0%	14%	74%

EOG and head (Acc.: 72%)
majority vote -> 100 %

- There were frequent characteristic patterns on EOG while reading or eating. EOG based approach could classify them.

EOG + Head motion based Activity Classification

		Predicted class			
		reading	typing	eating	talking
Actual class	reading	78%	8%	0%	12%
	typing	36%	19%	1%	42%
	eating	0%	6%	93%	0%
	talking	12%	37%	7%	41%

EOG (Acc.: 54%)
majority vote -> 72 %

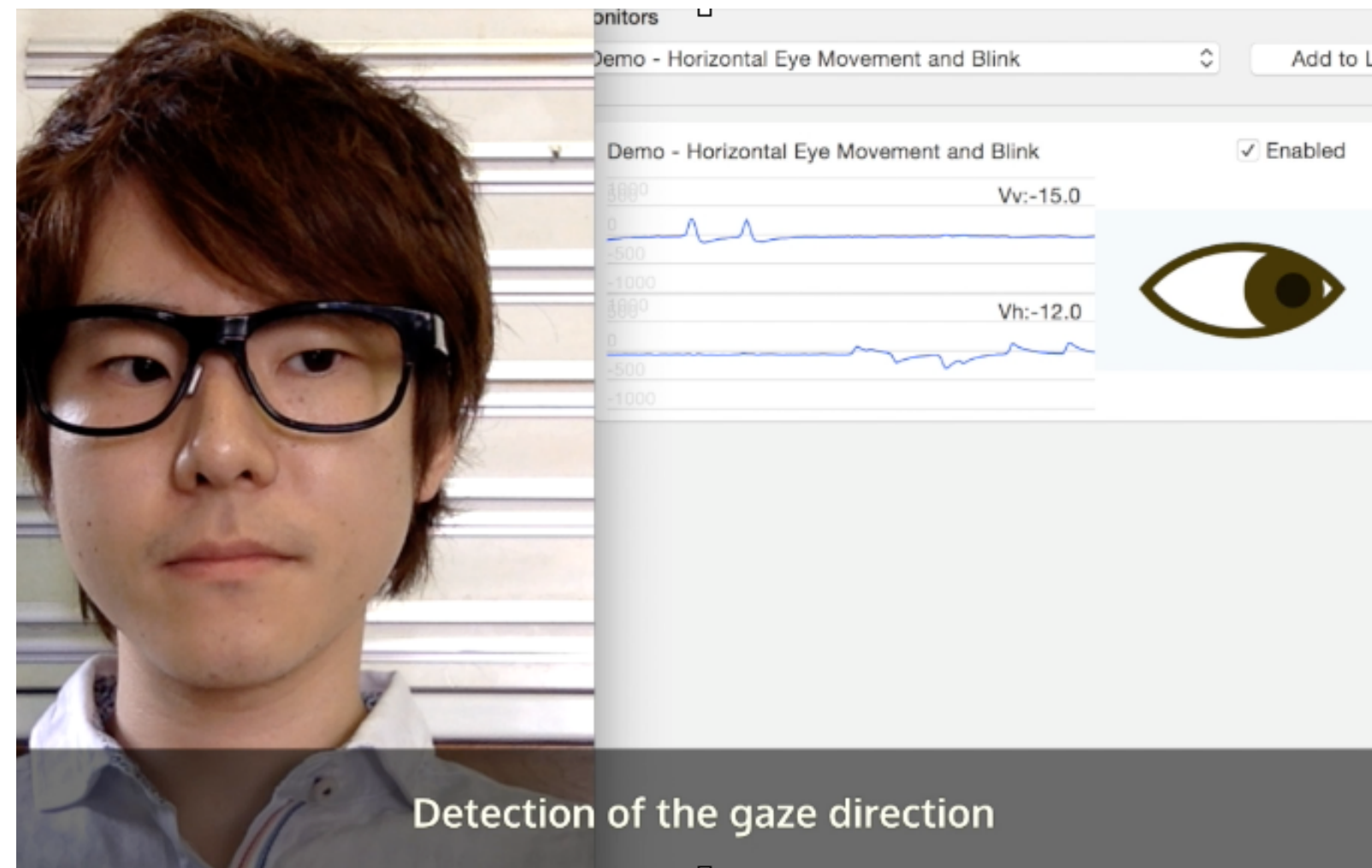
		Predicted class			
		reading	typing	eating	talking
Actual class	reading	69%	18%	0%	12%
	typing	23%	66%	0%	9%
	eating	3%	1%	77%	18%
	talking	10%	0%	14%	74%

EOG and head (Acc.: 72%)
majority vote -> 100 %

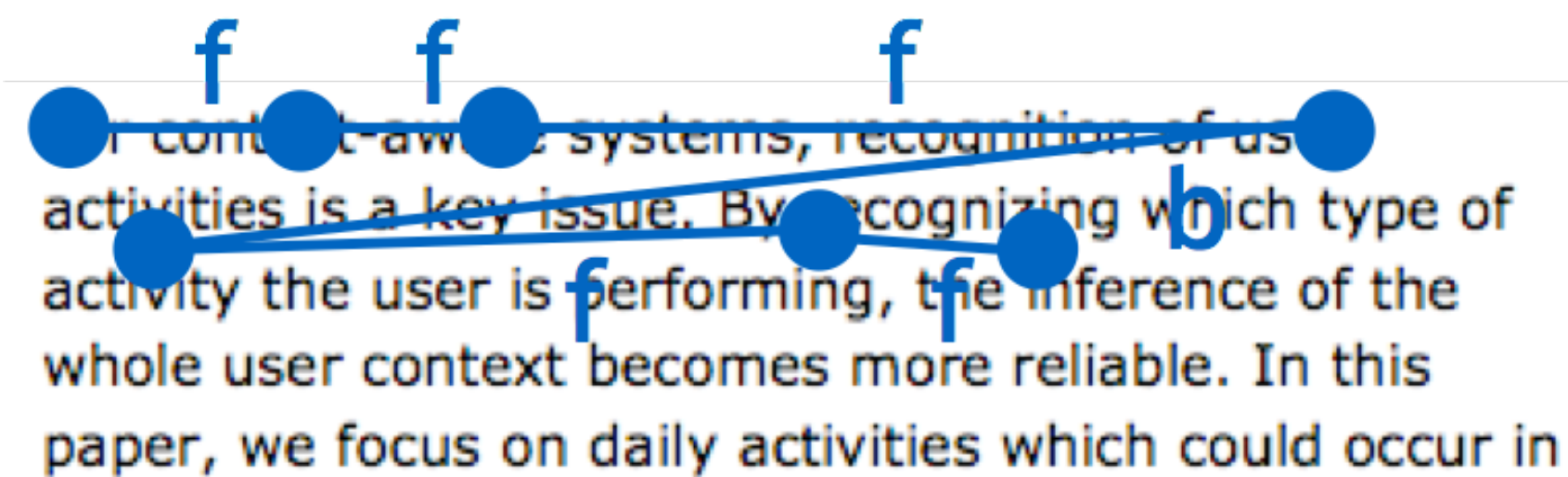
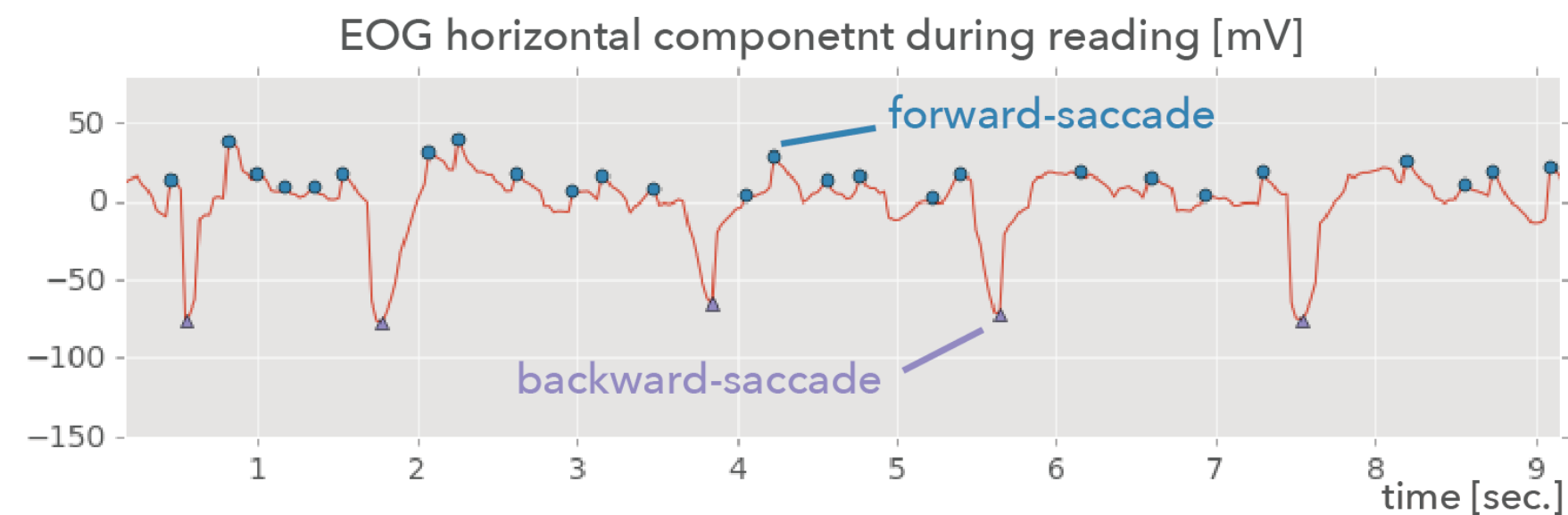
- There were frequent characteristic patterns on EOG while reading or eating. EOG based approach could classify them.
- Typing and talking were often misclassified in the EOG-based approach (similar eye move patterns). But the combination with head motion solved this problem.

Wordometer System Overview

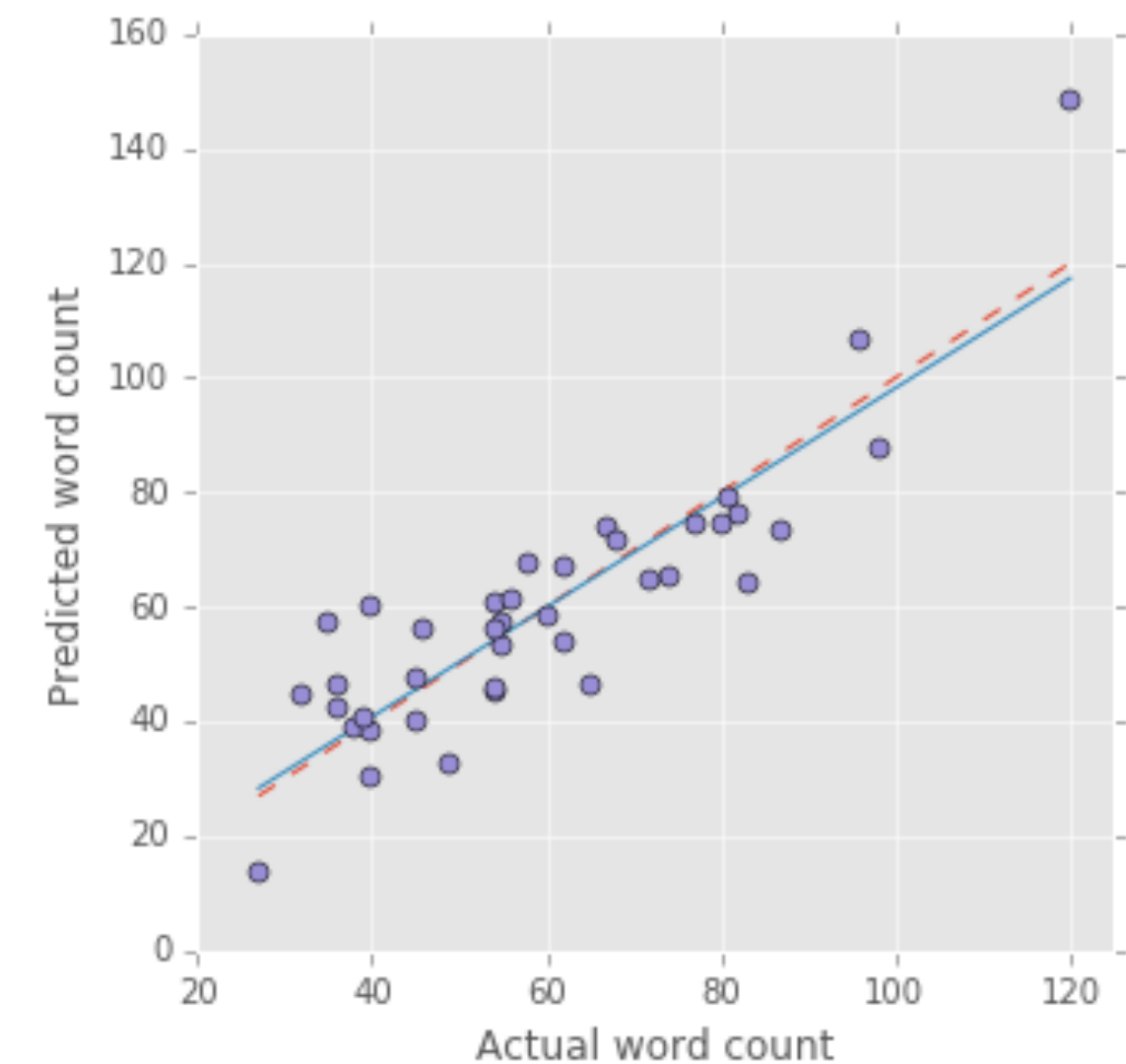
Recording EOG signals



Detecting forward/backward-saccades



Support Vector Regression



- Four features (the numbers of forward/ backward-saccades, mean EOG while the saccades) could estimate the number of read words.
- Word count estimation error: 3% in user-dependent, 11% in user-independent approach (evaluated in an experiment with 5 participants read 38 documents).

Wordometer: Estimating the number of read words



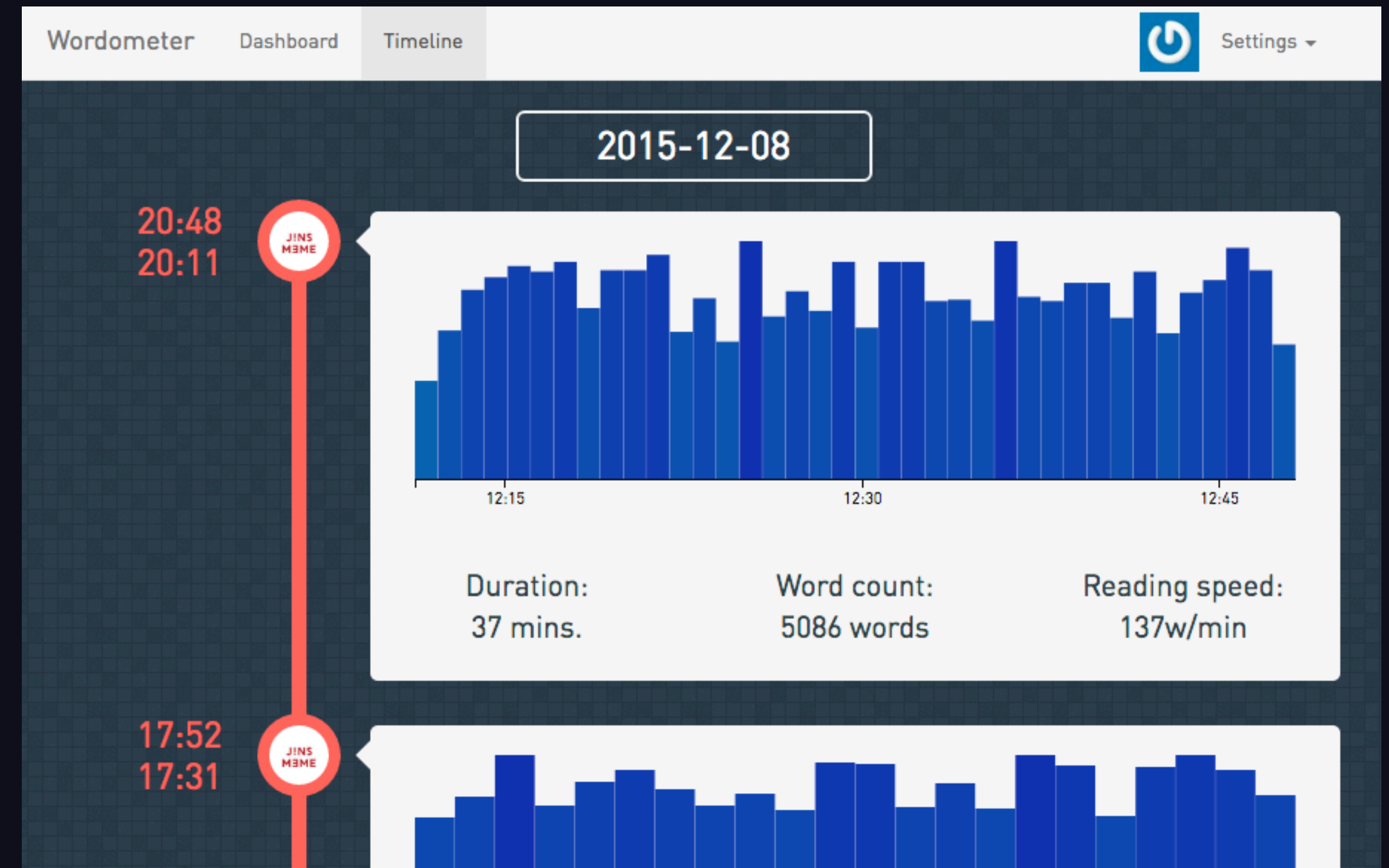
Explicit feedback

Nice condition. Read one paragraph more!

There are interesting articles for you.



+



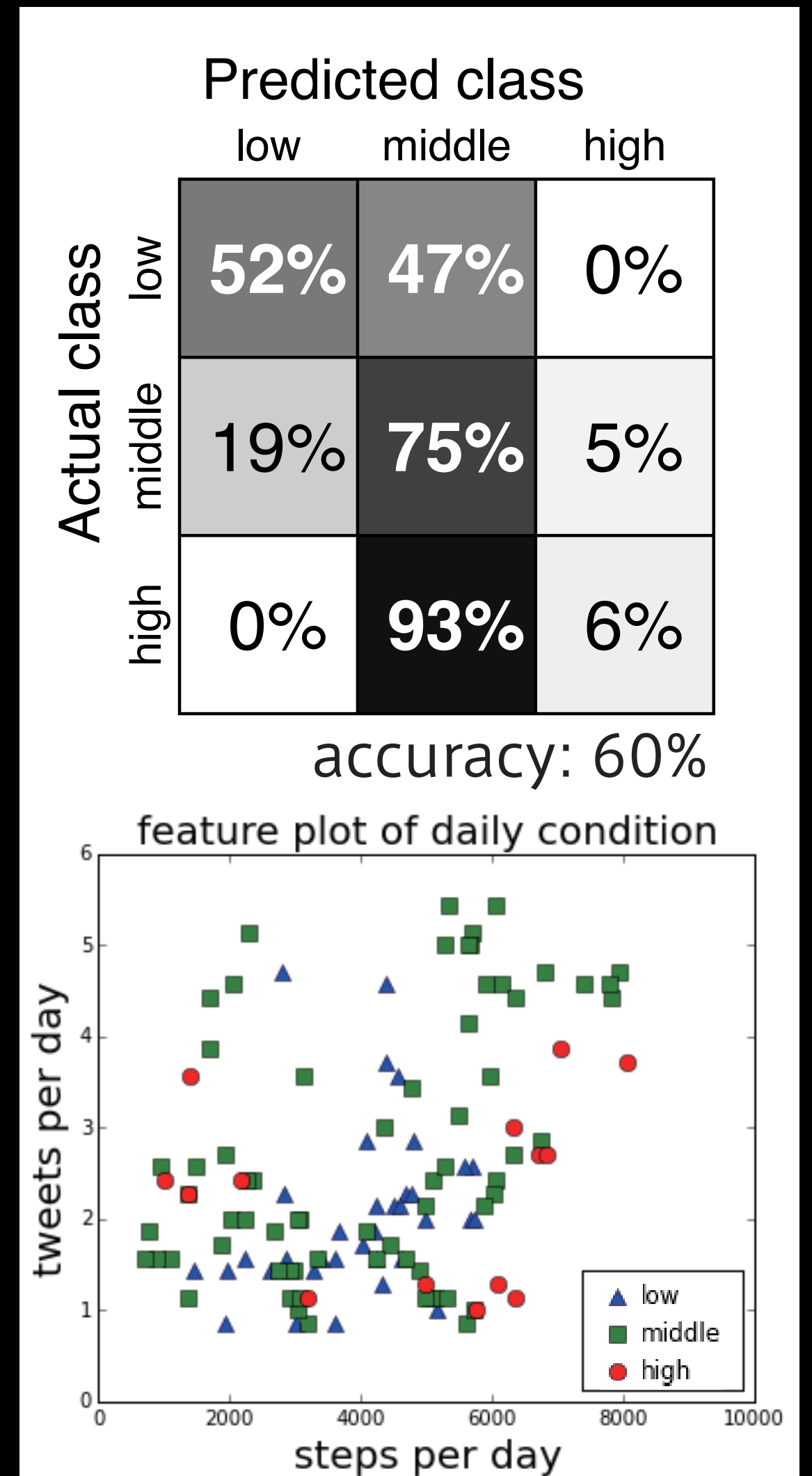
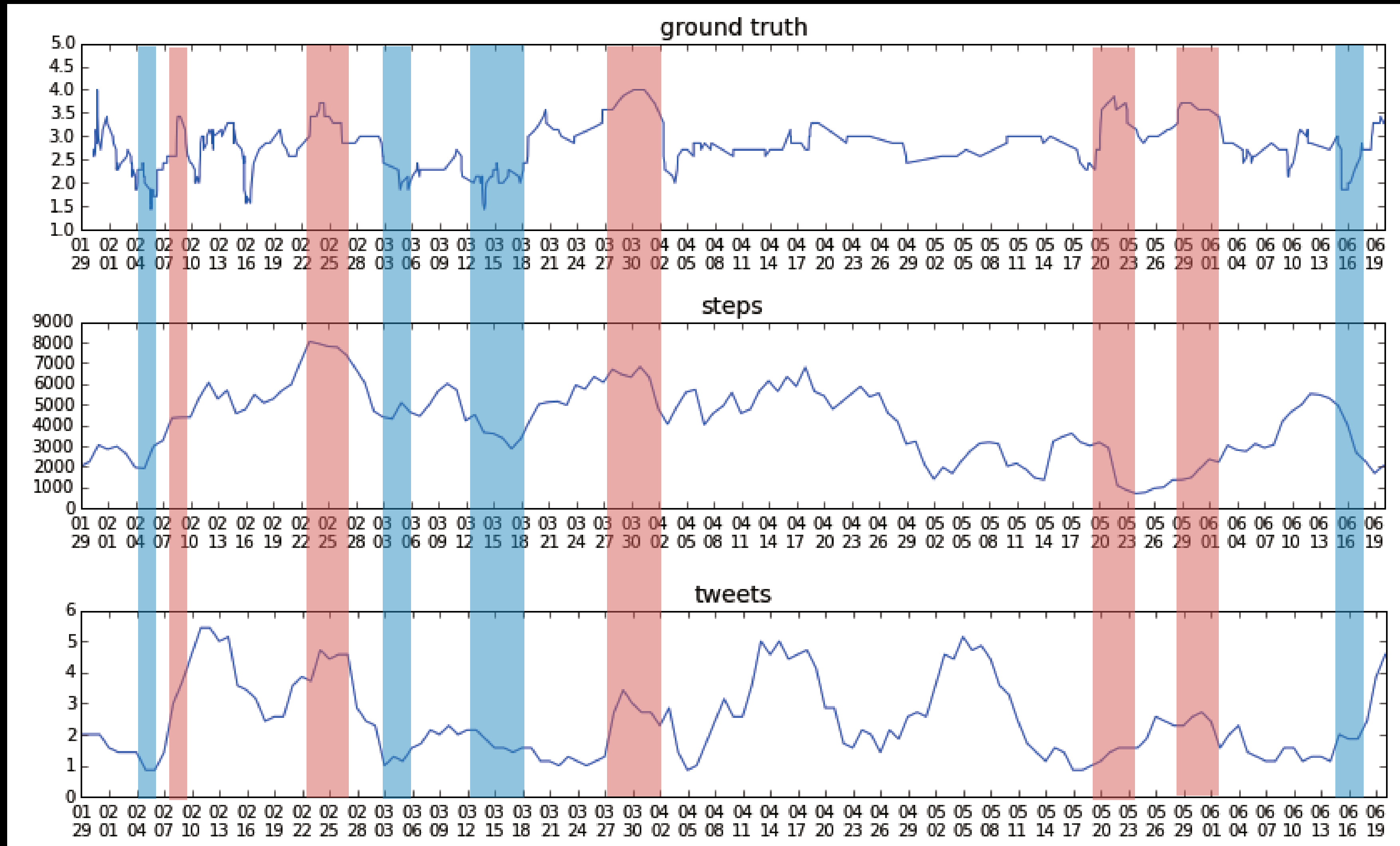
Implicit feedback

Shin'on kei (心温計)

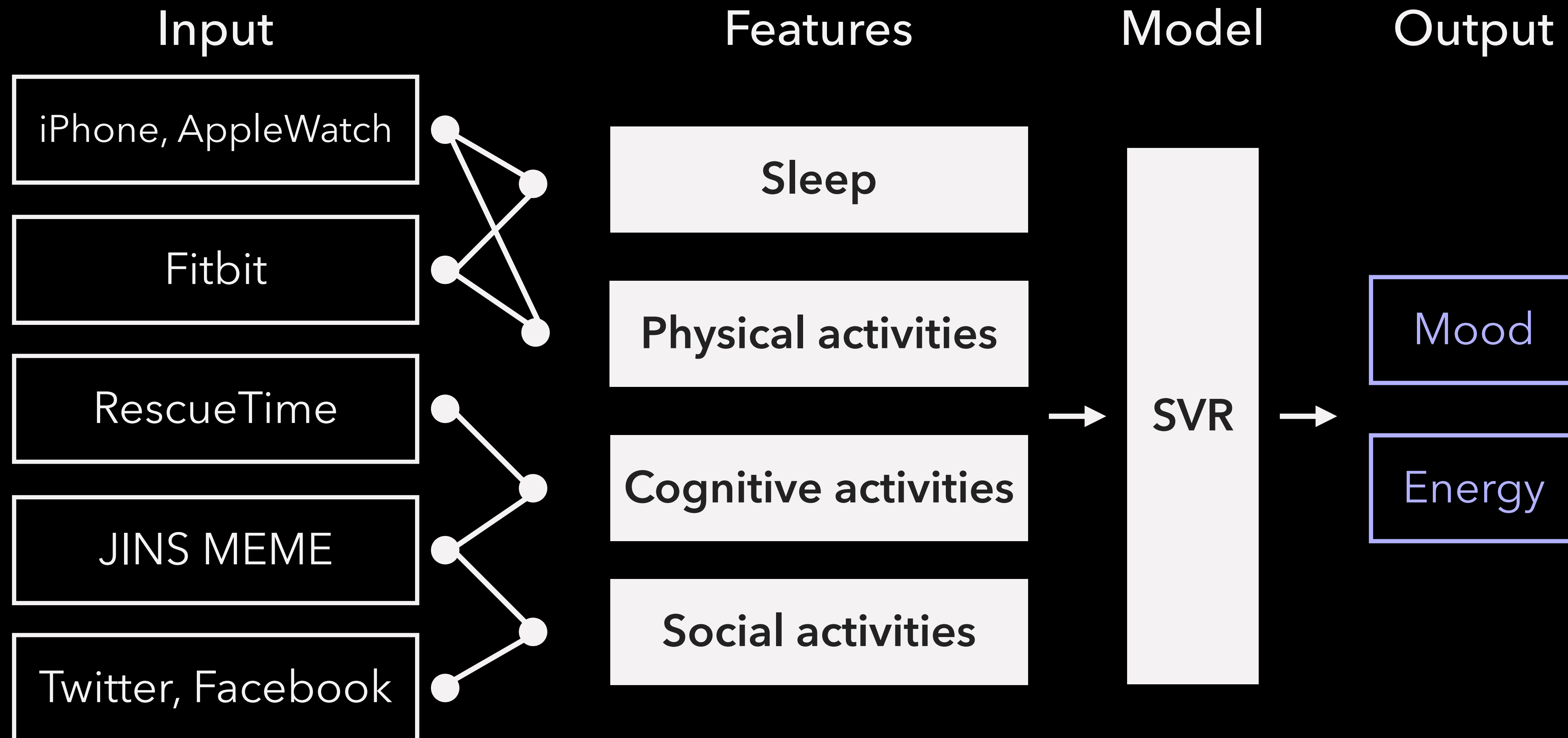
An app that quantifies and displays mental states based on daily activity logs recorded by sensors.



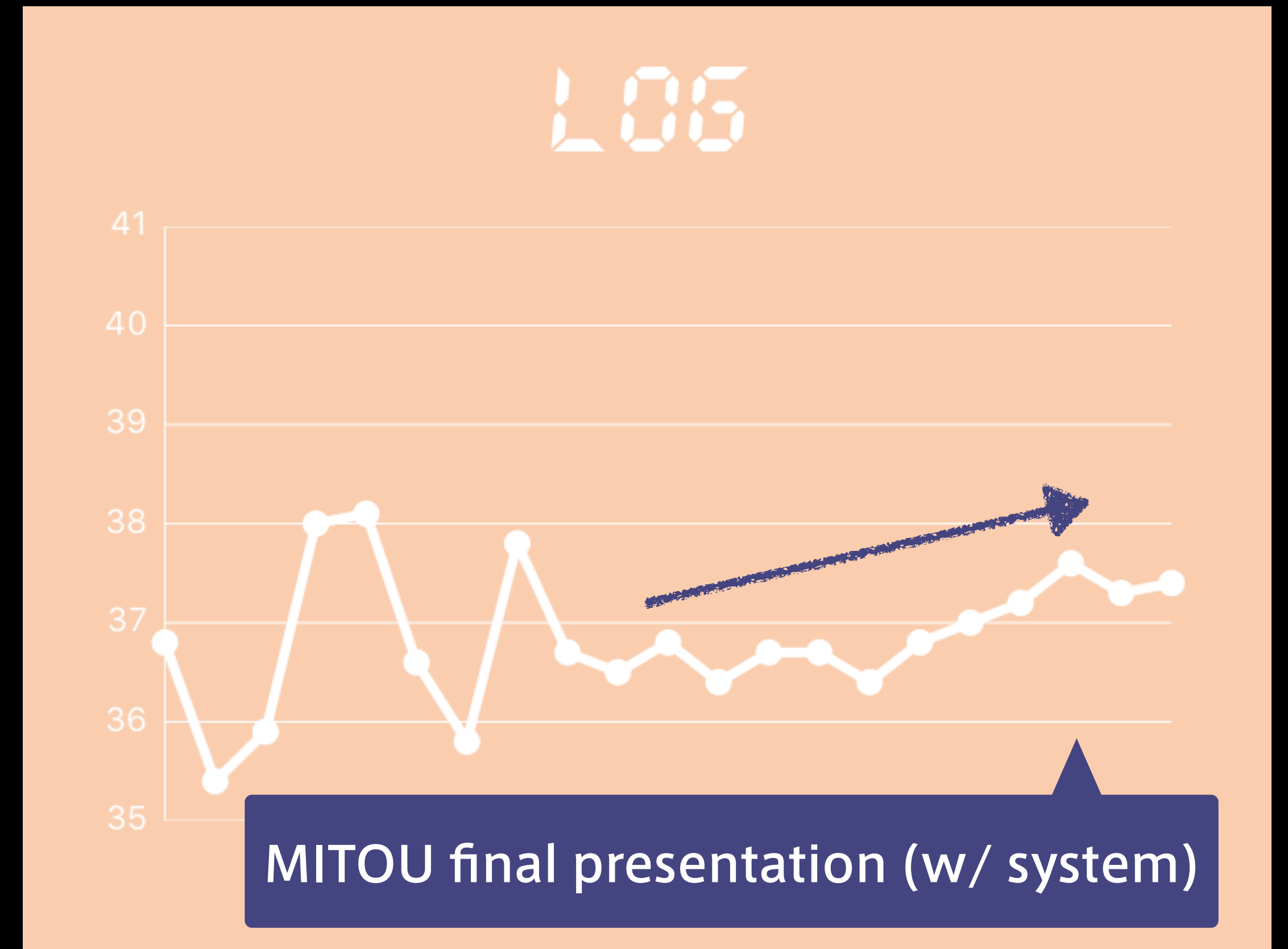
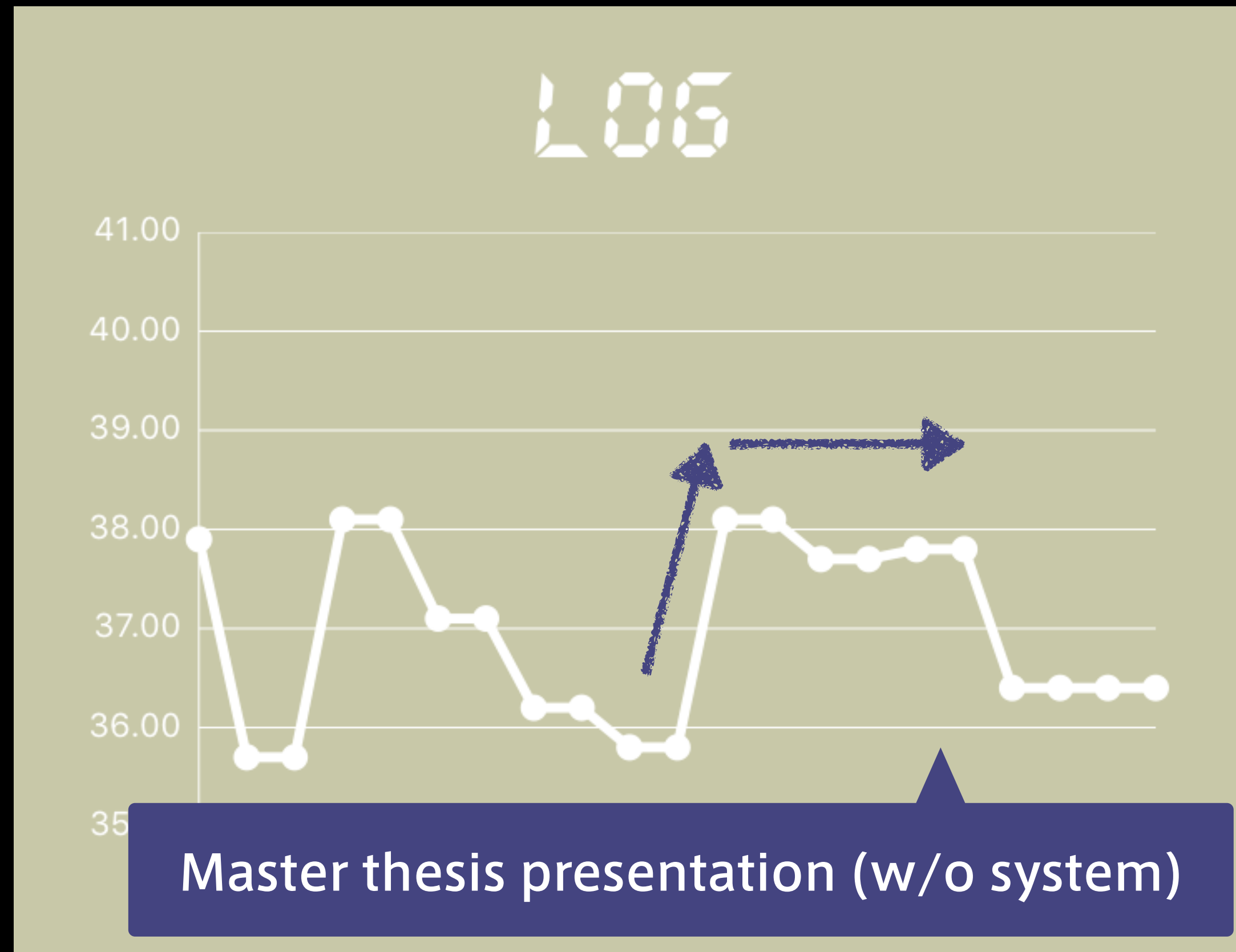
Correlation between depressed moods and sensor data



Estimating mental states based on changes in daily activity levels

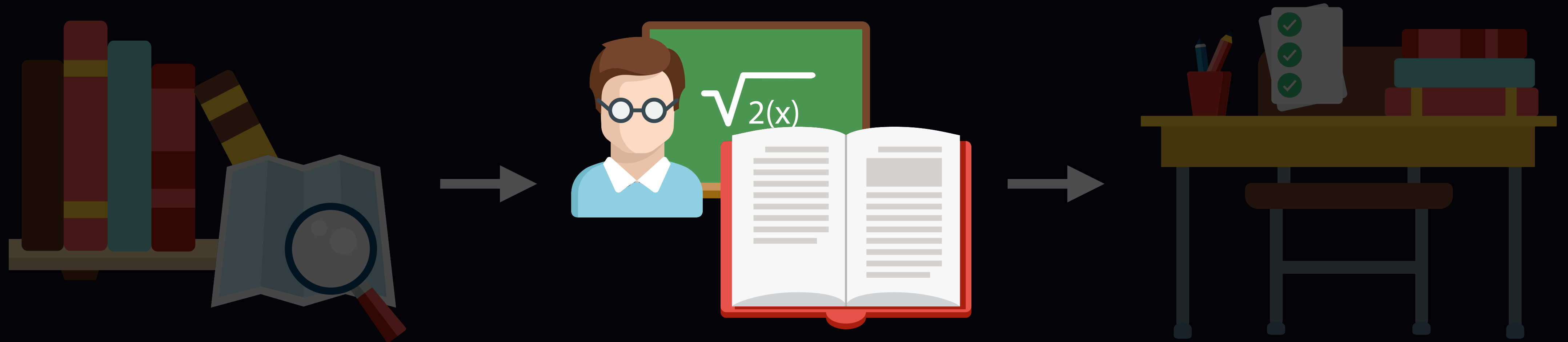


(Not-well-evaluated) experimental results



The user avoided high-stress by checking the system regularly.

Augmented Positive Learning Cycle



Daily habits and curiosity

Lecture and textbook

Questions (test)

Can AI understand interest of students and nudge them to learn voluntarily?

Which question should be reviewed carefully? (considering priority & leakage)

HyperMind

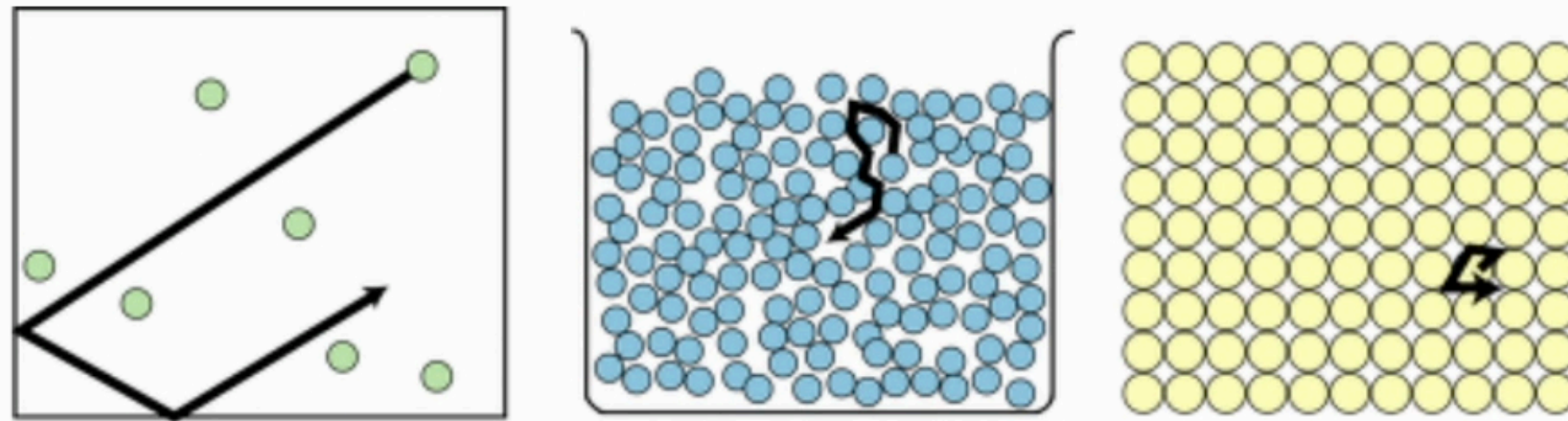


Figure b: Random motion of atoms in a gas, a liquid, and a solid.

2.4.1 Heat is kinetic energy.

What is heat really? Is it an invisible fluid that your bare feet soak up from a hot sidewalk? Can one ever remove all the heat from an object? Is there a maximum to the temperature scale?

The theory of heat as a fluid seemed to explain why colder objects absorbed heat from hotter ones, but once it became clear that heat was a form of energy, it began to seem unlikely that a material substance could transform itself into and out of all those other forms of energy like motion or light. For instance, a compost pile gets hot, and we describe this as a case where, through the action of bacteria, chemical energy stored in the plant cuttings is transformed into heat energy. The heating occurs even if there is no nearby warmer object that could have been leaking “heat fluid” into the pile.

An alternative interpretation of heat was suggested by the theory that matter is made of atoms. Since gases are thousands of times

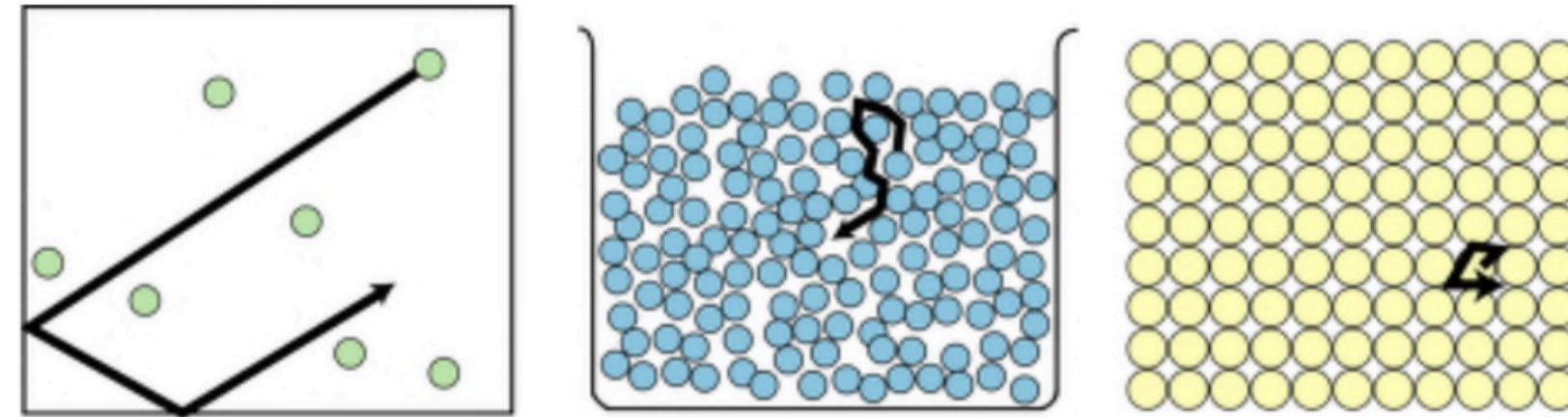


Figure b: Random motion of atoms in a gas, a liquid, and a solid.

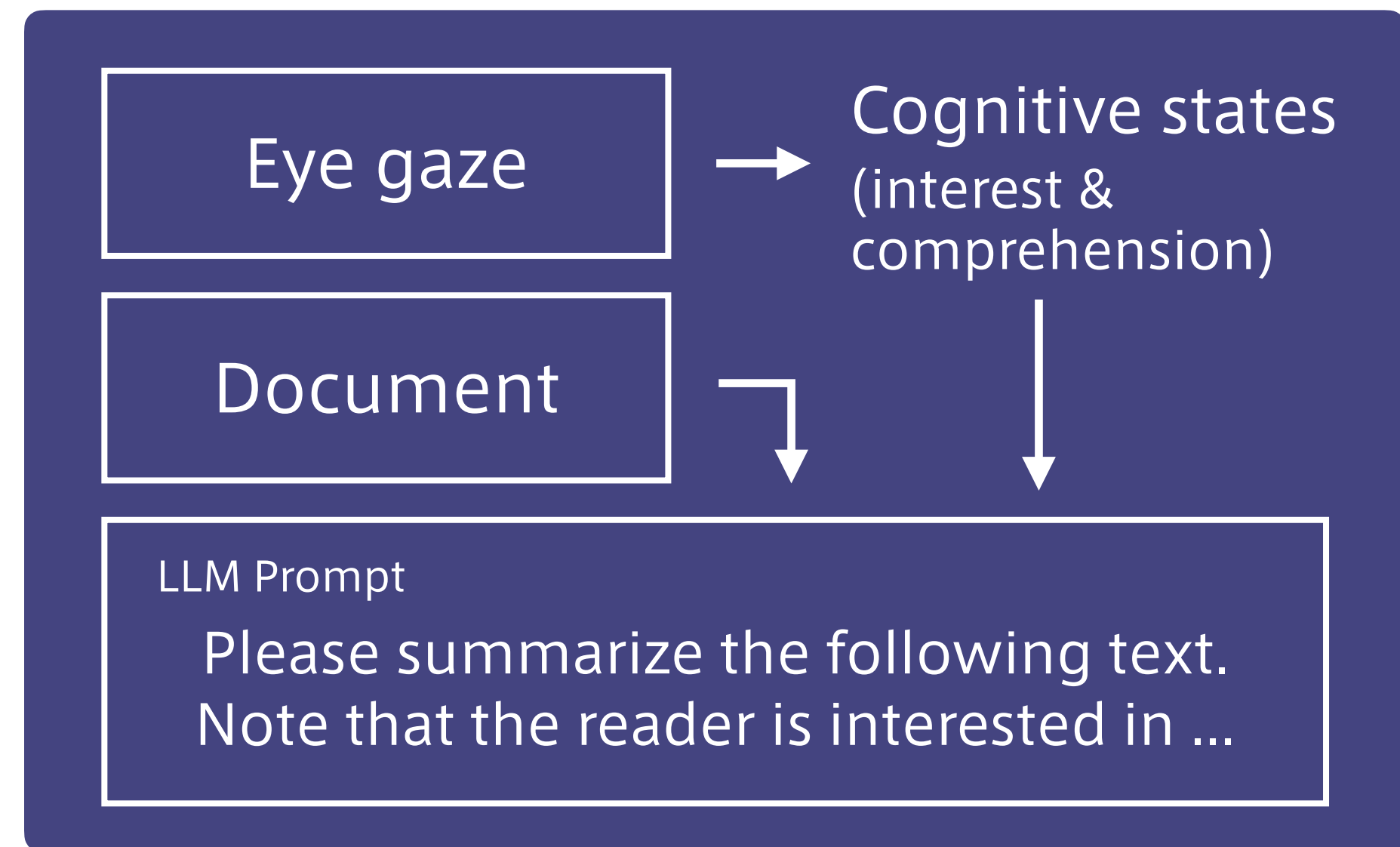
2.4.1 Heat is kinetic energy.

What is heat really? Is it an invisible fluid that your bare feet soak up from a hot sidewalk? Can one ever remove all the heat from an object? Is there a maximum to the temperature scale?

The theory of heat as a fluid seemed to explain why colder objects absorbed heat from hotter ones, but once it became clear that heat was a form of energy, it began to seem unlikely that a material substance could transform itself into and out of all those other forms of energy like motion or light. For instance, a compost pile gets hot, and we describe this as a case where, through the action of bacteria, chemical energy stored in the plant cuttings is transformed into heat energy. The heating occurs even if there is no nearby warmer object that could have been leaking “heat fluid” into the pile.

An alternative interpretation of heat was suggested by the theory that matter is made of atoms. Since gases are thousands of times less dense than solids or liquids, the atoms (or clusters of atoms called molecules) in a gas must be far apart. In that case, what is keeping all the air molecules from settling into a thin film on the

System Architecture



Heat was seen differently with the idea that matter is made of atoms. Gases have far apart atoms or molecules, but they don't settle because they're moving very fast and bouncing off surfaces. This explains how a gas can have a lot of force even though it's not dense.

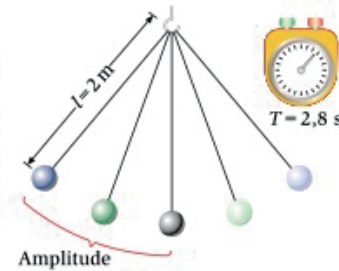
Experimental Design

Introduction
Definitions
Applications

Akustik – Töne und Schall (basierend auf Schwingungen)

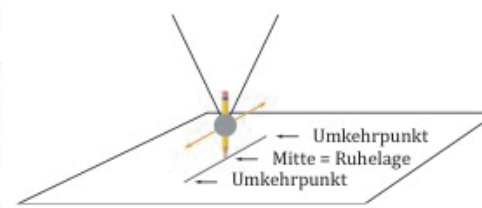
Schall entsteht durch Schwingungen eines *Schallerzeugers*. Sind die Schwingungen nicht zu langsam und nicht zu schnell, können wir sie als **Ton** wahrnehmen, also hören.

Wie kann man **Schwingungen** (und damit auch Schall) beschreiben? Der Bewegungsablauf eines **Fadenpendels** ist eine Schwingung, die bequem zu beachten ist. An diesem Beispiel möchten wir deshalb die Größen erklären, mit denen man eine Schwingungsbewegung beschreibt. Lenken wir die Kugel aus und lassen sie los, dann bewegt sie sich ständig hin und her.

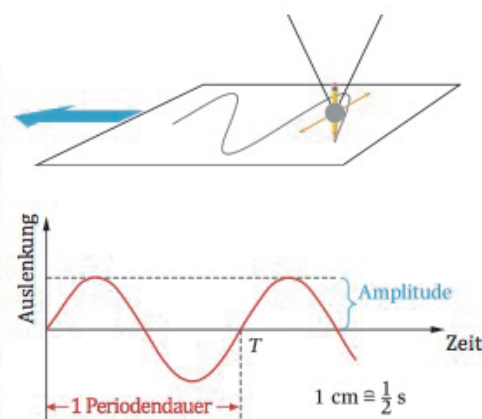


- Eine solche sich ständig wiederholende Bewegung heißt **periodische Bewegung**.
- Eine vollständige Hin- und Herbewegung der Pendelkugel ist eine **Periode**.
- Die Zeit, in der das Pendel eine Periode (also eine vollständige Hin- und Herbewegung) ausführt, heißt **Periodendauer T**.
- Während jeder Hin- und Herbewegung schwingt die Pendelkugel zu den selben zwei Umkehrpunkten und hat dort ihre größte Auslenkung. Die Auslenkung von der Mittellage zu einem Umkehrpunkt nennen wir die **Amplitude** der Schwingung.
- Die **Frequenz f** einer Schwingung beschreibt die Anzahl von Schwingungen pro Sekunde. Sie ist der Quotient aus der Anzahl *n* der Perioden und der dazu benötigten Zeit *t*: $f = n/t$. Ihre Einheit ist 1 Hz (sprich: Hertz), $1 \text{ Hz} = 1/s$. Je größer die Frequenz ist, desto schneller ist die Schwingungsbewegung.
- Mit zunehmender Frequenz steigt die **Tonhöhe** der Schwingung. Mit wachsender Amplitude nimmt die **Lautstärke** zu.
- Für die Periodendauer *T* und die Frequenz *f* gilt: $T = 1/f$; $f = 1/T$. Aus einer größeren Frequenz folgt also eine kleinere Periodendauer und umgekehrt.

Befestigen wir einen Schreibstift an einem Fadenpendel, so schreibt er während der Hin- und Herbewegung des Pendels eine geradlinige Spur (eine Linie) auf ein darunter liegendes Blatt Papier. Anfang und Ende der Spur sind die Umkehrpunkte der Schwingung, die Mitte ist die Ruhelage des Pendels. Wir erhalten die **Amplitude** der Schwingung, wenn wir die Entfernung der Umkehrpunkte von der Mitte messen.

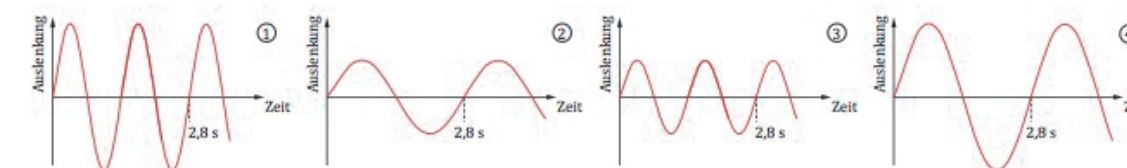
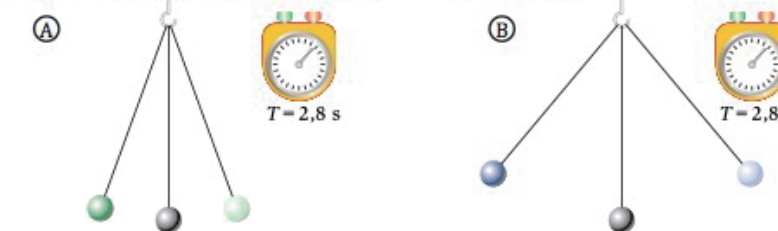


Ziehen wir das Blatt Papier gleichmäßig und senkrecht zur Schwingungsrichtung unter dem sich bewegenden Pendel weg, so schreibt der Stift eine gleichmäßige Schlangenlinie (wie auf dem Bild rechts). Es entsteht ein **Diagramm** der Schwingung auf dem Blatt Papier, mit der Auslenkung des Pendels in *y*-Richtung und der Zeit in *x*-Richtung. Die Schwingungen eines *Schallerzeugers* verlaufen ebenso wie die eines Pendels. Wir können also die Größen, die wir zur Beschreibung einer Pendelschwingung verwendet haben, auf Schallerzeuger und Töne übertragen.



Aufgaben

- Wie schwingen die Flügel einer Biene im Vergleich zu denen eines Adlers? Kreuze alle richtigen Aussagen an!
 - Die **Periodendauer T** ist bei der Biene größer.
 - Die **Amplitude** ist beim Adler größer.
 - Die **Perioden** sind beim Adler erkennbar.
 - Die **Frequenz** ist bei der Biene kleiner.
- Bestimme die Periodendauer, die ein Ton mit der Frequenz 50 Hz hat.
- Was versteht man unter einer Periode?
- Die beiden Grafiken A und B zeigen ein schwingendes Fadenpendel in seiner Ruhelage und in den beiden Umkehrpunkten. Außerdem ist die zugehörige Periodendauer angegeben. Unter den Grafiken siehst du vier verschiedene Diagramme einer Schwingung.
 - Welches Diagramm stellt die Schwingung aus Grafik A dar?
 - Welches Diagramm stellt die Schwingung aus Grafik B dar?



- Gib zu den folgenden Diagrammen jeweils die zugehörige Frequenz *f* an!
 - 5.A
 - 5.B
 - 5.C
- Ein Fadenpendel braucht für 9 Perioden 12 Sekunden. Berechne die Periodendauer *T*. Bestimme die Anzahl der Perioden in 1 Sekunde. Berechne die Frequenz des Pendels.

Attention Analysis

Introduction

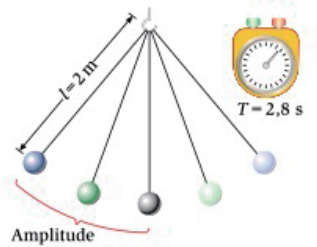
Definitions

Applications

Akustik – Töne und Schall (basierend auf Schwingungen)

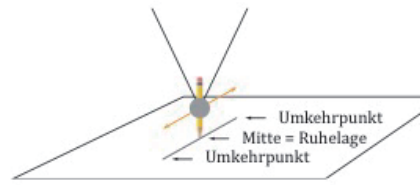
Schall entsteht durch Schwingungen eines Schallerzeugers. Sind die Schwingungen nicht zu langsam und nicht zu schnell, können wir sie als **Ton** wahrnehmen, also hören.

Wie kann man **Schwingungen** (und damit auch Schall) beschreiben? Der Bewegungsablauf eines **Fadenpendels** ist eine Schwingung, die bequem zu beachten ist. An diesem Beispiel möchten wir deshalb die Größen erklären, mit denen man eine Schwingungsbewegung beschreibt. Lenken wir die Kugel aus und lassen sie los, dann bewegt sie sich ständig hin und her.

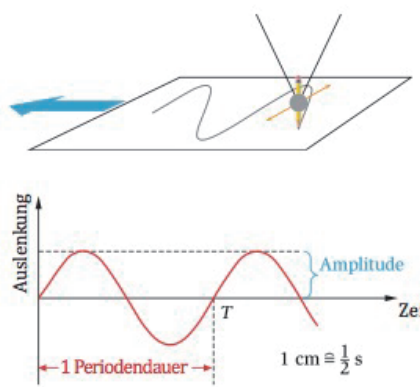


- Eine solche sich ständig wiederholende Bewegung heißt **periodische Bewegung**.
- Eine vollständige Hin- und Herbewegung der Pendelkugel ist eine **Periode**.
- Die Zeit, in der das Pendel eine Periode (also eine vollständige Hin- und Herbewegung) ausführt, heißt **Periodendauer T** .
- Während jeder Hin- und Herbewegung schwingt die Pendelkugel zu den selben zwei Umkehrpunkten und hat dort ihre größte Auslenkung. Die Auslenkung von der Mittellage zu einem Umkehrpunkt nennen wir die **Amplitude** der Schwingung.
- Die **Frequenz f** einer Schwingung beschreibt die Anzahl von Schwingungen pro Sekunde. Sie ist der Quotient aus der Anzahl n der Perioden und der dazu benötigten Zeit t : $f = n/t$. Ihre Einheit ist 1 Hz (sprich: Hertz), $1 \text{ Hz} = 1/\text{s}$. Je größer die Frequenz ist, desto schneller ist die Schwingungsbewegung.
- Mit zunehmender Frequenz steigt die **Tonhöhe** der Schwingung. Mit wachsender Amplitude nimmt die **Lautstärke** zu.
- Für die Periodendauer T und die Frequenz f gilt: $T = 1/f$; $f = 1/T$. Aus einer größeren Frequenz folgt also eine kleinere Periodendauer und umgekehrt.

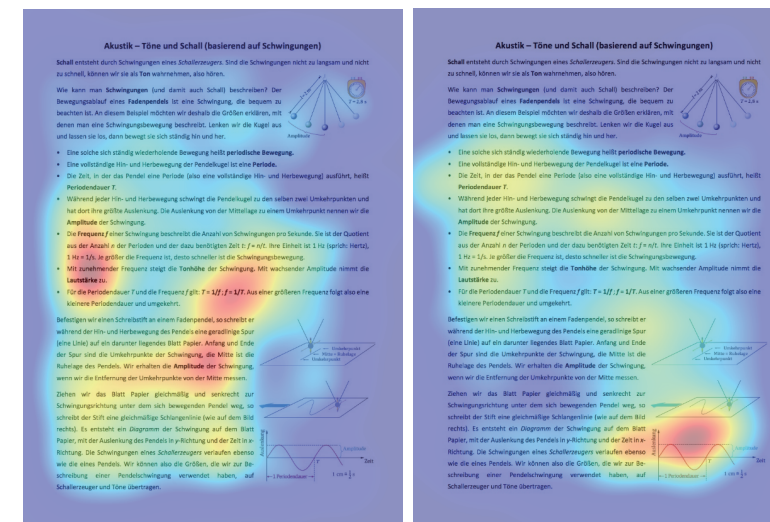
Befestigen wir einen Schreibstift an einem Fadenpendel, so schreibt er während der Hin- und Herbewegung des Pendels eine geradlinige Spur (eine Linie) auf ein darunter liegendes Blatt Papier. Anfang und Ende der Spur sind die Umkehrpunkte der Schwingung, die Mitte ist die Ruhelage des Pendels. Wir erhalten die **Amplitude** der Schwingung, wenn wir die Entfernung der Umkehrpunkte von der Mitte messen.



Ziehen wir das Blatt Papier gleichmäßig und senkrecht zur Schwingungsrichtung unter dem sich bewegenden Pendel weg, so schreibt der Stift eine gleichmäßige Schlangenlinie (wie auf dem Bild rechts). Es entsteht ein **Diagramm** der Schwingung auf dem Blatt Papier, mit der Auslenkung des Pendels in y -Richtung und der Zeit in x -Richtung. Die Schwingungen eines **Schallerzeugers** verlaufen ebenso wie die eines Pendels. Wir können also die Größen, die wir zur Beschreibung einer Pendelschwingung verwendet haben, auf Schallerzeuger und Töne übertragen.



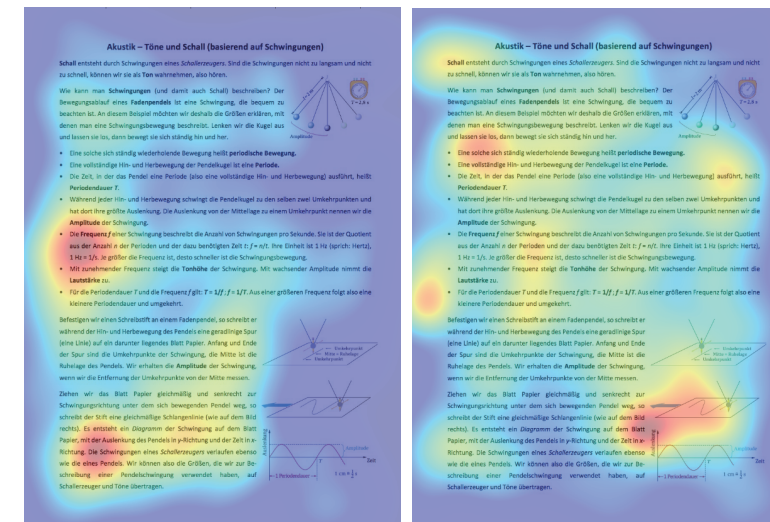
Novice



(a) reading

(b) solving

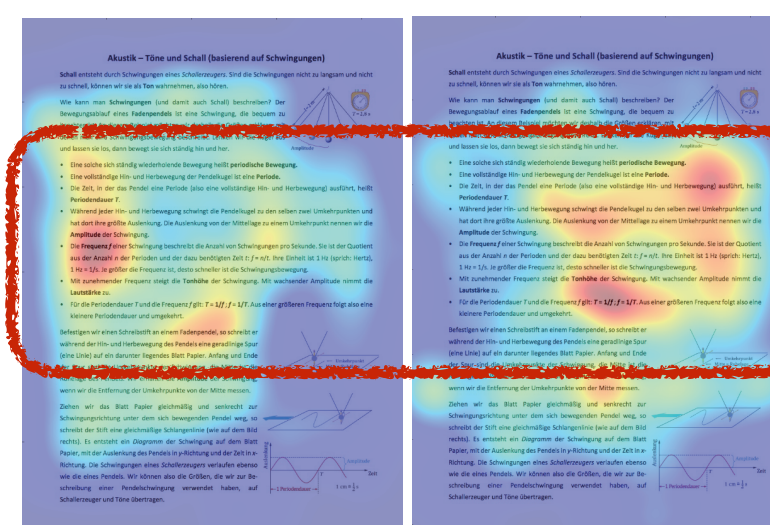
Intermediate



(c) reading

(d) solving

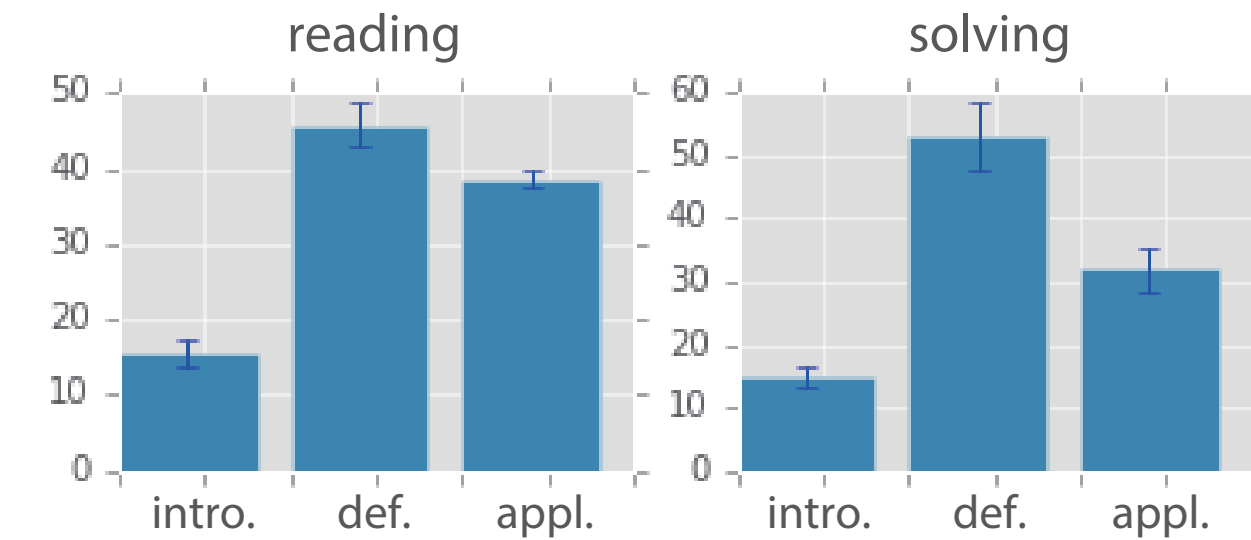
Expert



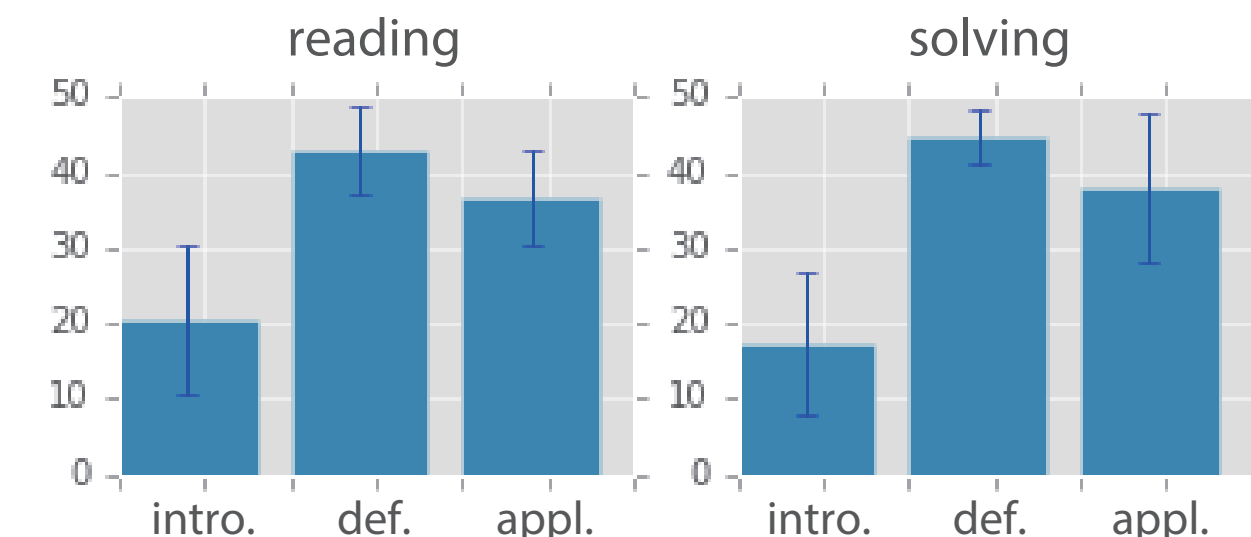
(e) reading

(f) solving

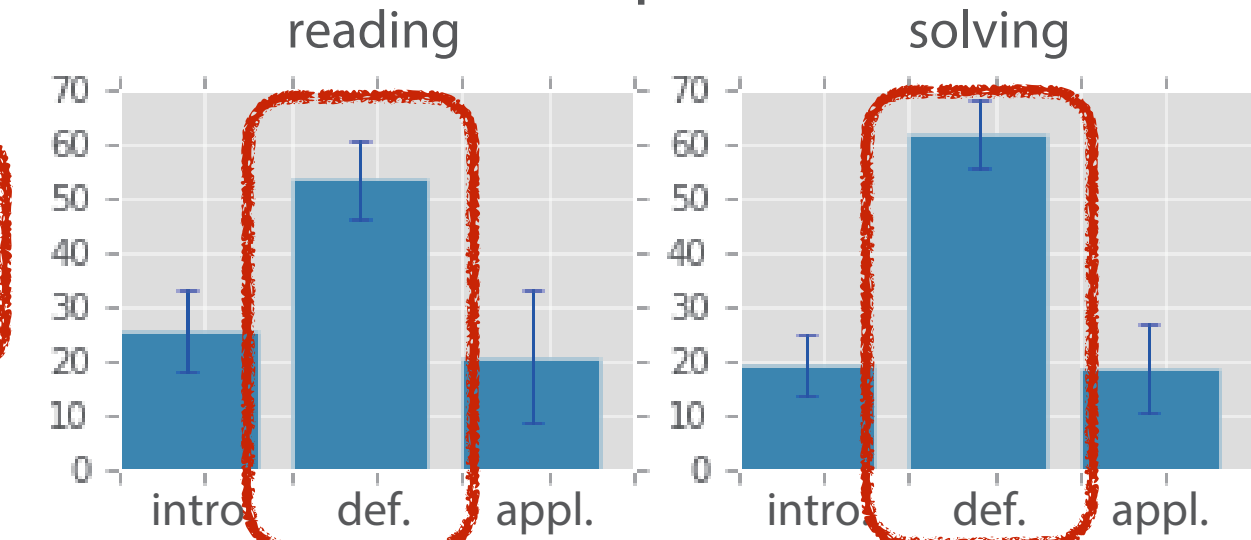
Novice



Intermediate

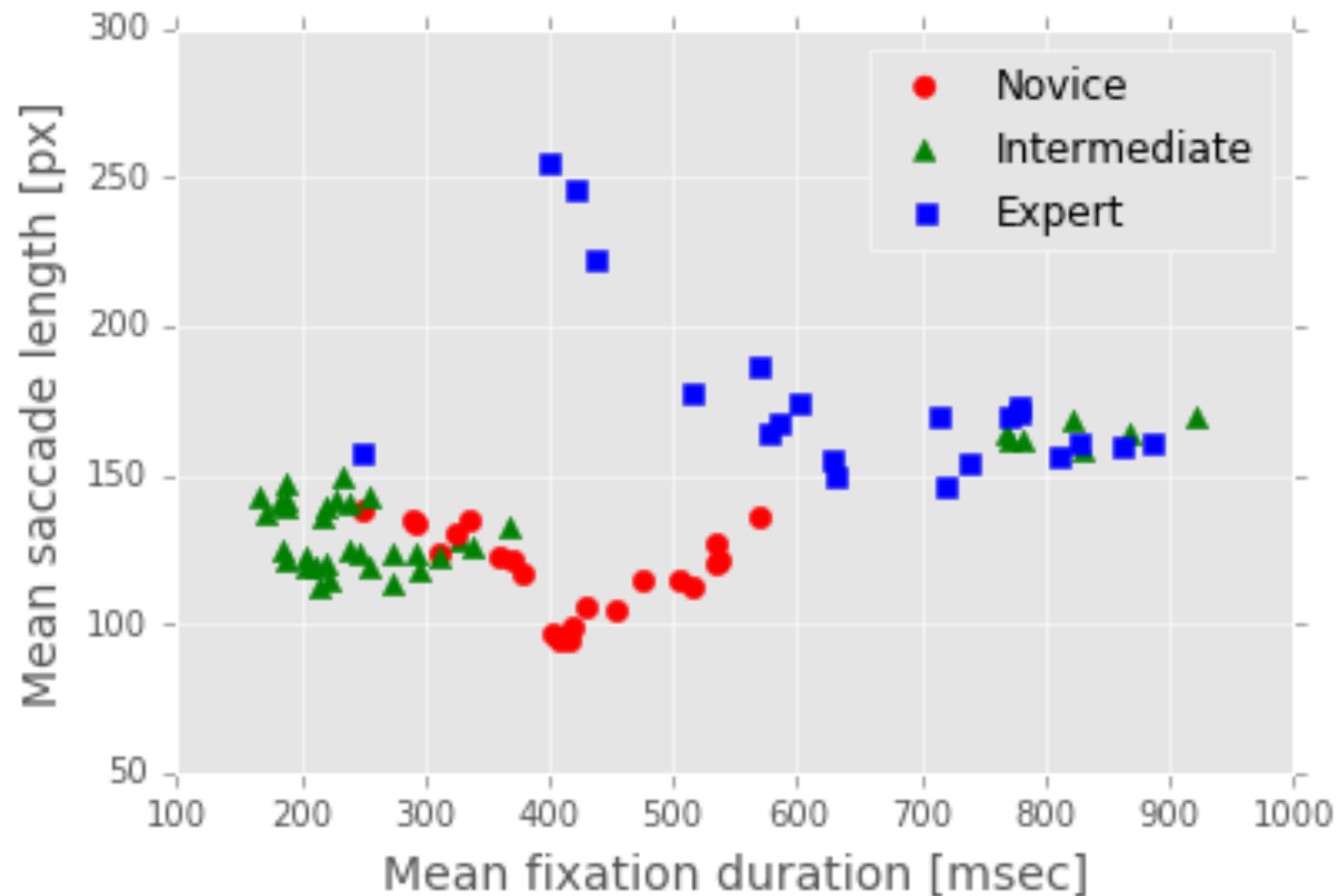


Expert



Comprehension Recognition

Feature representation of all participants' data in subsequences (1 min.) based approach.

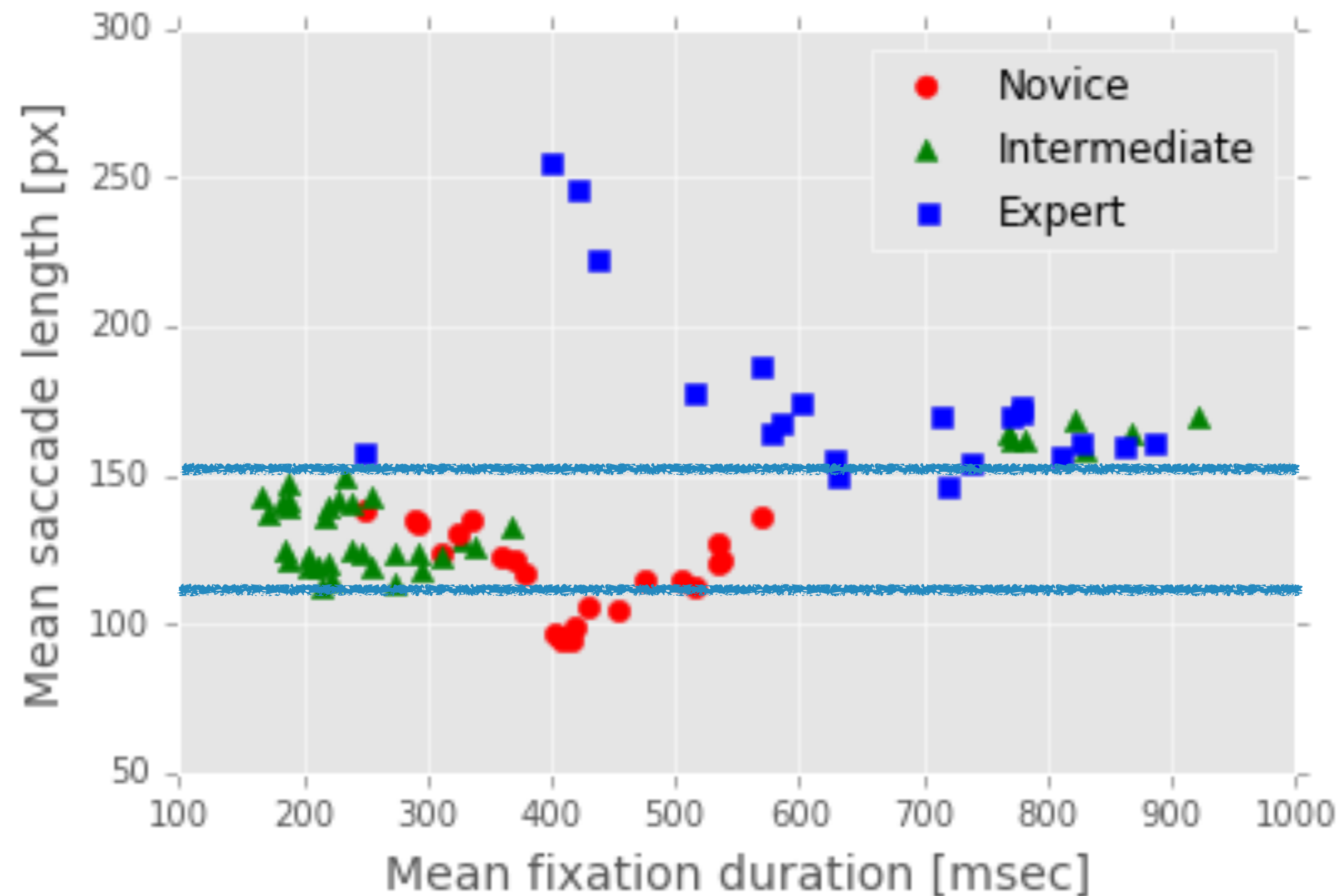


- 70% accuracy for classifying expertises into three classes (user-independent training)

		Predicted class		
		Nov.	Int.	Exp.
Actual class	Nov.	15	1	7
	Int.	4	27	7
	Exp.	5	1	15

Comprehension Recognition

Feature representation of all participants' data in subsequences (1 min.) based approach.

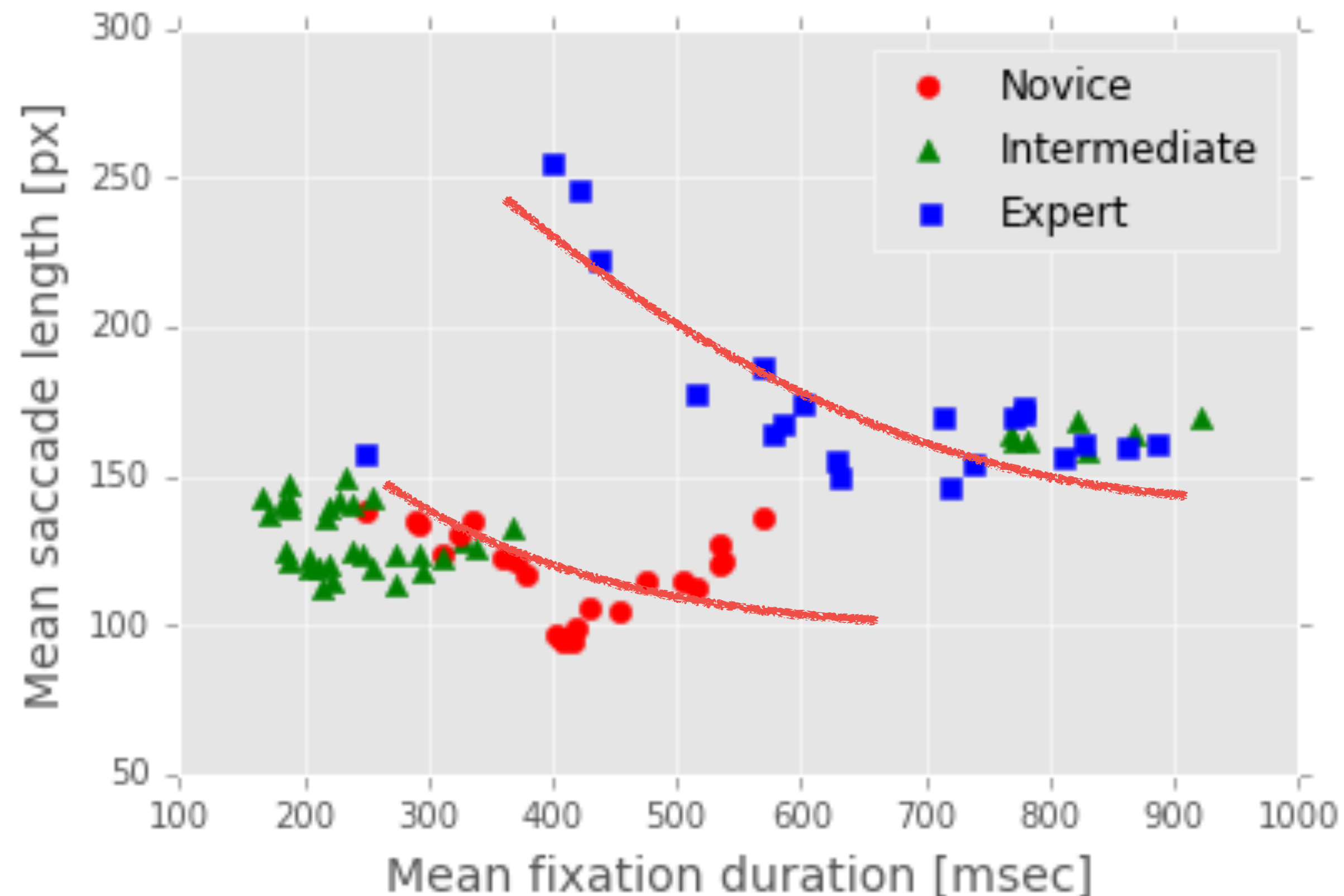


- 70% accuracy for classifying expertises into three classes (user-independent training)
- Novice students read a textbook slowly with small steps.

		Predicted class		
		Nov.	Int.	Exp.
Actual class	Nov.	15	1	7
	Int.	4	27	7
	Exp.	5	1	15

Comprehension Recognition

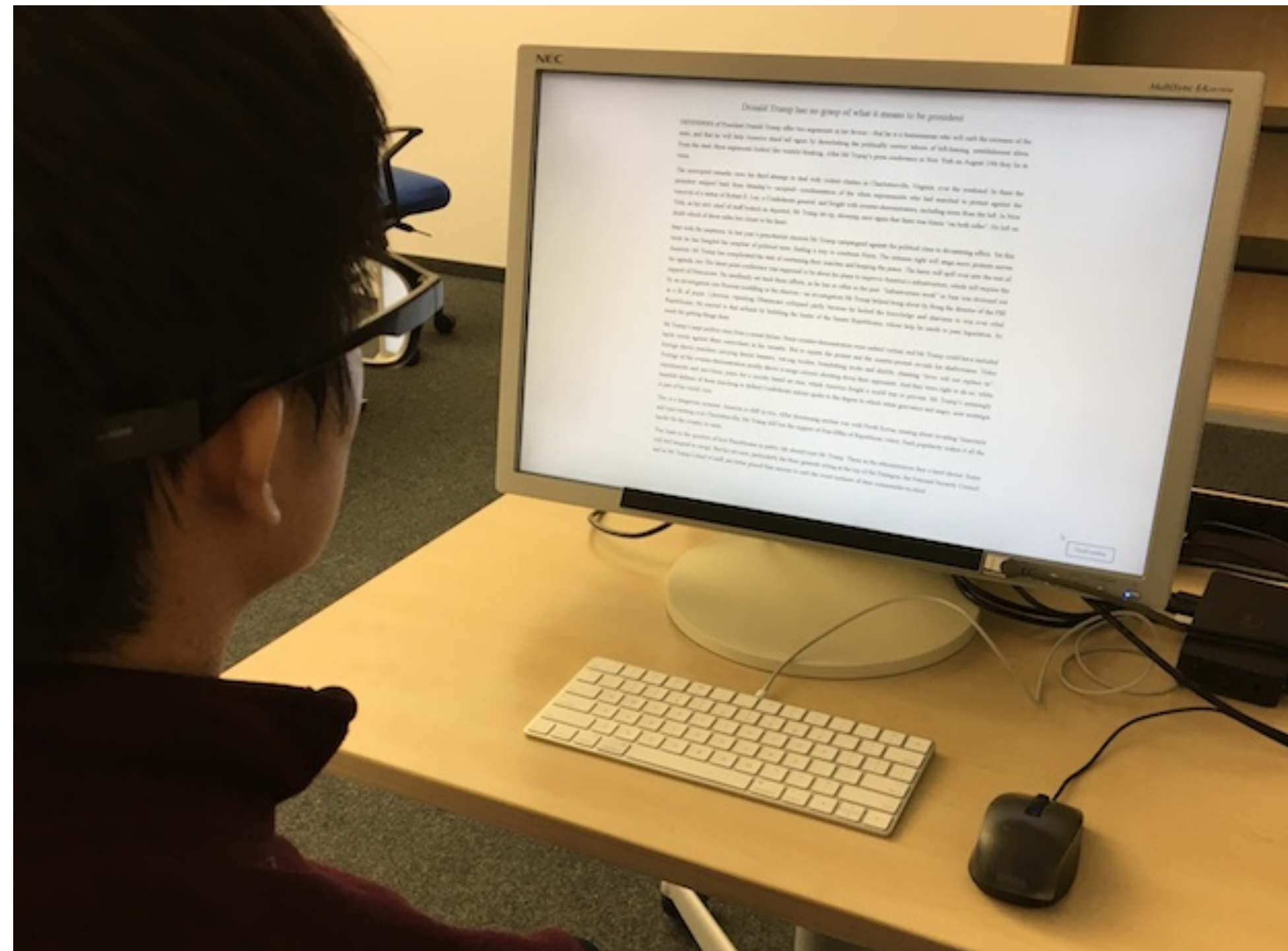
Feature representation of all participants' data in subsequences (1 min.) based approach.



- 70% accuracy for classifying expertises into three classes (user-independent training)
- Novice students read a textbook slowly with small steps.
- Expert students sometimes skip their eyes on the text and focus on the content they are interested in.

		Predicted class		
		Nov.	Int.	Exp.
Actual class	Nov.	15	1	7
	Int.	4	27	7
	Exp.	5	1	15

Experimental Design



The Enigma of Joseph Heller

But again, Daugherty is often perceptive about Heller's place in the larger culture, even if the novelist himself rarely comes into focus. For the human aspect, one turns to Erica Heller's frank but loving memoir of her father, "Yossarian Slept Here," which comes as close as possible, I dare say, to deciphering the enigma behind the obsessive, pitch-black fiction. Joseph Heller, the opposite of demonstrative, was given to oblique ways of showing affection, like pelting Erica and her brother with "gray, shriveled" snowballs he'd preserved in the freezer for summer fun, or by hiding behind a newspaper in the back of a bus so that Erica wouldn't know he was worriedly following her to school each morning. Such vignettes are all the more charming, and telling, because the author shares her subject's sense of humor, and is herself a good writer to boot. "I noticed the way my father was looking at Nixon alongside the way my mother was looking at my

Q1. How much of the article did you understand?

- None of it. 0%
- Some of it. 30%
- Most of it. 60%
- All of it. 100%

whom he's compulsively unamused, and a dreary 15-year-old daughter who likes to remark that she wouldn't mind terribly if he were dead. Recognizing real-life "verbatim conversations" in a chapter titled "My Daughter Is Unhappy," a "demolished" Erica Heller confronted her father. "What makes you think you're interesting enough to write about?" he replied.

The miracle of this memoir is that it never seems less than fair: Erica Heller's worst grievances are mentioned more in sorrow (or levity) than anger, and she's careful to give her own shortcomings their due. Delightfully — after her own, rather dismal tale is told — she steps aside to allow testimony from her father's colleagues and friends, like Christopher Buckley (who, she slyly reminds us, wrote an "incomparably restrained" memoir about his parents, "Losing Mum and Pup"): all are glowing, and persuasively so, since Heller was capable of being (as Updike would have it) very "sweet" indeed, such that even his family seemed to forgive him everything in the end.

Finish reading

- 13 university students read 18 newspaper articles (555 ± 70 words)
- Device: SMI REDn scientific 60Hz eye tracker
- Calibrate the device → Rest (20 sec.) → Read a document → Answer surveys

Interest Recognition Overview and Result

Feature calculation

No.	feature
1-2	{mean, standard deviation} of fixation duration
3-4	{mean, standard deviation} of forward saccade length
5-6	{mean, standard deviation} of forward saccade speed
7-8	{mean, standard deviation} of regression length
9-10	{mean, standard deviation} of regression speed
11-12	count of {forward saccades, regressions}
13	regression ratio
14-15	{mean, standard deviation} of pupil diameter
16	blink frequency
17	standard deviation of blink interval

and a tube of ringed wool ("brings bryaic").
 Again, in the Edda Gunnar one of the Beasts of Germany says, "My helmet and my white shield come from the Hall of Klars." These quotations are hardly sufficient evidence of it perhaps, but it seems as if in the tenth century white shields were borne by leudars and red ones by the common soldiers.—every one who possessed six marks.



SVM

		Predicted class			
		1	2	3	4
Actual class	1	36%	31%	26%	7%
	2	20%	36%	24%	20%
	3	19%	21%	38%	22%
	4	10%	14%	16%	59%

mean: 44%

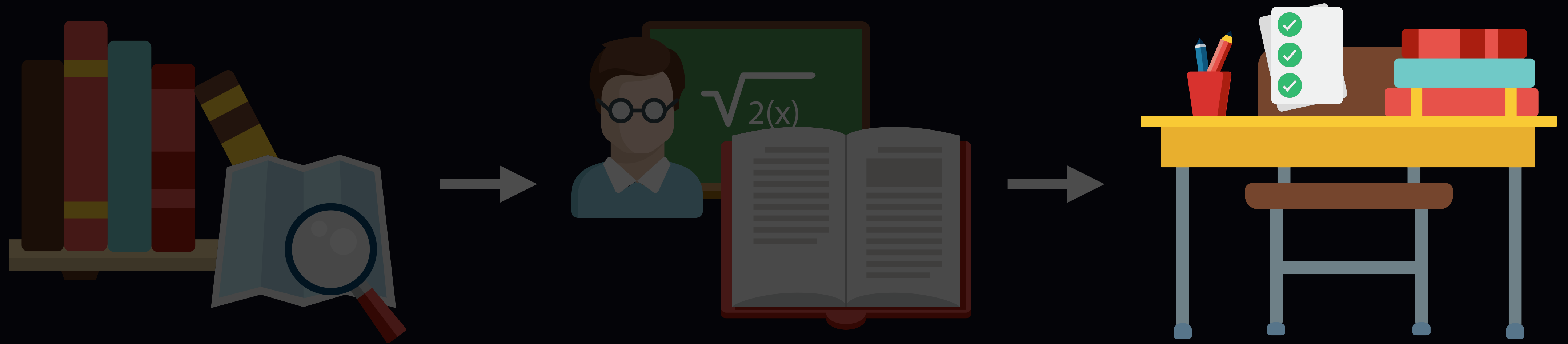
CNN-LSTM

		Predicted class			
		1	2	3	4
Actual class	1	62%	25%	6%	7%
	2	16%	66%	15%	3%
	3	2%	7%	83%	8%
	4	1%	0%	6%	93%

mean: 78%

In user-dependent evaluation (leave-one-document-out cross-validation in a participant), SVM classified interests into 4 levels with 44% accuracy, CNN-LSTM classified with 78% acc.

Augmented Positive Learning Cycle



Daily habits and curiosity

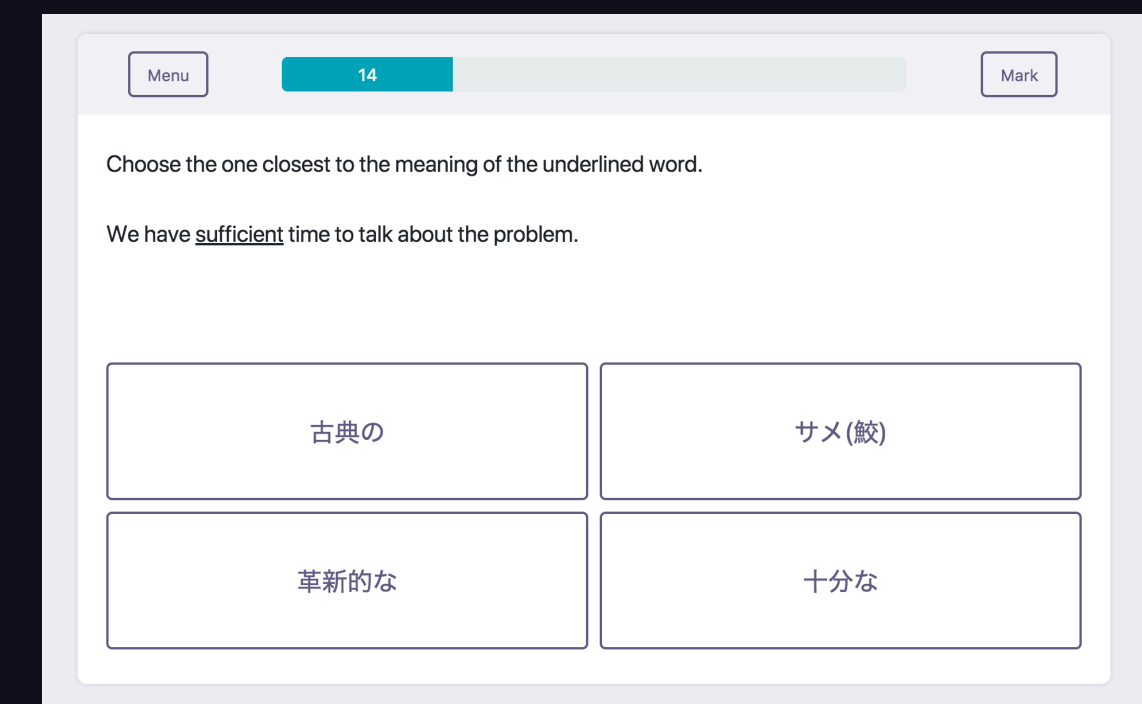
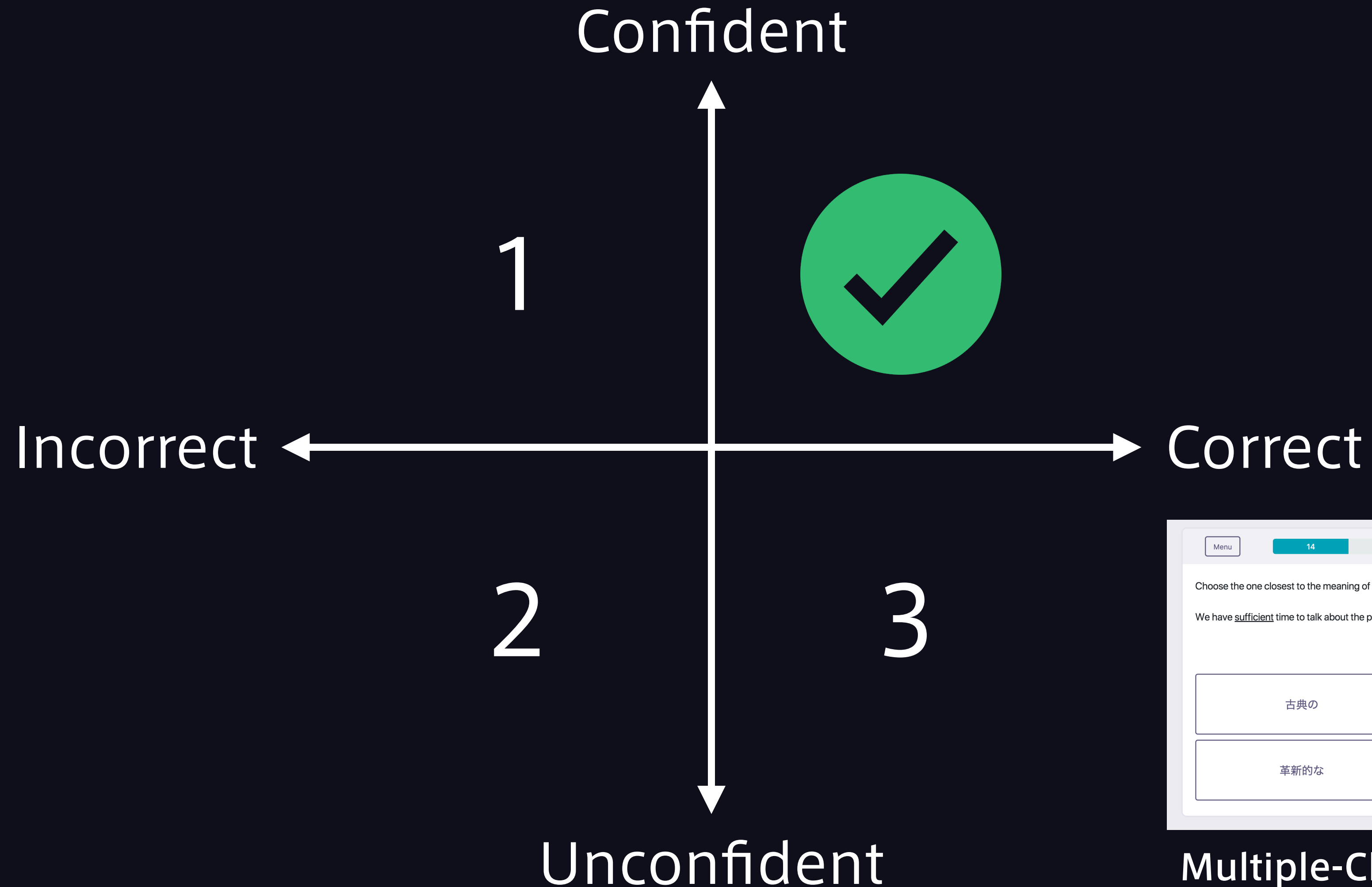
Lecture and textbook

Questions (test)

Can AI understand interest of students and nudge them to learn voluntarily?

Which question should be reviewed carefully? (considering priority & leakage)

CoALA: Confidence-Aware Learning Assistant



Multiple-Choice Question

Gaze-Based Self-Confidence Estimation

Start Q1 Q2 Q3 Q4 Q5 Q6 Q7 Finish

Q. Fill in the blank with the most appropriate answer.

This program is aimed at providing a solid ____ of numerous financial instruments for investors.

understand understood
understanding understandable

The screenshot shows a quiz interface with a progress bar at the top indicating the current question (Q2). The question asks for a word to fill in the blank. Below the question, a network diagram of related words is displayed. The word 'understood' is highlighted in a dark blue box, indicating it is the selected answer. Other words in the network include 'understand', 'understanding', and 'understandable'.

Decision with confidence

Start Q1 Q2 Q3 Q4 Q5 Q6 Q7 Finish

Q. Fill in the blank with the most appropriate answer.

Simon Motors requires that its employees ____ company-sponsored workshops at least twice a year.

attend attending
to attend attended

The screenshot shows a quiz interface with a progress bar at the top indicating the current question (Q3). The question asks for a word to fill in the blank. Below the question, a network diagram of related words is displayed. The word 'attend' is highlighted in a dark blue box, indicating it is the selected answer. Other words in the network include 'attending', 'to attend', and 'attended'.

Decision without confidence

スタディサプリ

Product Team Blog

2019-02-25

「目は口ほどに物を言う？」アイトラッキングを活用した英単語学習支援システムの試み

Data

はじめまして。データリサーチエンジニアの @bwtakacy です。私は、データ分析やAI技術を活用して学習をサポートする仕組みや機能の研究開発を行うチームに所属しています。今回は、研究開発の取り組みの一つをご紹介します。

英単語学習に新しい「目線」を

早速ですが、以下の二つの画像を見てください。英単語の意味を答える問題をPC上で解いてもらった時の目線の動きをヒートマップと経路として表示したものです。

一つは、あまり目移りせずに正解を選んだパターン。



CoALA



Q2 **【Incorrect】** Low confidence

Read the answer

Feedback as high conf.

Q3 **【Correct】** High confidence

Read the answer

Feedback as low conf.

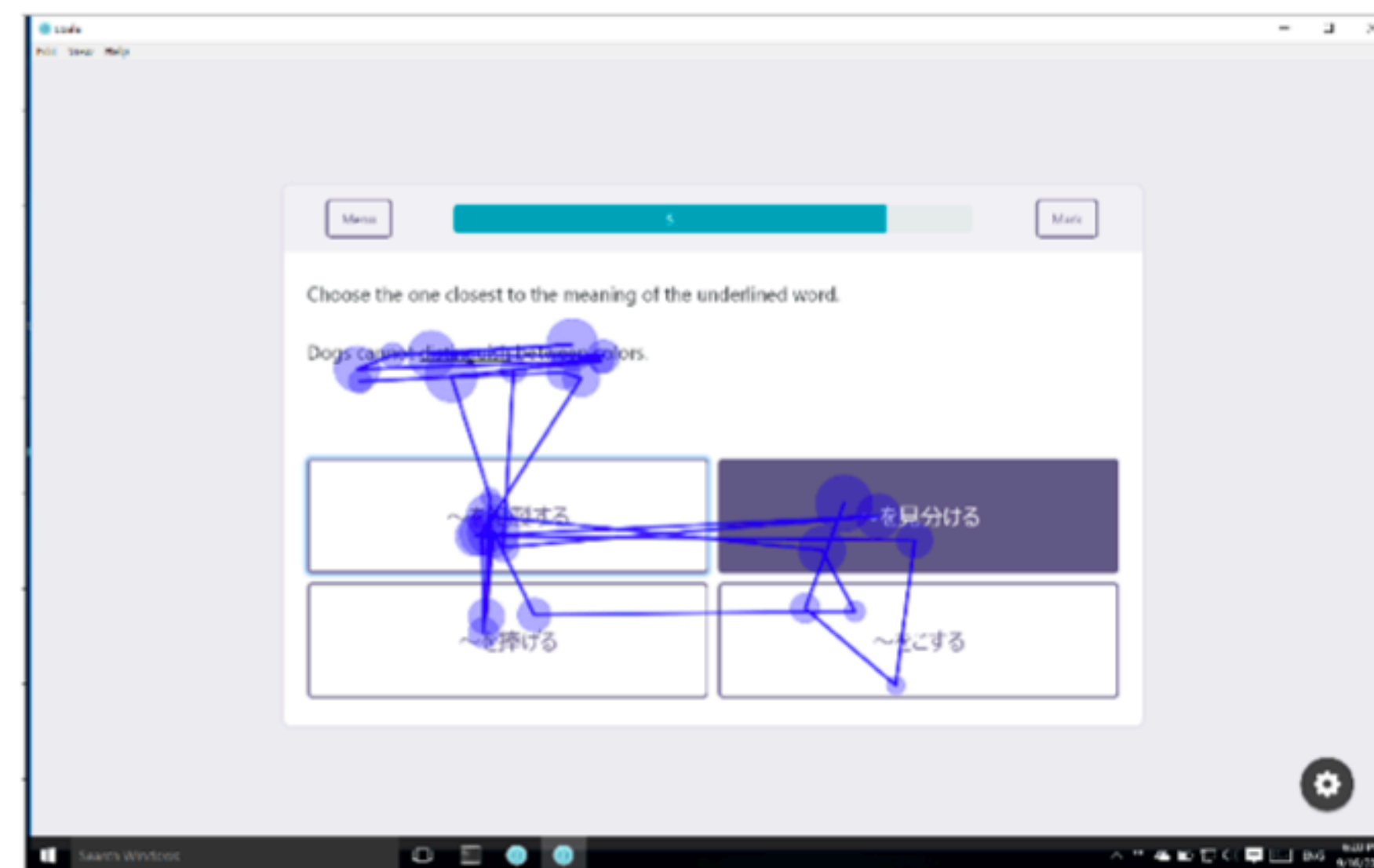
Q4 **【Correct】** Low confidence ☁

Read the answer

Feedback as high conf.

Dogs cannot distinguish between colors.

1. ～を複製する
2. ～を見分ける **Correct Selected**
3. ～を捧げる
4. ～をこする



Q5 **【Correct】** High confidence

Read the answer

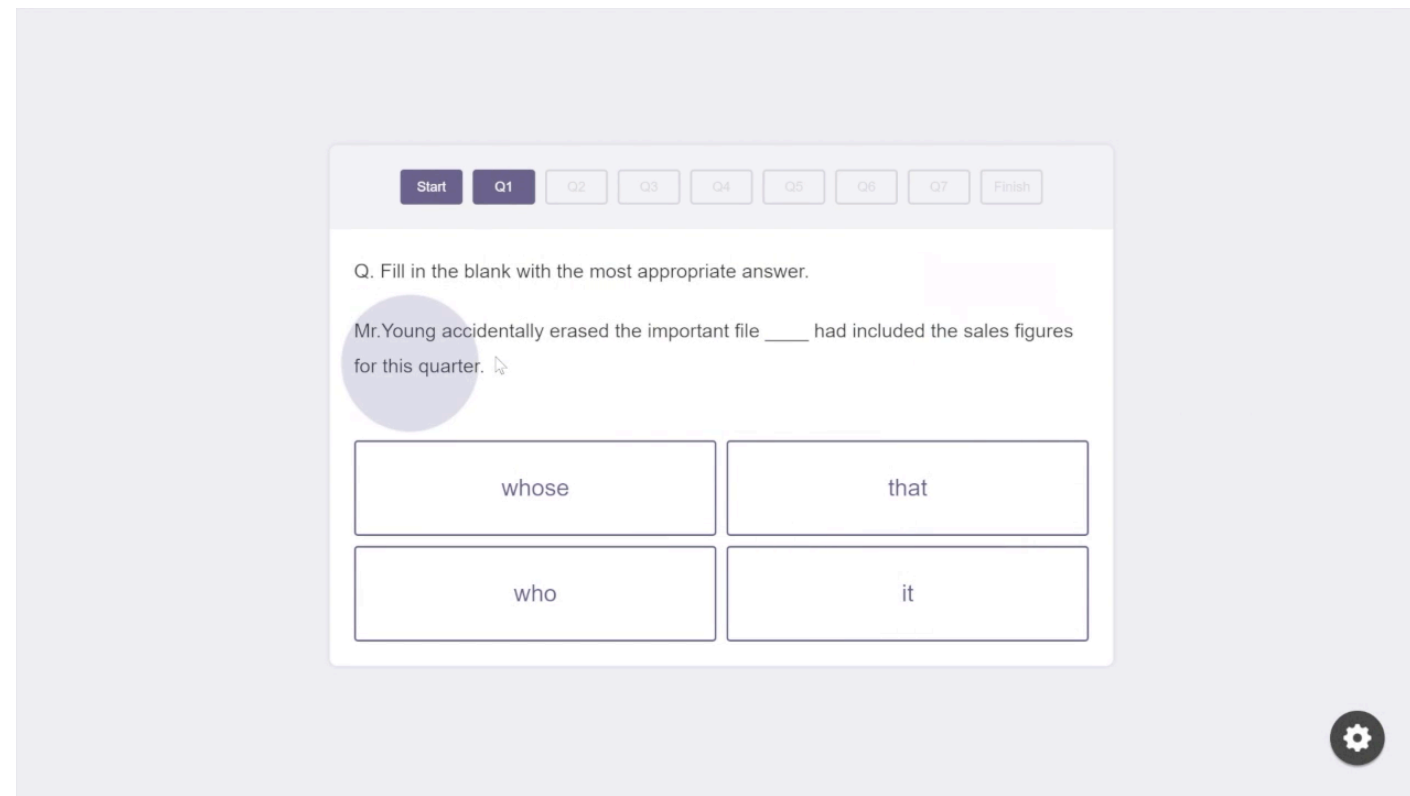
Feedback as low conf.

☂ : Incorrect answer with high confidence

☁ : Correct answer with low confidence

System Overview

Preprocessing fixation/saccade detection



Feature calculation

No.	Feature
1-2	fixation {count, ratio} on Choices
3-4	fixation {count, ratio} on Question
5-8	{sum, mean, max, min} of fixation durations on Choices
9-12	{sum, mean, max, min} of fixation durations on Question
13-14	variance of {x, y} coordinate of fixations
15-16	{sum, mean} of saccade length
17-20	saccade count: {all, on Question, between Choices, between Question and Choices}
21-24	{sum, mean, max, min} of saccade durations
25-28	{sum, mean, max, min} of saccade speeds
29	reading-time
30	correctness of the answer

Classification

- Binary classification
- Support Vector Machine (SVM)
 - RBF kernel,
 - $C=1, \gamma=0.125$

We utilize the following five effective features selected by Forward-stepwise feature selection.

f5 sum of fixation durations on Choices, **f13** variance of x coordinate of fixations

f19 the number of saccades between choice areas, **f21** sum of saccade durations, **f29** reading-time.

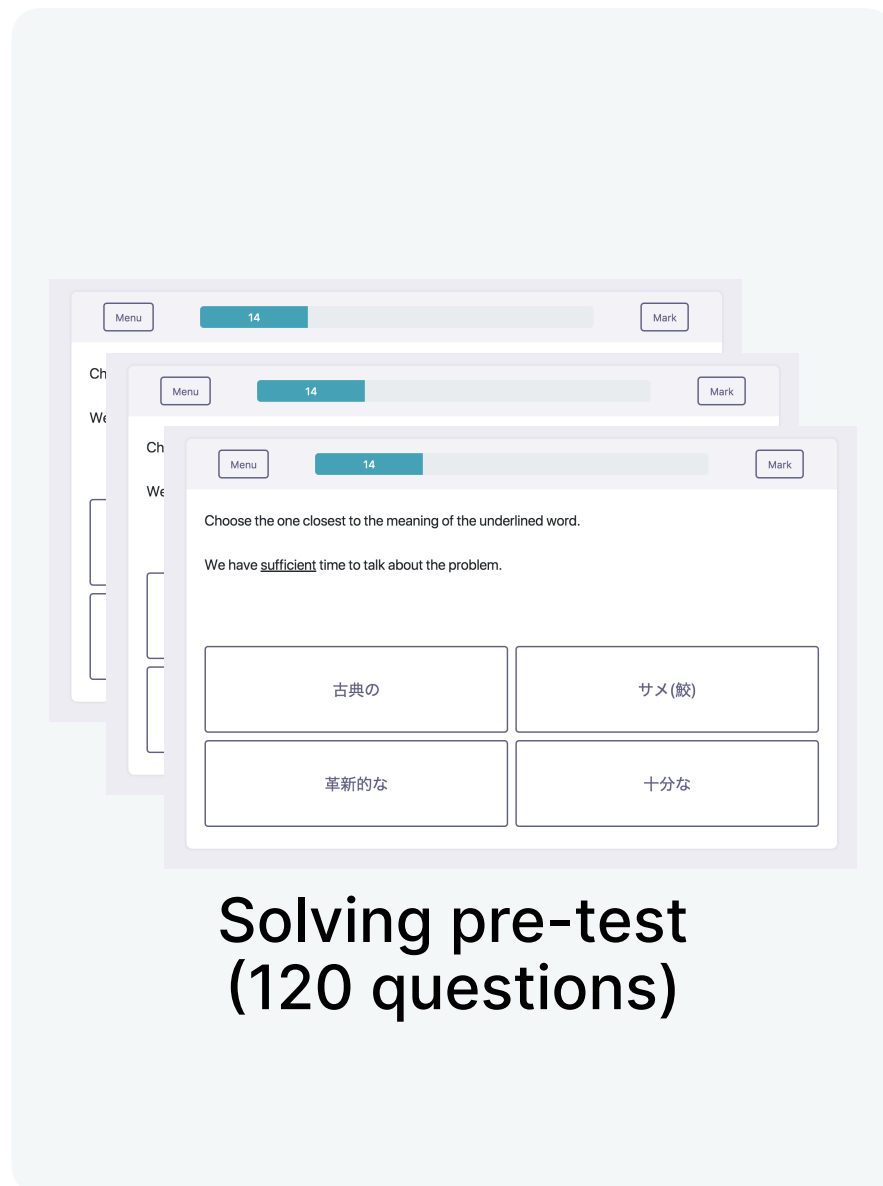
Research Hypotheses

- **RH1:** Correct answers without confidence (**vague knowledge**) tend to be incorrectly answered when asked again compared to correct answer with confidence.
- **RH2:** Incorrect answers with confidence (**misunderstanding**) tend to be mistaken again compared to incorrect answers without confidence.
- **RH3:** Estimating self-confidence from learning behaviors and giving feedback (e.g., adding questions to a review list, highlighting them while reviewing) avoids the scenarios outlined in RH1 and RH2.

Experimental Design

Day 1

Confidence estimation



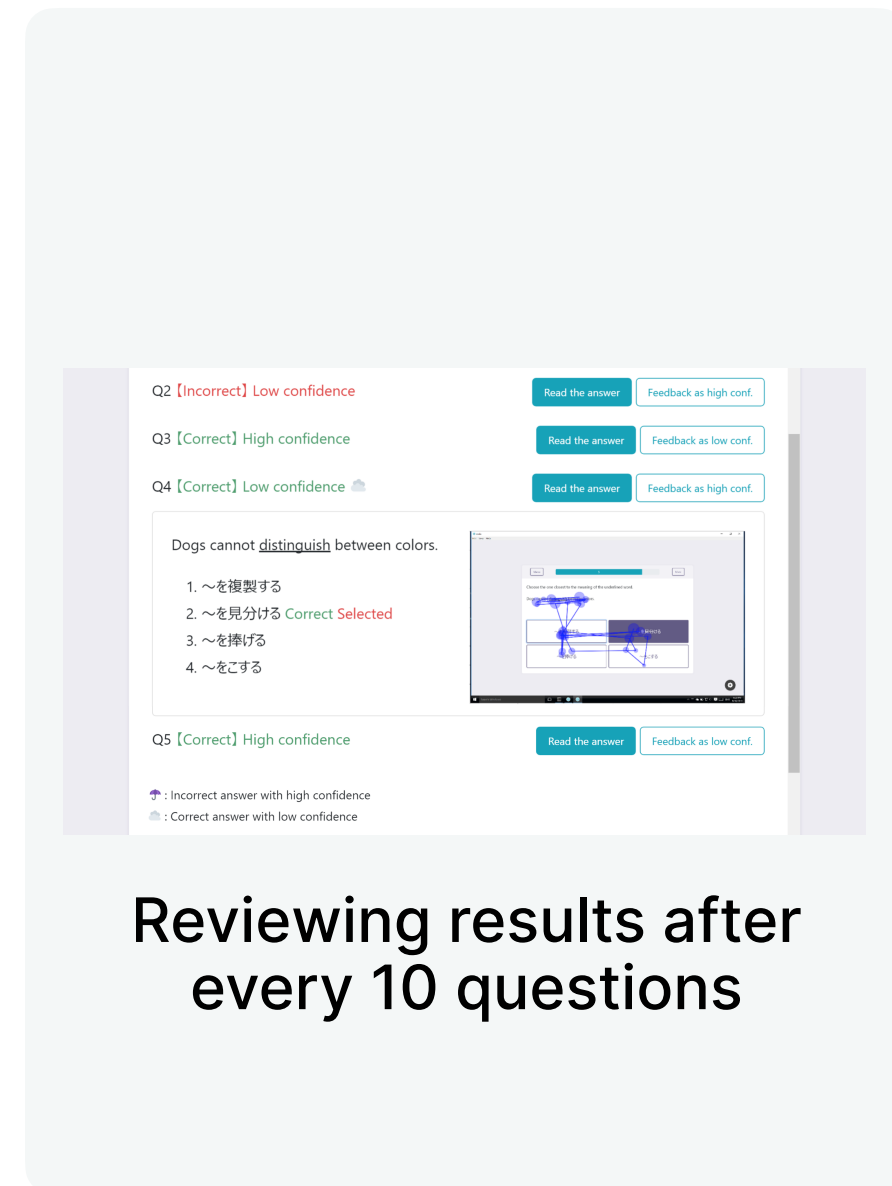
Solving pre-test (120 questions)

Correct w/ conf.

Correct w/o conf.

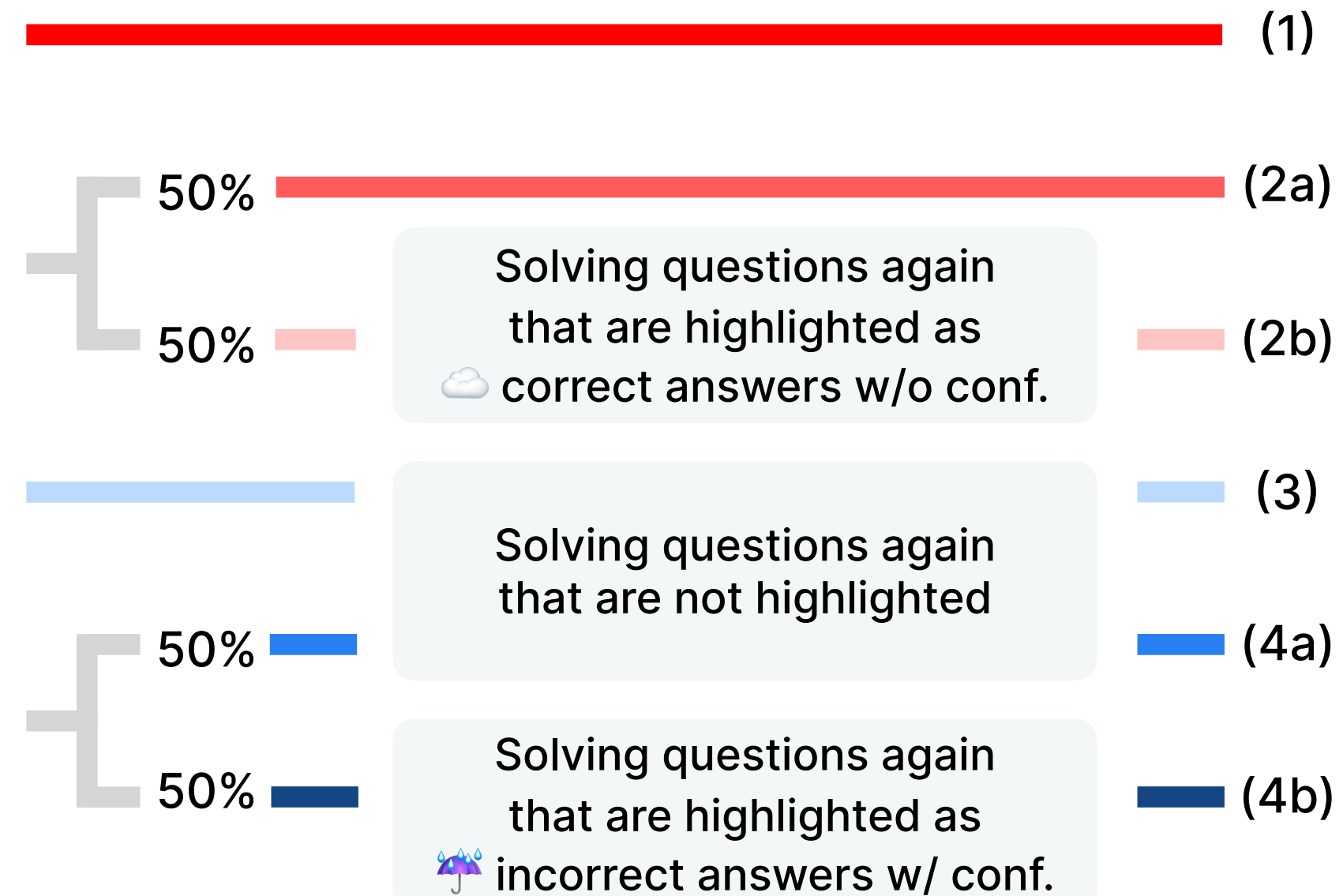
Incorrect w/o conf.

Incorrect w/ conf.



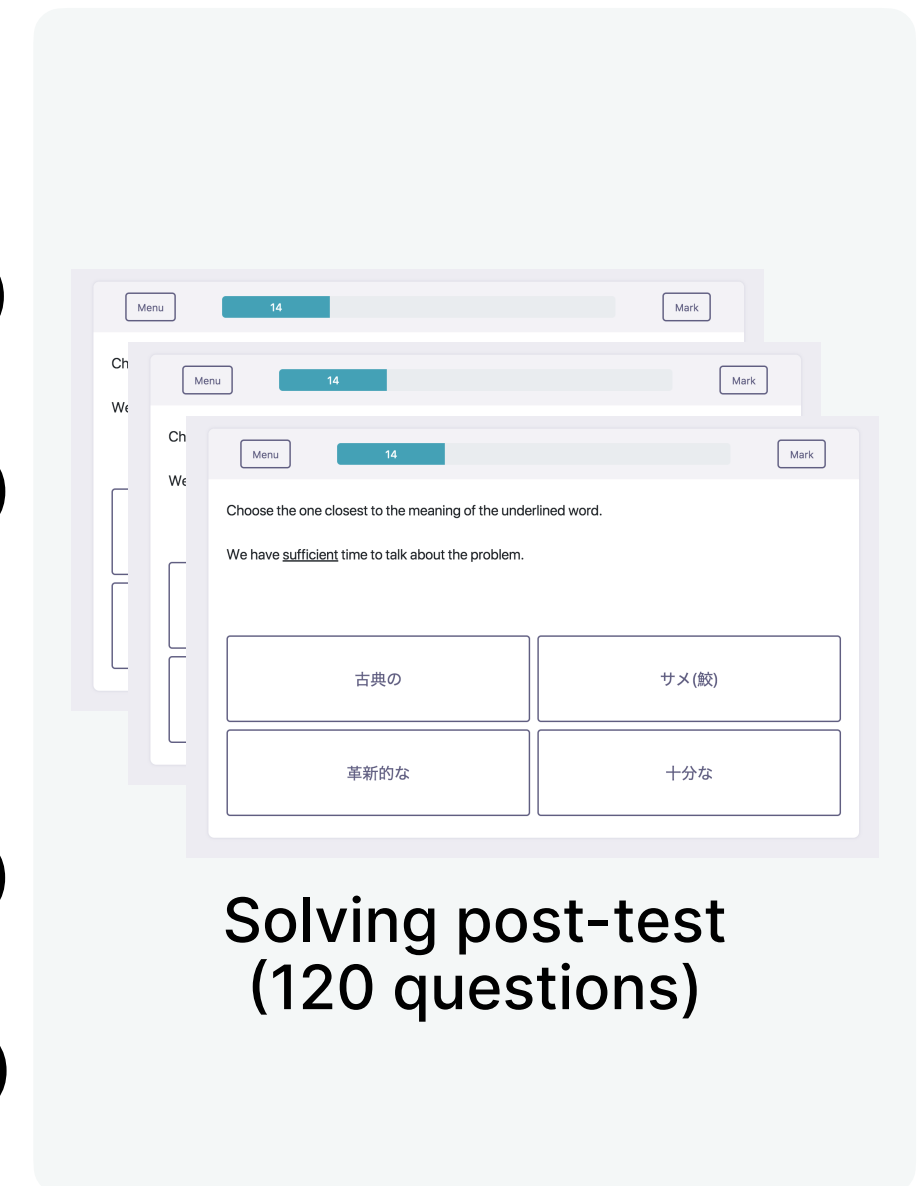
Reviewing results after every 10 questions

Experimental and control group division



Day 3

Day 5

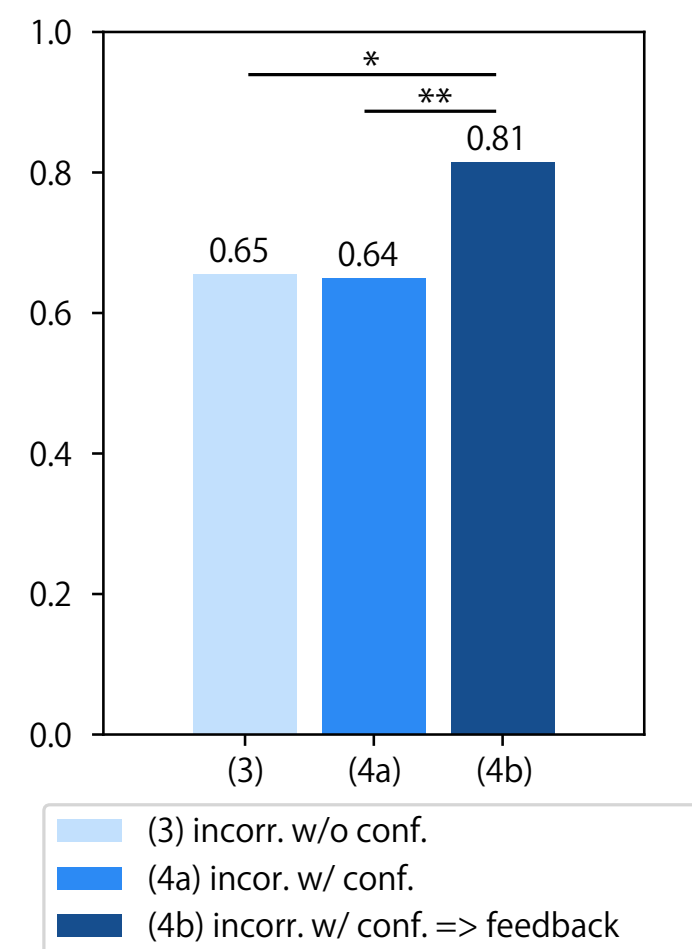
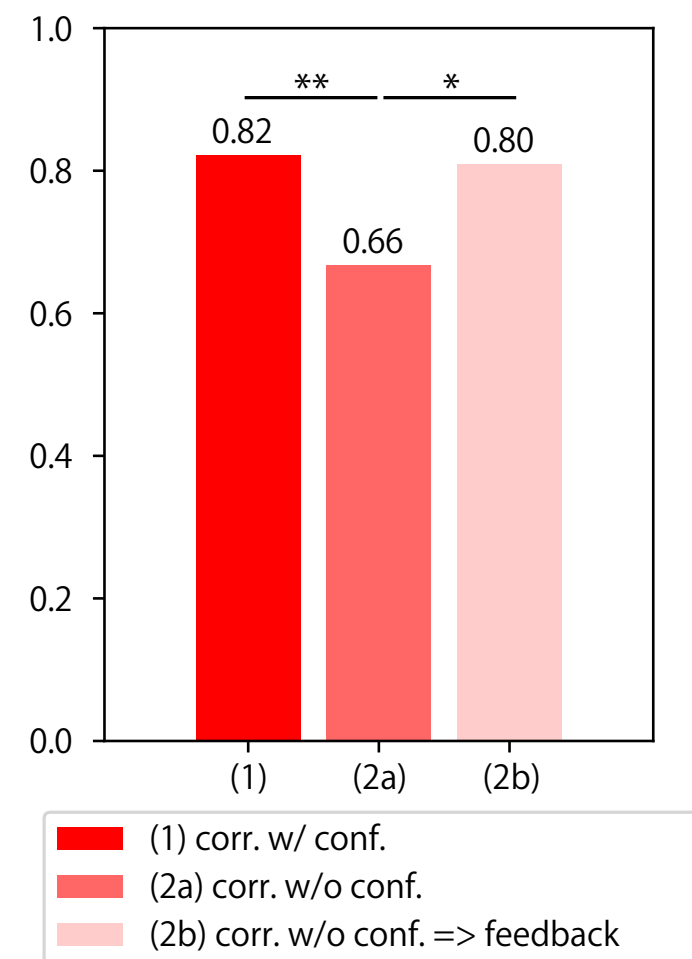


Solving post-test (120 questions)

Comparison of...

- (1) and (2a): How often **answers without confidence are forgotten**.
- (2a) and (2b): How CoALA is effective in a such condition.
- (3) and (4a): How often **answers with confidence are incorrectly answered again**.
- (4a) and (4b): How CoALA is effective in a such condition.

Experimental Results

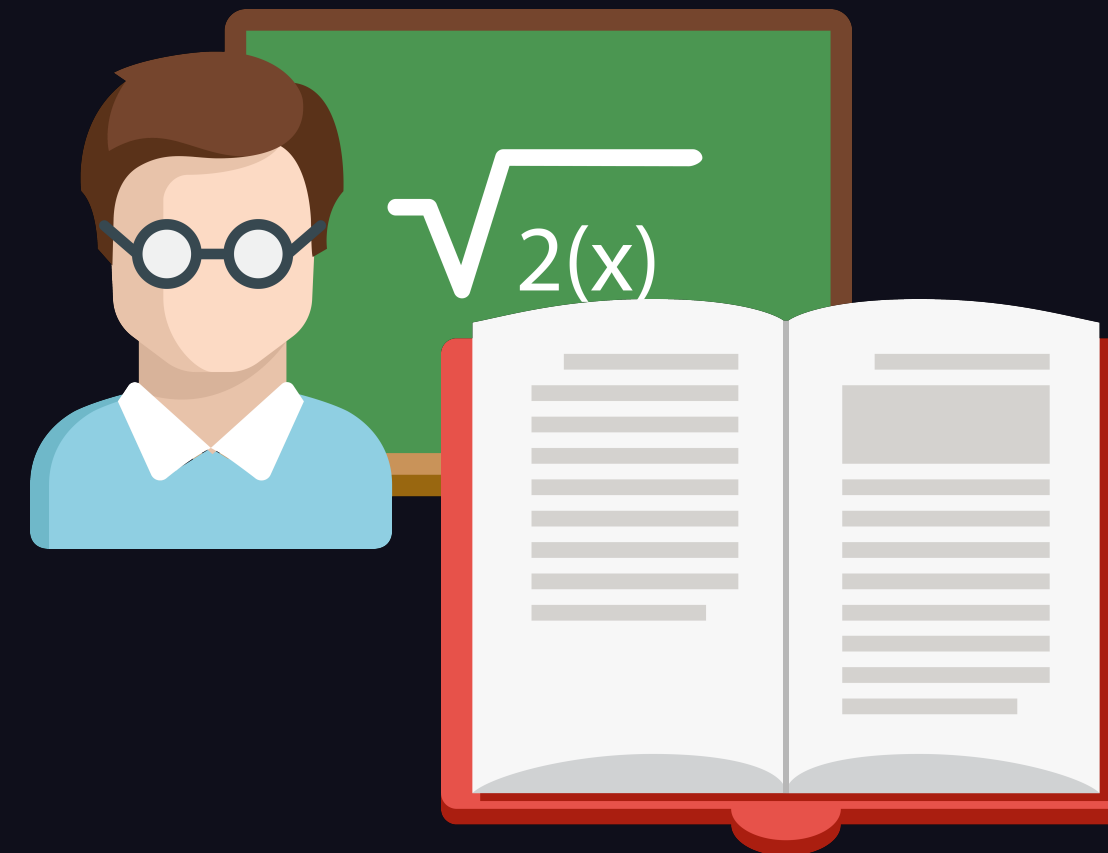


- **RH1:** Correct answers without confidence tend to be incorrectly answered when asked again compared to answer with confidence.
 - Supported. (2a) dropped 16% compared to (1) ($p < 0.01$)
- **RH2:** Incorrect answers with confidence tend to be mistaken again compared to incorrect answers without confidence.
 - Not supported. No significant difference between (3) and (4a)
- **RH3:** Estimating self-confidence from learning behaviors and giving feedback (e.g., adding questions to a review list, highlighting them while reviewing) avoids such scenarios.
 - Supported. Comparisons between (2a) vs (2b), (4a) vs (4b).

Augmented Positive Learning Cycle



Daily habits and curiosity



Lecture and textbook



Questions (test)

Can AI understand interest of students and nudge them to learn voluntarily?

Which question should be reviewed carefully? (considering priority & leakage)

Wordometer

HyperMind

CoALA

Making people smarter... What does "smart" mean?

Learning

Acquiring knowledge in a usable form

We propose intelligent learning assistants that recognize cognitive/affective states (e.g., attention, interest, workload, self-confidence) and optimize learning materials for individuals to improve their learning performance.

Reasoning

Making a decision based on knowledge

Smart sensing technologies enable people to track their daily activities and make rational decisions on the basis of activity logs and future projections. We also investigate nudging techniques that encourage change in their behavior.

Communicating

Sharing knowledge with others

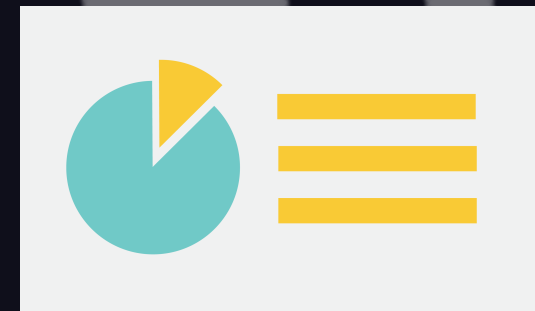
Communicating with others is an important activity for coming up with new ideas and transferring skills. We quantify social activities such as teaching, presenting, and discussing by sensors to support collaborative work.

PsyberLab defines intelligence as the ability of learning, reasoning, and communicating

Learning Cyclotron



Perceive
Lecture, Textbook



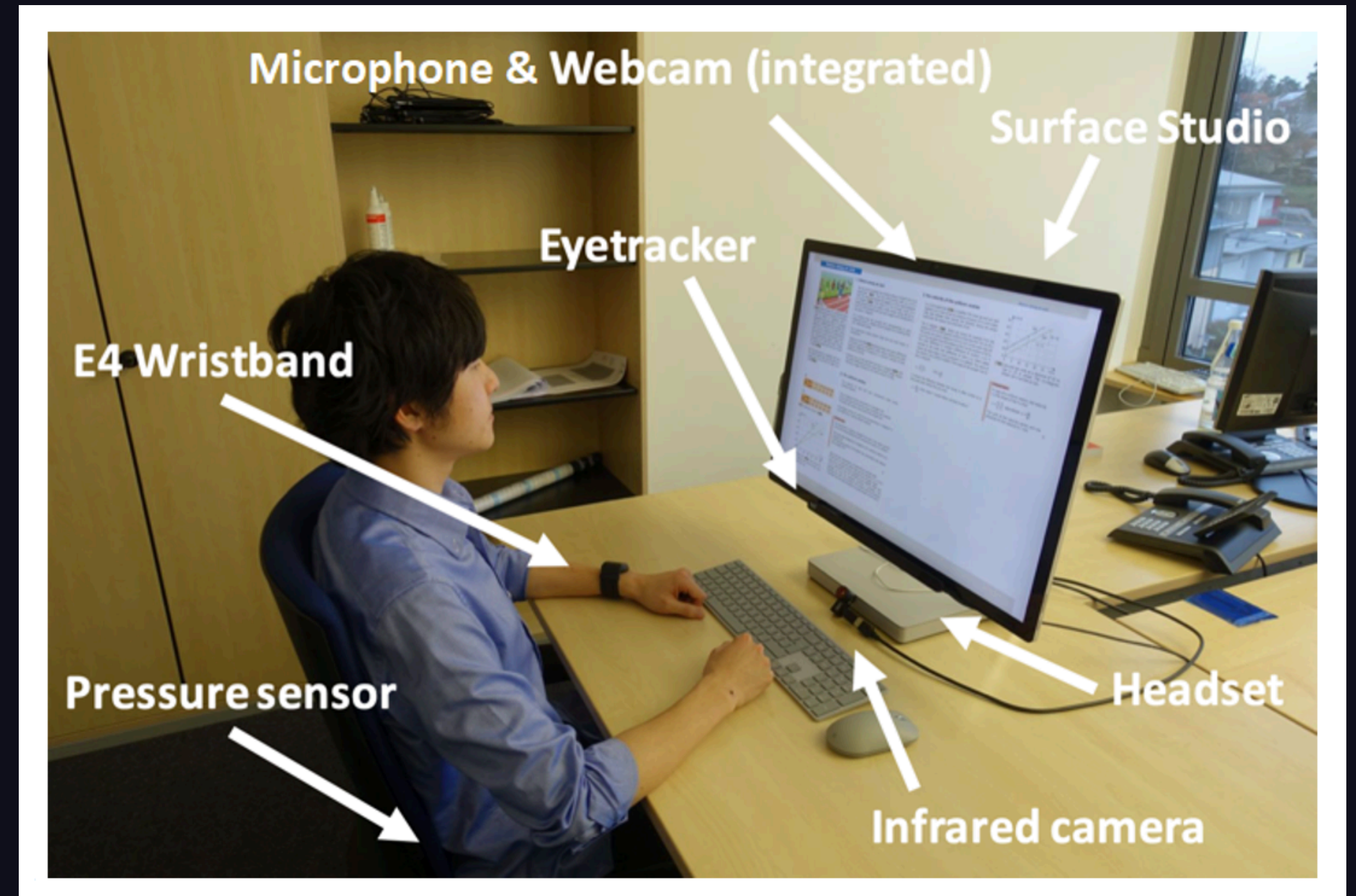
Transfer

Presentation, Discussion



Master

Exercises, Experiment

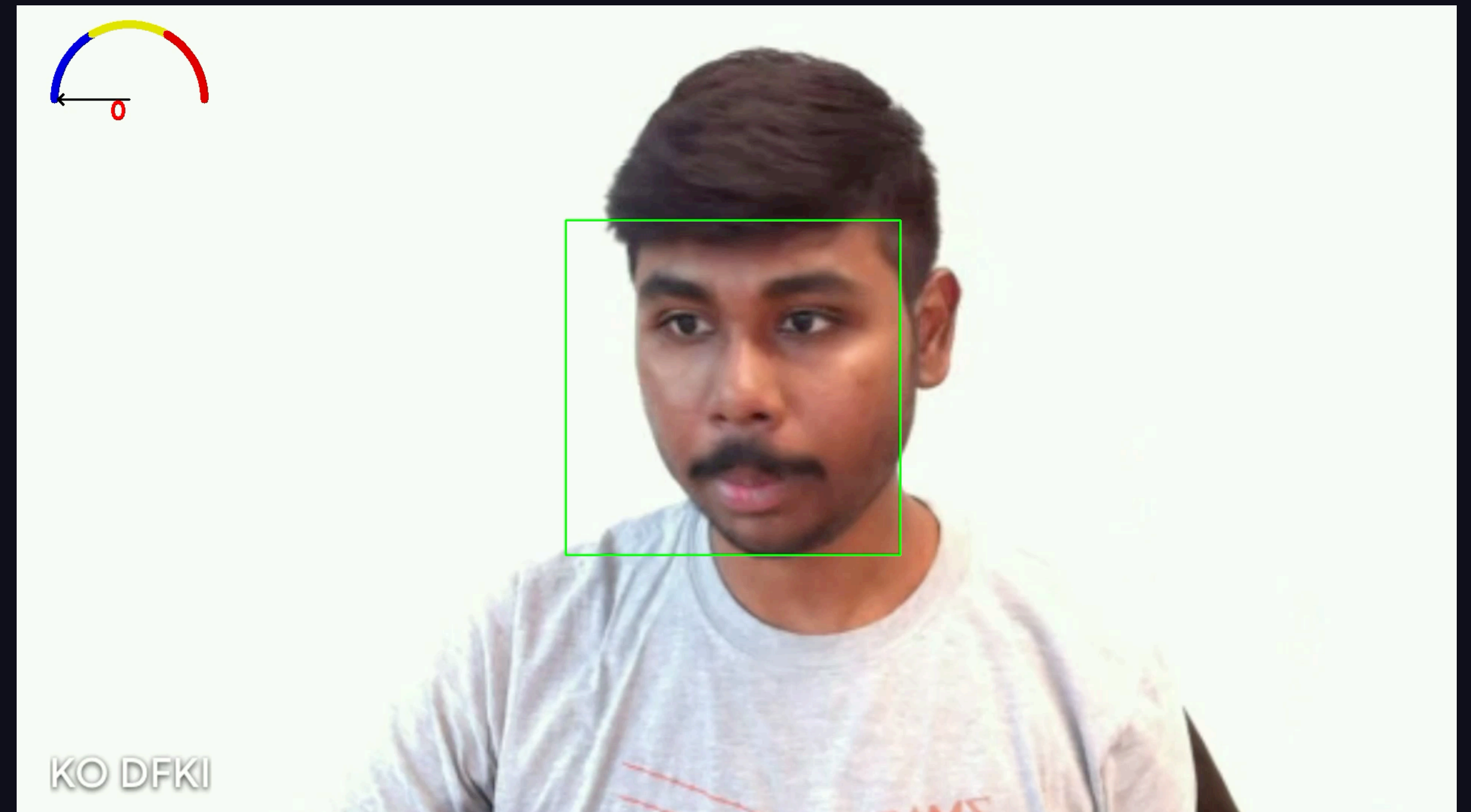


Typical experimental setup

Meeting Activity Analysis

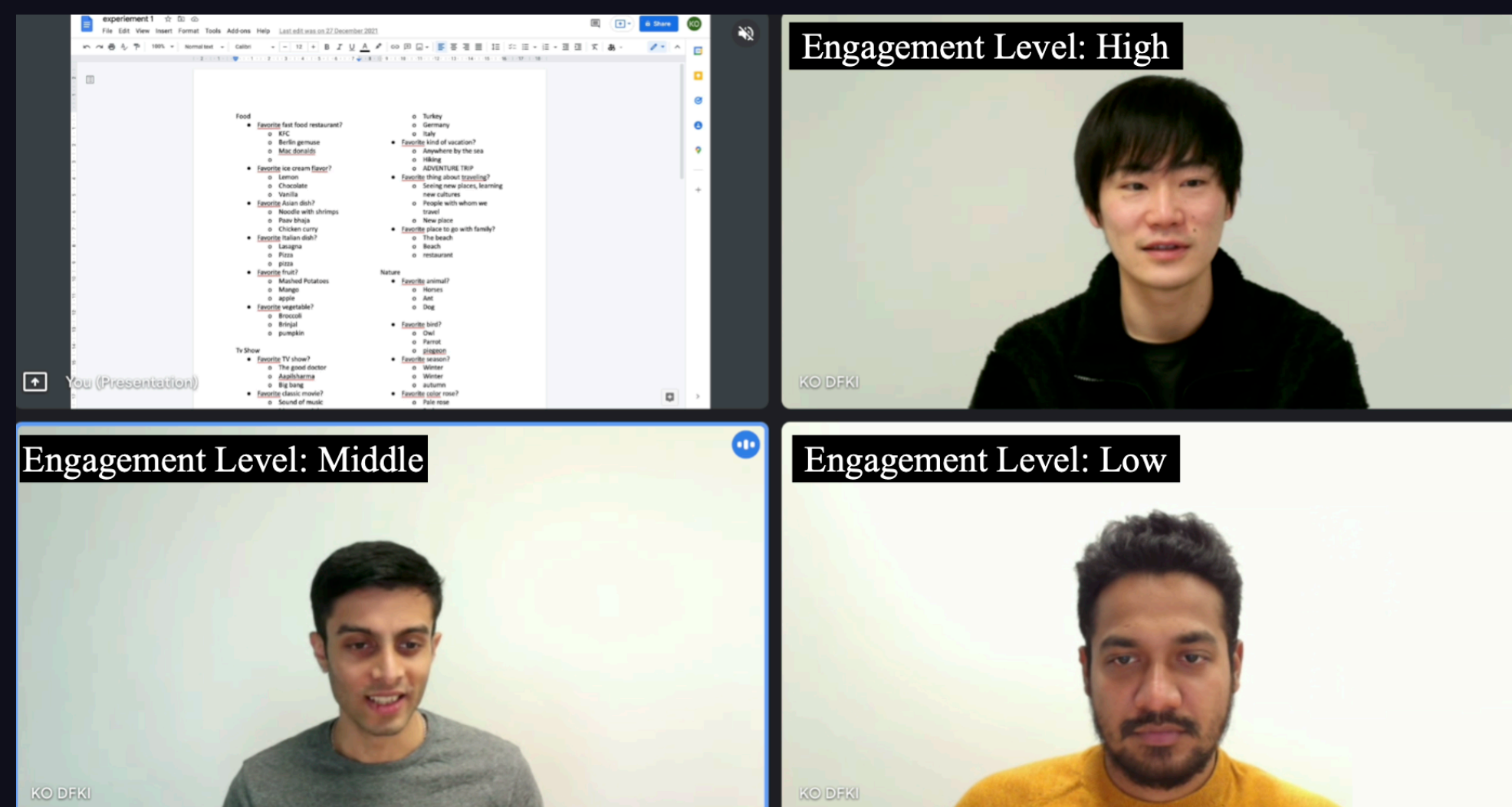


Detecting talking, nodding, smiling
(F1-score: 0.94, 0.79, 0.71)



The level of Engagement into three classes
(Mean f1-score: 0.92)

Estimating the level of engagement level in a video conference



High-engage role
without distractor

Mid-engage role
required to type meeting notes

Low-engage role
required to read a document
and type its summary

- We involved 24 participants and divided them into 8 groups (=3 members).
- We assigned 3 different roles to the participants: high, middle, low-engage
- We asked participants to rotate the roles and perform $3 * 2$ times in total.

Approaches and Results



Table 3: Comparisons of F1 scores in leave-one-group-out cross-validation (LOGOCV).

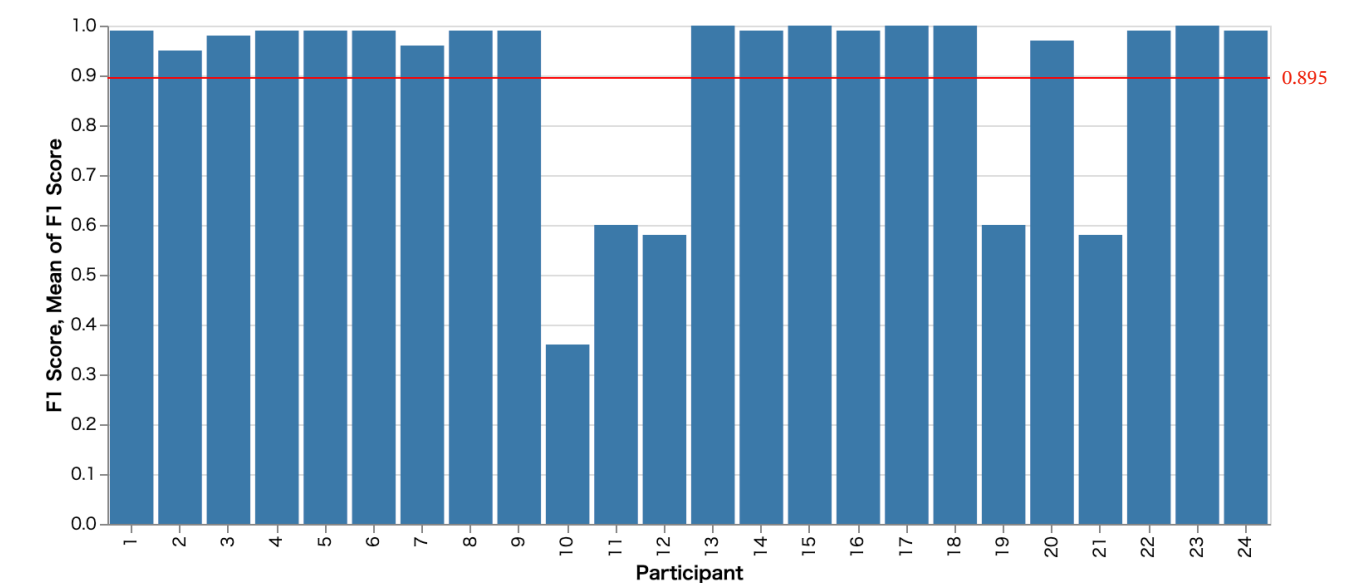
Test dataset (group)	1	2	3	4	5	6	7	8	Mean
Random Forest	0.264	0.464	0.521	0.476	0.520	0.519	0.429	0.546	0.467
VGG 16	0.347	0.562	0.623	0.545	0.612	0.602	0.512	0.589	0.549
Xception	0.895	0.975	0.902	0.748	0.957	0.961	0.821	0.981	0.905
MobileNetV1	0.898	0.981	0.911	0.755	0.982	0.980	0.833	0.996	0.917
MobileNetV2	0.901	0.999	0.915	0.761	0.985	0.985	0.835	0.999	0.923

Actual class	Predicted class		
	High	Middle	Low
High	50%	25%	23%
Middle	7%	47%	45%
Low	4%	53%	42%

Random Forest
(existing, F1: 0.46)

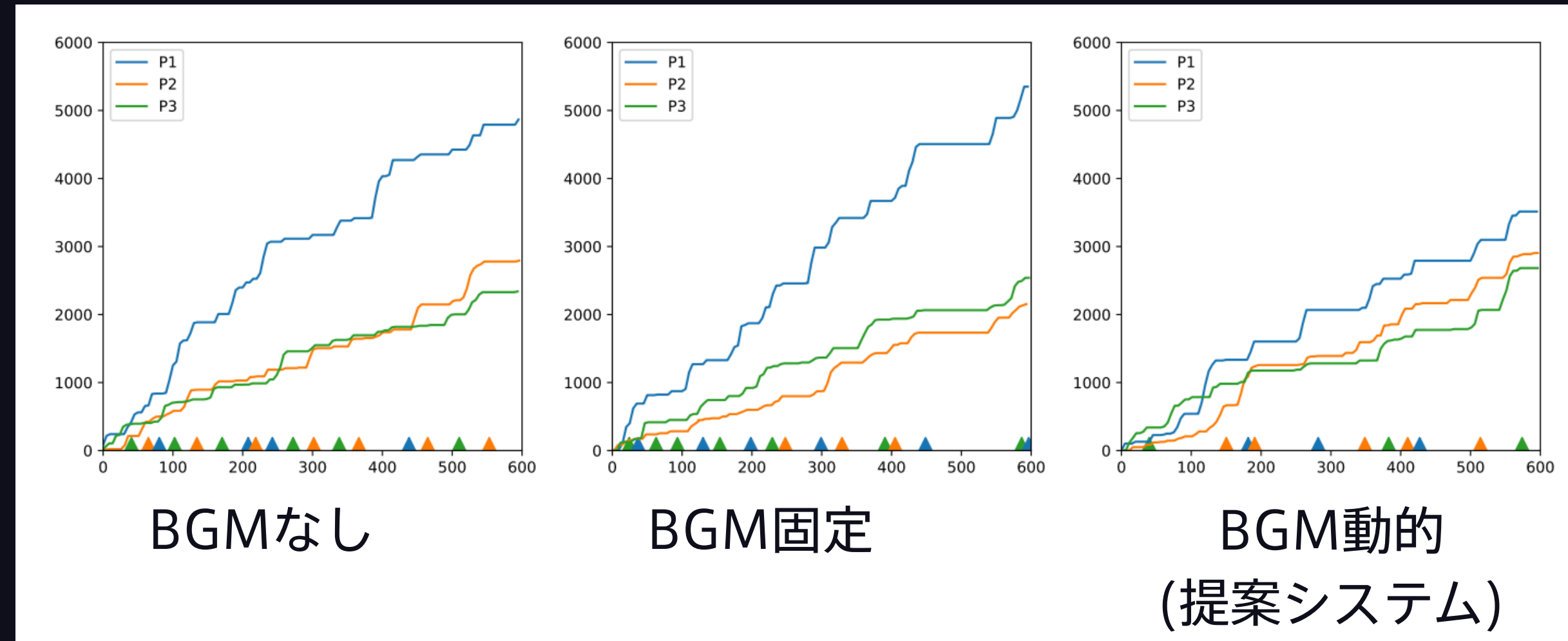
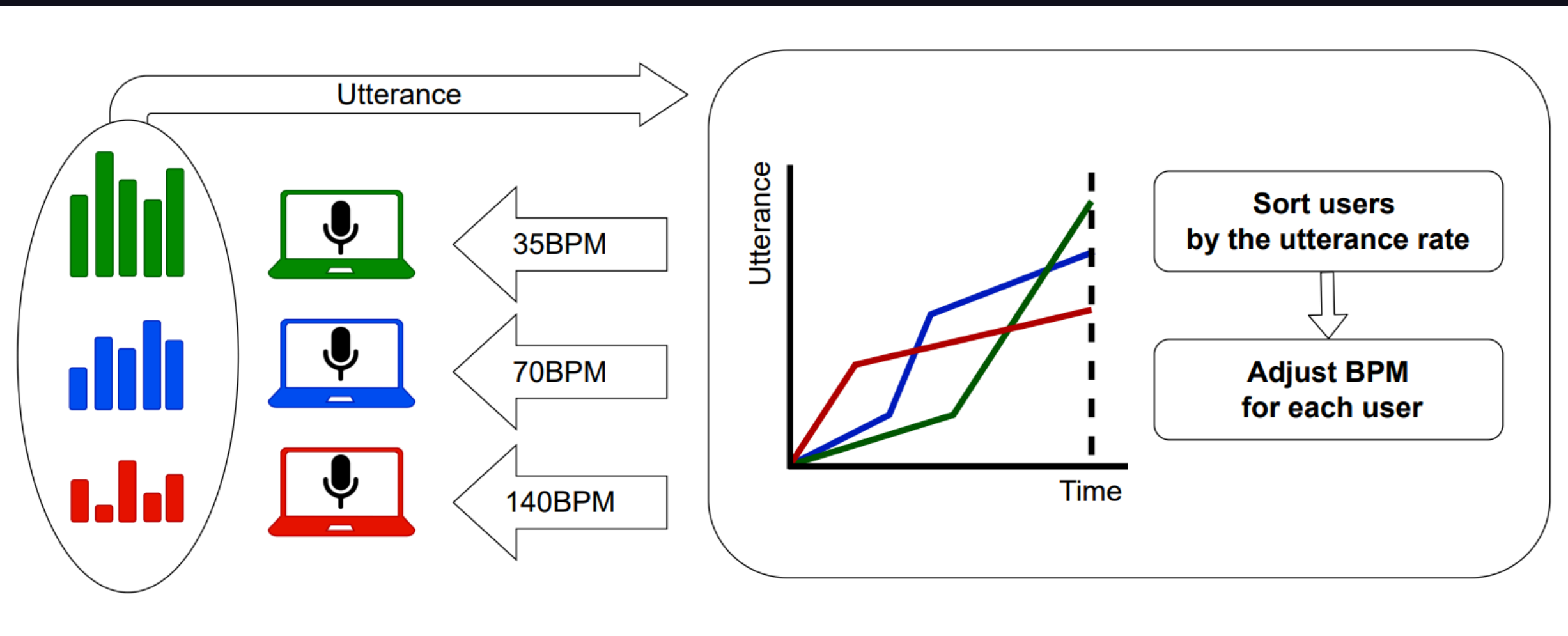
Actual class	Predicted class		
	High	Middle	Low
High	99%	0%	1%
Middle	1%	99%	0%
Low	1%	0%	99%

CNNベース深層学習
(ours, F1: 0.92)



- Existing approach: Facial landmarks + Random Forests
- Proposed approach : Convolutional Neural Network
- Result : F1-score increased from 0.46 to 0.92

Discussion Jockey (DJ)



We developed an online meeting system where each participant's PC BGM dynamically changes based on the amount of speak.

Experiments have demonstrated that our system can control the amount of speak implicitly.

Summary

- This talk presented some psyber systems that recognize human internal states and provide optimized actuations towards making people smarter ([Wordometer](#), [HyperMind](#), [CoALA](#), [EnGauge](#), etc.).
- Cognitive augmentation is still challenging field, by considering user-dependencies, hardware limitations, and data-labeling issues.

Contributions to AI-Physical Systems

- Investigating cognitive/affective layer (workload, engagement, etc.)
- LLM and AI-agent modeling

AI for *Psy*ber-Physical Systems

Prof. Dr. Shoya Ishimaru (Osaka Metropolitan University)
ASPIRE "AI-Physical Systems" Kick-Off Meeting, 04.03.2026



Contact: ishimaru@omu.ac.jp
<https://shoya.io>, @shoya140