# Earinter: A Closed-Loop System for Eating Pace Regulation with Just-in-Time Intervention Using Commodity Earbuds

JUN FANG[†], Tsinghua University, China
KA I CHAN[†], University of Michigan, United States
XIYUXING ZHANG[†], Tsinghua University, China
YUNTAO WANG[*], Tsinghua University, China
MINGZE GAO, Tsinghua University, China
LEYI PENG, Yale University, United States
JIAJIN LI[‡], Tsinghua University, China
ZIHANG ZHAN, Tsinghua University, China
ZHIXIN ZHAO, Tsinghua University, China
YUANCHUN SHI, Tsinghua University, China

Rapid eating is common yet difficult to regulate in situ, partly because people seldom notice pace changes and sustained self-monitoring is effortful. We present Earinter, a commodity-earbud-based closed-loop system that integrates in-the-wild sensing, real-time reasoning, and theory-grounded just-in-time (JIT) intervention to regulate eating pace during daily meals. Earinter repurposes the earbud's bone-conduction voice sensor to capture chewing-related vibrations and estimate eating pace as chews per swallow (CPS) for on-device inference. With data collected equally across sessions in quiet and noisy environments, Earinter achieves reliable chewing detection (F1 = 0.97) and accurate eating pace estimation (MAE: 0.18 ± 0.13 chews/min, 3.65 ± 3.86 chews/swallow), enabling robust tracking for closed-loop use. Guided by Dual Systems Theory and refined through two Wizard-of-Oz pilots, Earinter adopts a user-friendly design for JIT intervention content and delivery policy in daily meals. In a 13-day within-subject field study (N=14), the closed-loop system significantly increased CPS and slowed meal-level eating pace, with statistical signs of carryover on retention-probe days and generally favorable comfort and usability. These effects should be interpreted as changes in eating-pace regulation rather than evidence of weight loss. Our findings highlight how single-modality commodity earables can support practical, theory-driven closed-loop JIT interventions for regulating eating pace in the wild.

CCS Concepts: • **Human-centered computing** → **Ubiquitous and mobile computing design and evaluation methods**; • **Applied computing** → **Health informatics**.

[†]Co-author equal contribution.
[*]Corresponding author.
[‡]Visiting student at the time of this work.

Authors' Contact Information: Jun Fang, Tsinghua University, Beijing, Beijing, China, fangy23@mails.tsinghua.edu.cn; Ka I Chan, University of Michigan, Ann Arbor, Michigan, United States, chankai@umich.edu; Xiyuxing Zhang, Tsinghua University, Beijing, Beijing, China, zxyx22@mails.tsinghua.edu.cn; Yuntao Wang (corresponding author), Tsinghua University, Beijing, Beijing, China, Qinghai University, Xining, Qinghai, China, yuntaowang@tsinghua.edu.cn; Mingze Gao, Tsinghua University, Beijing, Beijing, China, gaomingze2022@gmail.com; Leyi Peng, Yale University, New Haven, Connecticut, United States, leyi.peng@yale.edu; Jiajin Li, Tsinghua University, Beijing, Beijing, China, lijiajin0516@gmail.com; Zihang Zhan, Tsinghua University, Beijing, Beijing, China, zhanzh22@mails.tsinghua.edu.cn; Zhixin Zhao, Tsinghua University, Beijing, Beijing, China, zhaozhix22@mails.tsinghua.edu.cn; Yuanchun Shi, Tsinghua University, Beijing, Beijing, China, Qinghai University, Xining, Qinghai, China, shiyc@tsinghua.edu.cn.



This work is licensed under a Creative Commons Attribution 4.0 International License.
© 2026 Copyright held by the owner/author(s).
ACM 2474-9567/2026/9-ART98
https://doi.org/10.1145/3831629

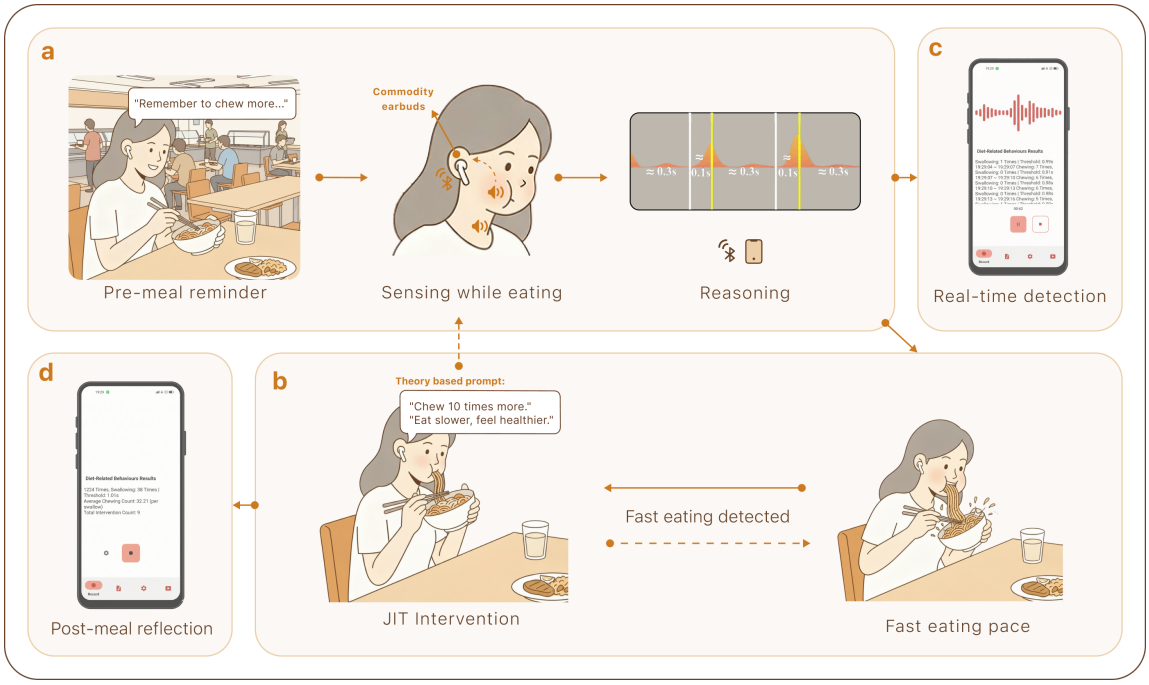



Fig. 1. Earinter is an earbud-based closed-loop system for monitoring and regulating eating pace in daily meals. **(a)** It has a pre-meal reminder, senses ingestive signals while eating in the wild, and performs real-time reasoning to estimate pace. **(b)** When a rapid eating pace is detected, it delivers theory-based audio prompts, subject to a cooldown interval. **(c)** It shows real-time ingestive events and pace estimates. **(d)** After the meal, it summarizes pace-related statistics to support reflection.

Additional Key Words and Phrases: Closed-loop system, Eating pace, Earbuds, Wizard-of-Oz, Intervention design, Just-in-time, Behavior change.

**ACM Reference Format:**
Jun Fang, Ka I Chan, Xiyuxing Zhang, Yuntao Wang, Mingze Gao, Leyi Peng, Jiajin Li, Zihang Zhan, Zhixin Zhao, and Yuanchun Shi. 2026. Earinter: A Closed-Loop System for Eating Pace Regulation with Just-in-Time Intervention Using Commodity Earbuds. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 10, 3, Article 98 (September 2026), 36 pages. https://doi.org/10.1145/3831629

## 1 INTRODUCTION

Eating is a foundational daily activity with downstream consequences for health, quality of life, and well-being [39, 67]. While prior work and public attention largely emphasize what people eat (e.g., diet quality [97, 98] and nutritional balance [4, 9, 24]), how people eat during a meal is equally consequential. We define **eating pace** as the temporal pattern of within-meal food intake, capturing how quickly and steadily a person progresses through eating. However, it is often overlooked in everyday life [101].

In everyday meals, regulating eating pace matters because persistent fast eating has been associated with delayed satiety and higher risks of overeating and obesity [33, 74, 113], increased gastrointestinal burden due to insufficient chewing [57, 64], and reduced awareness of the eating experience [84]. Yet it is difficult to regulate in situ. People often fail to notice when they speed up [66], cannot sustain attention across the meal [63], and face

strong habitual stability reinforced by daily and social contexts [19, 21, 30, 49]. As a result, eating pace becomes a largely automatic behavior with little immediate feedback, allowing fast-eating patterns to persist unchecked.

Recent advances in ubiquitous and wearable sensing have made it increasingly feasible to detect ingestive behaviors such as chewing and swallowing in the wild and even run online inference during daily meals [17, 42, 61, 91, 108, 110]. In practice, however, most existing work still has limited translation of these behavioral signals into actionable, acceptable, and timely support [25]. When interventions are explored, they are often either after the fact (e.g., post-meal visual summaries for reflection) [24, 36, 73] or workflow-bound in-meal cues (e.g., video playback changes [20], replayed eating sounds [21], vibrotactile prompts [34, 48]). A key reason is that eating pace is dynamic, context-dependent, and habit-driven – it fluctuates within meals and varies across people, foods, and social settings – making post-meal feedback too late and fixed rules brittle. What is still missing is an iterative closed-loop mechanism that continuously senses and infers momentary eating states and maps them to intervention timing and dose (when to intervene, how often, and for how long), enabling support that remains effective and acceptable in everyday meals.

To address this gap, we present Earinter, a commodity-earbud-based closed-loop system for eating-pace regulation during daily meals. We treat sensing and intervention as components of a single closed loop rather than as separate stages. We choose earables not only for sensing, but also because they co-locate sensing and private, eyes-free feedback on an already-adopted form factor, reducing setup friction and limiting competition with visual attention during meals. Earinter repurposes the bone-conduction voice-pickup sensor already embedded in commodity earbuds, a design pattern originally intended for call enhancement and now common in mainstream devices. This practical choice can capture chewing-related vibrations and infer chews per swallow (CPS) as an actionable indicator of oral processing intensity [31, 88]. Across data collected in both quiet and noisy settings, Earinter achieved reliable chewing detection (F1 = 0.97) and low error in pace-relevant estimation (MAE: 0.18 ± 0.13 chews/min; 3.65 ± 3.86 chews/swallow). Then Earinter uses these real-time estimates to trigger just-in-time support. To make support practical in everyday eating, we further design a theory-grounded intervention mechanism guided by Dual Systems Theory and refine content and delivery policy through two Wizard-of-Oz pilots. We then evaluated the full system in a 13-day within-subject field study (N=14), where Earinter significantly increased CPS and reduced food-consumption speed, with statistical signs of carryover on retention-probe days and overall comfort and usability. Together, these results show the promise of commodity earables as a practical platform for theory-driven, closed-loop JIT intervention for eating-pace regulation in the wild. In this paper, we make the following contributions:

- We present Earinter, a commodity-earbud-based closed-loop system that regulates eating pace in daily meals by continuously sensing and reasoning about ongoing ingestion, delivering JIT auditory interventions in situation and a post-meal summary report.
- We develop a theory-grounded, psychologically informed JIT intervention mechanism. A Wizard-of-Oz pilot study refines the final prompt set and calibrates delivery parameters, including frequency and sequencing, to preserve comfort while maintaining perceived usefulness.
- We evaluate Earinter in a 13-day within-subject field study with baseline, control, and experimental phases, showing significant improvements in eating pace and statistically early signs of carryover with favorable comfort and usability.

## 2 BACKGROUND AND RELATED WORK

### 2.1 Eating Pace as a Dynamic In-Meal Behavior

Eating pace is a continuous, within-meal behavior that reflects the moment-to-moment rhythm of ingestion. It is closely related to mindful eating in that both emphasize awareness of the eating process and attention to bodily cues [94]. Persistent fast eating pace has been linked to a range of adverse outcomes, including greater risks of

overeating and obesity [33, 75, 76], higher cardiometabolic burden [81, 107], and even acute safety hazards [51]. Mechanistically, fast eating pace often entails insufficient mastication and hurried swallowing, which can increase gastrointestinal load and reduce the effectiveness of digestive processing [57]. Meanwhile, chewing-related physiological feedback (e.g., appetite and glucose regulation) may lag behind rapid intake [83, 113], making satiety signals arrive too late to shape ongoing behavior and thereby encouraging excessive energy intake [3]. In everyday meals, eating pace is shaped by time pressure, situational norms, and social contexts, and fast eating is often acceptable [19, 49]. Eating is also often treated as a secondary companion activity during leisure-time, and people frequently engage in screen media or other distractions during meals [20]. This makes it difficult to sustain attention to eating pace throughout a meal, especially in the absence of immediate corrective feedback. Moreover, eating pace tends to exhibit habitual stability across contexts, which weakens the effectiveness of intention-based regulation strategies [21, 30].

Building on prior work, we operationalize eating pace through oral processing intensity, quantified as chews per swallow (CPS). CPS is defined as the number of chewing cycles between two consecutive swallowing events [31, 88]. As an intuitive behavioral indicator, CPS is more actionable than other chewing-related metrics, which are regulated by the rhythm generator in the central nervous system and therefore tend to vary only modestly in response to external influences [5, 62, 70]. In contrast, CPS can be adjusted through simple, in-the-moment strategies such as extending mastication before swallowing. Recent work increasingly adopts CPS to characterize eating pace and examine its links to health outcomes [12, 31, 113]. In this paper, CPS serves as the primary continuous indicator for monitoring and regulating eating pace during daily meals.

## 2.2 JIT Intervention Design for Ingestion Behavior Change

Just-in-time (JIT) interventions refer to methods that deliver support when an individual is most likely to benefit from and accept it [72]. Such moments are typically identified through real-time assessments of the user's behavior, physiology, or context [22, 72, 89]. Beyond timely delivery, effective JIT design also requires a principled account of why the target behavior occurs and is hard to change. Accordingly, prior researches in behavior change emphasize grounding JIT intervention in psychological theories to improve acceptability and increase the likelihood of sustained habit change [68, 100]. In digital well-being, Dual Systems Theory [41] frames problematic smartphone use as a conflict between habitual and deliberative control processes [79]. It motivates interventions that adapt to in-the-moment triggers such as fatigue, emotions, and context [103]. In medication adherence, Self-efficacy Theory [6] motivates feedback that reinforces success and supports recovery after lapses [54].

In eating-related interventions, recent systems have increasingly incorporated theories to strengthen long-term change in dietary behaviors. Self-efficacy Theory has motivated designs that provide positive reinforcement, small achievable actions, or peer support to encourage healthier food choice routines [46, 80]. Goal-oriented theories further strengthen goal-directed behavior in complementary ways – Goal-setting Theory operationalizes distal goals into proximal subgoals with progress feedback [78]; Control Theory frames self-regulation as discrepancy monitoring and closed-loop correction [35]; and Goal-framing Theory increases goal salience by shaping how rationales and benefits are communicated [1, 23]. However, these theory-driven work focuses on behavior change outside the eating episode, and theory is often used for after-the-fact interpretation rather than for in-the-moment intervention design [25]. When considering JIT interventions during eating episodes, existing approaches often rely on interactive artifacts to deliver timely cues, such as game-based installations [29, 40], interactive robots [85, 104], or transformable tableware [18, 45]. Other work probes behavioral effects by manipulating elements within a pre-defined routine, such as adjusting video playback speed [20] or replaying eating-related sounds [21].

However, to our knowledge, no prior system has grounded both in-meal intervention content and explicitly designed delivery policies (e.g., when to intervene, how often, and for how long) in a shared theory-driven rationale for habit-like behavior that fluctuates dynamically within meals [20, 21, 24, 25, 34, 36, 48, 73]. Motivated

by this gap, Earinter uses theory not only to shape what support is delivered, but also to inform when support is triggered and how intervention dose is regulated during the meal. We use theory to characterize the habitual behavior, guide the design of prompt content, and inform an adaptation logic that responds to real-time eating pace. This framing lays the foundation for more effective and personalized closed-loop support for eating-pace regulation.

### 2.3 Closed-loop System

The closed-loop system is inspired by the Personal Informatics Framework [56], mapping collection to sensing, integration to reasoning, and the reflection-action sequence to intervention [25]. It's an iterative cycle that uses continuous sensing and reasoning to deliver context-relevant support and update subsequent actions based on users' moment-to-moment responses. This paradigm has increasingly been adopted in ingestion health. However, systems for ingestion still face persistent challenges in achieving end-to-end closed-loop support. First, intervention outputs are often limited to lightweight cue modalities, including audio [21] and visual cues [48, 73, 101], whatever sensing pathways and modalities are. Consequently, sensing and intervention remain weakly integrated, with limited coupling between sensing outputs and feedback pathways. Second, many pipelines depend on multiple devices to realize the full loop, which increases setup complexity and users' cognitive burden. Third, systems often prioritize recognition accuracy while treating intervention as a downstream components rather than a theory-informed design problem. As a result, adaptation logic for when to intervene, how frequently, and how to respond to short-term behavioral changes remains under-specified.

For eating pace in particular, much prior work emphasizes sensing and reasoning, shifting from self-report to wearable based monitoring [37]. When interventions are included, the lack of theory grounding often leads to limited design choices. Some systems mainly support post-meal reflection with summaries [24, 90], while others use simple but information-sparse in-meal cues such as vibration [34, 48]. More expressive in-meal interfaces, including real-time visualizations [15, 47, 73] or transformative utensils [18, 112], can increase information load and potentially compete with food awareness. Collectively, these patterns motivate closed-loop eating-pace systems that tightly integrate sensing, reasoning, and theory-informed intervention, while specifying adaptation policies that balance effectiveness with practical acceptability in daily meals.

## 3 SYSTEM OVERVIEW

Before detailing the architecture, we clarify Earinter's position relative to representative strands of prior eating-related systems. Reflection-oriented systems primarily support awareness after the eating episode [24, 36]; fixed in-meal cueing systems deliver prompts during eating but often use predefined delivery or focus on workflow-bound cues [20, 21, 34, 48]; and sensing-centric ingestion systems mainly advance recognition and monitoring, with intervention typically treated as a downstream step [8, 10, 17, 42, 61, 91, 110]. Earinter extends these strands by integrating real-time CPS estimation from commodity-earbud sensing and a theory-grounded JIT intervention mechanism within a closed-loop system for daily meals. We use commodity earables as the system substrate because they can unify practical sensing and private, in-ear intervention within a single practical form factor. As shown in Fig. 2, Earinter is a closed-loop system using commodity earbuds, which comprises three stages: sensing, reasoning, and intervention.

*Sensing.* Earinter repurposes the bone-conduction sensor embedded in a commercial earbud (**Honor Earbuds 3 Pro**[1]) to capture body-conducted chewing and swallowing signals with reduced sensitivity to ambient noise. We describe the sensing setup and signal acquisition in Section 4.1.1.

[1] https://www.honor.com/global/audio/honor-earbuds-3-pro/

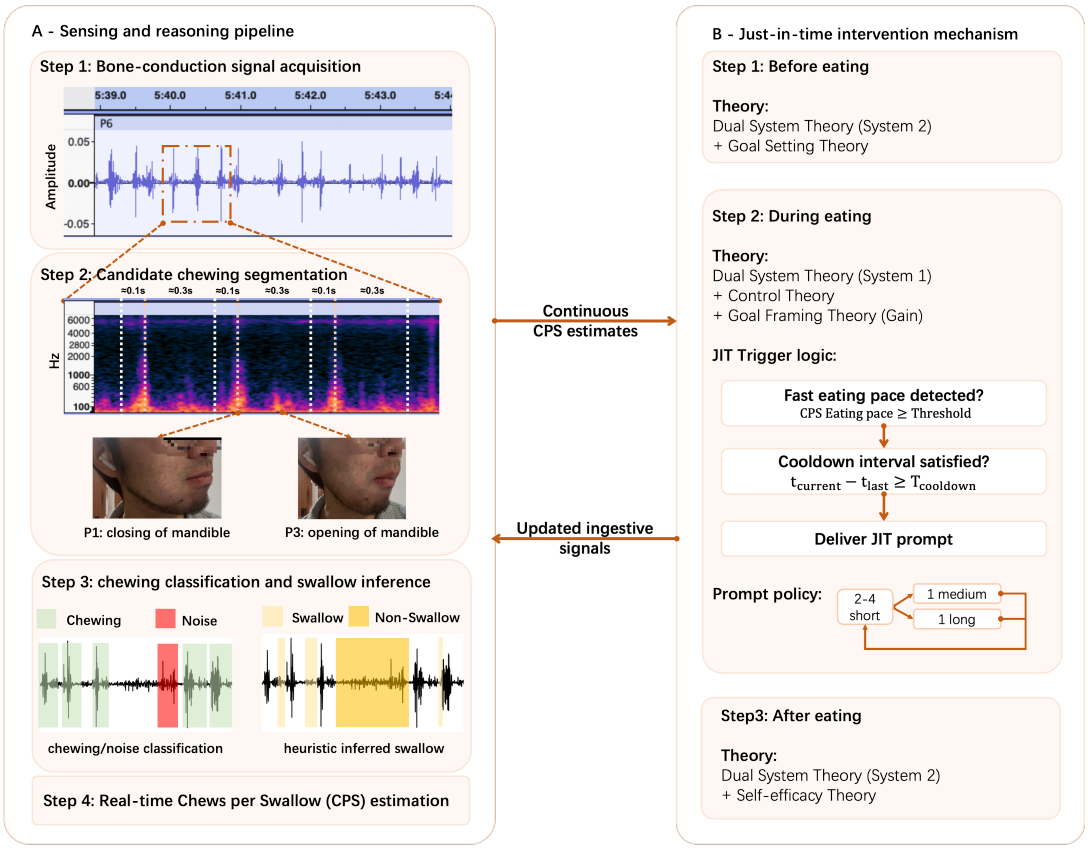



Fig. 2. Closed-loop system overview of Earinter for eating-pace regulation. (a) Sensing and reasoning pipeline. The commodity earbud captures bone-conducted ingestive signals during eating. Earinter first acquires the raw signal, segments candidate chewing events, classifies chewing versus noise, and infers swallow boundaries heuristically, producing real-time chews-per-swallow (CPS) estimates. (b) Just-in-time intervention mechanism. Continuous CPS estimates drive theory-grounded intervention across the eating episode, including pre-meal goal activation, in-meal JIT prompting, and post-meal reflection. During eating, prompts are triggered when fast eating pace is detected and the cooldown interval is satisfied, with diversified prompt lengths used to reduce burden. The resulting behavioral response generates updated ingestive signals, which are continuously re-sensed by Earinter to complete the closed loop.

*Reasoning.* Signals are processed on an Android smartphone (**OPPO Reno5 PRO**[2]) for real-time inference. To validate the sensing pipeline before field deployment, we first describe a separate annotated sensing dataset collected across complementary quiet and noisy contexts in Section 4.1.2, followed by the detection model in Section 4.2 and model validation in Section 4.3.

*Intervention.* When inferred CPS indicates fast eating, Earinter delivers JIT audio prompts through the earbud to encourage a slower pace. The intervention framework is grounded in Dual System Theory (Section 5.1), with prompt content informed by relevant physiological theories (Section 5.2), and the delivery policy including logic and frequency refined for practical use in daily meals (Section 5.3).

[2]https://www.oppo.com/en/smartphones/series-reno/reno5-pro-5g/

## 4 REAL-TIME EATING PACE SENSING STRATEGY ON EARABLES

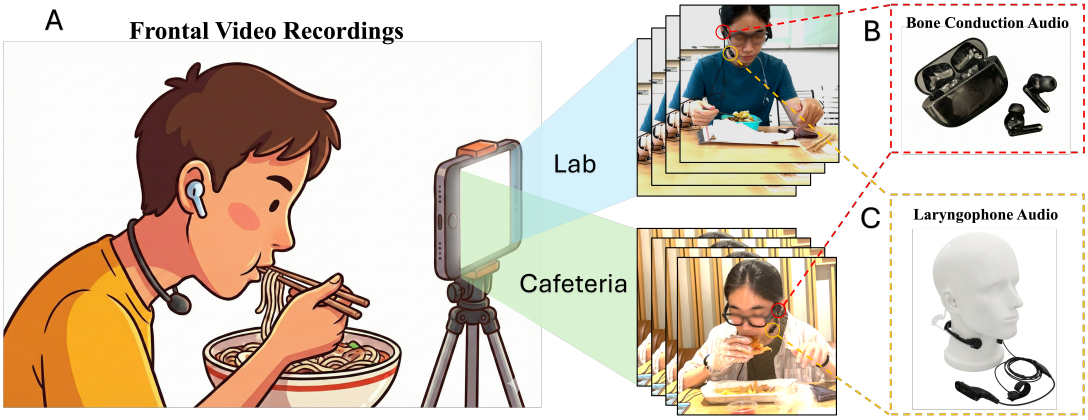



Fig. 3. Sensing-data collection setup and cross-context recording protocol used to validate the ingestion sensing pipeline. Eating behaviors were recorded across two environments (lab and cafeteria) using: **(A)** an eye-level iPhone for video ground truth, **(B)** integrated bone-conduction sensors in Honor Earbuds 3 Pro for capturing ear-canal vibrations, and **(C)** a laryngophone serving as a high-fidelity reference for precise chewing and swallowing annotation.

### 4.1 Data Collection

*4.1.1 Hardware Setup.* We designed a data collection platform that measures eating pace, which we define as the number of chewing cycles between two consecutive swallows. **Honor Earbuds 3 Pro** leverages integrated bone-conduction (BC) sensors to capture chewing-induced vibrations within the ear canal. The devices were provided through a collaborative research agreement with Honor Device Co., Ltd. To establish a reliable reference for identifying chewing and swallowing events, a **DZMIHKS6550 throat microphone** [3] was concurrently employed alongside an **iPhone 12 Pro** [4] serving as a rear-facing camera. Audio was sampled at 16 kHz with 16-bit resolution, while video was recorded at 1080p with a frame rate of 30 fps. Synchronizing these multi-modal streams enabled precise temporal calibration to ensure the accurate determination of ingestion events.

*4.1.2 Data Collection Protocol.* To validate the sensing pipeline before the main field deployment, we collected a separate annotated sensing dataset across two complementary meal-recording settings. Food choice was unrestricted in both settings, so the main contrast was the recording environment. The quiet-environment sessions took place in a lab conference room to approximate quiet individual eating scenarios, whereas the noisy cafeteria sessions introduced realistic variation in ambient sound and surroundings while keeping the sensing configuration comparable across settings.

This study was approved by our Institutional Review Board. Participants were recruited via word-of-mouth from our campus. Informed written consent was obtained from the human subjects participating in the study before any study procedures began. We recruited 18 participants (9 females, 9 males) aged 20 to 29 years ($M$ = 23.83, $SD$ = 2.66). All participants were healthy adults with no history of gastrointestinal or metabolic diseases. No participants reported any pain or discomfort while chewing or swallowing.

[3] https://www.amazon.com/Earpiece-Acoustic-Compatible-MMotorola-XiRP8268/dp/B016BI0CD4/
[4] https://www.amazon.com/Apple-iPhone-128GB-Pacific-Blue/dp/B08PMYLKVF?th=1

As shown in Fig. 3, each participant completed sessions in two settings: a quiet lab conference room and a noisy campus cafeteria chosen by the participant. To ensure ecological validity, sessions were conducted during regular meal times, and participants were free to select their food. Before each session, researchers assisted participants with device wearing. Participants wore earbuds in both ears to capture bone-conduction audio, while a throat microphone was worn on the neck to provide reference data. The eye-level iPhone 12 Pro recorded the meal discreetly to avoid interfering with natural behavior. To synchronize data streams, participants double-tapped the earbuds on the table at the start of the session. Participants were encouraged to eat at their natural pace and avoid any intentional control over their chewing or swallowing to ensure the data reflected their genuine habits.

The dataset comprises 10.07 total hours of data, including 5.65 hours from quiet and 4.42 hours from noisy settings. Bone conduction audio was manually annotated using Praat [11] and ELAN [65] as annotation tools. Through the annotation process with video and throat microphone audio as reference, we identified 24,521 and 21,061 chewing events, alongside 1,114 and 1,036 swallows for the sessions in quiet and noisy environments, respectively. Detailed statistical results are shown in Table 4.

To characterize food variation in this dataset, we further conducted a coarse meal-level texture annotation. Because meals often contained multiple components, categories were not mutually exclusive: all 36 meals included soft components, 23 included crunchy components, and 18 were annotated as clearly mixed-texture meals. Examples ranged from softer noodle-based dishes to crunchier meals such as fried chicken and leafy salads. These annotations indicate that the dataset covered heterogeneous everyday meals rather than a narrowly controlled single-food benchmark, better reflecting the practical meal contexts targeted by Earinter for daily eating-pace regulation.

## 4.2 Eating Pace Sensing Pipeline

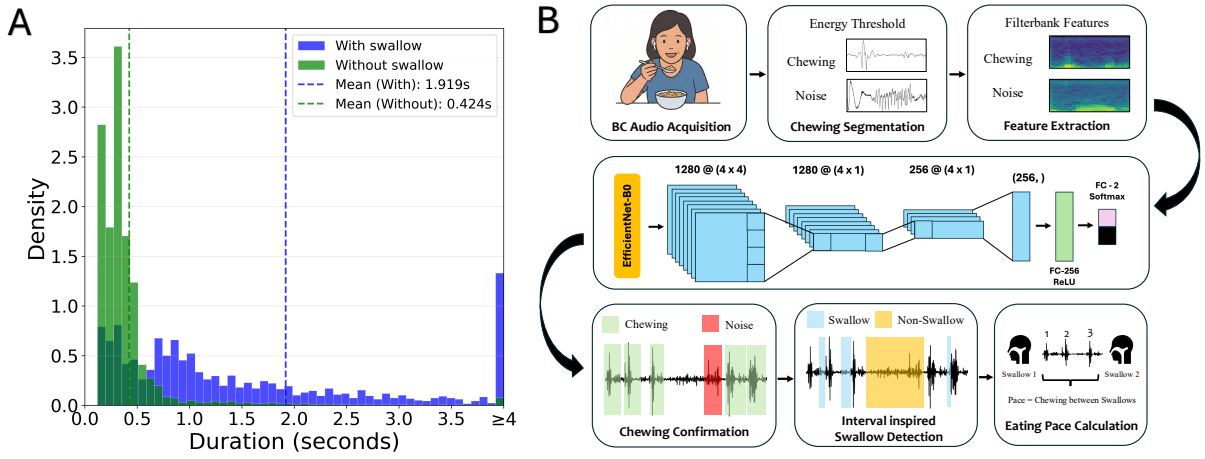



Fig. 4. Swallow detection foundation and the eating pace sensing pipeline. **(A)** Probability density distribution of durations between consecutive chewing cycles. The distinct temporal separation between intervals with swallows ($\mu = 1.919s$) and without swallows ($\mu = 0.424s$) validates the interval-inspired hypothesis for swallow detection. **(B)** The proposed sensing pipeline, which integrates candidate chewing segmentation, deep learning-based chewing confirmation, interval-inspired swallow detection to estimate the eating pace.

Fig. 4 depicts the sensing pipeline for eating pace monitoring on earables. We detail each stage below.

*4.2.1 Audio Pre-processing and Candidate Chewing Segmentation.* Chewing is modeled as a four-stage mandibular motion cycle: food compression during mandible closing (P1), a brief pause (P2), mandible opening (P3), and another pause before the next cycle (P4) [109]. P2 and P4 barely generate acoustic energy, and therefore they are not used for chewing segmentation. As shown in Fig. 2, P1 represents the high-energy phase of closing the lower chin to crush food compared to P3. We first implement a real-time state machine to extract candidates by partitioning audio into 50 ms frames and applying a 40 dB threshold. To align with physiological patterns, valid segments are limited to a duration between 100 ms and 400 ms with a maximum silence tolerance of 150 ms. These parameters were empirically determined to ensure that all chewing events within the dataset were captured without omission. This lightweight screening optimizes system efficiency by filtering non-eating noise and reducing the computational load for downstream deep-learning classification.

*4.2.2 Fine-grained Chewing Detection.* To filter out motion artifacts introduced during the initial screening, we implement a binary classification model. The dataset is partitioned into 14 participants (77.8%) for training, 2 for validation (11.1%), and 2 for testing (11.1%), which facilitates cross-user evaluation. Each candidate segment was normalized to a range between $[-1, 1]$ and padded to a uniform duration of 400 ms.

Inspired by the EarVAS-Net framework [111], the windowed audio is transformed into log-scaled Mel filter bank features of size 38 x 128 using a 25 ms Hanning window with a 10 ms stride. We adapt the EarVAS audio sub-module for single-channel processing by utilizing EfficientNet-B0 as the feature extraction backbone. This architecture incorporates Mean Pooling along the frequency axis, a $1 \times 1$ convolutional layer, and subsequent Mean Pooling along the time axis. This streamlined configuration maintains the robust feature extraction of EfficientNet-B0 while optimizing computational efficiency for real-time binary classification. To address the inherent class imbalance between genuine chewing events and non-eating artifacts, we adopted Focal Loss [58] as the loss function. The model was trained on an NVIDIA GeForce RTX 4090 GPU using a batch size of 128 and a learning rate of $1e^{-5}$.

*4.2.3 Heuristic Estimation of Eating Pace.* Eating pace was defined as the number of chewing events between two consecutive swallows. However, the commercial earbuds that we used often fail to capture swallowing events due to their subtle acoustic signatures. To estimate eating pace, we propose a heuristic algorithm based on the inter-chew interval. In the chewing cycle, P4 corresponds to the pause after mandibular opening and before the next chewing cycle. When swallowing occurs, this pause is often prolonged, creating a longer interval between consecutive chewing events. Specifically, our statistical analysis indicates that swallowing events are primarily localized within the prolonged intervals between chewing events, as shown in Fig. 4A. This observation aligns with the natural eating experience. Based on the observations, once the chewing detection algorithm locates chewing events, the presence of swallowing within the remaining temporal gaps is determined by the following heuristic algorithm:

$$isSwallowing = \begin{cases} yes, & \text{if } d > 0.8 \text{ s} \\ yes, & \text{if } d > 1.5 \times \bar{d}_{\text{chew}} \\ no, & \text{otherwise} \end{cases} \tag{1}$$

where $d$ represents the duration of the current interval between consecutive chewing events, while $\bar{d}_{\text{chew}}$ denotes the average duration of all chewing intervals recorded since the most recent swallow. The absolute threshold (0.8 s) captures prolonged intervals likely containing a swallow while limiting false positives. To account for individual- and food-dependent chewing tempo, we also apply an empirically chosen adaptive criterion ($1.5 \times \bar{d}_{\text{chew}}$), enabling swallow inference and CPS estimation without explicit swallow acoustics.

### 4.3 Evaluation and On-device Deployment

*4.3.1 Evaluation of Chewing Detection.* Since the system performs binary classification on extracted candidate segments, we assess the performance of the chewing detection module using four standard metrics. On the test dataset, the model achieves an Accuracy of 96.86%, Precision of 99.02%, Recall of 95.05% and an F1-score of 0.97. These results demonstrate the high reliability of the pipeline in distinguishing genuine chewing events from artifacts.

*4.3.2 Evaluation of Eating Pace Estimation.* To validate the temporal precision of the heuristic algorithm, we analyze eating pace through two complementary dimensions. At a macro-level, we measure the mean estimation error of the chewing frequency per minute to characterize the overall velocity of intake. In all experimental environments, the system achieves a mean absolute error (MAE) of $0.18 \pm 0.13$.

At a more granular level, we assess the mean estimation error of the number of chews per swallow, which serves to validate the effectiveness of heuristic segmentation in accurately demarcating discrete masticatory sequences between inferred swallowing events. For this fine-grained metric, the system achieves an MAE of $3.65 \pm 3.86$. In summary, these results demonstrate that the proposed system provides adequate monitoring of both long-term eating trends and detailed masticatory behavior.

*4.3.3 On-device Inference and Performance.* To enable the intervention of the proposed system, we deployed the sensing framework on an **Oppo Reno5 Pro**. We developed a dedicated Android application to facilitate real-time interaction and data visualization for the user. To ensure efficient execution on the mobile devices, the deep learning classification model was converted to the ONNX format. The entire pipeline processes incoming acoustic data from the earable device using a 3-second sliding window.

We measured power consumption of our system using Android batterystats during one hour of screen-off background execution. Power consumption is approximately 1.84% battery capacity per hour on Oppo Reno5 Pro with a 4350 mAh battery. According to the data collected, eating-episode duration ranged from 5.25 to 29.02 minutes. This suggests that a typical meal-length would consume approximately 0.16%-0.89% of the phone battery. As typical smartphone applications like music players consumes about 2% per hour [38], our algorithms bring limited battery burden for daily meal-time deployment.

To guage system performance, we measure CPU load and latency by running a specific code segment 1000 times. The system processes a 3-second audio stream in just $57.64 \pm 24.89$ ms, demonstrating its high efficiency for real-time eating pace estimation and just-in-time intervention.

## 5 PILOT STUDY FOR INTERVENTION DESIGN

Given that Earinter senses ingestive signals using commodity earbuds, we deliver intervention through auditory prompts to support modality-consistent closed-loop system. This design locates sensing and intervention on the same wearable channel, which reduces setup complexity and cognitive burden in everyday use. Auditory prompts are also hands-free and eyes-free, which better aligns with the attentional demands and social nature of daily meals. Moreover, language is the most common persuasion mechanism in behavior change methods [82, 92]. Concise reminders grounded in facts or goals can effectively support the in-the-moment self-regulation [43, 71]. Accordingly, we implement an earbud-based JIT audio intervention grounded in Dual System Theory. Building on this theory-driven pipeline, we further organize in-meal prompts using additional psychological theories, and conduct two Wizard-of-Oz (WoZ) pilot studies to refine (PS1) prompt content and (PS2) prompting frequency for practical, in-the-wild use.

### 5.1 Dual System Theory for Intervention Design

As described in Section 2.2, drawing on prior work on intervening in habitual behaviors [20, 103], we ground our intervention design in Dual System Theory [41]. Dual System Theory describes the structure of cognitive processing: fast, automatic/implicit processes (System 1) versus slow, controlled/deliberative processes (System 2). We conceptualize rapid eating pace as a largely automatic routine driven by System 1, whereas regulating pace requires re-emphasizing System 2 processes, such as attention re-allocation [20] and self-control [103].

Guided by Dual System Theory, we structured prompt content across the full eating episode to support both goal activation (System 2) and in-the-moment disruption of habitual rapid eating (System 1).

- **Before eating (System 2).** We generated prompts informed by Goal Setting Theory [60] to make the target behavior explicit and memorable. These prompts emphasize specifying actionable targets to initiate deliberate self-regulation, aligning with System 2 processing.
- **During eating (System 1).** We deliver brief prompts as intuitive cues that shape automatic pacing with minimal cognitive demand. We treat these prompts as System 1-oriented micro-nudges. They rely on simple, immediate signals rather than analytical reasoning or rule-based feedback, and can be acted upon without sustained deliberation.
- **After eating (System 2).** Drawing on Self-efficacy Theory [6], we provide a brief summary of pace-related statistics and supportive reflection cues to strengthen confidence in behavior change. This component supports sustaining behavior change beyond immediate prompts and thus constitutes part of System 2.

After defining this pipeline, we further drew on additional psychological theories summarized in prior behavior change syntheses [52, 100] to design candidate in-meal prompt families. These families vary in informational specificity, affective tone, and theoretical intent. We also incorporated a commonly cited chewing heuristic (approximately 25 chews per mouthful) as an intuitive reference in prompt wording [2, 94]. The full set of theory-grounded intervention content is summarized in Table 1.

### 5.2 PS1: Intervention Content

Table 1. Overview of the theory-grounded prompting pipeline and candidate in-meal prompt families *before* the WoZ refinement. We organize intervention content across three stages: pre-meal and post-meal draw on System 2 processes with additional psychological theories (Goal Setting Theory and Self-efficacy Theory), while in-meal prompts include System 1-compatible micro-nudges, non-verbal cues, and theory-informed messages (Control Theory, Goal Framing Theory).

| Timestamp | Theory Based | Example Content Design |
|---|---|---|
| Before Eating | Dual System Theory (System 2) + Goal Setting Theory [60] | *"Remember to chew more. Aim for at least 25 chews before swallowing."* |
| During Eating | Dual System Theory (System 1) | *"Slow down." "More chewing." "A bit fast."* |
| | None | Neutral chime cue (*"ding"*) |
| | None | Failure jingle cue |
| | Control Theory [16] | *"You are xxx chews away from the recommended 25 chews."* |
| | Goal Framing Theory (Gain) [59, 99] | *"Eating more slowly and chewing more can improve digestion and increase satiety, helping you reach your weight-loss goal sooner."* |
| | Goal Framing Theory (Loss) [99] | *"Eating too fast and swallowing with insufficient chewing may cause digestive discomfort and, contribute to weight gain over time."* |
| | Mixed | A chime is played upon each rapid-pace detection, and a spoken reminder is delivered after repeated detections |
| After Eating | Dual System Theory (System 2) + Self-efficacy Theory [6] | Summary of pace-related statistics |

We conducted a Wizard-of-Oz (WoZ) pilot study (N=6) across diverse daily meal contexts: two solo meals, two dyadic meals, and two family co-present meals. Participants also differed in auditory sensitivity. This allows us to assess perceived usefulness, comfort, and preference across prompt families under individual differences, and to refine wording and content selection prior to deploying the automated intervention.

At the beginning of each session, we explained the study purpose that participants would receive different audio prompts generated from the system during eating [44]. They are intended to encourage slower eating pace and more chewing. Participants were instructed to eat naturally as usual. During the meal, an experimenter manually triggered prompts in a randomized order, covering the candidate in-meal prompt families shown in Table 1. Prompts were delivered through earbuds, with an interval of approximately 60 seconds between consecutive prompts. After all prompts were presented, we sent participants the corresponding text versions and conducted a semi-structured interview. We asked participants to (1) identify prompts they felt were effective versus ineffective, (2) rank most and least preferred prompts, (3) comment on the mixed design, and (4) reflect on perceived annoyance.

All sessions were video-recorded, and interviews were audio-recorded and transcribed. The research team conducted a thematic analysis by coding participants' comments on effectiveness, acceptability, and preferences across prompt families. Codes were iteratively refined through team discussions until consensus was reached, and transcripts were revisited to ensure consistent and accurate coding.

Overall, participants preferred prompts that were concise, easy to interpret, and low-effort to act on during eating. They noted that longer, information-dense messages could become annoying if delivered frequently (S5[5], "Words all time takes a lot of effort, and interrupt the entertainment"). So they suggested short prompts for frequent delivery while reserving longer messages for occasional use. Participants also emphasized the importance of content variety to reduce habituation over time. Neutral chime cues alone were sometimes perceived as ambiguous without upfront explanation (S3, "initially thought it was a phone notification"), while failure jingles were viewed as uncomfortable. Moreover, the mixed design was considered acceptable if "an explanation and demonstration of the cue meaning were provided beforehand" (S6).

Based on these findings, we prioritized three in-meal prompt families for subsequent deployment: (i) System 1-oriented micro-nudges, (ii) control-theory-based progress feedback, and (iii) gain-framed rationales. To support practical deployment and reduce repetitiveness, we used GPT-4o to expand each selected family into a prompt library with varied phrasings and multiple lengths, enabling randomized selection at runtime. Generation prompt for GPT-4o and final audio content can be found in Appendix C.

### 5.3 PS2: Intervention Frequency

Well-crafted prompts can lose effectiveness and become irritating if delivered too often, as shown in PS1. We conducted a second WoZ pilot (N=10) to determine a comfortable maximum prompting frequency, and develop a prompt logic that balances intervention efficacy with user comfort in daily use.

Based on the three categories selected above, we designed three types of prompt lengths: short prompts (approximately 1-2 seconds), medium prompts (2-3 seconds), and relatively long prompts (7 seconds). Then we manipulated the minimum interval between consecutive prompts at three levels: 10 seconds, 20 seconds, and 30 seconds. These conditions were chosen to span a range from very high frequency (roughly one prompt per bite for a fast eater around 10s [34]) to moderate frequency. Each participant experienced all three interval conditions in a random order. In each condition, five prompts were delivered at the specified interval. To ensure consistency, we fixed the sequence of prompt lengths in each condition to two short prompts, one medium-length prompt, then two short prompts again. This consistent pattern exposed participants to a mix of brief and slightly longer messages at each frequency. Participants were instructed to eat normally and informed that the system would

[5]This is the serial number of the participant in the semi-structured interview for PS1.

give messages to encourage slower eating through the earbuds. A 60-second break was provided between each frequency condition to allow participants to rest and minimize carryover effects.

After each condition, participants rated their annoyance with the prompting frequency on a 5-point Likert scale (1 = not annoyed at all, 5 = extremely annoyed). We calculated the scores for each intervention frequency. The 10-second interval condition was the least favored (mean = 3.0) and only one participant chose it as prefer. The 20-second condition was slightly better (mean = 2.7), and 3 participants preferring it. The 30-second minimum interval was the most comfortable by far. It had the lowest annoyance rating (mean = 2.3) and was preferred by 6 out of 10 participants. Participants' feedback clearly indicated that more frequent prompts led to greater annoyance. Based on these findings, we set 30 seconds as the minimum interval between prompts in our system design, even if the user remains fast eating pace.

We also refined the prompt logic from users' feedback to incorporate the content length recommendations from PS1 with the frequency decision above. In the final intervention mechanism, we employ a hybrid prompting strategy to diversify the content and avoid overwhelming users with repetitive messages. The system randomly plays 2-4 short prompts, followed by either a medium-length prompt or a relatively long prompt, with equal probability. All these prompts are selected randomly from three families identified in PS1 which are expanded via GPT-4o (Appendix C).

## 6 USER STUDY

Building on the end-to-end pipeline described in Sections 4–5, we evaluated the full Earinter closed-loop system in a within-subject field deployment conducted entirely in participants' everyday in-the-wild eating contexts, including real-time eating-pace inference, theory-based JIT audio prompting with a cooldown policy, and post-meal summaries for reflection. We evaluated Earinter against comparison conditions that kept sensing and logging identical but disabled in-meal intervention and post-meal feedback to isolate the contribution of the closed-loop intervention.

### 6.1 Recruitment and Study Design

*6.1.1 Participants.* We recruited an initial pool of participants who self-reported fast eating pace using snowball sampling through an on-campus communication platform [32]. All candidates reported being in good general health, having no special dietary habits or dietary restrictions, and having a full set of natural teeth with no pain or discomfort during chewing. Screening was conducted primarily through self-report and online communication, and we did not knowingly enroll participants who self-reported a diagnosed eating disorder or current treatment for an eating-related clinical condition.

Because self-reports may not accurately reflect eating behavior in everyday settings, we conducted an in-the-wild screening procedure including two meal sessions to confirm eating pace. Specifically, during two free-living meals, we collected ground truth with frontal videos, and assessed eating pace using two commonly adopted criteria: (i) *chews per swallow* $< 20$ [2, 94] and (ii) *food consumed per minute* $> 50$ g/min [50, 102]. Candidates were eligible if they met both criteria in the screening. Based on this screening, we enrolled 14 participants (9 male, 5 female, $M = 23.50$, $SD = 2.68$) in the main study. These participants completed the full study and received \$200 compensation. The \$200 compensation was provided for completing the full field deployment. Participant effort was distributed across Day 0 onboarding, two free-living screening meals, daily indexed meals during the deployment, and the pre- and post-study interview.

*6.1.2 Study Design.* We conducted a 13-day within-subject field study to evaluate the performance of the JIT intervention and the short-term carryover of our closed-loop system in daily eating contexts. The study compared two conditions:

Table 2. Participant demographics and study assignment. Single earbud was worn on the dominant chewing side. All participants met our fast-eater screening criteria (chews per swallow < 20 and food consumed per minute > 50 g/min). The sequence order was counterbalanced across participants.

| Participant id | Gender | Age | Dominant chewing side | Baseline eating pace | | | Condition sequence |
|---|---|---|---|---|---|---|---|
| | | | | chews per swallow | food consumed per min | eating-episode duration | |
| P1 | F | 22 | Left | 11.78 | 66.05 g/min | 10 min 2s | Control → Experiment |
| P2 | M | 20 | Left | 14.23 | 67.55 g/min | 9 min 14s | Control → Experiment |
| P3 | M | 24 | Right | 13.68 | 59.35 g/min | 9 min 3s | Control → Experiment |
| P4 | F | 24 | Right | 16.72 | 53.56 g/min | 8 min 21s | Experiment → Control |
| P5 | M | 23 | Left | 17.77 | 68.29 g/min | 8 min 9s | Control → Experiment |
| P6 | M | 22 | Right | 11.22 | 70.85 g/min | 7 min 3s | Control → Experiment |
| P7 | M | 20 | Right | 15.41 | 114.93 g/min | 5 min 53s | Control → Experiment |
| P8 | F | 22 | Right | 9.04 | 139.18 g/min | 6 min 17s | Control → Experiment |
| P9 | F | 29 | Both | 14.75 | 64.88 g/min | 7 min 30s | Experiment → Control |
| P10 | M | 29 | Both | 15.09 | 79.59 g/min | 8 min 22s | Experiment → Control |
| P11 | F | 23 | Left | 16.56 | 72.53 g/min | 8 min 31s | Experiment → Control |
| P12 | M | 23 | Left | 14.23 | 83.29 g/min | 9 min 8s | Experiment → Control |
| P13 | M | 24 | Left | 14.21 | 78.23 g/min | 8 min 12s | Experiment → Control |
| P14 | M | 24 | Right | 17.26 | 100.60 g/min | 6 min 57s | Experiment → Control |

- **Control condition:** Participants received a pre-meal reminder at the beginning of each eating episode, with no intervention or feedback delivered during or after the meal.
- **Experiment condition:** Based on the closed-loop system of Earinter, participants received the pre-meal reminder and the JIT audio interventions delivered through the earbud during eating. After each meal, participants also received a post-meal summary report as feedback.

The Control condition represents a commonly used approach based on pre-meal self-cueing or external reminders. In contrast, the Experiment condition incorporates our JIT in-meal intervention design, and adds a post-meal summary to complete a per-meal closed loop. This design allows us to evaluate the full effect of Earinter, highlighting the value of a continuous intervention pipeline that intervenes during the meal and reinforces behavior change through post-meal reflection.

Beyond the condition design, we explicitly incorporated retention days at the end of each condition. This allowed us to quantify short-term behavioral persistence after withdrawing all support and reduce potential carryover contamination in our within-subject cross-over design. As shown in Fig. 5a, the overall study followed a three-stage pipeline:

- **Baseline (Day 1):** Participants ate normally with the sensing and reasoning part of Earinter enabled, but received no audio prompts and feedback throughout the eating episode, including before and after eating.
- **Phase 1 (Days 2–7):** Five consecutive days under the assigned condition (Days 2–6), followed by a retention day (Day 7) with no prompts or feedback.
- **Phase 2 (Days 8–13):** Five consecutive days under the other condition (Days 8–12), followed by a retention day (Day 13) with no prompts or feedback.

To mitigate potential order effects, participants were randomly assigned to one of two sequences (*Control* → *Experiment* or *Experiment* → *Control*), resulting in a counterbalanced design across participants. Moreover, on Day 0, participants provided informed consent and received instructions on wearing the earbuds and using the Earinter app. To ensure consistent sensing and intervention delivery, they identified their dominant chewing side via self-report or a brief chewing calibration task and wore a single earbud on that side throughout the study. We further selected earbud tips to ensure a tight fit with each participant's ear canal.

The study protocol was approved by the institutional review board. Demographic and study design for each participant are summarized in Table 2.

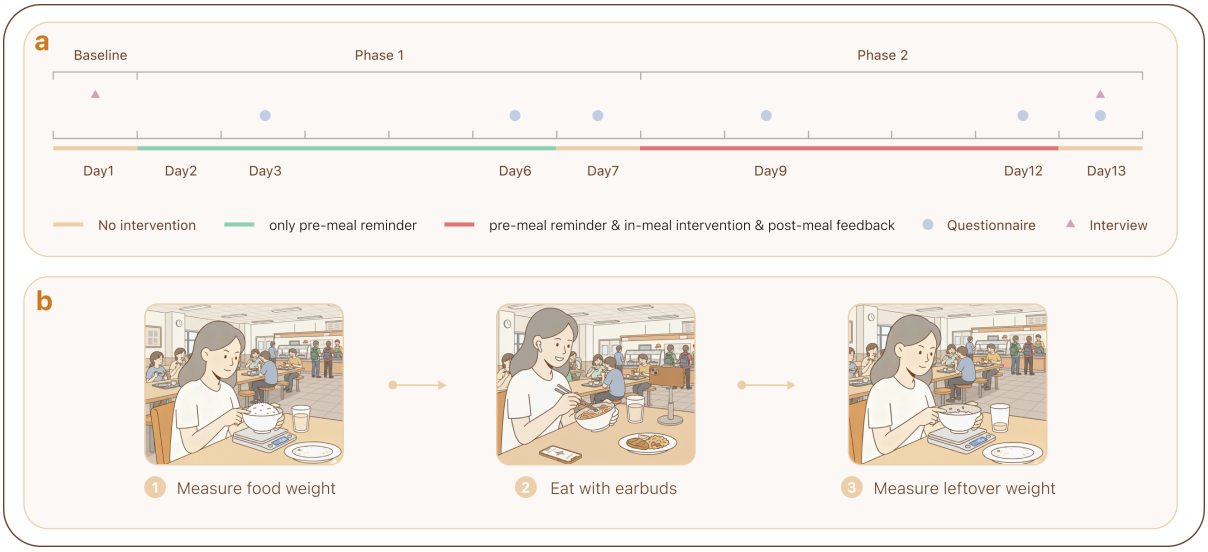



Fig. 5. User study overview for *Control → Experiment*. (a) Timeline of the 13-day within-subject field study: a baseline day (no intervention), followed by two 6-day phases with five active days under each condition (Control: only pre-meal reminder; Experiment: pre-meal reminder & in-meal JIT intervention & post-meal feedback) and a retention day with no support. Dots and triangles mark questionnaire and interview. (b) Daily procedure for data collecting: participants measured food weight before eating, ate with earbuds, and weighed leftovers to estimate intake.

## 6.2 Data Collection and Preprocess

*6.2.1 Measures and Data Collection.* During the whole study, we collected both objective behavioral measures and subjective feedback to evaluate system usability, intervention effects, and user experience in daily eating contexts.

Participants were asked to measure pre-meal food weights. During meals, to support validation, participants recorded a rear-facing video. We also collected bone-conducted audio signals, including chewing and swallowing, via acoustic vibrations from the earbud. Then, we calculated food intake (excluding water) using post-meal food weights after finishing eating. These meal-level data were used to derive eating-pace indicators and to characterize how eating patterns evolved across different conditions. Daily procedures are shown in Fig. 5b.

To capture subjective eating experience and system usability over time, participants completed questionnaires on Days 3, 6, 7, 9, 12, and 13. The questionnaire package included (1) the *Self-Reported Eating Experience Questionnaire* (Appendix D) and (2) the *System Usability Scale (SUS)* [14]. The eating-experience questionnaire assessed perceived changes in eating pace (e.g., chewing more [28, 94], slower eating), intake-related behaviors (e.g., reduced intake [93], smaller bites [87, 95, 114]) and overall meal experience.

Moreover, we conducted semi-structured interviews at the beginning of the study (Day 1) and after study completion (Day 13). The Day 1 interview focused on understanding how fast-eating behaviors are formed and maintained through the interaction of habit and context, and eliciting participants' perceived feasibility of intervention strategies under real-world constraints. The Day 13 interview focused on participants' eating-pace related behavioral changes across conditions, their subjective experience of the system, and recommendations for improving long-term use. The complete interview questions are shown in Appendix E.

*6.2.2 Data Preprocessing, Annotation, and Qualitative Coding.* All 14 remaining participants completed the 13-day protocol and we collected 364 meal sessions. Prior to pre-processing and analysis of daily recordings and interviews, we screened the dataset based on participant self-reports and author inspection. For food-consumption speed, the analyzed eating-episode duration was defined to reflect active ingestion time rather than the entire raw recording span. Pre-meal and post-meal non-eating segments (e.g., chatting before starting or after finishing the meal) were excluded. In the retained sessions, substantial non-ingestive interruptions during the meal were uncommon. When an individual session involved unusual circumstances, we reviewed the case and asked participants about the specific situation of that day during the post-study interview. We then determined whether the session should be treated as an anomalous meal and excluded, or whether a clearly non-ingestive segment should be removed during preprocessing. This procedure was intended to avoid inflating the denominator of food-consumption speed with time that did not reflect active ingestion. After excluding 11 sessions (3%) due to device issues (e.g., Bluetooth disconnections, corrupted audio files, or incomplete logs) and 11 sessions (3%) for participant-reported anomalous meals (e.g., unusually time-constrained meals or other individual reasons), 342 valid meal sessions remained. Across 14 participants over the 13-day study, this corresponds to an average of 24.43 valid sessions per participant, or approximately 1.88 valid sessions per participant per day. Our conclusions were not materially affected by these exclusions.

We used frontal video recordings as the primary ground-truth source for ingestion-related events [8, 10]. To derive meal-level chewing and swallowing counts, we recruited crowd annotators and performed manual annotation using ELAN [65]. Each video was independently annotated by two annotators. We computed Cohen's Kappa $\kappa$ for each video and required $\kappa \geq 0.75$ to indicate substantial agreement [53, 69]. For videos that did not meet this threshold, the first author reviewed the annotations, adjudicated disagreements, and finalized the labels in consultation with the research team. The finalized annotations were then aggregated to compute per-meal counts and eating-pace-related measures used in subsequent analyses.

We analyzed the semi-structured interviews using a bottom-up thematic analysis approach [13]. Given the relatively small sample size, the first author conducted the full coding process, iteratively developing and refining codes and candidate themes. The broader research team subsequently reviewed the coding structure and the resulting themes. Then, we refined the final theme set through group discussion until consensus was reached. We did not compute inter-rater reliability. Consistent with qualitative research guidance, we prioritized collaborative interpretation and reflexive discussion over purely statistical agreement measures [77].

*6.2.3 Statistical Analysis.* The statistical analyses were organized around two primary analysis questions derived from Earinter's intended intervention mechanism. First, does the active *Experiment* condition increase oral processing intensity, operationalized as chews per swallow (CPS), relative to the active *Control* condition? Second, does the active *Experiment* condition reduce food-consumption speed (g/min) relative to the active *Control* condition? Accordingly, the two primary behavioral outcomes were CPS and food-consumption speed, and the primary comparison for both outcomes was the within-participant contrast between active *Experiment* days and active *Control* days. Comparisons involving *Baseline* and retention-probe days were treated as secondary analyses intended to contextualize pre-intervention behavior and examine short-term persistence after support withdrawal. Subjective questionnaire factors were treated as secondary outcomes.

To make the within-subject nature of the design more transparent, we computed participant-level paired summaries comparing active Control and active Experiment days. For each participant, condition-level CPS was calculated as total chews divided by total swallows across valid meals within that condition, and condition-level food-consumption speed was calculated as total food consumed divided by total analyzed eating-episode duration across valid meals within that condition. We visualize these within-participant contrasts using paired plots and assess them using two-sided Wilcoxon signed-rank tests as robustness analyses.

Because the study followed a within-subject cross-over design, we used meal-level mixed-effects models that retain all valid sessions while accounting for repeated measures within participants, within-day clustering, and sequence effects. For CPS, we fit a Negative Binomial Generalized Linear Mixed Model (NB-GLMM) with chew count as the outcome and log(swallow count) as an offset, so that the model estimates chewing rate per swallow. Fixed effects included Condition (Baseline, Control, Control_retention, Experiment, Experiment_retention), Phase, and Condition sequence (Control→Experiment vs. Experiment→Control), with random intercepts for participants and participant-by-day random intercepts.

For food-consumption speed (g/min), our primary linear mixed-effects model (LMM) was fit on post-baseline meals to isolate differences among the study conditions after Day 1. Fixed effects included Condition (Control, Control_retention, Experiment, Experiment_retention), Meal type (lunch vs. dinner), Phase, and Condition sequence, with random intercepts for participants and participant-by-day random intercepts. Baseline food-consumption speed (Day 1) was included as a covariate. We report Type-III tests for omnibus effects and conducted post-hoc contrasts using estimated marginal means (EMMs) with Holm adjustment for multiple comparisons. We additionally report a complementary model that treats Baseline as an additional Condition level as a robustness check.

For self-reported eating-experience questionnaires, items were grouped into four theoretically derived factors: Behavioral Positive, Appetite, Fullness, and Comfort (see Appendix D). Behavioral Positive was computed as a composite factor, whereas Appetite, Fullness, and Comfort were analyzed as ordinal item-level factors. We assessed internal consistency using Cronbach's $\alpha$ and report reliability values in the Results. Behavioral Positive was analyzed using an LMM, whereas Appetite, Fullness, and Comfort were analyzed using cumulative link mixed models (CLMMs) to respect the ordinal response scale. These models included fixed effects of Condition (Control vs. Experiment), Phase (Active vs. Retention), and their interaction, with random intercepts for participants.

Because the active *Experiment*-versus-*Control* contrast was the primary comparison for the two behavioral outcomes, it was interpreted as the primary analysis contrast. Additional pairwise contrasts among condition states (e.g., involving *Baseline* or retention-probe days) were treated as secondary, estimated using EMMs, and adjusted using Holm correction. Pairwise contrasts for questionnaire outcomes were estimated using EMMs and adjusted using Tukey correction.

## 6.3 Results

To provide an intuition for how the Earinter system operates in situ, we present one representative day from the field study (P5, Day 2 in the *Experiment* condition) in Fig. 6. The corresponding eating-pace signals for all participants on that day are reported in Appendix F.

### *6.3.1 Quantitative Results.*

*Participant-level paired comparison of active conditions.* To make the within-subject cross-over design explicit, we first summarized each participant's active *Control* and active *Experiment* data. As shown in Fig. 7, CPS increased under *Experiment* for all participants. At the participant level, CPS increased from a median of 17.31 under *Control* to 24.20 under *Experiment*, with a median paired difference of 6.15 CPS (Wilcoxon signed-rank test: $V = 105$, $p = 0.000122$). Food-consumption speed decreased under *Experiment* for 10/14 participants. The participant-level median decreased from 58.36 g/min under *Control* to 40.52 g/min under *Experiment*, with a median paired difference of -9.87 g/min (Wilcoxon signed-rank test: $V = 18$, $p = 0.0295$). Although food-consumption speed showed more individual heterogeneity than CPS, all participants increased CPS under *Experiment*. This divergence is plausible in free-living meals because food-consumption speed depends not only on oral processing intensity, but also on meal content, bite size, and the amount of food consumed per unit time. We therefore interpret CPS and food-consumption speed as complementary indicators of eating-pace regulation. The paired summaries

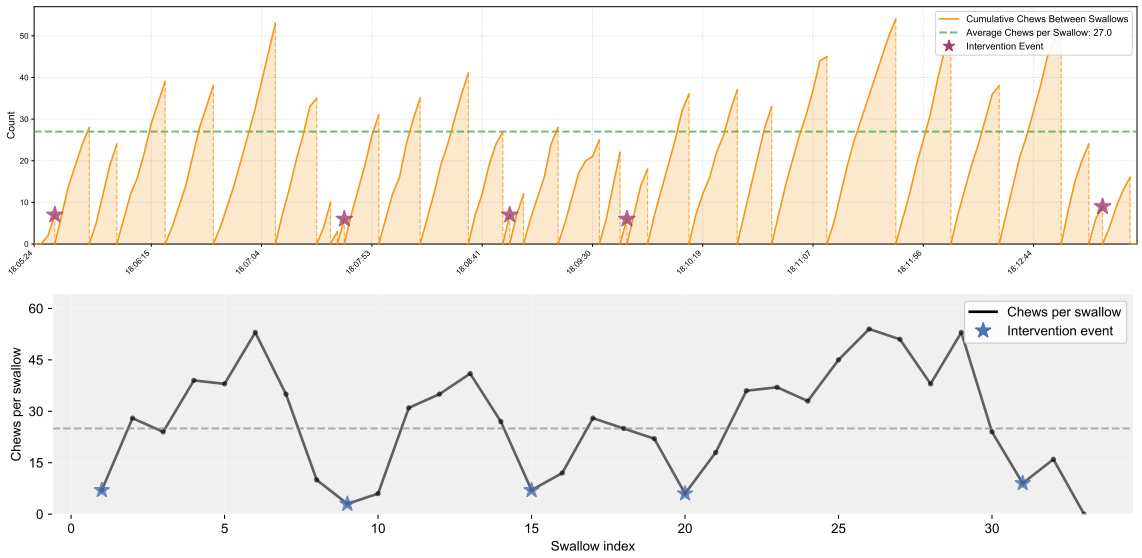



Fig. 6. Representative eating-pace signals for P5 from the second day in *Experiment* condition. The top plot shows time-based cumulative chewing activity, and the bottom shows chews per swallow. Star markers indicate moments when Earinter triggered JIT prompts based on the real-time estimated pace.

above aggregate each participant's active *Control* and active *Experiment* meals into one value per condition, whereas the following descriptives and mixed-effects models use meal-level observations.

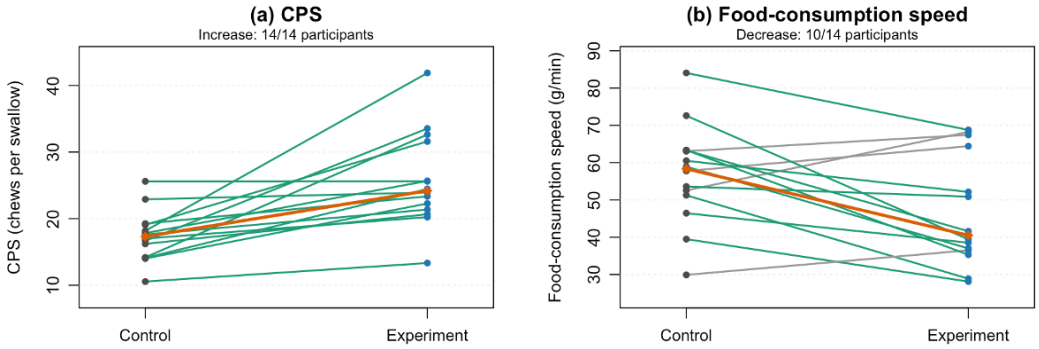



Fig. 7. Participant-level paired comparisons between active *Control* and active *Experiment* days. Each line represents one participant, with condition-level values aggregated across all valid active meals within that condition. (a) CPS, computed as total chews divided by total swallows. (b) Food-consumption speed, computed as total food consumed divided by total analyzed eating-episode duration. Green lines indicate changes in the expected direction (higher CPS or lower food-consumption speed under *Experiment*), gray lines indicate changes in the opposite direction, and the orange line indicates the median across participants within each condition.

*Eating pace – chews per swallow (CPS).* We analyzed meal-level CPS using the NB-GLMM described in Section 6.2.3, which retains all valid sessions while accounting for repeated measures, phase, and condition sequence. Meal-level descriptives, shown in Fig. 8, indicated that participants chewed on average 15.02 times per swallow at *Baseline* ($SD = 3.67$). Under the active *Control* condition, mean CPS was 17.88 ($SD = 5.32$), whereas under the active *Experiment* condition, mean CPS increased to 26.42 ($SD = 8.61$). During retention-probe days, CPS was 18.60 ($SD = 4.20$) after *Control* and 26.29 ($SD = 8.94$) after *Experiment*.

The primary analysis contrast showed that the active *Experiment* condition produced significantly higher CPS than the active *Control* condition (estimated effect = 0.37, $p < 0.0001^{****}$), indicating that closed-loop JIT prompting increased oral processing intensity relative to reminder-only support. In secondary retention-related contrasts, *Experiment_retention* remained higher than *Control* (estimated effect = 0.32, $p < 0.001^{***}$). Relative to *Baseline*, CPS increased significantly under *Experiment* both during active intervention (estimated effect = 0.53, $p < 0.001^{***}$) and on the retention-probe day (estimated effect = 0.53, $p < 0.001^{***}$), whereas *Control* and *Control_retention* showed only modest increases ($p < 0.05^{*}$). Finally, CPS did not significantly differ between *Experiment* and *Experiment_retention* (estimated effect = 0.0023, $p \approx 1$), suggesting that the increase in oral processing intensity was maintained after support withdrawal within the study window. Together, these results indicate that Earinter more strongly regulated eating pace than reminder-only support, with short-term persistence on no-support probe days.

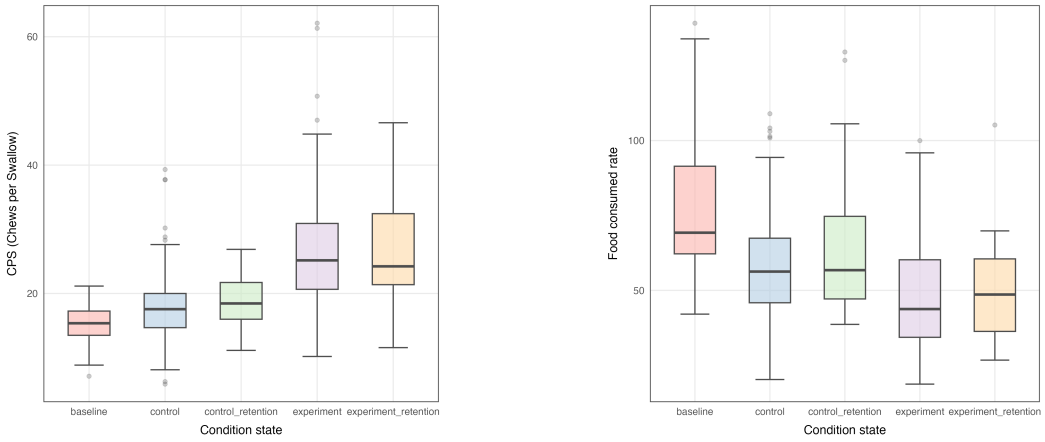



Fig. 8. Box plots showing the distribution of meal-level eating pace metrics across five conditions: *Baseline, Control, Control_retention, Experiment, and Experiment_retention*. The left plot shows the chews per swallow (CPS), indicating the oral process intensity. The right plot illustrates the food consumed rate (g/min), representing the speed of food intake.

*Eating pace – food consumed per minute.* We analyzed meal-level food-consumption speed using the LMM described in Section 6.2.3. The primary model showed a significant main effect of *Condition* on food-consumption speed ($F(3, 155.36) = 15.11$, $p < 0.001^{***}$). For the primary analysis contrast, the active *Experiment* condition showed a lower estimated food-consumption speed ($EMM = 47.82$, $95\%CI = [41.39, 54.25]$) than the active *Control* condition ($EMM = 58.31$, $95\%CI = [51.89, 64.73]$; $\Delta = -10.48$ g/min, $t = -5.31$, $p < 0.001^{***}$). This result indicates that closed-loop JIT prompting reduced food-consumption speed relative to reminder-only support. In secondary retention-related contrasts, *Experiment_retention* remained lower than *Control* ($EMM = 49.01$, $95\%CI = [40.69, 57.33]$; $\Delta = -9.29$ g/min, $t = -2.71$, $p = 0.02^{*}$). The difference between *Experiment* and

*Experiment_retention* was not significant ($\Delta = -1.19$ g/min, $t = 0.34$, $p = 0.73^{ns}$), suggesting that the reduction in food-consumption speed was maintained on no-support retention-probe days within the study window.

As a robustness check, we also fit a complementary LMM that treated *Baseline* as an additional *Condition* level. In this model, *Baseline* food-consumption speed ($EMM = 76.40$, $95\%CI = [67.47, 85.33]$) was significantly higher than all other conditions, including *Control* ($\Delta = 18.05$ g/min, $p < 0.0001^{****}$), *Control_retention* ($\Delta = 11.32$ g/min, $p = 0.0489^{*}$), *Experiment* ($\Delta = 28.55$ g/min, $p < 0.0001^{****}$), and *Experiment_retention* ($\Delta = 27.30$ g/min, $p < 0.0001^{****}$). Together, these results support the conclusion that Earinter reduced food-consumption speed beyond reminder-only support, with short-term persistence after support withdrawal.

*Self-reported eating experience.* We next report the self-reported eating-experience factors analyzed using the models described in Section 6.2.3. The questionnaire items were grouped into four factors: *Behavioral Positive* (positive perception of eating behavior), *Appetite*, *Fullness*, and *Comfort* (wearing comfort). A total of 84 valid observations were included (28 *Control*, 28 *Experiment*, 14 *Control_retention*, and 14 *Experiment_retention*). The overall scale showed good internal consistency (Cronbach's $\alpha = 0.866$), and the *Behavioral Positive* subscale showed excellent consistency ($\alpha = 0.929$). Results are summarized in Fig. 9; see Appendix D for the mapping of questionnaire items to factors.

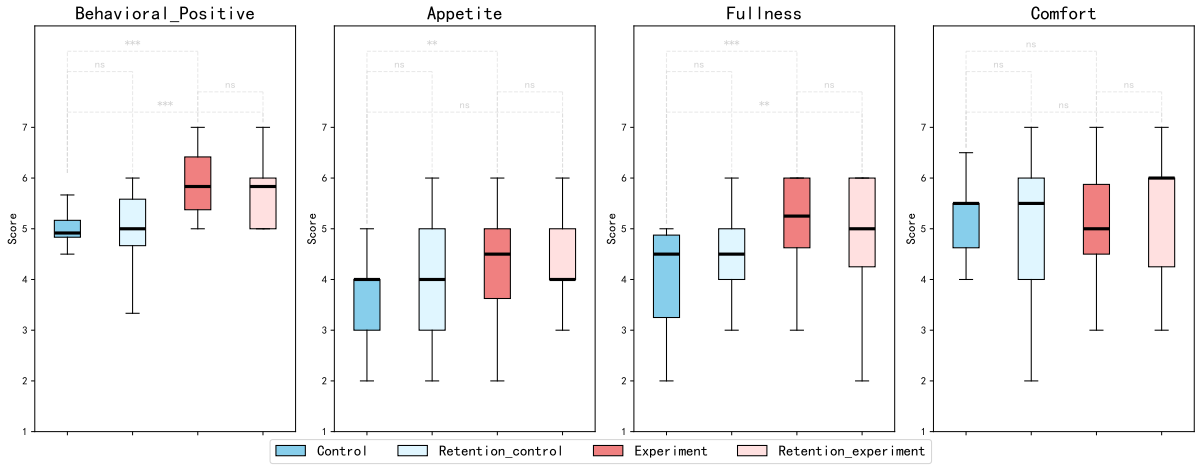



Fig. 9. Box plots of self-reported eating experience scores for the four factors, *Behavioral Positive*, *Appetite*, *Fullness*, and *Comfort*, across the *Control, Control_retention, Experiment, and Experiment_retention* conditions. Asterisks denote statistically significant pairwise differences (* indicates $p < 0.05$, ** indicates $p < 0.01$, *** indicates $p < 0.001$).

Behavioral Positive: Participants perceived their eating behavior as more positive during the *Experiment* condition ($M = 5.68$, $SD = 1.12$) than the *Control* condition ($M = 4.88$, $SD = 0.83$). An LMM analysis confirmed this difference ($\beta = 0.80$, $SE = 0.13$, $t = 6.12$, $p < .001^{***}$), indicating improved perceived eating control under the JIT intervention. After support withdrawal, *Experiment_retention* session still yielded higher scores ($M = 5.55$, $SD = 1.03$) than *Control* ($p = 0.0001^{***}$), consistent with short-term persistence of perceived behavioral improvement, although scores were negligibly lower than during the Intervention phase ($p = 0.415^{ns}$). These results indicate the short-term persistence of perceived behavioral improvement , with only mild attenuation after support withdrawal.

Appetite: Participants reported higher appetite-related scores during the *Experiment* condition ($M = 4.32$, $SD = 1.25$) than during *Control* ($M = 3.68$, $SD = 1.09$). The difference was confirmed with a CLMM analysis

($OR = 4.43$, $95\%CI = [1.54, 12.77]$, $p = 0.0059^{**}$), indicating improved perceived appetite regulation during the JIT intervention. No significant difference was observed between the *Experiment* and *Experiment_retention* conditions ($p = 0.593^{ns}$), suggesting that this perceived improvement did not substantially decline after support withdrawal. Scores in *Experiment_retention* were slightly higher than those in *Control*, although this comparison did not reach conventional significance ($p = 0.056$). Together, these results suggest that appetite-related perception improved during the JIT intervention in *Experiment* condition and showed signs of short-term persistence after support withdrawal.

Fullness: Compared with *Control* ($M = 4.07$, $SD = 1.12$), participants reported greater fullness during the *Experiment* condition ($M = 5.00$, $SD = 1.25$). The CLMM showed that this difference was significant ($OR = 12.62$, $95\%CI = [3.76, 42.32]$, $p < .001^{***}$). After support withdrawal, *Experiment_retention* did not differ significantly from the active *Experiment* condition ($p = 0.684^{ns}$), but remained significantly higher than *Control* ($p = 0.0012^{**}$). These results suggest short-term persistence of perceived fullness improvement following the JIT intervention.

Comfort: The comfort factor was reverse-coded prior to analysis. Ratings were comparable between *Control* ($M = 5.10$, $SD = 1.07$) and *Experiment* ($M = 4.96$, $SD = 1.37$). No significant condition effect was observed ($OR = 0.92$, $95\%CI = [0.34, 2.51]$, $p = 0.873^{ns}$). After support withdrawal, both comparisons remained non-significant (*Control_retention*: $p = 0.606^{ns}$; *Experiment_retention*: $p = 0.565^{ns}$). These results indicate that the presence of auditory feedback, whether JIT or not, did not reduce comfort relative to simply wearing the earbuds. Descriptively, the overall comfort rating ($M = 5.03$, $SD = 1.27$) suggests generally favorable comfort of commodity earbuds across phases.

*System Usability Scale.* We evaluated the system's perceived usability using the System Usability Scale (SUS) [14]. Given the documented advantage of 7-point Likert scales in enhancing response sensitivity [26], we adopted this format for the SUS questionnaire following Lee et al. [55]. Across *Experiment* phases, the mean SUS score was 70.0 ± 9.6 ($median = 70.0$, $IQR = 16.7$), placing the system within the "OK-Good" range on the interpretive scale proposed by Bangor et al. [7].

Across the ten SUS items of experiment, participants reported higher ratings on positively worded statements such as "I found the system easy to use" ($M = 5.58$) and "I felt very confident using the system" ($M = 5.69$). These indicated improved perceived ease of use and confidence with the system. In contrast, lower scores appeared on negatively worded items such as "I found inconsistencies in the system" ($M = 2.84$) and "I needed to learn a lot of things before I could get going with this system" ($M = 2.78$), suggesting that users in the experimental condition perceived fewer complexities and a smoother learning process. These differences reflect a generally more intuitive and user-friendly experience in the experimental interface.

Overall, participants perceived the earbuds-based JIT intervention system as usable and well-designed, indicating promising usability and user acceptance for real-world behavioral intervention contexts.

*6.3.2 Qualitative Findings.* We report qualitative findings from two rounds of semi-structured interviews, analyzed using an inductive thematic analysis approach (Section 6.2.2). Baseline interviews (Day 1) characterized participants' fast-eating routines and constraints that limit intentional change, while exit interviews (Day 13) captured experiences across conditions, perceived behavior change (including retention-probe days), and design suggestions. We synthesize themes across both rounds to explain why fast eating persists and how participants experienced our interventions in daily ingestion.

*Theme 1: Habit- and context-driven fast eating makes self-regulation difficult, motivating low-burden external support.* In baseline interviews, participants commonly described fast eating pace as an ingrained and automatic habit that developed and reinforced by some constraints over time. For instance, P12 noted that he had "always chewed less and eaten faster since childhood". Others attributed their pace to long-term time scarcity and efficiency-seeking (e.g., P5:"I want to save time."), including routines shaped by short breaks earlier in life. Beyond

Table 3. Summary of participants' self-perceived behavioral changes and design preferences reported in semi-structured interviews (N=14). Changes were aggregated per participant (✓ = explicitly mentioned; – = not mentioned). "Condition preference" indicates overall preference among baseline, control and experiment (closed-loop) conditions; "Preference for JIT prompt type" refers to the prompt framing participants favored (see Section 5.2).

| Participant id | Self-perceived behavioral changes | | | | | | Design preferences and attitudes | |
|---|---|---|---|---|---|---|---|---|
| | Slower eating pace | Increased CPS | Smaller bite size | Earlier fullness | Reduced food intake | Heightened food awareness | Condition preference | JIT prompt preference |
| P1 | ✓ | ✓ | – | – | – | – | Experiment | Control Theory |
| P2 | ✓ | ✓ | – | – | – | – | Experiment | Control Theory |
| P3 | ✓ | ✓ | – | ✓ | ✓ | ✓ | Experiment | Goal Framing Theory |
| P4 | ✓ | ✓ | – | – | – | ✓ | Experiment | Control Theory |
| P5 | ✓ | ✓ | ✓ | – | – | ✓ | Experiment | Control Theory |
| P6 | ✓ | ✓ | – | – | – | ✓ | Experiment | Goal Framing Theory |
| P7 | ✓ | ✓ | – | – | – | ✓ | Experiment | Control Theory |
| P8 | ✓ | ✓ | – | ✓ | ✓ | – | Experiment | Control Theory |
| P9 | – | – | – | – | – | ✓ | Control | Control Theory |
| P10 | ✓ | ✓ | ✓ | – | – | ✓ | Experiment | Control Theory |
| P11 | ✓ | ✓ | ✓ | ✓ | – | ✓ | Experiment | Control Theory |
| P12 | ✓ | – | ✓ | – | – | – | Baseline | Dual System Theory (System 1) |
| P13 | ✓ | ✓ | – | – | – | – | Experiment | Control Theory |
| P14 | ✓ | ✓ | ✓ | ✓ | – | ✓ | Experiment | Goal Framing Theory |

habit, participants emphasized that pace was strongly context-dependent, varying with workload, social settings, and concurrent activities. P7 reported speeding up to match peers when eating together, while P14 described time-compressed lunch breaks that made slow eating feel unrealistic. Although some participants occasionally recognized fast eating pace, they reported that self-regulation was difficult to sustain. Because maintaining self-reminders felt cognitively demanding (e.g., P3) and sometimes physically uncomfortable during busy or stressful meals (e.g., P1, P8). These barriers motivated a preference for external and low-burden support, so participants expressed interest in prompts triggered by objective eating signals. They also cautioned that messages which are overly long or delivered too frequently could increase perceived burden and become easier to ignore (e.g., P3, P6, P8), underscoring the importance of pilot study in intervention design (Section 5).

*Theme 2: Participants attributed fast eating to delayed satiety. Slowing down increased perceived fullness, while intake was constrained by portion norms.* Participants described a perceived mismatch between fast eating pace and internal satiety awareness. Baseline interview suggested heterogeneity in how pace relates to intake. Some participants described relatively fixed portions, whereas others (e.g., P3, P8, P11) linked fast eating to delayed satiety cues and a tendency to overeat. Several participants described this experience as "satiety-related signals cannot keep up when eating too fast". In exit interviews, multiple participants reported that when the closed-loop condition helped them chew more and slow down, they felt full earlier or found it harder to finish their usual portion (e.g., P3, P8). However, regardless of increased fullness, many participants reported little change in total intake. They often cite plate-clearing norms (avoiding food waste) and fixed meal expectations as constraints that decouple pace from intake. Taken together, these results suggest that pace interventions may more reliably change chewing dynamics and perceived satiety than absolute intake for users whose portions are socially or habitually fixed.

*Theme 3: JIT interventions were most effective when actionable and tightly coupled to eating pace, supporting moment-to-moment correction.* Baseline interviews indicated that participants were generally optimistic about our closed-loop system including JIT intervention. For exit interview, as shown in Table 3, most participants preferred closed-loop system and reported that the closed-loop condition effectively helped slow their moment-to-moment eating pace to a stable level. The most frequently mentioned behavioral change was increased chewing per swallow. For example, P1 explicitly stated that "the most noticeable positive effect was chewing more". Some participants also reported learning-related changes over time, such as "gradually adopting smaller bites later in the final period" (P10) and "eating in smaller and more even bites" (P5). Participants further described carryover

in retention day and after the user study. For instance, P10 noted that even without continuing to use the system after the study, "this experience really made me eat more slowly". These accounts are consistent with our intended mechanism, in which brief in-the-moment interruptions redirect attention to slower eating pace. Together, they suggest early signs of behavioral persistence beyond active prompting, aligning with the long-term aspirations of closed-loop system design.

Participants also differentiated between prompt styles. Across interviews, many expressed a preference for prompts that provided actionable, control-oriented guidance, because such prompts supported immediate self-correction. In contrast, messages derived from more general behavioral theories were often described as generic, and goal-framing prompts were sometimes perceived as "too common-sense" (P11) or "not relevant to me" (P4).

*Theme 4: Participants desired positive reinforcement, transparency, and personalization to support long-term use.* Participants suggested incorporating more positive reinforcement to enhance motivation and willingness to persist. For example, P2 and P5 recommended that the system acknowledge improvements when they returned to a healthier pace. In addition, about half of participants explicitly requested greater transparency and personalization. Transparency referred to helping users better understand their ongoing eating process through richer, real-time indicators (e.g., P4, P5: a progress-bar-like visualization) rather than a "black-box" experience. Personalization requests focused on adapting targets to meal content and individual differences. They noted that different food properties likely require different chewing expectations (e.g., P12), and suggested allowing users to set pre-meal intentions or goals (P1). Together, these suggestions point to design directions that could improve acceptability and support sustained behavior change beyond the study period.

## 7 DISCUSSION

### 7.1 Closed-Loop System for Dynamic Ingestive Behavior Regulation

During a meal, eating pace can fluctuate dynamically without the eater's awareness, making purely open-loop strategies inherently misaligned with moment-to-moment changes [24, 36]. Feedback may arrive too early or too late, or without sufficient context to be actionable, which can weaken perceived usefulness and lead users to disengage. We therefore adopt a closed-loop design that tightly couples sensing, real-time reasoning, and intervention [25]. In Earinter, continuous CPS estimation enables behavior-contingent support: the system intervenes only when inferred pace drifts toward fast eating and remains silent when pace is already within a desired range, reducing unnecessary interruptions while keeping feedback directly relevant to the user's immediate state. This design further enables an iterative *Observe* $\rightarrow$ *Correct* $\rightarrow$ *Observe* cycle: Earinter cues correction at the moment of deviation, re-assesses the subsequent response through continued sensing, and re-prompts if the pace drifts again.

The closed-loop structure, as Earinter illustrates, maps naturally onto *Prompt*, *Ability*, and *Motivation* [27]. First, cues serve as timely *Prompt*, triggered specifically at deviation moments rather than on a fixed schedule. Second, the closed-loop cycle strengthens perceived *Ability* in two complementary ways. By externalizing sustained self-monitoring, the system lowers the cognitive effort required to notice and correct pace changes, making regulation easier in the moment. In addition, when users correct their behavior and the system subsequently remains silent, this immediate feedback serves as confirmation that their action led to measurable improvement, reinforcing a sense of self-efficacy and control. Third, the system supports *Motivation* through a lightweight intervention policy designed that reduces frustration and makes sustained regulation more likely, and through repeated, attainable corrections that help users experience progress. Although a two-week deployment cannot establish durable long-term habit change, we observed significant retention on no-support retention-probe days within the study window, suggesting early transfer from external prompting to self-regulation. As P10 reflected a week after the study, "I found the intervention really effective. I genuinely ate more slowly!" These findings suggest that closed-loop systems can iteratively reshape habitual eating pace through repeated, contextually

grounded feedback. We encourage future HCI work on habit change to explicitly consider these iterative benefits and examine how repeated closed-loop systems may support cognitive-level internalization and persistence over longer time horizons.

## 7.2 Earables Enable End-to-End Regulation as a Practical Platform

Earables offer a practical substrate for closed-loop systems not only because they co-locate sensing and feedback, but also because ear-worn devices are already integrated into everyday life. Prior work characterizes earables as building on an existing consumer platform and notes the large-scale adoption of wireless earphones [86].

Beyond convenience, earables provide several advantages as an intervention channel. In-ear audio cues are private and eyes-free, and therefore do not compete with visual attention during meals [86, 96]. Compared with screen-based prompts, visible wearables, or tabletop systems, earables can deliver subtle guidance with relatively low social intrusiveness [105]. Earinter further limits disruption through behavior-contingent triggering, cooldown intervals, and theory-grounded prompts, which together aim to preserve intervention salience while maintaining comfort and usability.

The ear is also a favorable site for robust, fine-grained ingestive sensing [91]. Compared with air-conducted microphones that are more susceptible to environmental noise in real meals, bone-conducted sensing captures chewing-related vibrations transmitted through the ear canal and skull, making the signal less sensitive to ambient sounds and more privacy-preserving. While prior ear-based systems often focus on coarse meal or eating-episode detection [8, 10, 91], Earinter targets within-meal pacing support and therefore requires chew-level reliability for real-time intervention.

Building on these properties, Earinter repurposes a commodity earbud's bone-conduction voice sensor to capture chewing-related vibrations while using the same device to deliver in-ear prompts. This dual use enables a closed-loop system on an already-adopted form factor, reduces setup friction relative to multi-device configurations, and lowers the barrier to sustained daily use. Our evaluation provides further evidence for the practical value of this design. Comfort remains comparable to the control condition, with a mean rating of $5.03 \pm 1.27$, and the mean SUS score is $70.0 \pm 9.6$. Participants also report strong willingness to use the system frequently in the SUS ($M = 5.14, SD = 1.37$), suggesting that the earable form factor is generally comfortable and usable across diverse in-the-wild meals. Qualitative feedback from Wizard-of-Oz pilot sessions across solo, dyadic, and family co-present meals further informs the selected prompting strategy, favoring interventions that participants perceive as acceptable and minimally interruptive, with few reported negative reactions. Moreover, during post-study interviews, most participants indicate that the earbuds "do not impose noticeable burden during in-the-wild eating" (P2). Some participants routinely wear earbuds during meals (e.g., P6, P13) and report no perceived burden. In addition, several participants express a desire to continue using the system on their own earbuds (P9, P13), further supporting the practical advantages of earables. Taken together, these findings suggest that this earable form factor can support low social and cognitive interference in many everyday meal contexts.

## 7.3 Towards the Just-In-Time Adaptive Intervention

Just-in-time (JIT) shifts intervention from static advice on what to do to situated support about what to do now, and JIT adaptive intervention (JITAI) further emphasizes that effective support requires not only timely triggering but also adaptive decisions as context conditions change [72]. This framing is particularly suitable for habitual behaviors that fluctuate within daily episodes, such as eating pace, where a user's vulnerability may emerge intermittently and is hard to notice through sustained self-monitoring.

We found that adaptivity is central to both effectiveness and acceptability. Earinter therefore treats the delivery policy as part of the intervention: pacing-contingent triggering, a cooldown interval, and varied prompt lengths together adjust the intervention dose to preserve relevance while reducing interruption. This is consistent with

evidence that interventions embedded at in-situ decision points and requiring lightweight engagement can outperform static notifications in real-world behavior change [106]. In our study, participants also preferred actionable, control-theory-oriented prompts (e.g., "chew xx more times to reach the goal"), suggesting that JIT content should emphasize concrete, immediately executable actions. This aligns with a control-theoretic view of regulation – users benefit from feedback that makes the current deviation legible and specifies how to correct it.

Looking forward, a promising direction is to extend earable-based pacing support toward fuller JITAI capabilities by incorporating richer contextual signals, such as social context, concurrent media use, time pressure, and personalization. These directions point toward JIT interventions that behave less like fixed reminders and more like adaptive companions for self-regulation of habitual behaviors.

### 7.4 Practical Reliability under Non-Ingestive Daily Activities

Beyond the meal-based evaluation, we also examined how the trained chewing-detection pipeline behaves under several non-ingestive activities during daily earbud use, with full task-level statistics reported in Appendix B. We recruited 10 participants to complete continuous no-eating blocks covering walking-related activities, speech and social activities, head motion, and audio playback. The results suggest that the detector is not indiscriminately activated by routine earbud use. Passive external speech, phone checking, and audio playback were rarely confused with chewing. More specifically, within the speech/social block, self-generated speech introduced measurable false positives – loud reading produced 3.851 FP/min, dyadic conversation produced 2.474 FP/min, and listening produced 0 FP/min.

At the same time, the additional test reveals clear boundary conditions. Continuous loud self-generated speech, repeated head motion, and stronger body movement such as jogging or fast stair walking produced non-negligible false positives, likely because bone-conduction sensing can also capture self-vocalization and motion-related vibrations. These results help define Earinter's practical operating envelope. Because Earinter is a user-initiated meal-time intervention rather than a passive all-day detector, users are expected to initiate it during typical meals rather than during high-motion activities or socially intensive meal contexts. Though, loud speech remains a realistic boundary condition. Future systems should therefore incorporate richer contextual suppression so that closed-loop support can remain effective and socially acceptable across a wider range of daily situations.

### 7.5 Limitations and Future Work

*Deployment scope and participant diversity.* Our current study deployment lasted 13 days and primarily served to validate the feasibility and effectiveness of an in-the-wild closed-loop system for eating-pace regulation. Although we observed significant improvements in CPS and slower eating pace, together with short-term retention on no-support retention-probe days within the study window, this duration is still insufficient to establish durable habit change. Future work should conduct longer deployments to examine how effects evolve over time, including patterns of persistence, attenuation, and habituation. Importantly, these short-term effects were limited to eating-pace regulation and should not be interpreted as evidence of weight loss. In an exploratory participant-level analysis, the reduction in food-consumption speed was not significantly correlated with short-term body-weight change, irrespective of participants' starting body weight. Longer studies with pre-specified intake and weight outcomes are needed to examine this relationship. Our participant sample is also limited in demographic diversity. Although participants included undergraduate students, graduate students, and campus workers, all enrolled participants were young adults aged 20–29. In addition, we did not conduct formal clinical screening for eating disorders using a validated assessment instrument, and instead relied on self-report during recruitment. Future work should evaluate the system with more diverse populations and incorporate more structured screening procedures when appropriate.

*Pace-estimation precision and meal-dependent targets.* The current CPS estimation also has room for improvement. Because swallowing signatures are relatively weak in commodity earbud bone-conduction signals, Earinter infers swallow boundaries from inter-chew intervals rather than directly sensing each swallow. This limitation is reflected in the MAE. When prolonged inter-chew gaps reflect non-ingestive pauses rather than true swallows, the heuristic may insert false swallow boundaries and consequently underestimate CPS. In the analyzed field-study sessions, substantial non-ingestive interruptions were uncommon. The intervention policy also operates on continuously updated CPS estimates with pace thresholds and cooldown control, reducing the likelihood that isolated pause-related errors lead to persistent or excessive prompting. Advances in hardware or signal processing could enable more precise behavioral regulation while preserving a similarly portable and practical form factor. Healthy eating pace is also food-dependent. We used commonly cited chewing heuristics as an intuitive reference to standardize prompt design and evaluate system effects [2, 94]. A promising direction for future work is to incorporate meal characteristics, such as food texture or type, to adapt CPS goals and improve robustness in daily use.

*Situational fit and social acceptability.* Although our intervention policy is designed to limit interruption, it may still be misaligned with users' situational preferences in some contexts. In particular, in conversation-centered social meals, the acceptability of wearing earphones and receiving auditory prompts may depend on individual preferences, relationship norms, and the salience of the social setting. Future systems could therefore incorporate context-aware suppression for conversation-heavy moments, richer personalization of when and how prompts are delivered, or alternative device forms that are more robust to social sensitivity.

## 8 CONCLUSION

In this paper, we present Earinter, a closed-loop system using commodity earbuds for regulating eating pace during daily meals. Earinter repurposes an earbud's bone-conduction voice sensor to detect chewing and infer swallow boundaries for real-time chews-per-swallow (CPS) tracking. With hours of data collected equally across sessions in quiet and noisy environments, Earinter achieved reliable chewing detection ($F1 = 0.97$) and accurate pace estimation ($MAE = 0.18 \pm 0.13$ chews/min; $3.65 \pm 3.86$ chews/swallow). Guided by Dual Systems Theory, Earinter delivers theory-based just-in-time audio prompts with an adaptive delivery policy designed to preserve comfort and limit interruption. We conducted a 13-day within-subject field study with 14 participants to evaluate the effectiveness and acceptability of the closed-loop intervention in the wild. Our results show that Earinter significantly increases CPS ($15.02 \pm 3.67 \rightarrow 26.42 \pm 8.61$) and slowed meal-level eating pace compared to self-cueing or external reminder, with significant retention on no-support retention-probe days and favorable comfort and usability. However, these short-term pace-related changes were not evidence of body-weight loss. Together, these results highlight the promise of commodity earables as practical platforms for theory-driven closed-loop JIT interventions that regulate eating pace in real-world settings.

## Acknowledgments

This work is supported by the Natural Science Foundation of China under grant No. 62472244, 62132010, Key R&D Program of Shandong Province of China under grant No. 2026CXGC010116, the Qinghai University Research Ability Enhancement Project under grant No. 2025KTSA05, Institute for Artificial Intelligence of Tsinghua University (THUAI), Undergraduate / Graduate Education Innovation Grants of Tsinghua University, and Tsinghua University Initiative Scientifc Research Program.

## References

[1] Palakorn Achananuparp, Ee-Peng Lim, Vibhanshu Abhishek, and Tianjiao Yun. 2018. Eat & tell: a randomized trial of random-loss incentive to increase dietary self-tracking compliance. In *Proceedings of the 2018 International Conference on Digital Health*. 45–54.

[2] R McNeill Alexander. 1998. News of chews: the optimization of mastication. *Nature* 391, 6665 (1998), 329–329.

[3] Wathik Alsalim and Bo Ahrén. 2019. Insulin and incretin hormone responses to rapid versus slow ingestion of a standardized solid breakfast in healthy subjects. *Endocrinology, Diabetes & Metabolism* 2, 2 (2019), e00056.

[4] Nabil Alshurafa, Haik Kalantarian, Mohammad Pourhomayoun, Jason J Liu, Shruti Sarin, Behnam Shahbazi, and Majid Sarrafzadeh. 2015. Recognition of nutrition intake using time-frequency decomposition in a wearable necklace using a piezoelectric sensor. *IEEE sensors journal* 15, 7 (2015), 3909–3916.

[5] Megumi Aoshima, Kanako Deguchi, Risako Yamamoto-Wada, Chihiro Ushiroda, Eri Hiraiwa, Miyuki Yokoi, Chisato Ono, Mitsuyoshi Yoshida, and Katsumi Iizuka. 2025. Greater Numbers of Chews and Bites and Slow External Rhythmic Stimulation Prolong Meal Duration in Healthy Subjects. *Nutrients* 17, 6 (2025), 962.

[6] Albert Bandura. 1997. *Self-efficacy: The exercise of control*. Macmillan.

[7] Aaron Bangor, Philip T Kortum, and James T Miller. 2008. An empirical evaluation of the system usability scale. *Intl. Journal of Human–Computer Interaction* 24, 6 (2008), 574–594.

[8] Abdelkareem Bedri, Richard Li, Malcolm Haynes, Raj Prateek Kosaraju, Ishaan Grover, Temiloluwa Prioleau, Min Yan Beh, Mayank Goel, Thad Starner, and Gregory Abowd. 2017. EarBit: using wearable sensors to detect eating episodes in unconstrained environments. *Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies* 1, 3 (2017), 1–20.

[9] Abdelkareem Bedri, Yuchen Liang, Sudershan Boovaraghavan, Geoff Kaufman, and Mayank Goel. 2022. Fitnibble: A field study to evaluate the utility and usability of automatic diet monitoring in food journaling using an eyeglasses-based wearable. In *Proceedings of the 27th International Conference on Intelligent User Interfaces*. 79–92.

[10] Shengjie Bi, Tao Wang, Nicole Tobias, Josephine Nordrum, Shang Wang, George Halvorsen, Sougata Sen, Ronald Peterson, Kofi Odame, Kelly Caine, et al. 2018. Auracle: Detecting eating episodes with an ear-mounted sensor. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 3 (2018), 1–27.

[11] Paul Boersma and David Weenink. 2026. Praat: doing phonetics by computer [Computer program]. Retrieved January 25, 2026, from https://www.praat.org/. https://www.praat.org/

[12] Montinee Borvornparadorn, Vasana Sapampai, Chunya Champakerdsap, Wipawee Kurupakorn, and Suwimol Sapwarobol. 2019. Increased chewing reduces energy intake, but not postprandial glucose and insulin, in healthy weight and overweight young adults. *Nutrition & Dietetics* 76, 1 (2019), 89–94.

[13] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101.

[14] John Brooke et al. 1996. SUS-A quick and dirty usability scale. *Usability evaluation in industry* 189, 194 (1996), 4–7.

[15] Veerle Buntsma. 2023. *Using an interactive diner table to influence eating rate in a social setting*. Thesis. University of Twente, Enschede. https://purl.utwente.nl/essays/96524

[16] Charles S Carver and Michael F Scheier. 1982. Control theory: A useful conceptual framework for personality–social, clinical, and health psychology. *Psychological bulletin* 92, 1 (1982), 111.

[17] Yang Chen, Zhitong Cui, and Ching Chiuan Yen. 2021. Chewpin: a wearable acoustic device for chewing detection. In *Adjunct Proceedings of the 2021 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2021 ACM International Symposium on Wearable Computers*. 11–12.

[18] Yang Chen, Katherine Fennedy, Anna Fogel, Shengdong Zhao, Chao Zhang, Lijuan Liu, and Chingchiuan Yen. 2022. SSpoon: a shape-changing spoon that optimizes bite size for eating rate regulation. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 3 (2022), 1–32.

[19] Yang Chen, Anna Fogel, Yue Bi, and Ching Chiuan Yen. 2023. Factors associated with eating rate: a systematic review and narrative synthesis informed by socio-ecological model. *Nutrition Research Reviews* (2023), 1–20.

[20] Yang Chen, Felicia Fang-Yi Tan, Zhuoyu Wang, Xing Liu, Jiayi Zhang, Yun Huang, Shengdong Zhao, and Ching Chiuan Yen. 2025. ViFeed: Promoting Slow Eating and Food Awareness through Strategic Video Manipulation during Screen-Based Dining. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. 1–24.

[21] Yang Chen and Ching Chiuan Yen. 2022. SLNOM: Exploring the sound of mastication as a behavioral change strategy for rapid eating regulation. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts*. 1–6.

[22] Woohyeok Choi, Sangkeun Park, Duyeon Kim, Youn-kyung Lim, and Uichin Lee. 2019. Multi-stage receptivity model for mobile just-in-time health intervention. *Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies* 3, 2 (2019), 1–26.

[23] Judith Irene Maria De Groot. 2022. The effectiveness of normative messages to decrease meat consumption: The superiority of dynamic normative messages framed as a loss. *Frontiers in Sustainability* 3 (2022), 968201.

[24] Sarah Faltaous, Simon Janzon, Roman Heger, Marvin Strauss, Pedram Golkar, Matteo Viefhaus, Marvin Prochazka, Uwe Gruenefeld, and Stefan Schneegass. 2021. Wisdom of the IoT crowd: Envisioning a smart home-based nutritional intake monitoring system. In *Proceedings of Mensch und Computer 2021*. 568–573.

[25] Jun Fang, Yanuo Zhou, Ka I Chan, Jiajin Li, Zeyi Sun, Zhengnan Li, Zicong Fu, Hongjing Piao, Haodong Xu, Yuntao Wang, et al. 2025. A Review of Behavioral Closed-Loop Paradigm from Sensing to Intervention for Ingestion Health. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 9, 4 (2025), 1–50.

[26] Kraig Finstad. 2010. Response interpolation and scale sensitivity: Evidence against 5-point scales. *Journal of usability studies* 5, 3 (2010), 104–110.

[27] Brian J Fogg. 2019. Fogg behavior model. *Behav. Des. Lab., Stanford Univ., Stanford, CA, USA, Tech. Rep* (2019).

[28] Hideki Fukuda, Toshiyuki Saito, Masato Mizuta, Shunji Moromugi, Takakazu Ishimatsu, Shinobu Nishikado, Hiroshi Takagi, and Yoshihiko Konomi. 2013. Chewing number is related to incremental increases in body weight from 20 years of age in Japanese middle-aged adults. *Gerodontology* 30, 3 (2013), 214–219.

[29] Sangita Ganesh, Paul Marshall, Yvonne Rogers, and Kenton O'Hara. 2014. FoodWorks: tackling fussy eating by digitally augmenting children's meals. In *Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational*. 147–156.

[30] Ai Ting Goh, Georgia Chatonidi, Michelle Choy, Shalini Ponnalagu, Markus Stieger, and Ciarán G Forde. 2021. Impact of individual differences in eating rate on oral processing, bolus properties and post-meal glucose responses. *Physiology & Behavior* 238 (2021), 113495.

[31] Ai Ting Goh, Jie Ying Michelle Choy, Xin Hui Chua, Shalini Ponnalagu, Chin Meng Khoo, Clare Whitton, Rob Martinus van Dam, and Ciarán Gerard Forde. 2021. Increased oral processing and a slower eating rate increase glycaemic, insulin and satiety responses to a mixed meal tolerance test. *European Journal of nutrition* 60, 5 (2021), 2719–2733.

[32] Leo A Goodman. 1961. Snowball sampling. *The annals of mathematical statistics* (1961), 148–170.

[33] Takaharu Goto, Atsuko Nakamich, Megumi Watanabe, Kan Nagao, Miwa Matsuyama, and Tetsuo Ichikawa. 2015. Influence of food volume per mouthful on chewing and bolus properties. *Physiology & Behavior* 141 (2015), 58–62.

[34] Roel CJ Hermans, Sander Hermsen, Eric Robinson, Suzanne Higgs, Monica Mars, and Jeana H Frost. 2017. The effect of real-time vibrotactile feedback delivered through an augmented fork on eating rate, satiation, and food intake. *Appetite* 113 (2017), 7–13.

[35] Eclaire AG Hietbrink, Anouk Middelweerd, Pepijn van Empelen, Katharina Preuhs, Annemieke AJ Konijnendijk, Wendy Oude Nijeweme-d'Hollosy, Laura K Schrijver, Gozewijn D Laverman, and Miriam MR Vollenbroek-Hutten. 2023. A digital lifestyle coach (E-Supporter 1.0) to support people with type 2 diabetes: participatory development study. *JMIR human factors* 10, 1 (2023), e40017.

[36] S Hori, K Hori, S Yoshimura, F Uehara, N Sato, Y Hasegawa, K Akazawa, and T Ono. 2023. Masticatory behavior change with a wearable chewing counter: A randomized controlled trial. *Journal of Dental Research* 102, 1 (2023), 21–27.

[37] Delwar Hossain, J Graham Thomas, Megan A McCrory, Janine Higgins, and Edward Sazonov. 2025. A Systematic Review of Sensor-Based Methods for Measurement of Eating Behavior. *Sensors* 25, 10 (2025), 2966.

[38] Changshuo Hu, Thivya Kandappu, Jake Stuchbury-Wass, Yang Liu, Anthony Tang, Cecilia Mascolo, and Dong Ma. 2024. Detecting Foot Strikes during Running with Earbuds. In *Proceedings of the Workshop on Body-Centric Computing Systems* (Minato-ku, Tokyo, Japan) *(BodySys '24)*. Association for Computing Machinery, New York, NY, USA, 35–40. doi:10.1145/3662009.3662023

[39] Emil Jovanov, Edward Sazonov, and Carmen Poon. 2014. Sensors and systems for obesity care and research. In *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 3188–3191.

[40] Azusa Kadomura, Cheng-Yuan Li, Koji Tsukada, Hao-Hua Chu, and Itiro Siio. 2014. Persuasive technology to improve eating behavior using a sensor-embedded fork. In *Proceedings of the 2014 acm international joint conference on pervasive and ubiquitous computing*. 319–329.

[41] Daniel Kahneman. 2011. *Thinking, fast and slow*. macmillan.

[42] Haruka Kamachi, Takumi Kondo, Tahera Hossain, Anna Yokokubo, and Guillaume Lopez. 2021. Automatic segmentation method of bone conduction sound for eating activity detailed detection. In *Adjunct Proceedings of the 2021 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2021 ACM International Symposium on Wearable Computers*. 310–315.

[43] Maurits Kaptein, Boris De Ruyter, Panos Markopoulos, and Emile Aarts. 2012. Adaptive persuasive systems: a study of tailored persuasive text messages to reduce snacking. *ACM Transactions on Interactive Intelligent Systems (TiiS)* 2, 2 (2012), 1–25.

[44] Alan E Kazdin. 1982. Observer effects: Reactivity of direct observation. *New Directions for Methodology of Social & Behavioral Science* (1982).

[45] Rohit Ashok Khot, Jung-Ying Yi, and Deepti Aggarwal. 2020. SWAN: Designing a companion spoon for mindful eating. In *Proceedings of the fourteenth international conference on tangible, embedded, and embodied interaction*. 743–756.

[46] Heewon Kim, Meara Faw, and Andreas Michaelides. 2017. Mobile but connected: harnessing the power of self-efficacy and group support for weight loss success through mHealth intervention. *Journal of Health Communication* 22, 5 (2017), 395–402.

[47] Joohee Kim and Byung-Chull Bae. 2018. An Animated Emoji Feedback System for Eating Rate Guidance. In *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers*. 388–391.

[48] Joohee Kim, Kwang-Jae Lee, Mankyung Lee, Nahyeon Lee, Byung-Chull Bae, Genehee Lee, Juhee Cho, Young Mog Shim, and Jun-Dong Cho. 2016. Slowee: A smart eating-speed guide system with light and vibration feedback. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. 2563–2569.

[49] Joohee Kim, Mankyung Lee, Kwang-Jae Lee, TaeYang Lee, Byung-Chull Bae, and Jun-Dong Cho. 2016. An eating speed guide system using a wristband and tabletop unit. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*. 121–124.

[50] HARRY R Kissileff, GARY Klingsberg, and THEODORE B Van Itallie. 1980. Universal eating monitor for continuous recording of solid or liquid consumption in man. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology* 238, 1 (1980), R14–R22.
[51] MV Pradeep Kumar, VT Venkatesh, and SR Jagannatha. 2008. Fast eating syndrome. *Medicine, Science and the Law* 48, 1 (2008), 78–81.
[52] Dominika Kwasnicka, Stephan U Dombrowski, Martin White, and Falko Sniehotta. 2016. Theoretical explanations for maintenance of behaviour change: a systematic review of behaviour theories. *Health psychology review* 10, 3 (2016), 277–296.
[53] J Richard Landis and Gary G Koch. 1977. The measurement of observer agreement for categorical data. *biometrics* (1977), 159–174.
[54] Matthew L Lee and Anind K Dey. 2014. Real-time feedback for improving medication taking. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2259–2268.
[55] Sunjae Lee, Minwoo Jeong, Daye Song, Junyoung Choi, Seoyun Son, Jean Y Song, and Insik Shin. 2024. FLUID-IoT: Flexible and Granular Access Control in Shared IoT Environments via-UI-Level Control Distribution. In *2024 CHI Conference on Human Factors in Computing Sytems, CHI 2024*. Association for Computing Machinery, 848.
[56] Ian Li, Anind Dey, and Jodi Forlizzi. 2010. A stage-based model of personal informatics systems. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 557–566.
[57] Kong-Ling Li, Ji-Hong Chen, Qian Zhang, Jan D Huizinga, Shawn Vadakepeedika, Yu-Rong Zhao, Wen-Zhen Yu, and He-Sheng Luo. 2013. Habitual rapid food intake and ineffective esophageal motility. *World Journal of Gastroenterology: WJG* 19, 14 (2013), 2270.
[58] Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollár. 2017. Focal loss for dense object detection. In *Proceedings of the IEEE international conference on computer vision*. 2980–2988.
[59] Siegwart Lindenberg and Linda Steg. 2007. Normative, gain and hedonic goal frames guiding environmental behavior. *Journal of Social issues* 63, 1 (2007), 117–137.
[60] Edwin A Locke. 1968. Toward a theory of task motivation and incentives. *Organizational behavior and human performance* 3, 2 (1968), 157–189.
[61] Roya Lotfi, George Tzanetakis, Rasit Eskicioglu, and Pourang Irani. 2020. A comparison between audio and IMU data to detect chewing events based on an earable device. In *Proceedings of the 11th Augmented Human International Conference*. 1–8.
[62] James P Lund and Arlette Kolta. 2006. Generation of the central masticatory pattern and its modification by sensory feedback. *Dysphagia* 21, 3 (2006), 167–174.
[63] Michail Mantzios. 2025. "Eating isn't just about paying attention—It's about the self-regulation of sensory attention while eating!": Exploring mindful eating by examining sensory attention and non-judgmental awareness in the context of eating cessation. *Nutrition and Health* 31, 3 (2025), 1237–1242.
[64] Flávia Luciana Pinheiro de Souza Pinto Martins, Millena Borges Inete, Yasmym Dannielle do Espírito Santo Souza, Rafaela Lorena Viana Costa, Rafaelle Dias Gabbay, Tainá Martins Moraes, Vanessa Vieira Lourenço Costa, Carla Cristina Paiva Paracampo, Luiz Carlos de Albuquerque, and Daniela Lopes Gomes. 2024. Association between Self-Perception of Chewing, Chewing Behavior, and the Presence of Gastrointestinal Symptoms in Candidates for Bariatric Surgery. *Nutrients* 16, 8 (2024), 1096.
[65] Max Planck Institute for Psycholinguistics. 2025. ELAN (Version 7.0) [Computer software]. https://archive.mpi.nl/tla/elan
[66] Kathleen J Melanson, Carolyn N Matsumoto, and Geoffrey W Greene. 2023. Eating pace instruction is effective in slowing eating rate in women with overweight and obesity. *Eating Behaviors* 48 (2023), 101701.
[67] AE Mesas, M Muñoz-Pareja, E López-García, and F Rodríguez-Artalejo. 2012. Selected eating behaviours and excess body weight: a systematic review. *Obesity Reviews* 13, 2 (2012), 106–135.
[68] Susan Michie, Marie Johnston, Jill Francis, Wendy Hardeman, and Martin Eccles. 2008. From theory to intervention: mapping theoretically derived behavioural determinants to behaviour change techniques. *Applied psychology* 57, 4 (2008), 660–680.
[69] J Mordal, Ø Gundersen, and JG Bramness. 2010. Norwegian version of the Mini-International Neuropsychiatric Interview: feasibility, acceptability and test-retest reliability in an acute psychiatric ward. *European Psychiatry* 25, 3 (2010), 172–177.
[70] Philippe Morquette, Raphaël Lavoie, Mitch-David Fhima, Xavier Lamoureux, Dorly Verdier, and Arlette Kolta. 2012. Generation of the masticatory central pattern and its modulation by sensory feedback. *Progress in neurobiology* 96, 3 (2012), 340–355.
[71] Robert R Morris, Kareem Kouddous, Rohan Kshirsagar, and Stephen M Schueller. 2018. Towards an artificially empathic conversational agent for mental health applications: system design and user perceptions. *Journal of medical Internet research* 20, 6 (2018), e10148.
[72] Inbal Nahum-Shani, Shawna N Smith, Bonnie J Spring, Linda M Collins, Katie Witkiewitz, Ambuj Tewari, and Susan A Murphy. 2016. Just-in-time adaptive interventions (JITAIs) in mobile health: key components and design principles for ongoing health behavior support. *Annals of behavioral medicine* (2016), 1–17.
[73] Yugo Nakamura, Rei Nakaoka, Yuki Matsuda, and Keiichi Yasumoto. 2023. Eat2pic: an eating-painting interactive system to nudge users into making healthier diet choices. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 7, 1 (2023), 1–23.
[74] Theresa A Nicklas, Tom Baranowski, Karen W Cullen, and Gerald Berenson. 2001. Eating patterns, dietary quality and obesity. *Journal of the American college of nutrition* 20, 6 (2001), 599–608.

[75] T Ohkuma, Y Hirakawa, Udai Nakamura, Y Kiyohara, T Kitazono, and T Ninomiya. 2015. Association between eating rate and obesity: a systematic review and meta-analysis. *International journal of obesity* 39, 11 (2015), 1589–1596.

[76] Rei Otsuka, Koji Tamakoshi, Hiroshi Yatsuya, Chiyoe Murata, Atsushi Sekiya, Keiko Wada, Hui Ming Zhang, Kunihiro Matsushita, Kaichiro Sugiura, Seiko Takefuji, et al. 2006. Eating fast leads to obesity: findings based on self-administered questionnaires among middle-aged Japanese men and women. *Journal of epidemiology* 16, 3 (2006), 117–124.

[77] Cliodhna O'Connor and Helene Joffe. 2020. Intercoder reliability in qualitative research: Debates and practical guidelines. *International journal of qualitative methods* 19 (2020), 1609406919899220.

[78] Elena Patra, Anna Kokkinopoulou, Saskia Wilson-Barnes, Kathryn Hart, Lazaros P Gymnopoulos, Dorothea Tsatsou, Vassilios Solachidis, Kosmas Dimitropoulos, Konstantinos Rouskas, Anagnostis Argiriou, et al. 2024. Personal goals, user engagement, and meal adherence within a personalised AI-based mobile application for nutrition and physical activity. *Life* 14, 10 (2024), 1238.

[79] Charlie Pinder, Jo Vermeulen, Benjamin R Cowan, and Russell Beale. 2018. Digital behaviour change interventions to break and form habits. *ACM Transactions on Computer-Human Interaction (TOCHI)* 25, 3 (2018), 1–66.

[80] Mashfiqui Rabbi, Min Hane Aung, Mi Zhang, and Tanzeem Choudhury. 2015. MyBehavior: automatic personalized health feedback from user behaviors and preferences using smartphones. In *Proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing*. 707–718.

[81] Lina Radzevičienė and Rytas Ostrauskas. 2013. Fast eating and the risk of type 2 diabetes mellitus: a case-control study. *Clinical nutrition* 32, 2 (2013), 232–235.

[82] Christof Rapp. 2022. Aristotle's rhetoric. (2022).

[83] NW Read, I McL Welch, CJ Austen, C Barnish, CE Bartlett, AJ Baxter, G Brown, ME Comption, KE Hume, I Storie, et al. 1986. Swallowing food without chewing; a simple way to reduce postprandial glycaemia. *British journal of nutrition* 55, 1 (1986), 43–47.

[84] Eric Robinson, Paul Aveyard, Amanda Daley, Kate Jolly, Amanda Lewis, Deborah Lycett, and Suzanne Higgs. 2013. Eating attentively: a systematic review and meta-analysis of the effect of food intake memory and awareness on eating. *The American journal of clinical nutrition* 97, 4 (2013), 728–742.

[85] Nicole L Robinson, Jennifer Connolly, Gavin Suddery, Madeliene Turner, and David J Kavanagh. 2021. A humanoid social robot to provide personalized feedback for health promotion in diet, physical activity, alcohol and cigarette use: a health clinic trial. In *2021 30th IEEE international conference on robot & human interactive communication (RO-MAN)*. IEEE, 720–726.

[86] Tobias Röddiger, Christopher Clarke, Paula Breitling, Tim Schneegans, Haibin Zhao, Hans Gellersen, and Michael Beigl. 2022. Sensing with earables: A systematic literature review and taxonomy of phenomena. *Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies* 6, 3 (2022), 1–57.

[87] Rianne MAJ Ruijschop, Nicolien Zijlstra, Alexandra EM Boelrijk, Annereinou Dijkstra, Maurits JM Burgering, Cees de Graaf, and Margriet S Westerterp-Plantenga. 2011. Effects of bite size and duration of oral processing on retro-nasal aroma release–features contributing to meal termination. *British Journal of Nutrition* 105, 2 (2011), 307–315.

[88] Alfonso Sánchez-Ayala, Arcelino Farias-Neto, Nara Hellen Campanha, and Renata Cunha Matheus Rodrigues Garcia. 2013. Relationship between chewing rate and masticatory performance. *CRANIO®* 31, 2 (2013), 118–122.

[89] Hillol Sarker, Moushumi Sharmin, Amin Ahsan Ali, Md Mahbubur Rahman, Rummana Bari, Syed Monowar Hossain, and Santosh Kumar. 2014. Assessing the availability of users to engage in just-in-time intervention in the natural environment. In *Proceedings of the 2014 ACM international joint conference on pervasive and ubiquitous computing*. 909–920.

[90] Jenna L Scisco, Eric R Muth, Yujie Dong, and Adam W Hoover. 2011. Slowing bite-rate reduces energy intake: an application of the bite counter device. *Journal of the American Dietetic Association* 111, 8 (2011), 1231–1235.

[91] Jaemin Shin, Seungjoo Lee, Taesik Gong, Hyungjun Yoon, Hyunchul Roh, Andrea Bianchi, and Sung-Ju Lee. 2022. Mydj: Sensing food intakes with an attachable on your eyeglass frame. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–17.

[92] Herbert W Simons. 1976. Persuasion: Understanding, practice, and analysis. *(No Title)* (1976).

[93] Hendrik Jan Smit, E Katherine Kemsley, Henri S Tapp, and C Jeya K Henry. 2011. Does prolonged chewing reduce food intake? Fletcherism revisited. *Appetite* 57, 1 (2011), 295–298.

[94] Kazuo Sonoki, Masanori Iwase, Yutaka Takata, Tetsuji Nakamoto, Chihiro Masaki, Ryuji Hosokawa, Shigeki Murakami, Kaoru Chiwata, and Hiromasa Inoue. 2013. Effects of thirty-times chewing per bite on secretion of glucagon-like peptide-1 in healthy volunteers and type 2 diabetic patients. *Endocrine journal* 60, 3 (2013), 311–319.

[95] TA Spiegel. 2000. Rate of intake, bites, and chews—the interpretation of lean–obese differences. *Neuroscience & Biobehavioral Reviews* 24, 2 (2000), 229–237.

[96] Dennis Stanke, Tim Duente, Kerem Can Demir, and Michael Rohs. 2023. Can You Ear Me? A Comparison of Different Private and Public Notification Channels for the Earlobe. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 7, 3 (2023), 1–23.

[97] Zhida Sun, Sitong Wang, Wenjie Yang, Onur Yürüten, Chuhan Shi, and Xiaojuan Ma. 2020. " A Postcard from Your Food Journey in the Past" Promoting Self-Reflection on Social Food Posting. In *Proceedings of the 2020 ACM Designing Interactive Systems Conference*. 1819–1832.

[98] Lie Ming Tang, Judy Kay, Farahnaz Yekeh, and Bob Kummerfeld. 2015. Introducing SAL, simple, situated, ambient loggers, for personal goals. In *Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers*. 997–1002.

[99] Amos Tversky and Daniel Kahneman. 1981. The framing of decisions and the psychology of choice. *science* 211, 4481 (1981), 453–458.

[100] Thomas L Webb, Falko F Sniehotta, and Susan Michie. 2010. Using theories of behaviour change to inform interventions for addictive behaviours. *Addiction* 105, 11 (2010), 1879–1892.

[101] Laura L Wilkinson, Danielle Ferriday, Matthew L Bosworth, Nicolas Godinot, Nathalie Martin, Peter J Rogers, and Jeffrey M Brunstrom. 2016. Keeping pace with your eating: Visual feedback affects eating rate in humans. *PLoS One* 11, 2 (2016), e0147603.

[102] Eilis Woodward, Jillian Haszard, Anna Worsfold, and Bernard Venn. 2020. Comparison of self-reported speed of eating with an objective measure of eating rate. *Nutrients* 12, 3 (2020), 599.

[103] Ruolan Wu, Chun Yu, Xiaole Pan, Yujia Liu, Ningning Zhang, Yue Fu, Yuhan Wang, Zhi Zheng, Li Chen, Qiaolei Jiang, et al. 2024. MindShift: leveraging large language models for mental-states-based problematic smartphone use intervention. In *Proceedings of the 2024 CHI conference on human factors in computing systems*. 1–24.

[104] Yahan Xie, Hanhui Yang, Jing Huang, Jing Wang, Hanxuan Li, and Yijie Guo. 2023. CHIBO: A Pneumatic Robot that Provides Real-time Intervention for Parent-child Feeding Behavior. In *Companion of the 2023 ACM/IEEE International Conference on Human-Robot Interaction*. 725–729.

[105] Xuhai Xu, Haitian Shi, Xin Yi, WenJia Liu, Yukang Yan, Yuanchun Shi, Alex Mariakakis, Jennifer Mankoff, and Anind K Dey. 2020. Earbuddy: Enabling on-face interaction via wireless earbuds. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–14.

[106] Xuhai Xu, Tianyuan Zou, Han Xiao, Yanzhang Li, Ruolin Wang, Tianyi Yuan, Yuntao Wang, Yuanchun Shi, Jennifer Mankoff, and Anind K Dey. 2022. TypeOut: leveraging just-in-time self-affirmation for smartphone overuse reduction. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–17.

[107] Takayuki Yamaji, Shinsuke Mikami, Hiroshi Kobatake, Koichi Tanaka, Yukihito Higashi, and Yasuki Kihara. 2017. Slow down, you eat too fast: fast eating associate with obesity and future prevalence of metabolic syndrome. *Circulation* 136, suppl_1 (2017), A20249–A20249.

[108] Bo Zhang, Kaiwen Deng, Jie Shen, Lingrui Cai, Bohdana Ratitch, Haoda Fu, and Yuanfang Guan. 2022. Enabling eating detection in a free-living environment: Integrative engineering and machine learning study. *Journal of medical Internet research* 24, 3 (2022), e27934.

[109] Hao Zhang, Guillaume Lopez, Masaki Shuzo, Jean-Jacques Delaunay, and Ichiro Yamada. 2012. Mastication counting method robust to food type and individual. In *International Conference on Health Informatics*, Vol. 2. SCITEPRESS, 374–377.

[110] Ruidong Zhang, Jihai Zhang, Nitish Gade, Peng Cao, Seyun Kim, Junchi Yan, and Cheng Zhang. 2022. EatingTrak: Detecting fine-grained eating moments in the wild using a wrist-mounted IMU. *Proceedings of the ACM on Human-Computer Interaction* 6, MHCI (2022), 1–22.

[111] Xiyuxing Zhang, Yuntao Wang, Yuxuan Han, Chen Liang, Ishan Chatterjee, Jiankai Tang, Xin Yi, Shwetak Patel, and Yuanchun Shi. 2024. The EarSAVAS Dataset: Enabling Subject-Aware Vocal Activity Sensing on Earables. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 8, 2, Article 83 (May 2024), 26 pages. doi:10.1145/3659616

[112] Zuoyi Zhang. 2020. *Exploring a Novel Design to Support Healthy Eating*. Master's thesis. University of Manitoba. http://hdl.handle.net/1993/34959

[113] Yong Zhu, Walter H Hsu, and James H Hollis. 2013. Increasing the number of masticatory cycles is associated with reduced appetite and altered postprandial plasma concentrations of gut hormones, insulin and glucose. *British Journal of Nutrition* 110, 2 (2013), 384–390.

[114] Nicolien Zijlstra, Rene De Wijk, Monica Mars, Annette Stafleu, and Cees de Graaf. 2009. Effect of bite size and oral processing time of a semisolid food on satiation. *The American journal of clinical nutrition* 90, 2 (2009), 269–275.

## A STATISTICAL DETAILS OF THE DATASET FOR EVALUATING REAL-TIME EATING PACE SENSING METHODS

We summarized the eating acoustic events for each participant in both quiet (lab conference room) and noisy (cafeteria) scenarios, as shown in Table 4, to evaluate our proposed real-time eating pace sensing method.

Table 4. Per-participant Acoustic Events for Evaluating Real-time Eating Pace Sensing

| User ID | Environment | Duration (s) | Chewing Counts | Swallowing Counts |
|---|---|---|---|---|
| D1 | Quiet | 1296.53 | 2514 | 109 |
| | Noisy | 1410.59 | 2392 | 112 |
| D2 | Quiet | 1620.08 | 1743 | 72 |
| | Noisy | 641.52 | 888 | 36 |
| D3 | Quiet | 829.65 | 997 | 51 |
| | Noisy | 465.15 | 516 | 45 |
| D4 | Quiet | 1128.75 | 879 | 48 |
| | Noisy | 1018.62 | 983 | 48 |
| D5 | Quiet | 1611.70 | 1911 | 69 |
| | Noisy | 903.26 | 1250 | 48 |
| D6 | Quiet | 952.64 | 1167 | 76 |
| | Noisy | 1028.56 | 1201 | 89 |
| D7 | Quiet | 742.69 | 777 | 34 |
| | Noisy | 583.13 | 757 | 35 |
| D8 | Quiet | 1174.82 | 1305 | 97 |
| | Noisy | 649.35 | 807 | 58 |
| D9 | Quiet | 1688.16 | 2399 | 70 |
| | Noisy | 1315.94 | 2037 | 84 |
| D10 | Quiet | 1696.93 | 2257 | 76 |
| | Noisy | 1618.92 | 2266 | 89 |
| D11 | Quiet | 1136.75 | 1087 | 58 |
| | Noisy | 519.95 | 683 | 39 |
| D12 | Quiet | 1304.85 | 1371 | 63 |
| | Noisy | 1185.19 | 998 | 52 |
| D13 | Quiet | 971.70 | 874 | 63 |
| | Noisy | 1167.12 | 942 | 76 |
| D14 | Quiet | 798.33 | 1034 | 43 |
| | Noisy | 316.04 | 484 | 17 |
| D15 | Quiet | 817.94 | 1092 | 62 |
| | Noisy | 773.53 | 949 | 58 |
| D16 | Quiet | 442.34 | 325 | 18 |
| | Noisy | 370.64 | 308 | 9 |
| D17 | Quiet | 1022.70 | 1132 | 47 |
| | Noisy | 1198.47 | 1236 | 50 |
| D18 | Quiet | 1105.16 | 1657 | 58 |
| | Noisy | 1750.69 | 2364 | 91 |

## B NON-INGESTIVE ACTIVITY STRESS TEST

To provide task-level detail for the practical reliability discussion in Section 7.4, we conducted an additional non-ingestive negative-control stress test. We recruited 10 participants (6 female, 4 male; ages 21–31, $M = 24.60$, $SD = 3.34$). Participants wore the sensing earbud and completed continuous no-eating blocks covering walking-related activities, speech and social activities, head motion and phone checking, and audio playback. Participants were instructed not to eat, chew gum, or intentionally chew during these sessions. Therefore, the entire recording served as non-chewing ground truth, and every detected chewing event was counted as a false positive.

We applied the trained chewing-detection pipeline without retraining or adding these activities as negative examples. For each task, we report false chewing detections per minute (FP/min), computed as the number of detected chewing events divided by the valid recording duration. We also report the median and interquartile range (IQR) of session-level FP/min values, together with the number of sessions with zero false positives. Category-level statistics aggregate session-level FP/min values across tasks within the same activity category.

Table 5. False chewing detections under non-ingestive confound stress tests. All detected chewing events were counted as false positives. Category-level statistics aggregate the session-level FP/min values within each category.

| Category | Total min | FP/min, median [IQR] | Category zero-FP | Task | Sessions | False chewing detections per min | FP/min, median [IQR] | Task zero-FP |
|---|---|---|---|---|---|---|---|---|
| Walking | 56.12 | 0.63 [0.00, 1.93] | 10/30 | Normal walking | 10 | 0.196 | 0.00 [0.00, 0.36] | 7/10 |
| | | | | Light jogging | 10 | 3.222 | 0.73 [0.12, 3.49] | 3/10 |
| | | | | Stair walking | 10 | 3.686 | 1.62 [0.70, 4.23] | 0/10 |
| Speech/social | 117.38 | 0.35 [0.00, 2.91] | 12/30 | Loud reading | 10 | 3.851 | 2.91 [0.45, 6.40] | 2/10 |
| | | | | Listening to another speaker | 10 | 0.000 | 0.00 [0.00, 0.00] | 10/10 |
| | | | | Dyadic conversation | 10 | 2.474 | 1.34 [0.34, 3.47] | 0/10 |
| Head motion | 21.35 | 0.00 [0.00, 0.00] | 16/20 | Continuous head motion | 10 | 1.544 | 0.00 [0.00, 1.73] | 6/10 |
| | | | | Phone checking/reading | 10 | 0.000 | 0.00 [0.00, 0.00] | 10/10 |
| Playback | 15.72 | 0.00 [0.00, 0.00] | 10/10 | Audio playback | 10 | 0.000 | 0.00 [0.00, 0.00] | 10/10 |

## C GENERATION PROMPT (GPT-4O)

This appendix documents the generation prompt used with GPT-4o.

> I currently have three theory-based intervention content designs: (1) "A bit fast" based on Dual System Theory (System 1); (2) "You are xxx chews away from reaching the recommended 25 chews" based on Control Theory; and (3) "Eating more slowly and chewing more can improve digestion and increase satiety, helping you reach your weight-loss goal sooner" based on Goal Framing Theory using a gain frame. Please help me generate sentences for these three types, and for each type include three sentence lengths: a short sentence (within 10 Chinese characters), a medium sentence (10-20 Chinese characters), and a long sentence (more than 20 Chinese characters). Thank you.

## D SELF-REPORTED EATING EXPERIENCE QUESTIONNAIRE AND FACTOR MAPPING

All the questionnaire questions were asked using a 7-point Likert scale (1 = strongly disagree, 7 = strongly agree). *Behavioral Positive*, *Appetite*, *Fullness*, and *Comfort* represent four theoretically derived factors. Q6 was reverse-coded before reliability and significance analysis.

Table 6. Mapping of questionnaire items (Q1-Q6) to factors in the Self-Reported Eating Experience scale.

| ID | Item | Factor |
|---|---|---|
| Q1 | I felt that my eating pace became slower. | Behavioral Positive |
| Q2 | I felt that I chewed more times per bite. | Behavioral Positive |
| Q3 | I felt that my portion size became smaller. | Appetite |
| Q4 | I felt that I became full more easily. | Fullness |
| Q5 | I felt that the intervention had a positive impact on my eating process. | Behavioral Positive |
| Q6 | When wearing earbuds during eating, I felt more uncomfortable than usual. | Comfort |

## E SEMI-STRUCTURED INTERVIEW GUIDE

Interviews were semi-structured. The interviewer used the questions below as a guide and asked follow-up questions as needed. (Interviews were conducted in Chinese and translated to English for reporting.)

### E.1 Baseline Interview (Day 1)

(1) **Typical eating routine.** Could you walk me through a typical lunch/dinner in your daily life (where, with whom, and how long it usually takes)?
(2) **Perceived fast-eating episodes.** In the past week, how often did you feel you ate too fast? What situations triggered it?
(3) **Drivers of fast eating.** What do you think are the main reasons you tend to eat fast?
(4) **Consequences.** When you eat fast, what changes do you notice (e.g., eating more than intended, satiety, discomfort)?
(5) **Prior attempts to slow down.** Have you tried to intentionally slow down your eating pace before? If so, what strategies did you use, and did they work?
(6) **Acceptable interventions.** If a system provides reminders/feedback during meals, what forms of prompting would you find acceptable?
(7) **When not to interrupt.** In what situations would you *not* want to receive any intervention during eating?

### E.2 Post-study Interview (Day 13)

(1) **Contextual confounds.** During the study, were there any unusual events or constraints that affected your eating behavior?
(2) **Overall perceived change.** Looking back across the study, do you feel your eating pace changed?
(3) **Comparison across conditions.** Please compare your experiences across the *baseline*, *control*, and *experimental (closed-loop)* phases.
(4) **Retention days.** On the retention days (Days 7 and 13) when the system delivered no audio prompts or feedback, how did your eating behavior compare to the days immediately before?

(5) **Control condition experience.** In the control condition , did you notice any changes in your eating pace, attention, or intake?
(6) **Closed-loop intervention experience.** In the experimental condition, with in-meal JIT interventions and post-meal summary report, did you notice any changes in your eating pace, attention, or intake?
(7) **Perceived effects of JIT prompts.** Which in-meal JIT audio prompts (or types of prompts) had the most noticeable positive effect on you?
(8) **Negative effects and burden.** Did you experience any negative reactions or unintended effects from the interventions?
(9) **Design suggestions.** Do you have any suggestion for the system and intervention design?

## F PARTICIPANT-LEVEL EATING PACE TRAJECTORIES

This appendix visualizes participant-level eating-pace trajectories derived from Earinter on-device inference logs to make individual variability explicit. For consistency and to avoid selection bias, we fixed the same reference meal in the second day of the experiment condition. Also, we included all participants with valid data, regardless of meal choice or other factors reported in semi-interview sessions. The x-axis indexes time-ordered segments within a single meal, and dashed lines indicate the within-meal median chews-per-swallow (CPS). As expected in in-the-wild settings, trajectories vary substantially: some participants show more stable or higher CPS over the meal, while others appear noisier, often when foods required minimal chewing or due to habitual chewing styles. These panels are descriptive mainly for the logic and performance of Earinter using in wild contexts.

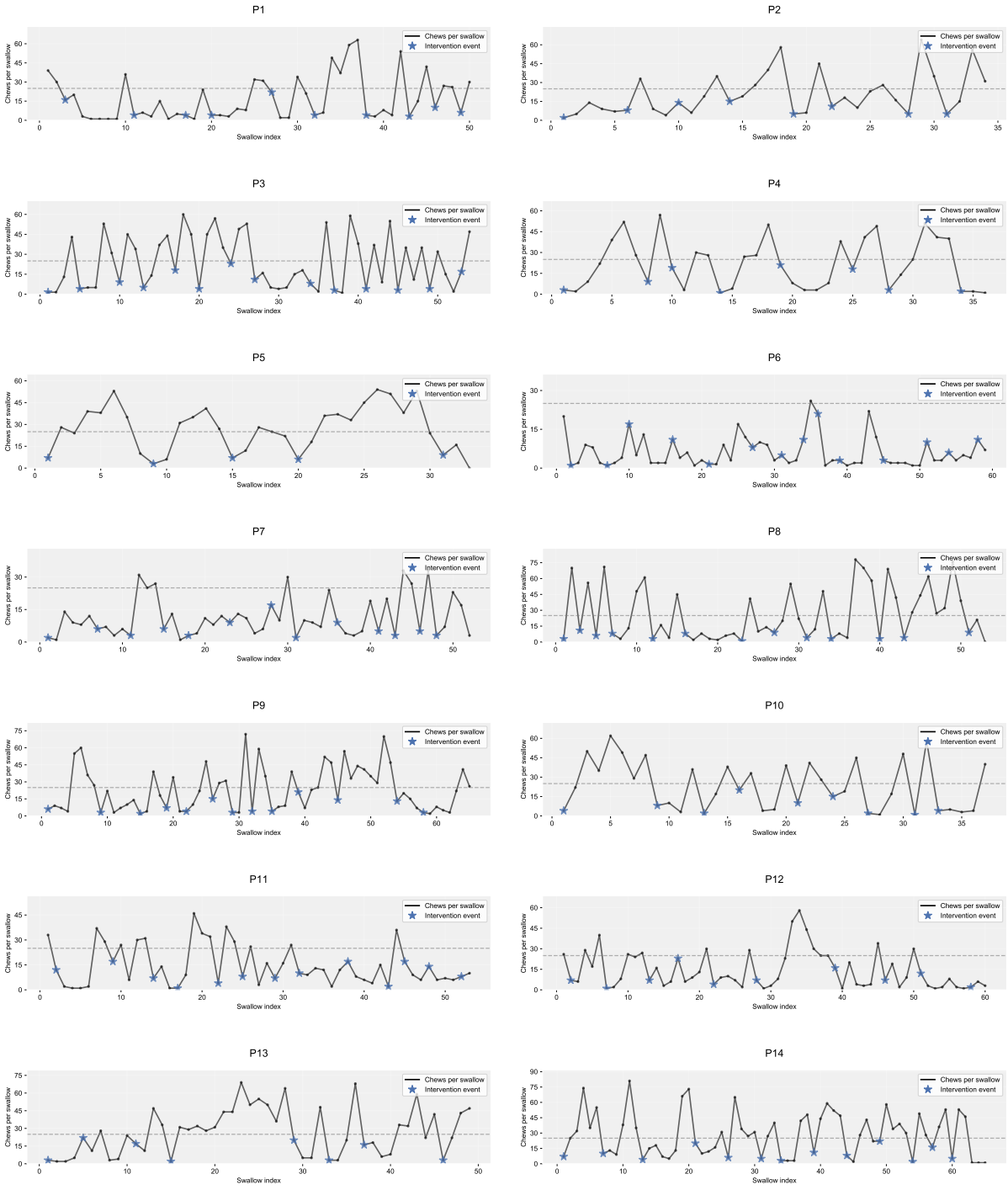



Fig. 10. Participant-level trajectories of eating pace (chews-per-swallow) during a representative Experiment meal.