# Rationale and Design of the SenseWhy Project: A Passive Sensing and Ecological Momentary Assessment Study on Characteristics of Overeating Episodes

Nabil I Alshurafa[1], Ph.D.

Tammy K Stump[2], Ph.D.

Christopher S Romano[1], B.S.

Angela Fidler Pfammatter[1], Ph.D.

Annie W. Lin[3], Ph.D.

Josiah Hester[4,5], Ph.D.

Evan Forman[6], Ph.D.

Bonnie Spring[1], Ph.D.

[1] Department of Preventive Medicine, Feinberg School of Medicine, Northwestern University, 680 N Lake Shore Drive, Chicago, IL, USA

[2] Department of Dermatology, University of Utah, 2000 Circle of Hope Drive, Salt Lake City, UT, USA

[3] Department of Nutrition, Benedictine University, 5700 College Road, Lisle, IL, USA

[4] Department of Electrical and Computer Engineering, McCormick School of Engineering and Applied Science, Northwestern University, 2145 Sheridan Road, Evanston, IL, USA

[5] Department of Computer Science, Northwestern University, 2233 Tech Drive, Third Floor, Evanston, IL, USA

[6] Department of Psychological and Brain Sciences, Drexel University, 3201 Chestnut Street, Philadelphia, PA, USA

## Abstract

**Objectives**: Overeating interventions and research often focus on single determinants and use subjective or non-personalized measures. We aim to (1) identify automatically detectable features that predict overeating and (2) build clusters of eating episodes that identify theoretically meaningful and clinically known problematic overeating behaviors (e.g., stress eating) as well as new phenotypes based on social and psychological features.

**Method:** Up to 60 adults with obesity in the Chicagoland area will be recruited for a 14-day free-living observational study. Participants will complete ecological momentary assessments (EMAs) and wear three sensors designed to capture features of overeating episodes (e.g., chews) that can be visually confirmed. Participants will also complete daily dietitian-administered 24-hour recalls of all food and beverages consumed.

**Analysis:** Overeating is defined as caloric consumption exceeding 1 standard deviation of an individual's mean consumption per eating episode. To identify features that predict overeating, we will apply two complementary machine learning methods: Correlation-based Feature Selection and Wrapper-based Feature Selection. We will then generate clusters of overeating types and assess how they align with clinically meaningful overeating phenotypes.

**Conclusions:** This study is the first to assess characteristics of eating episodes in situ over a multiweek period with visual confirmation of eating behaviors. An additional strength of this study is the assessment of predictors of problematic eating during periods when individuals are not on a structured diet and/or engaged in a weight loss intervention. Our assessment of overeating episodes in real-world settings is likely to yield new insights regarding determinants of overeating that may translate into novel interventions.

**Keywords**

Wearable sensors, overeating, machine learning, behavioral phenotypes, detection, prediction.

## Introduction

More than a third (42.4%) of US adults have obesity,[1] which can contribute to increased risk of chronic diseases and their associated healthcare costs.[2] Frequent overeating (i.e., consistent intake of excess kilocalories) relative to need is a risk factor for obesity.[3] Behavioral studies on overeating often focus on a single putative causal mechanism or proximal determinant (e.g., stress, emotion, environment).[4, 5] However, recent advances in wearable sensing have enabled researchers to identify more complex feature patterns that characterize episodes of excess calorie intake.[6-11] This approach increases researchers' ability to detect excess energy intake with respect to various behavioral phenotypes (e.g., emotional eating, mindless eating) that will have different treatment implications.[12, 13] Through the study design described in this paper, we will apply a new objective sensing system to characterize overeating phenotypes and identify in-the-moment predictors of overeating. We lay groundwork to detect, predict, and intervene in a manner that adapts to an individual's problematic overeating profile, thus paving the way toward personalized behavioral medicine interventions.

Previous sensing systems advanced the field of overeating in that they were designed to capture proxies to eating episodes such as feeding gestures,[14] chewing activity,[9] swallows.[10, 15] These activities were captured using varying sensing modalities ranging from the use of sound waves around the ear,[11, 16] visual cues of the behavior, all the way to the movement of the hand and vibrations at the temporalis muscle.[11, 17-19] Yet these devices: 1) were only tested in a controlled lab setting which limits external validity;[20] 2) required heavy participant burden (e.g., participants taped sensors on their bodies[21] or had to turn the system on and off during eating[22]); and 3) did not capture meaningful and interpretable features about the physical and psychological contexts of eating episodes.[23] To address these challenges, we use the SenseWhy passive eating detection system to gain new insight into overeating episodes. SenseWhy comprises a well-tested customizable neck-worn sensor (NeckSense),[10, 17, 24, 25] location through GPS, Ecological Momentary Assessments (EMAs), and an advanced privacy-sensitive video camera with infrared sensing to enable heat signatures (chest-mounted to collect visual confirmation of diet, behavior, and the environment). Our current observational study will apply this system to all eating episodes in a free-living setting, in hopes of identifying overeating episodes within the context of an individual's daily life.

A rapidly growing body of research has begun to characterize predictors of problematic eating behaviors, including overeating episodes, using EMA. In these studies, self-reported affect,[26-28] environmental triggers (e.g., social cues, presence of palatable foods),[27] and cravings/hunger have been found to be associated with dietary lapses,[27, 29] eating unhealthy foods,[28] and/or overeating.[26] By virtue of its collection of real-time contextual data surrounding overeating episodes, the SenseWhy system will yield new insights about the predictors and characteristics of overeating episodes. The present study will extend previous research by evaluating a wide range of passively detected features of eating episodes and EMA reports of eating behavior, specifically in the contextual factors of overeating with respect to obesity status.

We will also use the SenseWhy system to detect problematic eating phenotypes that are identified by causal mechanisms or proximal determinants. Many people engage in *emotional eating* (or stress eat) as a means of dealing with psychological stress, as opposed to satisfying hunger. *Eating in the absence of hunger (EAH*) refers to the susceptibility to eat despite being satiated, often in response to the presence of palatable snack foods; which is associated with weight gain over two months.[30] EAH has been shown to be highly related to *hedonic eating*,[31] which refers to one's desire to consume food for pleasure, in the absence of caloric need. *Cognitive restraint* refers to conscious restriction of food intake in order to control body weight or to promote weight loss.[32] and *uncontrolled eating* refers to the tendency to eat more than the usual because of loss of control. *Unplanned (over)eating* refers to an individual eating more than what was originally intended, and can be due to many reasons mentioned above including positive social interactions, negative emotions, or physiological craving.[33] *Mindless eating* often occurs when the mind is distracted and the person is not aware of what or how much food he or she is consuming. Mindless eating is associated with environmental factors, such as screen time and presence of food advertisements, and internal factors, such as disinhibition (due to boredom), lack of awareness (the mind is racing about what one needs to do), and emotional eating.[34] *Night eating* is a condition where people eat large amounts of food after dinner, often waking up during the night to eat. As described above, several of these problematic eating types overlap. However, it remains unknown how much they overlap, whether they share common causes, and which phenotypes are automatically detectable. Through the SenseWhy system, we will be able to assess characteristics of eating episodes that are indicative of these phenotypes as well as predictors of these various types of eating episodes. For instance, perhaps mindless eating is also accompanied by high amounts of stress. Redefining overeating behavioral phenotypes in a way that is automatically detectable will pave the way for developing and testing timely and adaptive interventions.

Within a 14-day observational study of eating episodes of adults with obesity, the primary aim of this project is to identify a subset of automatically detectable features that predict overeating episodes. EMA reports of affect, stress, hunger, and contextual factors, along with passively detected features (e.g., time, feeding gestures) will be analyzed via machine-learning algorithms to determine which features, both individually and jointly, predict overeating. These analyses with be performed using both isolated and combined data collected from EMA and passive detection. This integrated approach will allow us to determine whether and how predictions of overeating episodes change based on data sources. The secondary aims are to use the data to build clusters of overeating and eating episodes that identify problematic overeating behaviors (e.g., stress eating, social eating, and night eating), and to explore new phenotypes based on social (with friends, family, alone, etc.) and psychological (affect, craving) features that make up the overeating clusters.

## METHODS

### *Participants*

We are recruiting up to 60 participants. Eligible participants are adults with obesity (BMI$\geq$ 30 kg/m$^2$), between the ages of 18 and 65 years, who reside in the Chicago Metropolitan Area, own a smartphone, and have wireless internet at their residence. Participants are excluded if they are: currently dieting with the goal of losing weight; have lost 15 or more pounds in the previous three months; are, or plan to become, pregnant; have received, or plan to receive, bariatric or lap band surgery; currently take any medication that causes weight loss; have a genetic weight loss disorder (e.g., Prader-Willi, Bardet-Biedl, Cohen Syndrome); or have an active eating disorder. Who also exclude participants found to be susceptible to loss-of-control (LOC) eating when screened via interview because LOC eating is strongly connected to binge eating and other mental health disorders, and as a result warrants a separate study and likely requires more intensive intervention than we anticipate building through our mobile eating detection systems.

*Recruitment*

Participants are recruited on a rolling basis using online advertisements hosted on Craigslist.com, ResearchMatch, and TheNewNormal. Respondents who complete the REDCap[35] web screener are contacted by the study team if their provided information aligns with the study's eligibility criteria. Flow from recruitment to all other phases of the study is depicted in Figure 1.

**<INSERT FIGURE 1>**

*Telephone screening*

Respondents are assessed for disordered eating behaviors with multiple validated measures. The eating disorder module of the Patient Health Questionnaire (PHQ)[36] assesses the presence of active eating disorders.[37] Additionally, a trained investigator administers the loss of control assessment of the Eating Disorder Examination (EDE).[38] The combination of PHQ self-report items and investigator follow-on questions drawn from the EDE is considered the gold-standard approach to the assessment of disordered eating behaviors.[39] If any disordered eating behaviors are present during screening, the respondent is excluded from participation in the study. Otherwise, participants are scheduled for a baseline appointment.

*In-person baseline appointment*

Upon participants' arrival to the lab, informed consent is conducted as approved by the institutional review board (IRB). The study coordinator collects one weight and three height measurements using a standard analog scale (*Detecto mechanical eye-level physician scale with height bar)*. The weight measurement and the mean of the height measurements are used to calculate BMI, confirming that the participant indeed meets the study criteria of BMI $\geq$ 30 kg/m$^2$, at which point they become officially enrolled in the study.

Once enrolled, participants are instructed on the use of all study technology.

*Foodtrck app.* After downloading the Foodtrck eating-recording app from its appropriate public listing (see Appendix), participants log in and complete an entry of a meal that the study team gives them in lab. Participants receive teaching and supervision as they use the FoodTrck app before and after the in-lab eating episode (described below and in Figure 2).

**<INSERT FIGURE 2>**

*Study devices.* Participants are introduced to the four study devices (i.e., necklace, smartwatch, infrared activity-oriented device [IR- AOD] that provides visual confirmation of the meals, and study phone) and instructed on their usage, including turning on each device, verifying device functioning, wearing the devices at the appropriate on-body position, ending recording, removing devices, and connecting the devices to their chargers. Participants complete a 'role-play' teaching exercise in which they practice the start-collection and end-collection procedures 3 times each. Lastly, participants are informed of the circumstances in which they are allowed to remove the study devices (during water exposure, vigorous exercise, etc.).

*Sync event.* The sync event is intended to introduce an explicit synchronization point which is identifiable across sensor modalities to make sure all devices are capturing the same moment. Participants are also instructed on how to perform a sync event,[40] which they are asked to complete daily, immediately following the start of each recording period. To establish ground truth for the time-of-day, participants hold a smartphone application with a digital clock interface in the field of view of the IR-AOD, which captures the displayed time. Participants then perform three gestures. The first gesture involves participants covering the necklace with their hand three times, for 1 second each time. This creates a series of three near-zero proximity measurements that can be easily spotted in the proximity signal output of the necklace as well as the synchronized IR-AOD video, which provides visual confirmation of the hand gesture. The second activity involves participants clapping their hands in front of themselves (as if applauding) three times. These claps, as fast wrist motions followed by sudden stops, can be spotted in the smartwatch IMU signals as well as in the synchronized IR-AOD video, which provides visual confirmation of the clapping gesture. The final gesture involves participants either drinking from a glass or miming the action of drinking from a glass. This activity appears in the output of all three sensors: appearing visually to the IR-AOD, appearing in the lean-forward-angle and proximity signal of the necklace, and appearing in the accelerometer (hand-to-mouth) and gyroscope (glass tilting) signals of the smartwatch.

*In-wild phase*

After the in-lab session is completed, the in-wild phase of the study begins the following morning (see Figure 3). Participants are instructed to wear the devices throughout all waking hours, answer questions and record all food and beverage consumption with the Foodtrck app and complete an interviewer-administered 24-hour diet recall each day (beginning on Day 2). Recalls are initiated by the interviewer within the windows/window of availability provided to the study team by participants during the initial visit.

**<INSERT FIGURE 3>**

The experimenter monitors participant adherence to these instructions by checking heartbeat notifications (an indication of what time each device was last turned on), Foodtrck entries, and diet recall completion. If adherence indicators are absent, the experimenter contacts the participant to determine the cause of the lapse and provide reminders and/or technical support as necessary. If a participant is unable to perform all 3 tasks (wear devices, record meals, complete recalls) for at least 6 of the first 7 “run-in” in-wild days; the run-in period is considered failed, and the participant is withdrawn from the study. After 14 days of in-wild collection conclude, participants are asked to drop off the devices, have a final weight measurement taken, and complete a series of questionnaires (listed in Table 2 under "post survey" study element).

## Measures

### *Devices*

*Infrared Activity-Oriented Device (IR-AOD).* The IR-AOD is a wearable camera developed by the HABits Lab to maximize information collection and minimize user discomfort (both physical and psychological)[41, 42] and risks to privacy.[43] The IR-AOD is worn on a lanyard and secured to the upper-center area of the chest by a magnetic pad that connects to the back of the device behind the shirt/outermost layer of clothing. The top surface of the IR-AOD, a flat plane parallel to the ground, contains an upward-facing sensor array including a thermal infrared sensor, a red-green-blue (RGB) camera with a 180° fisheye lens, and a photoresistor-triggered infrared light emitting diode (IR LED) as a source of illumination for night vision. Each sensor is directed toward the face and upper torso of the wearer. When the IR-AOD is turned on, all three sensors record continuously to an onboard memory card.

*NeckSense.* NeckSense is a neck-worn sensor device that is principally designed to capture and quantify chewing. NeckSense is worn on a short, necklace-type lanyard, and is positioned at the base of the neck. The top surface of the necklace, a flat plane parallel to the ground, hosts an upward-facing Infrared proximity sensor that records at 20Hz (20 measurements a second). The positioning of NeckSense at the base of the neck allows the proximity sensor to capture the mouth open/close status of the wearer by recording the distance between the device and the bottom of the chin. In addition to the proximity signal, NeckSense continuously records ambient light (LUX), lean-forward angle (LFA), and triaxial accelerometer data, saving all collected data to an onboard memory card.

*Wrist-Based Recording System.* The wrist-based recording system consists of a commercial Fossil Smartwatch loaded with a customized data collection app that makes use of the watch's onboard sensors. The watch is worn on the dominant wrist, positioned two finger-lengths below the bottom of the palm. The wrist collection system records triaxial accelerometer data, triaxial gyroscope data, and photoplethysmography (PPG) data. Each of these measurements are recorded at 20Hz and saved to the watch's onboard memory card. Each time the watch is charged, it transmits its data to a paired cell phone, which then transmits the data to a secure cloud server maintained by the research team.

*App-delivered EMA surveys*

The FoodTrck app is installed on participant's personal phones and facilitates both dietary self-reporting and the delivery of targeted EMAs. To record a meal, snack, or beverage; the user creates a new entry on the FoodTrck home screen and begins completing the three surveys:

*Decided To.* Participants are instructed to complete the "Decided to" survey as soon as possible after they have decided to eat/drink and know what they will eat/drink (but does not necessarily have the food/drink in front of them, ready to consume. The "Decided to" survey contains EMA items on food source, affect, and presence of biologic hunger or hedonic eating.

*About to.* Participants are instructed to complete the "About to Eat" activity immediately prior to eating/drinking by taking a photo of all food/beverage items to be consumed and entering a descriptive name.

*After.* Participants are instructed to complete the "After" survey immediately after they finish eating/drinking. The "After" survey involves taking a photo of the remaining food/drink or empty food/drink containers, entering a descriptive name, answering questions about affect, cognitive restraint, uncontrolled eating, contextual or environmental factors (e.g., eating alone or with others, watching T.V., whether at home or work).

*Photo-assisted 24-hour diet recall*

Starting on the second day of the free-living period, a 24-hour diet recall using the 5-step Automated Multiple-Pass Method will be administered via telephone by a trained interviewer with nutrition expertise.[44] The objective of the 24-hour recall is to collect a complete account of everything the participant consumed or drank the previous day. Before the call, dietitians can access the photographs and food/drink descriptions recorded by participants using the FoodTrck app. During the interview, the dietitian will complete steps that has been shown to enhance the accuracy of the recall, including 1) developing a quick list of foods and beverages consumed, 2) probing for commonly forgotten food items, 3) collecting time and name of each eating occasion, 4) collecting and reviewing details of food and beverages, and 5) completing a final probe.[45] Images from the FoodTrck app will assist with probes during steps 2 to 4. Information gathered during 24-hour recalls is analyzed using the Nutritional Data System for Research (NDSR), a research-based nutrition software developed to collect and analyze 24-hour diet recalls.[46]

## Data Processing

*Alignment of multiple data sources*

Data collected by wearable devices are evaluated for synchronization by referencing the Sync Event at the beginning of each day. The time displayed on the cell phone (visible in the video data) provides the ground truth timestamp to which the data streams are compared and offsets are corrected. If a Sync event is missing for a given wear period, the evaluator attempts to identify other natural occurrences of gestures that can be recognized across sensor modalities or compare the timestamp of a

FoodTrck entry to the timestamp of the video frame in which the participant's completion of the entry appears. Once video timestamps are known to be aligned with the phone, the video timestamps are then used as ground truth. These methods allow us to establish accurate start and end times of activities, which is critical for building reliable machine-learning algorithms that detect human behaviors (e.g., feeding gestures, bites [see Table 1]) by learning from training data in which these behaviors are visually confirmed using multiple video annotators.

**<INSERT FIGURE 4>**

*Labeling procedures*

Labels of events captured by the IR-AOD (video data) are generated by trained annotators using a labeling guide. The annotators identify eating episodes in the video data and assign the appropriate label to each of the episode's preselected fine-grained elements. Duration and point labels indicate the time and date of the labeled event. These labels capture participant behaviors and define the start and end times of eating episodes. Table 1 includes of list and description of each of these features.

**Table 1.** Features of meals collected by study measures.

| ***Feature*** | ***Description*** | ***Label*** | ***Data*** | ***Source*** |
|---|---|---|---|---|
| **Behavioral** | | | | |
| Feeding Gesture | Hand brings food to mouth and returns to rest | Start and end time of each feeding gesture | Count and frequency of feeding gestures per meal | Visual (IR-AOD) |
| Bite | Jaw open/close as food enters mouth | Timestamp of the bite | Count and frequency of bites per meal | Visual (IR-AOD), Proximity (NeckSense) |

| | | | | |
|---|---|---|---|---|
| Chew | Consecutive jaw open/close sequence while food in mouth | Timestamp of the chew | Count and frequency of chews per meal | Visual (IR-AOD), Proximity (NeckSense) |
| Swallowing | Chin tilt up/down, throat vibrations as food is ingested | Inferred by cluster of consecutive chews | Count and frequency of swallows per meal | Visual (IR-AOD), Proximity (NeckSense) |
| Eating Duration | End time of episode - start time of episode | Start and end time of each feeding gesture | Length of feeding gesture per meal Min/Max/Average/STD | Visual (IR-AOD), Self-report (Phone, Recall) |
| **Physiological** | | | | |
| Heart Rate Variability | Interpulse-interval variability | Interpulse-interval between systolic peaks from PPG | Physiological Stress level | Photoplesthmographic (Wrist) |
| Heart Rate | Heart beats per minute | Count of systolic peaks from PPG | Heart rate (average)<br>1) Before meal (30s)<br>2) After meal (30s)<br>3) During the meal | Photoplesthmographic (Wrist) |
| **Social** | | | | |
| Time of Day | HH:MM:SS | Start and end time of each meal | Time of day of meal (morning/afternoon/night) | Visual (IR-AOD) Self-report (Phone, Recall) |
| Environment (social) | eating alone, eating with others | Social eating (FoodTrck) | Social eating (Binary) | Self-report (Phone) |
| Location (type) | work, home, school, other | Social eating (FoodTrck) | Locations of social eating (Categorical) | Self-report (Phone) |
| Screen Time | eating in presence of screen (TV, computer, phone, etc.) | Social eating (FoodTrck) | Presence of screen (Binary) | Visual (IR-AOD), Self-report (Phone) |
| Bystander | other person/s present in area | Social eating (FoodTrck) | Scales from social eating (Categorical) | Visual (IR-AOD) |
| **Psychological** | | | | |
| Stress | Stress level during eating | Stress (FoodTrck) | Perceived Stress from FoodTrck self-report (Likert Scale)<br>1) Before meal<br>2) During meal<br>3) After meal | Self-report (Phone) |
| Cognitive Restraint | Stopped eating to manage calorie intake | Cognitive Restraint (FoodTrck) | After meal Cognitive Restraint (Likert scale) | Self-report (Phone) |
| Uncontrolled Eating | Perceived inability to stop oneself from eating | Uncontrolled Eating (FoodTrck) | LOC (Likert scale) | Self-report (Phone) |
| Overeating | Eating more than intended | Overeating (FoodTrck) | Subjective overeating (Likert scale) | Self-report (Phone) |

| | | | | |
|---|---|---|---|---|
| Biological Hunger (Craving) | Desire for food motivated by hunger | Biological Hunger (FoodTrck) | Hedonic Eating vs Biological Hunger (Likert scale) | Self-report (Phone) |
| Hedonic Eating | Desire for food motivated by taste/pleasure | Hedonic Eating (FoodTrck) | Hedonic Eating vs Biological Hunger (Likert scale) | Self-report (Phone) |
| Emotion - Stressed/Anxious | Degree of feeling "Stressed/Anxious" | Stressed/Anxious (FoodTrck) | Emotional (Likert scale) | Self-report (Phone) |
| Emotion - Down/Lonely | Degree of feeling "Down/Lonely" | Down/Lonely (FoodTrck) | Emotional (Likert scale) | Self-report (Phone) |
| Emotion - Upbeat/Excited | Degree of feeling "Upbeat/Excited" | Upbeat/Excited (FoodTrck) | Emotional (Likert scale) | Self-report (Phone) |
| Emotion - Calm/Peaceful | Degree of feeling Calm/Peaceful | Calm/Peaceful (FoodTrck) | Emotional (Likert scale) | Self-report (Phone) |
| Social Eating | Presence of other people during meal | Social Eating (FoodTrck) | Social eating (Binary) | Self-report (Phone) |
| Activity co-occurrence | Other activities performed while eating | Other (FoodTrck) | Other activities (Categorical) | Self-report (Phone) |

*Foodtrck data processing*

Foodtrck entries are assigned to eating episodes based on corresponding timestamps. Foodtrck provides contextual, psychological, and geospatial information about a given eating episode (see Table 2). Contextual features of an eating episode collected by EMA surveys in Foodtrck are stored in an SQL database and attached to a specific eating episode as attributes.

**Table 2.** Measures employed in the study.

| **Study element** | **Construct** | **Description** | **Measure (reference)** |
|---|---|---|---|
| Web screener | Demographics | Age, self-reported height and weight, ethnicity, race, residence location/type | Developed by HABits Lab |
| In-Lab | Anthropometrics | Height, weight | Detecto Mechanical Metric-only Eye-Level Physician Scale with Height Rod |
| FoodTrck — Pre-meal | Pre-meal stress | Perceived stress before a meal | Developed by HABits Lab |

| | | | |
|---|---|---|---|
| | Affect | Perceived affective state before a meal | Adapted from Emotion Circumplex model of Affect[47] |
| | Eating in Absence of Hunger | Desire for food before a meal | Adapted from Emotion Circumplex model of Affect[47] |
| | Hedonic Eating | Degree to which *pleasure seeking* motivates upcoming meal | Forman et al. 2017[27] |
| | Biological Hunger | Degree to which *hunger satisfaction* motivates upcoming meal | Forman et al. 2017[27] |
| | Pre-meal food description | Self-reported name of food/drink items to be consumed | Forman et al. 2017[27] |
| — Post-meal | Post-meal stress | Subjective stress score (1-5) assessed before each meal via FoodTrck app | Forman et al. 2017[27] |
| | Cognitive Restraint | Whether participant purposely stopped eating to avoid weight gain/cut calories | Manasse et al. 2018[13] |
| | Uncontrolled Eating | Perceived inability to stop oneself from eating | Manasse et al. 2018[13] |
| | Overeating | Whether participant ate more than intended (subjective overeating) | Manasse et al. 2018[13] |
| | Hedonic Eating | Amount of pleasure experienced while eating | Forman et al. 2017[27] |
| | Social Eating | Presence/absence of other people during meal | Forman et al. 2017[27] |
| | Activity co-occurrence | Other activities performed while eating | Forman et al. 2017[27] |
| | Location type | Type of location (work, home, school, etc.) | Developed by HABits Lab |
| | Post-meal food description | Self-reported name of food/drink items remaining after meal | Forman et al. 2017[27] |
| 24-hour Diet Recall | Eating episode time and Location | Time meal began and location consumed | Developed by HABits Lab |
| | Meal content | Food/drink types and amounts consumed during meal | Nutrition Data System for Research (NDSR)[46] |
| | Nutritional content | Breakdown of caloric and nutrient values of food/drink items consumed | Nutrition Data System for Research (NDSR)[46] |
| Post-survey | User Burden | User experience and judgement of each wearable device | User Burden Scale (UBS)[48] |
| | Hedonic Eating | Propensity to eat for pleasure (in general) | Power of Food Scale (PFS)[49] |
| | Emotional Eating | Propensity to eat when experiencing certain emotions (in general) | Dutch Eating Behavior Questionnaire (DEBQ)[50] |
| | Overeating | Propensity to overeat and/or binge eat (in general) | Binge-Eating Disorder Screener (BEDS-7)[51] |
| | Intuitive Eating | Propensity to eat in response to sensory cues (in general) | Three Factor Eating Questionnaire (TFEQ)[52] |

*NDSR Data Processing and Merging with Foodtrck Data*

Data on each eating episode are downloaded from the NDSR software. Data extracted for each eating episode includes time, calories, and macronutrients. These data are merged with the Foodtrck datasets using both the participant ID and eating episode time variables. We then apply the following exclusion criteria for this merged dataset:

1) exclude participants who dropped out of the study or did not pass the 7-day run-in period

2) exclude eating episodes with missing or incomplete Foodtrck questionnaires

3) exclude eating episodes with 0 calories (e.g., non-caloric beverages only)

*Defining overeating*

In this study, overeating episodes is operationalized as the eating episodes for which calories consumed is more than 1 z-score greater than the participant's average calories consumed. This definition essentially operationalizes overeating in a manner that is personalized based on the individual's prior caloric intake distribution, where the calories per eating episode are validated by a dietitian. The method, involves capturing an individual's eating across two weeks providing a representation of each individual's eating pattern, yielding percentiles and cut points for overeating. In our previous work, we showed that our definition of an overeating episode strongly correlated with subjective overeating (Cohen's Kappa = 0.90, showing high agreement), validated by a dietitian.[53] In exploratory analyses, we will also test one other definition used in the literature, which defines an eating episode to be overeating if it is 1000 calories or more.[54, 55]

*Planned analyses*

Parallel analyses will be conducted using the three differing sets of features below:

*Passive sensing only.* Using machine learning, we will initially build an overeating predictive model using passive-sensing only features (representing behavioral and physiological features), such as number and frequency of feeding gestures, number and frequency of swallows, number and frequency of chews, eating duration, heart rate and heart-rate variability (HRV) before the eating episode, heart rate and HRV during the eating episode, and time-of-day.

*EMA –only.* We will then build a predictive model including EMA social and emotional state (expanding the features based on training). This allows us to identify which self-reported features predict overeating. See Table X for features captured.

*Passive sensing + EMA.* In a third set of analyses, we will then combine passive sensing features and EMA when predicting overeating. Because these analyses will include the most complete set of features, these analyses will be treated as the main analyses for making conclusions about predictors of overeating and characteristics of overeating clusters. The preceding analyses are conducted to determine the extent to which passive sensing vs. EMA-only analyses lead to different conclusions about overeating and whether one type of data may be sufficient.

*Identifying Predictive Features of Overeating*

*Feature Selection.* To avoid overfitting of the model to the training data, a subset of the features for modeling overeating will be selected using two complementary methods: 1) Correlation-based Feature Selection (CFS)[56] to find the optimal non-correlated feature set, independent of machine learning algorithm, and 2) Wrapper-based Feature Selection (WFS),[57] to find an optimal feature subset for the Random Forest machine learning algorithm, accounting for the possibility of CFS discarding potentially useful features that are useful for a specific machine learning algorithm. Random Forest is a prominent ensemble-based model that combines a large number of weak simple models to obtain a stronger ensemble prediction by averaging (resulting in decreased variance) models.[58] The output of other feature selection algorithms will also be compared with the output of CFS and WFS to ensure highly predictive features are included in the model.

*Machine-learned Classifiers*.[59-64] Discriminative classifiers generate machine-learned models that directly distinguish boundaries through observed data. Once features are selected, a discriminative supervised machine learning model such as Gradient Boosting Machines (GBMs) will be used. GBMs are a family of machine learning methods that have shown success in a wide range of applications in machine-learning challenges[57-59] producing competitive, highly robust, interpretable procedures for both regression and classification.[60] The principal idea is to construct new base-learners to be maximally correlated with the negative gradient of a chosen loss function (for example: Adaboost loss functions are used for categorical outcome variables). We will further compare GBMs to other combinations of discriminative classifiers such as logistic regression, support vector machine, random forest, and neural networks, and generative classifiers such as Bayesian networks and hidden Markov models. Generative classifiers are more indirect in their approach, and often deploy statistical models and probability theory, sometimes requiring more a priori knowledge that is often unknown, to estimate the probability of each outcome given the observed data.

*Evaluation.* Based on our sample size, we will generate a train, validation, and test set with a 60:20:20 train:validation:test split ratio. Each classifier will be trained on 60% of the data, and the hyperparameters of the classifier will be fine-tuned on the 20% validation set. We will also perform a ten-fold cross validation (averaging results across 10 runs with a 90:10 split of the data) will be used when building a model. We will report on the Receiver Operator Characteristic Area Under the Curve (ROC-AUC), and select the best model generated by the algorithm based on the average F-measure (more precise measure of performance that captures precision of the algorithm and recall of both overeating and non-overeating episodes).

*Identifying Problematic Overeating Phenotypes through Clustering*

*Generate Clusters.* To ensure that a cluster represents most overeating episodes and not regular eating episodes, the resulting optimal features that detect and predict overeating will be used to generate well-separated clusters of regular eating and overeating. Our goal will be to use the features to attain high intra-cluster similarity (eating episodes in a cluster represent the same label of overeating or non-overeating), and low inter-cluster similarity (samples from different clusters are dissimilar) between overeating and non-overeating clusters. We will test two algorithms. First, we will use k-means,[43, 65] which is the most widely used partitional clustering algorithm owing to its versatility and efficiency time and space complexity. Every aspect of it (initialization, distance function, termination criterion, etc.) can be modified; it is guaranteed to converge[66] at a quadratic rate;[67] and it is invariant to data ordering (random shuffling of data points). The optimal k-value for the number of clusters is determined by calculating the silhouette score,[68] which is computed for a given range of k-values and evaluates which k value yields clusters that are most representative of the data that comprise them. The silhouette score quantifies (between -1 and +1) how similar a given data point is to its own cluster, and how dissimilar it is to the other clusters. A silhouette score closer to +1 indicates that the data point is well matched to its own cluster and poorly matched to its neighbors. Second, we will also compare its performance with an agglomerative hierarchical clustering algorithm (which does not require us to prespecify the number of clusters) combined with Ward's minimum variance method (and variations of Ward's method) to minimize the total within-cluster variance. Hierarchical clustering techniques will be used to identify well-separated clusters based on their similarity matrix, a technique we tested in predicting models for hospital readmission.[69]

To determine the best clustering method, we will estimate purity, normalized mutual information, rand index, and F measures. We will weight each metric equally, when selecting the optimal set of clusters used. The purity of each cluster (a transparent evaluation measure), which given a set of K clusters $\Omega = \{\omega_1, \omega_2, \dots, \omega_K\}$, and a set $\mathbb{C} =$

$\{c_1, c_2\}$, where $c_1$ represents overeating, and $c_2$ represents non-overeating episodes, and given N total combined episodes: $purity(\Omega, \mathbb{C}) = \frac{1}{N}\sum_k \max_j |\omega_k \cap c_j|$. High purity signifies high intra-cluster similarity, but it can result in a high number of clusters (each sample being defined by its own cluster achieves high purity), resulting in the need for complementary metrics. To finalize our selection of a clustering method, we will additionally consider: 1) the Normalized mutual information (NMI) - an information-theoretical driven approach that uses maximum likelihood estimates and entropy (a measure of disorder or uncertainty in the clusters); 2) the Rand index, which penalizes both false positives and false negative decisions during clustering; and 3) the F measure, which differentially weights these two types of errors.

*Identifying Overeating and Non-overeating Clusters.* We will see whether theoretically meaningful problematic eating phenotypes occur, by analyzing the feature makeup of each cluster (e.g., level of stress, time of day, alone or with friends/family). We will further discover new phenotypes by analyzing the relationship between the features within each cluster with a significant proportion of overeating episodes. Analyzing levels of overeating in the clusters (average value of outcome) indicates risk in relation to other clusters.

We will provide statistics of each cluster, including purity of cluster, percent of overeating and regular eating, percent of stress eating, eating in absence of hunger, hedonic eating vs. biological hunger, cognitive restraint, uncontrolled eating, and overeating. This will allow us to identify how clusters co-occur with each other. We will use a personalized definition for each question, where we calculate the average of a person's response, and a value that exceeds the average will be assigned a high and others will be assigned a low value.

## Discussion

The ability to predict and detect overeating episodes creates new possibilities for answering pressing research questions and developing state-of-the-art weight management interventions. This project aims to produce a rich contextual characterization of overeating episodes by combining the use of wearable sensing devices, meal-triggered EMAs, and a series of 24-hour recalls of food and beverage intake.[44] Together, these measures are designed to quantify participants' diets and provide insight into the physiological, contextual, social, and psychological factors that surround eating. Through our planned analyses, we will detect which time-varying factors are associated with overeating. We will additionally use these factors to identify clusters of overeating episodes, potentially leading to new insights regarding the predictors of overeating and co-occurrence of clinically meaningful eating phenotypes.

Strengths of this study include the use of passive sensing and EMA measures, which provide a rich description of in-the-moment behavior while minimizing recall biases. The passive sensing system used in this study also has many advantages that reduce user burden and promote the collection of high-quality data. This system uses multiple sensors to detect multiple proxies to eating to ensure reliable detection and characterization of eating behavior. We report the F-measure (a more precise measure of performance than accuracy). Our study is the first of its kind because it enables visual confirmation in-the-real world over longer periods of time (prior work typically use cameras for short durations in the wild), which enables validation and confirmation of eating and other user behaviors as well (e.g., screen time, presence of others, and presence of secondary activities).[70, 71] By combining multiple sensing modalities, long device battery life (at least 48 hours for the necklace,[25] 16 hours for the camera and 24 hours for the wrist-sensor while collecting data continuously), high customizability,[72] and minimal degradation of performance in challenging eating environments (i.e., those confounded by various activities),[14, 17, 25, 73] we are able to provide the community with a realistic longitudinal dataset containing multiple characterizations of the eating episode against which to build and advance machine learning models for eating detection.

While several wearable sensors have shown promise in detecting eating in the real-world, they have predominantly been validated within a convenience sample, primarily in student populations. For systems to truly be generalizable additional data is needed from diverse populations, primarily people with varying body mass indeces, that is not only student-based or focused on healthy-people. Our recent work has shown that models trained on people without obesity perform poorly when applied to people with obesity.[9] Moreover, people with varied body shapes may experience the system differently, varying in reported comfort levels. Our dedication to studying people with obesity using wearable sensors enables deeper insight and translation of research to practice. To the best of our knowledge we are among the very few to explicity validate our automated detection systems in people with obesity.

Many prior works with sensors in the real-world setting focus on studying people while being enrolled in a weight-loss intervention. We are one of the first to study people with obesity while telling them to “be themselves as much as possible.” This allows us to capture their current problematic eating habits to redefine problematic eating behaviors through sensing and EMA. Until we are able to properly understand what constitutes an overeating episode, we will not know what to detect and ultimately prevent overeating relative to need, a main cause of obesity. This study is not without limitations. To characterize overeating, we rely on a recently developed operationalization of overeating episodes (i.e., eating episodes with energy intake that are 1 *SD* above the average caloric intake across all eating episodes). This operationalization was selected over past measures, including 1) a 1,000-calorie threshold, which is a crude indicator that is not sensitive to individual differences in BMI and eating pattern, and 2) whether or not a meal had been planned – a label which can then be applied to relatively low-calorie meals, requires user input to define overeating, and can only be used when individuals are on a diet. Our personalized definition of overeating at the episode level is consistent with our goal to understand when people consume more than they typically consume, paving the way for Just-In-Time Adaptive Interventions for overeating, and it strongly agrees with participants' subjective perceptions[53] of which eating episodes constitute overeating. This definition assumes that overeating episodes will occur among all individuals and at about the same rate. Although this feature of our definition does result in some conceptual ambiguity, we note that within a population with obesity understanding and predicting eating episodes with the highest energy intake will nevertheless be greatly beneficial for designing weight loss interventions.

**Acknowledgements:** The authors would like to thank Rawan Alharbi, Shibo Zhang, Qiuyang Xu, Tim Truty, Linda Van Horn, Don Hedeker, Santosh Kumar, Alexandra Karr, Danielle Ward, and Sheema Hussani for their support of this project.

**Contributorship:** NA, JH, and BS participated in the design of the system. NA, JH, TKS, CR, and EF participated in the development of the system. NA, TKS, AFP, AL, and BS were involved in the study design. NA and CR were involved in the deployment and testing of the system. NA and TKS participated in the analysis planning. NA, TKS, and CR drafted the manuscript, and all authors reviewed and edited the manuscript and approved the final version.

**Declaration of conflicting interests:** The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

**Ethical approval:** This study was approved by the Northwestern University Institutional Review Board (STU00204564).

**Funding:** This work was supported by the National Institute of Diabetes and Digestive and Kidney Diseases (NIH-NIDDK Project # 1K25DK113242-01A1).

**Guarantor:** Nabil Alshurafa

## References


1. Hales CM, Carroll MD, Fryar CD, et al. Prevalence of Obesity and Severe Obesity Among Adults: United States, 2017-2018. *NCHS Data Brief* 2020: 1-8.
2. Pollack A. AMA Recognizes Obesity as a Disease. http://www.nytimes.com/2013/06/19/business/ama-recognizes-obesity-as-a-disease.html: New York Times, 2013.
3. Liu JJ, Xu W, Huang M-C, et al. A dense pressure sensitive bedsheet design for unobtrusive sleep posture monitoring. *2013 IEEE International Conference on Pervasive Computing and Communications, PerCom 2013* 2013: 207-215. DOI: 10.1109/PerCom.2013.6526734.
4. LaCaille L. Eating Behavior. In: Gellman MD and Turner JR (eds) *Encyclopedia of Behavioral Medicine*. New York, NY: Springer New York, 2013, pp.641-642.
5. Larson N and Story M. A Review of Environmental Influences on Food Choices. *Annals of Behavioral Medicine* 2009; 38: s56-s73. DOI: 10.1007/s12160-009-9120-9.
6. Meegahapola L, Ruiz-Correa S, Robledo-Valero VdC, et al. One More Bite? Inferring Food Consumption Level of College Students Using Smartphone Sensing and Self-Reports. *Proc ACM Interact Mob Wearable Ubiquitous Technol* 2021; 5: Article 26. DOI: 10.1145/3448120.
7. Zhang S, Stogin W and Alshurafa N. I sense overeating: Motif-based machine learning framework to detect overeating using wrist-worn sensing. *Information Fusion* 2018; 41: 37-47. DOI: https://doi.org/10.1016/j.inffus.2017.08.003.
8. Sharma S and Hoover A. Top-Down Detection of Eating Episodes by Analyzing Large Windows of Wrist Motion Using a Convolutional Neural Network. *Bioengineering (Basel)* 2022; 9 20220211. DOI: 10.3390/bioengineering9020070.
9. Zhang S, Zhao Y, Nguyen DT, et al. NeckSense: A Multi-Sensor Necklace for Detecting Eating Activities in Free-Living Conditions. *Proc ACM Interact Mob Wearable Ubiquitous Technol* 2020; 4. DOI: 10.1145/3397313.
10. Alshurafa N, Kalantarian H, Pourhomayoun M, et al. Recognition of nutrition intake using time-frequency decomposition in a wearable necklace using a piezoelectric sensor. *IEEE Sensors Journal* 2015; 15: 3909-3916. DOI: 10.1109/JSEN.2015.2402652.
11. Bedri A, Li R, Haynes M, et al. EarBit: Using Wearable Sensors to Detect Eating Episodes in Unconstrained Environments. *Proc ACM Interact Mob Wearable Ubiquitous Technol* 2017; 1: Article 37. DOI: 10.1145/3130902.
12. Kral TVE, Moore RH, Chittams J, et al. Identifying behavioral phenotypes for childhood obesity. *Appetite* 2018; 127: 87-96. DOI: https://doi.org/10.1016/j.appet.2018.04.021.
13. Manasse SM, Schumacher LM, Goldstein SP, et al. Are individuals with loss-of-control eating more prone to dietary lapse in behavioural weight loss treatment? An ecological momentary assessment study. *Eur Eat Disord Rev* 2018; 26: 259-264. 20180227. DOI: 10.1002/erv.2583.
14. Zhang S, Alharbi R, Stogin W, et al. Food-Watch: Detecting and Characterizing Eating Episodes through Feeding Gestures. BEST PAPER AWARD. In: *BODYNETS 2016 - 11th International Conference on Body Area Networks* 2016.
15. Kalantarian H, Alshurafa N, Le T, et al. Non-invasive detection of medication adherence using a digital smart necklace. *2015 IEEE International Conference on Pervasive Computing and*

*Communication Workshops, PerCom Workshops 2015* 2015: 348-353. DOI: 10.1109/PERCOMW.2015.7134061.
16. Bi S, Wang T, Tobias N, et al. Auracle: Detecting Eating Episodes with an Ear-mounted Sensor. *Proc ACM Interact Mob Wearable Ubiquitous Technol* 2018; 2: Article 92. DOI: 10.1145/3264902.
17. Cohen E, Alharbi R, Pfammatter A, et al. Random Forests detects eating in a neck-worn sensor: Results from a simulated real-life eating setting. Best Paper Award. In: *2016 Obesity Week* New Orleans, LA, 2016.
18. Merck CA, Maher C, Mirtchouk M, et al. Multimodality sensing for eating recognition. In: *PervasiveHealth* 2016, pp.130-137.
19. Hossain D, Imtiaz MH and Sazonov E. Comparison of wearable sensors for estimation of chewing strength. *IEEE Sens J* 2020; 20: 5379-5388. 20200120. DOI: 10.1109/jsen.2020.2968009.
20. Amft O and Tröster G. Recognition of dietary activity events using on-body sensors. *Artificial Intelligence in Medicine* 2008; 42: 121-136. DOI: http://dx.doi.org/10.1016/j.artmed.2007.11.007.
21. Sazonov E, Schuckers S, Lopez-Meyer P, et al. Non-invasive monitoring of chewing and swallowing for objective quantification of ingestive behavior. *Physiological measurement* 2008; 29: 525-541. DOI: 10.1088/0967-3334/29/5/001.
22. Fontana JM, Farooq M and Sazonov E. Automatic Ingestion Monitor: A Novel Wearable Device for Monitoring of Ingestive Behavior. *IEEE transactions on bio-medical engineering* 2014; 61: 1772-1779. DOI: 10.1109/TBME.2014.2306773.
23. Rahman T, Czerwinski M, Gilad-Bachrach R, et al. Predicting "About-to-Eat" Moments for Just-in-Time Eating Intervention. *Proceedings of the 6th International Conference on Digital Health Conference*. Montréal, Québec, Canada: ACM, 2016, p. 141-150.
24. Alshurafa N. Improving Healthcare for Populations: Building Smart and Integrated Remote Health Monitoring Systems. 2015.
25. Cohen E, Stogin W, Kalantarian H, et al. SmartNecklace: Designing a Wearable Multi-sensor System for Smart Eating Detection. In: *BODYNETS 2016 - 11th International Conference on Body Area Networks* Turin, Italy, 2016.
26. Heron KE, Scott SB, Sliwinski MJ, et al. Eating behaviors and negative affect in college women's everyday lives. *Int J Eat Disord* 2014; 47: 853-859. 2014/05/07. DOI: 10.1002/eat.22292.
27. Forman EM, Schumacher LM, Crosby R, et al. Ecological Momentary Assessment of Dietary Lapses Across Behavioral Weight Loss Treatment: Characteristics, Predictors, and Relationships with Weight Change. *Ann Behav Med* 2017; 51: 741-753. DOI: 10.1007/s12160-017-9897-x.
28. Jeffers AJ, Mason TB and Benotsch EG. Psychological eating factors, affect, and ecological momentary assessed diet quality. *Eat Weight Disord* 2020; 25: 1151-1159. 2019/08/08. DOI: 10.1007/s40519-019-00743-3.
29. Crochiere RJ, Abber SR, Taylor LC, et al. Momentary predictors of dietary lapse from a mobile health weight loss intervention. *J Behav Med* 2021 2021/11/23. DOI: 10.1007/s10865-021-00264-4.

30. Feig EH, Piers AD, Kral TVE, et al. Eating in the absence of hunger is related to loss-of-control eating, hedonic hunger, and short-term weight gain in normal-weight women. *Appetite* 2018; 123: 317-324. DOI: https://doi.org/10.1016/j.appet.2018.01.013.
31. Espel-Huynh HM, Muratore AF and Lowe MR. A narrative review of the construct of hedonic hunger and its measurement by the Power of Food Scale. *Obes Sci Pract* 2018; 4: 238-249. 20180228. DOI: 10.1002/osp4.161.
32. Lowe MR and Kleifield EI. Cognitive restraint, weight suppression, and the regulation of eating. *Appetite* 1988; 10: 159-168.
33. Schlundt DG, Sbrocco T and Bell C. Identification of high-risk situations in a behavioral weight loss program: application of the relapse prevention model. *Int J Obes* 1989; 13: 223-234.
34. Cline R. Mindful vs. Mindless Eating : NCHPAD - Building Inclusive Communities. *National Center on Health, Physical Activity and Disability (NCHPAD)*.
35. Harris PA, Taylor R, Thielke R, et al. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 2009; 42: 377-381. 20080930. DOI: 10.1016/j.jbi.2008.08.010.
36. Spitzer RL, Kroenke K and Williams JB. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary Care Evaluation of Mental Disorders. Patient Health Questionnaire. *JAMA* 1999; 282: 1737-1744. DOI: 10.1001/jama.282.18.1737.
37. Goldschmidt AB, Loth KA, MacLehose RF, et al. Overeating with and without loss of control: Associations with weight status, weight-related characteristics, and psychosocial health. *The International journal of eating disorders* 2015; 48: 1150-1157. 20150914. DOI: 10.1002/eat.22465.
38. Cooper Z and Fairburn C. The eating disorder examination: A semi-structured interview for the assessment of the specific psychopathology of eating disorders. *International Journal of Eating Disorders* 1987; 6: 1-8. DOI: https://doi.org/10.1002/1098-108X(198701)6:1<1::AID-EAT2260060102>3.0.CO;2-9.
39. Wilfley DE, Schwartz MB, Spurrell EB, et al. Using the eating disorder examination to identify the specific psychopathology of binge eating disorder. *The International journal of eating disorders* 2000; 27: 259-269. DOI: 10.1002/(sici)1098-108x(200004)27:3<259::aid-eat2>3.0.co;2-g.
40. Plotz T, Chen C, Hammerla NY, et al. Automatic Synchronization of Wearable Sensors and Video-Cameras for Ground Truth Annotation -- A Practical Approach. In: *2012 16th International Symposium on Wearable Computers* 18-22 June 2012 2012, pp.100-103.
41. Gimhae G-N. Six human factors to acceptability of wearable computers. *International Journal of Multimedia and Ubiquitous Engineering* 2013; 8: 103-114.
42. Alharbi R, Stump T, Vafaie N, et al. I Can't Be Myself: Effects of Wearable Cameras on the Capture of Authentic Behavior in the Wild. *Proc ACM Interact Mob Wearable Ubiquitous Technol* 2018; 2: Article 90. DOI: 10.1145/3264900.
43. Alharbi R, Vafaie N, Liu K, et al. Investigating Barriers and Facilitators to Wearable Adherence in Fine-Grained Eating Detection. *IEEE Int Conf on Prevasive Computing and Communications (PerCom)* 2017.

44. Salvador Castell G, Serra-Majem L and Ribas-Barba L. What and how much do we eat? 24-hour dietary recall method. *Nutr Hosp* 2015; 31 Suppl 3: 46-48. 20150226. DOI: 10.3305/nh.2015.31.sup3.8750.
45. USDA FSRG. AMPM - USDA Automated Multiple-Pass Method, https://www.ars.usda.gov/northeast-area/beltsville-md-bhnrc/beltsville-human-nutrition-research-center/food-surveys-research-group/docs/ampm-usda-automated-multiple-pass-method/ (2021).
46. Schakel SF, Sievert YA and Buzzard IM. Sources of data for developing and maintaining a nutrient database. *J Am Diet Assoc* 1988; 88: 1268-1271.
47. Posner J, Russell JA and Peterson BS. The circumplex model of affect: an integrative approach to affective neuroscience, cognitive development, and psychopathology. *Dev Psychopathol* 2005; 17: 715-734. DOI: 10.1017/S0954579405050340.
48. Suh H, Shahriaree N, Hekler EB, et al. Developing and Validating the User Burden Scale: A Tool for Assessing User Burden in Computing Systems. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. San Jose, California, USA: Association for Computing Machinery, 2016, p. 3988–3999.
49. Lowe MR, Butryn ML, Didie ER, et al. The Power of Food Scale. A new measure of the psychological influence of the food environment. *Appetite* 2009; 53: 114-118. 20090612. DOI: 10.1016/j.appet.2009.05.016.
50. van Strien T, Frijters JER, Bergers GPA, et al. The Dutch Eating Behavior Questionnaire (DEBQ) for assessment of restrained, emotional, and external eating behavior. *International Journal of Eating Disorders* 1986; 5: 295-315. https://doi.org/10.1002/1098-108X(198602)5:2
51. Herman BK, Deal LS, DiBenedetti DB, et al. Development of the 7-Item Binge-Eating Disorder Screener (BEDS-7). *Prim Care Companion CNS Disord* 2016; 18 20160428. DOI: 10.4088/PCC.15m01896.
52. Stunkard AJ and Messick S. The three-factor eating questionnaire to measure dietary restraint, disinhibition and hunger. *J Psychosom Res* 1985; 29: 71-83. DOI: 10.1016/0022-3999(85)90010-8.
53. Alshurafa N, Alharbi R, Pfammatter A, et al. Detecting Real Time Episodic Overeating for Just In Time Interventions. *Annals of Behavioral Medicine* 2018; 52: S726-S726.
54. Mitchell JE, Crow S, Peterson CB, et al. Feeding laboratory studies in patients with eating disorders: a review. *Int J Eat Disord* 1998; 24: 115-124. DOI: 10.1002/(sici)1098-108x(199809)24:2<115::aid-eat1>3.0.co;2-h.
55. Engel SG, Kahler KA, Lystad CM, et al. Eating behavior in obese BED, obese non-BED, and non-obese control participants: a naturalistic study. *Behav Res Ther* 2009; 47: 897-900. 20090707. DOI: 10.1016/j.brat.2009.06.018.
56. Hall MA. Correlation-based Feature Selection for Discrete and Numeric Class Machine Learning. *Proceedings of the Seventeenth International Conference on Machine Learning*. Morgan Kaufmann Publishers Inc., 2000, p. 359-366.
57. Guyon I, Andr, #233, et al. An introduction to variable and feature selection. *J Mach Learn Res* 2003; 3: 1157-1182.
58. Breiman L. Random Forests. *Machine Learning* 2001; 45: 5-32. journal article. DOI: 10.1023/a:1010933404324.

59. Hutchinson RA, Liu L-P and Dietterich TG. Incorporating boosted regression trees into ecological latent variable models. *Proceedings of the Twenty-Fifth AAAI Conference on Artificial Intelligence*. San Francisco, California: AAAI Press, 2011, p. 1343-1348.
60. Pittman SJ and Brown KA. Multi-Scale Approach for Predicting Fish Species Distributions across Coral Reef Seascapes. *PLoS ONE* 2011; 6: e20583. DOI: 10.1371/journal.pone.0020583.
61. Johnson R and Zhang T. Learning Nonlinear Functions Using Regularized Greedy Forest.
62. Hastie T, Tibshirani R and Friedman J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition*. Springer, 2009.
63. Schapire RE. The Boosting Approach to Machine Learning: An Overview. In: Denison DD, Hansen MH, Holmes CC, et al. (eds) *Nonlinear Estimation and Classification*. New York, NY: Springer New York, 2003, pp.149-171.
64. Jiang W. On Weak Base Hypotheses and Their Implications for Boosting Regression and Classification. *The Annals of Statistics* 2002; 30: 51-73.
65. Alshurafa N, Xu W, Liu JJ, et al. Designing a robust activity recognition framework for health and exergaming using wearable sensors. *IEEE journal of biomedical and health informatics* 2014; 18: 1636-1646. DOI: 10.1109/JBHI.2013.2287504.
66. Selim SZ and Ismail MA. K-means-type algorithms: a generalized convergence theorem and characterization of local optimality.
67. Bottou L and Bengio Y. Convergence properties of the k-means algorithms. In: *Advances in Neural Information Processing Systems 7* 1995, Citeseer.
68. Shahapure KR and Nicholas C. Cluster Quality Analysis Using Silhouette Score. In: *2020 IEEE 7th International Conference on Data Science and Advanced Analytics (DSAA)* 6-9 Oct. 2020 2020, pp.747-748.
69. Sideris C, Shahbazi B, Pourhomayoun M, et al. Using electronic health records to predict severity of condition for Congestive Heart Failure patients. *UbiComp 2014 - Adjunct Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing* 2014: 1187-1192. DOI: 10.1145/2638728.2638815.
70. Sharma S, Jasper P, Muth E, et al. The Impact of Walking and Resting on Wrist Motion for Automated Detection of Meals. *ACM Trans Comput Healthcare* 2020; 1: Article 24. DOI: 10.1145/3407623.
71. Sharma S and Hoover A. The Challenge of Metrics in Automated Dietary Monitoring as Analysis Transitions from Small Data to Big Data. In: *2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)* 16-19 Dec. 2020 2020, pp.2647-2653.
72. Kalantarian H, Alshurafa N and Sarrafzadeh M. A Survey of Electronic Methods for Non-Invasive Detection of Eating Behavior. *IEEE Pervasive Computing* 2016.
73. Alharbi R, Zhang S, Spring B, et al. Machine learning algorithms applied to detect feeding gestures: Results from an eating study. In: *2016 Obesity Week* New Orleans, LA, 2016. Submitted.

**Appendix**

FoodTrck App public download

**<INSERT IMAGE 1>** **<INSERT IMAGE 2>**