Auditory Seasoning Filters: Altering Food Perception via Augmented Sonic Feedback of Chewing Sounds

Citation: Kleinberger, Rebecca, Van Troyer, Akito and Wang, Qian Janice. 2023. "Auditory Seasoning Filters: Altering Food Perception via Augmented Sonic Feedback of Chewing Sounds."

As Published: https://doi.org/10.1145/3544548.3580755

Publisher: ACM|Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems

Persistent URL: https://hdl.handle.net/1721.1/150616

Version: Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

Auditory Seasoning Filters: Altering Food Perception via Augmented Sonic Feedback of Chewing Sounds

Rebecca Kleinberger
Northeastern University
& MIT Media Lab
MA, USA
rebklein@mit.edu

Akito van Troyer
Berklee School of Music,
& MIT Media Lab
MA, USA
avantroyer@berklee.edu

Qian Janice Wang
Aarhus University,
& University of Copenhagen
Denmark
qjw@food.ku.dk

ABSTRACT

The experience of what we eat depends not only on the taste of the food, but also on other modalities of sensory feedback. Perceptual research has shown the potential of altering visual, olfactory, and textural food cues to affect flavor, texture, and satiety. Recently, the HCI community has leveraged such research to encourage healthy eating, but the resulting tools often require specialised and/or invasive devices. Ubiquitous and unobtrusive, audio feedback-based tools could alleviate those drawbacks, but research in this area has been limited to food texture. We expand on prior psychology research by exploring a wide range of auditory feedback styles to modify not only flavor attributes but also appetite-related measures. We present Auditory Seasoning, a mobile app that offers various curated audio modes to alter chewing sounds. In a Pringles-tasting experiment (N=37), this tool significantly influenced food perception and eating behavior beyond texture alone. Based on these results, we discuss design implications to create custom real-world flavor/satiety-enhancing tools.

CCS CONCEPTS

• **Human-centered computing → Auditory feedback**; *Sound-based input / output*; Ubiquitous and mobile devices.

KEYWORDS

food, auditory feedback, crossmodal correspondences, closed-loop system

ACM Reference Format:
Rebecca Kleinberger, Akito van Troyer, and Qian Janice Wang. 2023. Auditory Seasoning Filters: Altering Food Perception via Augmented Sonic Feedback of Chewing Sounds. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23), April 23–28, 2023, Hamburg, Germany.* ACM, New York, NY, USA, 15 pages. https://doi.org/10.1145/3544548.3580755

1 INTRODUCTION

The experience of eating is shared daily amongst most humans as a social, cultural, pleasurable and live-sustaining activity. A crucial part of body weight management, healthy eating involves considerations both in terms of food choice and food intake [37]. In terms of food choice, avoiding excessive intake of salt [20], sugar [53], and fat [10] is an important part of dietary recommendations from the World Health Organization. Behavioral interventions for healthier eating practices include eating at a slower rate [14, 66] and increasing the number of chews before swallowing [120]. When we eat, our experience of the food is not only dictated by our taste buds but also by many of our other senses. Perception research has demonstrated the important role of visual, olfactory, and tactile cues that can affect our perception of what we eat, in terms of texture, flavor, and sense of fullness. For instance, brighter or more contrasted colors make food look more appealing or sweeter [91] resulting in tools such as projection mapping [59] or AR [79] to affect eaters' experience. Satiety has also been shown to be influenced by changing the perceived size of food [76]. Smell can also shape judgements of food identity [106], taste intensity [42], and satiety [117]. HCI researchers have designed mixed reality experiences like the MetaCookie [77] or virtual donut [68] to explore how the addition of aroma alters the eating experience and perceived fullness. Textural cues also affect our senses of taste and satiety, with harder food making us feel more full [121]. This has led to the development of haptic devices to change our experience, including vibratory devices which attach to the tongue [85] or teeth [52]. Recent fabrication technologies have also opened the door to new types of food designs informed by perceptual research [69].

However, many of the aforementioned interventions require the use of cumbersome, invasive external hardware or the use of especially produced foods. These issues reduce the adoption factor of the technology and limit their potential as just-in-time interventions. Therefore, crucial questions arise: How can we design easily deployable, scalable, and ubiquitous interventions that can help users regulate their eating behavior in real-time, without being solely based on deliberate self-control? Can the food itself *tell* us what to eat and when to stop? Audio-based feedback may offer a key. In this paper, we lay down the groundwork for developing audio-based mobile interventions that focus on manipulating - and bringing attention to - eating sounds, where users are able to regulate their food perception and cravings, without the need for additional devices or deliberate self-regulation.

Indeed, contrary to visual, textural, or olfactory approaches, audio-based applications only require ubiquitous and rather unobtrusive hardware such as a simple pair of earphones. Research indicates that young adults habitually use their phones while eating [101], further indicating that a phone-based approach could have a higher adherence rate [87]. Besides, the various sounds we

CHI '23, April 23–28, 2023, Hamburg, Germany

ACM ISBN 978-1-4503-9421-5/23/04.
https://doi.org/10.1145/3544548.3580755

produce when eating (chewing, crunching, swallowing, etc.) play an important role in our perception of the food we eat as well as our response to it in terms of consumption rates, flavor assessment, and quantity consumed [89, 92, 93]. The purpose of this work is to introduce and evaluate the potential of "Auditory Seasoning Filters" (i.e., changing auditory feedback of chewing sound in real-time while eating) in influencing eating behavior and flavor perception. Closed-loop modified auditory feedback has been shown to have strong subconscious neurological effects during various behavioral tasks, including speaking [41], singing [56], breathing [43] and eating [118]. Adding effects such as time delay while speaking has been associated with slowed speech, slurred speech, reduced stuttering, or increased disfluencies depending on the chosen time delay. Promising work in the context of food has shown that applying a high pass filter on chewing sound in real-time has been associated with a higher sense of crunchiness or freshness in foods [118].

Although previous literature has demonstrated the undeniable links between sound and food perception, these works have often focused on single experiential aspects of food perception (flavor or texture) and have included, for the most part, external auditory stimuli such as background music. In this study and tool implementation, we aim to expand on previous crossmodal psychological research (e.g. [118]) by take a broader view of how modulating food-intrinsic auditory feedback parameters can influence different aspects of eating beyond flavor (e.g., deliciousness, fullness, craving, and eating speed). Our goal is to create a tool that can be easily used in real-life to augment specific eating parameters. Such a tool could then be used in various contexts, such as inducing the user to eat less salty foods, eat more slowly or in smaller quantities, and to be more present and aware of the food to avoid mindless eating.

Inspired by recent advances in elucidating the role of sound in food perception, we aimed to fill existing knowledge gaps around the effect of auditory feedback manipulations on eating sounds and the impact of sound augmentation on appetite and eating behavior. From a HCI perspective, we aimed to address the lack of non-intrusive off-the-shelf eating sound modification tools by designing an easy-to-use mobile app containing Auditory Seasoning Filters. After reviewing relevant literature and running a pilot experiment on eating sounds to inform our audio filter design choices, we specifically hypothesized that:

- H1: Volume amplification enhances food texture and perceived taste intensity,
- H2: Time delays reduce people's eating speed, therefore increasing their sense of fullness
- H3: Frequency filtering affects texture, but also flavor and satiety
- H4: Altering room reverberations affects the sense of fullness and self-awareness.

The rationale for these four hypotheses is first introduced in the Background section, then presented in detail in section 3.2. To test our assumptions, we built a custom iOS app called *Auditory Seasoning* that synthesizes various delays, altered room response reverbs, and filters with various parameters. Using this mobile app, we ran a pringle-tasting study on 37 subjects to collect self-reported effects on eating behavior, texture and taste. Our results show that altered auditory feedback of eating sounds can significantly influence subjective self-reported ratings as well as different aspects of eating experience.

Prior works in HCI, HFI (Human Food Interactions), and crossmodal perception have demonstrated the potential of using various senses to influence taste and eating experience [108]. From a tool-building perspective, audio is a particularly promising and under-explored medium [107], in terms of practicality and accessibility. Moreover, recent research in neurology and assistive technology on closed-loop audio systems suggests it may present advantages for intervention compared to open-loop systems. In this paper, we connect these fields and present new contributions. Our overall aim is to assess the feasibility of — and elucidate design parameters for — feedback-based auditory interventions that can implicitly help users regulate their eating behaviour. Building on this, our three main contributions are the following:

- First we present a survey of the fields of psychology, neurology and HCI with a particular focus on the potential of audio-based closed-loop tool building for food perception. To our knowledge, no such review has been done in a way that specifically bridges food perception, HFI, and HCI from the lens of auditory feedback .
- Second, informed by previous work, we ran a closed-loop Pringle-tasting experiment evaluating the effects of novel Auditory Seasoning Filters using a custom-made iOS app. Those filters include modes that have never been tested on food perception, as well as control modes with more expected outcomes. The design of the modes derives from the literature and provides a new perspective in applying knowledge from environmental auditory influences to food-intrinsic sound design.
- Third, we present design implications of this work, including potential intervention tools using custom auditory feedback filters based on the desired taste or satiety.

There have not been any prior iOS tools developed or tested to enable users to experience the effect of modulated auditory feedback on their everyday eating sounds. By providing an easy, scalable, and accessible tool, we are able to better test and understand the effect of modulated auditory feedback on the eating experience. We developed this iOS application containing 9 auditory feedback modes to test our hypotheses.

2 BACKGROUND

In this section, we review previous work on 1) food-related tools in the HCI community, 2) psychological research on sound and food perception, and 3) closed-loop auditory feedback systems and opportunities for crossmodal closed-loop tool design. Our current work offers a unique perspective by connecting and applying established research dealing with food-extrinsic environmental sounds to the novel area of food-intrinsic sound augmentation. Moreover, we situate our research in the theoretical framework of closed-loop auditory feedback systems, thus extending existing research on voice feedback to include eating sounds.

2.1 CHI & Food tools

Recently, the field of technologically-mediated human-food interaction (HFI) has emerged out of the general area of human-computer interaction, focusing on how technology can best support activities around food purchase, preparation, consumption, and disposal [5, 19, 32, 44]. HFI encompasses different communities of researchers from food and interaction design, multisensory human-food interaction, and AI-food/cooking interactions [5]. More specifically, developments in the area of eating can be broadly divided into three overlapping categories, a behavior-focused approach whereby technology is used to change the way people eat, a social approach that focuses on human-human/robot interactions, and a multisensory experience approach whereby the sensory experience of food is altered or augmented.

In the behavioral research approach, tools have been designed to encourage healthy eating, such as regulating eating speed. Some examples in this area are the Slowee, a wearable device that delivers light and vibration cues to the user if they eat too quickly [58]. Another example is the Sensing Fork [54], which vibrates to alert the user, and has been demonstrated in a small study to increase awareness and reduce eating speed in self-labeled "fast eaters" [50].

In contrast, social approaches to human-food interaction facilitate the user's interaction with other human diners or robot companions. In the first category, many interactive dining tables have been proposed over the years which either allow users to eat remotely together, such as the CoDine system [115], or enable them to interact with other co-located diners in new ways [30, 72]. For example, the Sensory Interactive Table uses embedded LEDs to allow diners to pass messages and play games with each other, with games designed to encourage vegetable eating or reduce eating speed [47]. Since loneliness is an increasing concern in the western world [49], an increasingly popular area of development is artificial eating companions, such as the myKeepon toy robot which mimics proper gaze behavior to either the users head, hand, or the food [70], or Fobo, a robot companion that mimics the user's eating movements [57].

Finally, the third category involves sensory augmentation. As eating is one of the most multisensory events that we all experience on a daily basis, human-food interaction is an ideal setting for the development of multisensory technologies [107]. Existing HCI research has, for the most part, focused on manipulating the visual appearance of food, such as color, shape, and visual texture. For instance, the AR "food changer" by Okajima and colleagues can identify and modify the appearance of food either using a projector [79] or a head-mounted display [102]. The findings of their experiments revealed, for example, a correlation between the color saturation and the rated sweetness in cake [79]; and between the visual texture and mouthfeel of sashimi [102]. Smell interfaces have also been developed to alter the flavor of foods, since smell is the largest determinant of flavour [90]. For example, the MetaCookie+ system, which combines a Virtual Reality Head Mounted Display (HMD) with an aroma delivery system, can alter the appearance and scent of a cookie in real time [77]. The researchers showed that, without changing the chemical composition of the food itself, 79% of participants experienced a change in the cookie taste using the pseudo-gustatory display. Going beyond taste perception, AR has also been shown to modify people's level of perceived satiety by altering the apparent size of the food consumed using real-time shape deformation [76].

It should be noted that all the devices mentioned above, although effective, often require invasive or additional custom devices. Given these restrictions, it is difficult for these interactions to be deployed in the real world on a wide scale [108]. Ubiquitous and unobtrusive, audio feedback-based tools could alleviate those drawbacks, but research in this area is limited. A notable example is the Chewing Jockey [62], a system that uses a bone conduction microphone and photo-reflector to detect chewing sounds, then plays back transformed sounds via bone conduction speakers. One scenario involved using a high pass filter to enhance the crispiness of foods, based on research demonstrating that potato chips are rated to be fresher and crispy when the chewing sounds from a higher frequency range are amplified [118]. Such a system has been implemented via EMG electrodes to make soft foods appear stiffer, which could potentially be beneficial for those on texture-modified diets to enjoy the eating experience more [39]. While promising, the Chewing Jockey system requires a specialized lab-based setup and cannot be easily deployed in the real world.

2.2 Food and Sounds

To fully capitalize on the potential of sound to alter the eating experience in HCI, it is important to review psychological research regarding how sound can influence food perception. The majority of the work in this area has revolved around sound and tactile properties, since food-related sound acts as an important guide in determining food texture [22, 35, 95, 112]. Mastication produces sounds which are then transmitted via both air and bone conduction, both of which impact the final texture assessment of food [29]. In a classic study by Zampini and Spence, potato chips were rated as more crispy and fresh when participants' chewing sounds were manipulated, either by raising the overall sound level or only amplifying the high-frequency components [118]. Similar effects have also been observed regarding the hardness and crispness of apples [31] and the expected carbonation level of soda water [119].

In recent years, a body of work has uncovered an array of sound attributes that correspond with basic tastes [92, 94]. For example, sourness is associated with high pitch and dissonant harmonies, whereas saltiness is associated with staccato articulation and auditory roughness [73, 111]. Going one step further, researchers and sound designers alike have explored the so-called "sonic seasoning effect", by which sounds congruent with a specific taste or flavor can enhance said taste in the food. For example, soundtracks with sweetness associations, which feature high pitch and consonant harmonies, have been shown to enhance the sweetness of various foods, such as toffee [26], chocolate [113], and beer [21].

Beyond taste or texture, sound can also influence people's cognitive and behavioral responses to food. For instance, eye-tracking research has shown that listening to music associated with eating a healthy meal (jazz, piano, etc.) can lead people to make healthier food choices, compared to music associated with unhealthy eating (rock, brass, etc.) [82]. There is also evidence to suggest that sonic cues can influence appetite; for example, real-world studies have shown that consumers at a pub tend to drink more and at a faster rate, when the volume of background music is louder [45]. In contrast, listening to slower tempo music leads to longer eating time [71] and presumably greater satiation [64].

Many of the above-cited studies, however, involve playing background music, which people are explicitly aware of. There is a lack of research on how findings from such food-extrinsic cues - like the effect of pitch or timbre - can be translated to auditory manipulations of eating sounds. This is an important area of research as this will broaden our knowledge of sound-food interactions and give designers more flexibility to create targeted food-related auditory experiences. At the same time, aside from food texture studies [31, 118, 119], there has been little research focusing on how augmenting natural eating sounds might influence other food-related factors such as flavor, appetite, and eating behavior. From a consumer psychology perspective, there is an opportunity to create auditory feedback tools that manipulate such naturally-occurring eating sounds in order to promote implicit behavior change.

2.3 Closed-Loop Altered Auditory Feedback

Beyond the use of open-loop systems which produces auditory cues independent of the user's action (sonic backgrounds, notes, music, or soundscapes), the use of closed-loop systems using real-time Digital Signal Processing (DSP) to directly transform self-produced sounds, have shown high potential with perception researchers and across the HCI community to affect different levels of behavior and deep internal processes. The added value of the closed-loop approach may come from reaching the right level of incongruence between internally predicted expected signals, and externally modified feedback signals. For most voluntary motor commands resulting in sound-producing activities, those two signals - the internal efference copy and external sensory feedback - are constantly compared in the brain. This comparison is hypothesized to take place in the basal ganglia [15], a group of subcortical nuclei responsible for motor control and learning, executive functions and behaviors, and emotions [4, 11]. Although this has not been formally established for eating sounds, the critical role of basal ganglia in complex movement [81] and its connection to motivated behaviors [80] and food-related disorders [67, 86] open potential avenues for merging cross-modal and closed-loop auditory feedback research. Examples of auditory false feedback include cardiac, breathing, body tapping/stepping sounds, and vocal/buccal/eating sounds. False heart rate feedback can affect subjects cognitively [104], emotionally [38] and has been shown to manipulate sexual arousal [16]. The HCI community has leveraged these findings in building tools for emotion regulation for anxiety [24], stage fright [75], or affection facilitation [78]. In addition to heart rate-based systems, HCI tools based on altering breathing sound feedback have been used to increase the sense of calm [43]. Auditory feedback of self-produced sounds has also been shown to contribute to body representation and control. Spatial modifications of tapping sounds can influence subjects' perception of body lengths [99] and altered footstep sound can change body weight perception [97] and has been used in tools to improve walking abilities [8, 23].

In the specific context of vocal sound manipulation, previous work had shown a wide range of potential effects of changing the way we hear our own voice, and the HCI community has leveraged them to create new tools and technologies. These effects range from mechanical control disturbances, local sensory mechanisms, affect, emotion, and higher-order neural modulation. Indeed, modulated voice feedback can cause distraction and has documented effects on speech control (speed, articulation, and fluency). For instance, a short delay added to the voice can lead to prolongation of vowels, repetition of consonants, higher utterance intensity, and other articulatory changes [41, 116]. A delay longer than 200ms often leads to jammed speech [40]. The speech jammer is a well-known use of these findings [63]. However, a delay of 20 to 150ms can increase fluency for people who stutter [55] In the case of stuttering, the neural basis for this effect is not completely understood. It is possible that the creation of large errors between expectations and modified feedback may alter reliance on the feedback signal and reduce the excessive motor speech repair mechanisms thought to underlie speech disfluencies [46] New technologies and devices have been developed leveraging this effect using simple delays and shifts in pitch [6] but also more complex musical modulations [60].

Besides speech control and expectation regulation, false voice feedback has shown potential for emotion and affect regulation. Study participants whose voices were modified to sound calmer during couple conflicts reported feeling less anxious [25]. Similarly, covert voice manipulation can significantly affect emotional states [7] and might even offer potential as a tool to improve negotiation outcomes [9]. Making the voice feedback sound more musical has also shown potential in affecting the emotional valence of the speaker and semantic content of speech as well as prosodic parameters [61].

Anchored in HCI best practices for building food augmentation tools, this project thus leverages psychological research from both the fields of multisensory perception and closed-loop auditory systems.

3 DESIGN CHOICES AND PRELIMINARY OBSERVATIONS OF PRINGLES EATING SOUNDS

This section covers the motivations behind our design decisions, including 1) the choice of food sample to use for tasting/testing, 2) the design of auditory feedback modes, and 3) initial insights collected from Pringles eating sound analysis.

3.1 Choice of Food

In this testing experiment, we used sour cream and onion-flavored Pringles as our unique food sample. This choice was motivated by several factors. We choose to focus on a crispy food sample because crispiness is rated as a highly desirable property [88] and is perceived through a combination of auditory and tactile feedback [36]. Various research studies have used chips as food samples [118], and Pringles offer a particular advantage in their consistency and global availability. We choose to use the sour cream and onion flavor as it allows us to decouple the aspects of taste perception into four parts: texture, saltiness, sourness, and overall flavor intensity (onion). The sourness was particularly relevant as previous crossmodal correspondences studies have explored the relationship between sound and sourness perception as well as saltiness perception [73, 111].

Table 1: Description of auditory modes in terms of implementation and perceptual effects. Literature-based hypotheses refer to findings based on environmental sounds that we aim to apply to chewing sound modifications.

Label	Mode Name	Implementation	Perception	Literature-based hypotheses
A, J, Off	No sound	- No sound output from the system	Sounds slightly dampened from wearing earphones	Less sensory intensity (taste, mouthfeel, flavor) [31, 118]
B	Raw Voice	- Passthrough 0dB (Unity Gain) - Delay 16ms	Playback, quite similar to real-world hearing experience	baseline
C	Amplified	- Passthrough 6dB - Delay 16ms	Four times louder than the Raw Voice mode	Greater sensory intensity (taste, mouthfeel, flavor) [22, 31, 118]
D	Short Delay	- A single delayed copy of mic signal - Delay 200ms	Barely perceptible delay	Slightly greater fullness and less craving [64]
E	Long Delay	- A single delayed copy of mic signal - Delay 400ms	Clearly perceptible delay	Greater fullness and less craving [64]
F	Small Room Reverberation	- Open Source Freeverb library - Room Size = 0 - Damping = 0.5 - Dry/Wet Level = 0.25/0.75 - Stereo width = 0.5	Akin to being in a small cupboard	More crispiness and saltiness [111]
G	Large Room Reverberation	- Open Source Freeverb library - Room Size = 0.9 - Damping = 0.5 - Dry/Wet Level = 0.25/0.75 - Stereo width = 0.5	Akin to being in a very large empty room (e.g., cathedral), with long reverberation time	Less crispiness and saltiness [111]
H	Low Pass Filter (LPF)	- 2-pole low pass filter - Slope of -12 dB per octave - Cutoff frequency of 2000 Hz	Only low-frequency sounds are heard	Less sourness and crispiness, greater fullness [27, 84, 118]
I	High Pass Filter (HPF)	- 2-pole high pass filter - Slope of -12 dB per octave - Cutoff frequency of 2000 Hz	Only high-frequency sounds are heard (high pitch noise, cracking sound, etc)	More sourness and crispiness [27, 118]

3.2 Choice of auditory feedback modes

Amongst the latent space of possible auditory transforms, we focused on four types of sound transformations: volume, delay, room response modulation, and frequency filtering. Hypotheses for each mode are summarized in Table 1.

The influence of **volume** was motivated by Drake's suggestions that the increase in sound amplitude is more apparent to determine food crispiness than any concurrent changes in the frequency spectrum [34]. Beyond louder volume enhancing crispiness [118], we also expect a louder volume to enhance other food characteristics including flavor intensity. This is based on the theory of magnitude framework [110], which suggests that increased magnitude in one sense modality can lead to higher intensity perceptions in another modality. For instance, eating in a room with brighter illumination enhances taste sensitivity [96] and leads to stronger overall flavor intensity [105].

The potential effect of **frequency filtering** was motivated by Darcemont's observation of frequency characteristics of low-moisture crisp products [28] and Zampini's 2004 study [118]. However, the above-mentioned studies only tested high-frequency amplification and did not investigate the effect of low-pass filtering. Moreover, given that sourness is associated with high pitched sounds [114], we would expect high-frequency filtering to emphasize the sour taste in the chips. On the other hand, since low pitch is associated with heaviness [109], which is in turn linked with greater food satisfaction [84], we have a good reason to believe that low-frequency filtering could also enhance satiety.

Testing the effects of **room response** was motivated by its common real-life occurrence and its role in auditory scene analysis [17]. Room response modification has been studied extensively in the context of speech understandability [12] cognitive performance [33], musical hall design [3] and is often used in music production [103]. Real-time modification of room reverberation has recently raised the interest of the HCI community and cognitive neurology researchers for its potential to affect unconscious neuronal processes. Artificial room response manipulations have indeed been associated with emotional responses, with acoustics of smaller rooms being considered more pleasant, calmer, and safer than bigger rooms [98]. By creating a sense of a larger or smaller space, where sounds are reverberating and we are more or less likely to be heard and observed, we suspect that we can affect the sense of fullness and self-awareness.

We chose to evaluate the effects of various **Delayed** Auditory Feedback (DAF) because of its known effects on speech and movement control. Depending on the delay length, DAF can lead to disturbed articulation, slowed speech, and greater sound pressure [41]. The disturbance caused by DAF depends on various factors such as individual variability, delay length [51], action type, and action trajectory [83]. Although the time ranges for DAF are well established for speech sound (effects but no control disturbance below

50ms, effects and control disturbance and slowed speech between 50 and 200ms, dissociation between feedback and motor action after 200ms) [46], our prior observations suggest that those times thresholds are different for the perceptual processing of eating sounds. This may be because of the more chaotic aspects of eating sounds versus organized speech sounds, and also the lower importance to monitor these sounds in the brain, however, this would require further validation. Based on our pilot testing (see section 3.3) we found that 200ms, although very large for speech, is only barely noticeable when eating. We chose this threshold as our short delay parameter, where the delayed eating sound still perceptually seems self-produced. Moreover, we chose 400 as our high delay parameter because the delayed chewing sounds were clearly disassociated from the motor movements. We expect that hearing delayed chewing sounds will reduce people's eating speed, therefore leading to more fullness [64].

3.3 Chewing sound analysis

To better understand how auditory alteration changes the perceptual experience of chip eating sounds [35], we first created a database of Pringles eating sounds from 5 people. The database was used to analyze the diversity of sonic eating behaviors and frequency composition between users and throughout the eating experience.

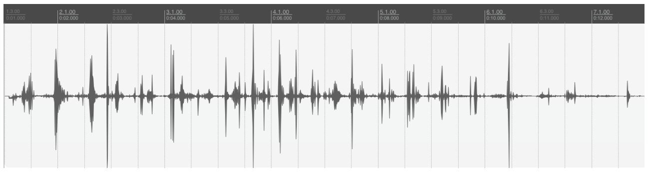


Figure 1: An audio example of eating a single Pringles chip. The recording is about 13 seconds long and each vertical line represents 1/2 second. The average RMS power is -40.2 dB. A rhythmic chewing pattern can be observed.

The recordings were gathered in a quiet room, using a pair of wired earphones with an embedded microphone (EarPods with Lightning Connector from Apple) and by asking the subjects to manually hold the microphone at a distance of 1cm in front of their mouth while eating one chip at a time. Figure 1 shows an example of such a recording. We gained several design insights from these experiments:

- The setup resulted in perceptually similar quality and amplitude recordings confirming that the instruction of holding the microphone could lead to comparable experiences.
- We validated the choice of 2kHz as a suitable cutoff frequency for our frequency filter and also observed interesting effects of low pass filtering under 2kHz of eating sound. Although less crispy, the chip seemed to have more presence in the mouth.
- Frequency filtering appears to not only affect texture but also flavor and satiety
- We noticed low intra-subjects variability in loudness, eating time, number of chews, and chewing speeds between the trials. (Intra-subjects CV= 6.% for chewing times, 6.8% for the number of chews, and 3.6% for chewing speed). People seem to always eat at roughly the same speed. This suggests each individual is quite consistent and that if a difference is individually observed, it would likely result from sound intervention rather than natural variability.
- In terms of inter-subjects consistency, we observed a good consistency in loudness. The time it takes to eat one chip ranges from 9 to 20 seconds, and the tempos range from 1 to 2 chews per second.

4 THE AUDITORY SEASONING APP

To test the effects of various sound alterations on food perception and eating behaviors, we designed a mobile app, called Auditory Seasoning, that incorporates the four audio feedback categories presented in section 3, each with different parameters. The use of mobile phones and wired earphones makes the system available to a very large number of people without requiring additional hardware.

4.1 Microphone and earphones

Auditory Seasoning is a mobile app designed to apply Digital Signal Processing (DSP) to incoming microphone signals and deliver an augmented audio output to the earphones. For this reason, the app assumes and recommends the users to be wearing wired earphones with an embedded microphone to keep tight control over delay and allow for near-real-time processing. We used EarPods with Lightning Connector from Apple.

4.2 Mobile phone

The study required the use of an Apple iOS device (iPhone or iPad). The Auditory Seasoning app designed for these iOS devices can run on an iOS version 9.3 or above. The app was distributed through Apple's App Store so subjects could download it onto their phone for the experiments. We provided the iOS device for some of the subjects who did not already own one. We tested the app to run on iPhone 8 and above.

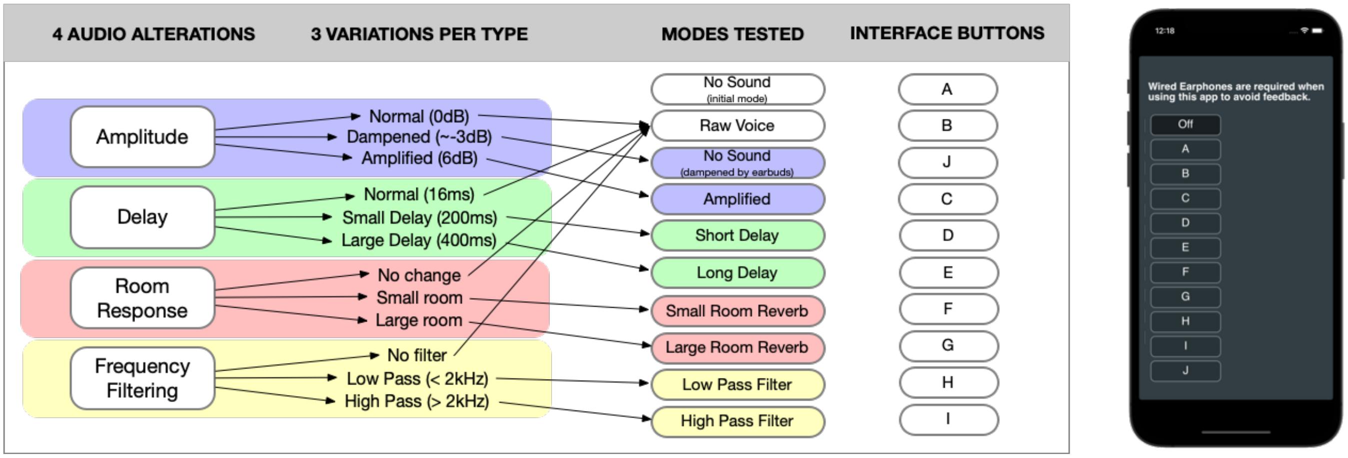
4.3 Audio processing

The app provided a total of 9 options that represented various audio transformation modes to the subjects with different parameters. Table 1 presents a brief description of each mode in terms of how they were implemented as well as perceptually in regard to how they sound to the user and how they relate to prior literature. The overall latency between the sound signal captured by the inline microphone to the audio signal going out of the earphones was 16 milliseconds across all modes.

4.4 Interface

To design the app, we first implemented each mode with interactive parameters and had the app tested by three audio experts. After the experts approved our chosen parameters and the quality of the audio output, we implemented a blinded version of the tool for use in our study.

Figure 2 shows the user interface that the subjects used during the experiment. The interface for the mobile app was kept minimal to avoid potential visual distraction during the experiments and to keep subjects blinded as to which mode they were using. For



Wired Earphones are required when using this app to avoid feedback.

Figure 2: Modes mapping and iPhone Interface: We choose and mapped audio feedback effects that could all be combined into 3 connected modes for each of our four hypotheses (amplitude, delay, room response, frequency filtering). However, the interface was meant not to reveal the mapping and the order was presented in randomized order to avoid bias.

this reason, the interface is composed of 12 buttons at the left of the screen to select a current task along with a text at the top encouraging the subjects to wear earphones while using the app.

5 USER STUDY

5.1 Participants

We recruited 37 subjects (19 females, mean age = 30.9 years, SD = 8.6) via social media and mailing lists advertised to several universities. Inclusion criteria included a normal sense of hearing, no active dental issues, and no dietary restrictions or allergies to the Pringles flavor used (sour cream and onion). The study protocol received exempt approval from our institution's ethics review board (Exempt ID: E-3343) The study was carried out according to the Declaration of Helsinki, and all participants gave informed consent at the beginning of the study. No compensation was offered to participants.

5.2 Protocol

During the study, subjects were guided through a series of eating tasks under different auditory feedback testing conditions, each followed by a short taste-perception self-assessment. The modes were presented in random order, but the initial mode (A) was always the same so that all participants had a practice baseline session with no sound, to get them used to eating with a microphone and answering questions.

For each mode, participants had to 1) select the mode in the app as specified by the online questionnaire, 2) eat one chip while holding the microphone directly in front of their mouth, 3) rate their eating experience and 4) drink a sip of water before passing to the next mode. The study concluded with a short demographics questionnaire with an open comment section. This protocol is described in Figure 3

The eating experience was evaluated in the following way: Participants were asked to first evaluate their tasting experience in terms of: overall deliciousness, overall flavor intensity, sourness, saltiness, crispiness, mouth-filling, level of fullness after tasting the chip, and level of craving for eating another chip. These were presented on a 1-9 scale (1 = not at all, 9 = extremely). The questions were asked in the form of: "Please evaluate your tasting experience: Overall deliciousness" and they were presented with a slider from 1 to 9. Next, participants gave a qualitative free-text description of how the specific mode changed their eating behavior, if at all (e.g. eating speed, number of chews). Finally, subjects rated how the auditory mode enhanced their attention in four ways: attention to the food in general, attention to the flavor of the food, attention to the texture of the food, and self-consciousness of their eating sounds. These four measures had good internal consistency (Cronbach's alpha = 0.84) so a single averaged index of attentiveness was used for data analysis.

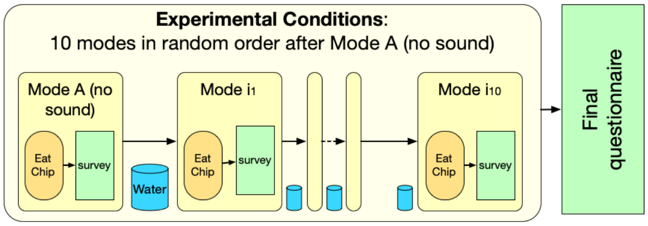


Figure 3: User Study Protocol

5.3 Setup and procedure

Subjects were asked to participate in a 30-minute session consisting of eating one Pringle chip at a time while using various modes of the Auditory Seasoning app. The subjects were guided by explicit instructions from a Qualtrics survey running on a different device (laptop or tablet) that also collected their answers. The subjects were asked to fast at least 2 hours prior to the experiment. The setup, illustrated in Figure 4, consisted of:

(1) an iOS device (version 9.3 or above) charged at least 50% running the Auditory Seasoning Filter app downloaded from the App Store, volume set up in maximum.
(2) a pair of wired earphones with an embedded inline microphone (EarPods with Lightning Connector from Apple). During eating tasks, subjects were instructed to manually hold the microphone close to their mouths while eating (approximately 1 cm).
(3) a laptop/desktop computer connected to the internet to access the online questionnaire.
(4) a large glass/bottle of still water
(5) an unopened can/cup of sour cream and onion-flavored Pringles.


[redacted]

Figure 4: User study setup describing the different elements needed including earphones, microphone, Pringles, laptop, iPhone and water. Participants were instructed by the online questionnaire on the laptop to select a mode on the app, then eat a potato chip while holding the microphone close to their mouths.

6 RESULTS

6.1 Effect of auditory modes on potato chip evaluation

According to our hypotheses, we tested the effect of different modes grouped thematically: the effect of modulating volume (modes J, B, C), delay (modes B, D, E), reverberation (modes B, F, G), and frequency range filtering (modes B,H, I). Mode B (raw voice) was always included as a baseline comparison. For each group, we applied a one-way repeated measures multivariate analysis of variance (rm-MANOVA) with the relevant modes as the within-subject factor, and evaluation ratings as dependent measures (deliciousness, crispiness, mouth-filling, flavor intensity, saltiness, sourness, fullness, and craving). A MANOVA was used to correct for multiple tests among the dependent measures

6.2 The effect of volume modulation

Figure 5A shows the chip evaluation ratings for the three loudness levels (no sound, baseline, amplified). We found a significant overall effect of mode (F(16,21) = 3.55, p = .004, partial η^2 = 0.73), and follow-up univariate ANOVAs revealed the effect of mode on the ratings of deliciousness (F(2,72) = 5.10, p = .008), crispiness (F(2,72) = 22.82, p = .008), mouth-filling (F(2,72) = 7.79, p = .001), overall flavor intensity (F(2,72) = 18.92, $p < .001$), saltiness (F(2,72) = 6.84, p = .002), sourness (F(2,72) = 12.30, $p < .001$), and fullness (F(2,72) = 3.81, p = .027). Significant differences via post-hoc Bonferroni-corrected tests can be seen in Figure 5A. Overall, compared to the baseline, amplifying eating sounds leads to higher intensity of overall flavor and specific tastes (salty, sour), more crispiness and mouth-filling sensations, and a higher degree of fullness ($p < .05$ for all comparisons).

6.3 The effect of delay modulation

Figure 5B shows the chip evaluation ratings for the different delay times (baseline = 0 ms, 200 ms, 400 ms). We did not observe a significant main effect of mode (F(16,21) = 3.55, p = .004) and thus did not perform any further follow-up univariate ANOVA tests.

6.4 The effect of reverberation modulation

Figure 5C shows the chip evaluation ratings for the three room response reverberation settings (baseline, small, and big room). There was a significant overall effect of mode (F(16,21) = 2.22, p = .044, partial η^2 = 0.63) on ratings of deliciousness (F(2,72) = 3.26, p = .044), sourness (F(2,72) = 3.36, p = .040), and craving (F(2,72) = 10.04, $p < .001$). Significant differences via post-hoc Bonferroni-corrected tests can be seen in Figure 5C. Compared to the baseline, applying large room reverberation effects to chewing sounds leads to decreased sour intensity (p = .049) and decreased craving (p = .001).

6.5 The effect of frequency filtering

Figure 5D shows the chip evaluation ratings for the three frequency filtering settings (baseline, low, and high pass filters). There was a significant overall effect of mode (F(16,21) = 2.17, p = .048, partial η^2 = 0.62) on ratings of crispiness (F(2,72) = 13.47, $p < .001$), mouthfill (F(2,72) = 4.94, p = .014), salty (F(2,72) = 4.25, p = .018), sour (F(2,72) = 3.50, p = .035), and fullness (F(2,72) = 7.36, p = .001). Significant differences via post-hoc Bonferroni-corrected tests can be seen in Figure Figure 5D. Compared to the baseline, applying a low pass filter (<2k Hz) to chewing sounds leads to decreased crispiness ($p < .001$), increased mouth-filling sensations (p = .033), and increased fullness (p = .020). In addition, applying a high pass filter (>2k Hz) also leads to increased fullness (p = .002) compared to the baseline condition.

6.6 Effect of auditory modes on self-reported eating behavior

Table 2 shows the number of subjects who reported seeing no change, a reduction or an increase in their eating speed in the open comment section. Not all participants reported in this section. We observed several trends suggesting that in the no sound mode, most

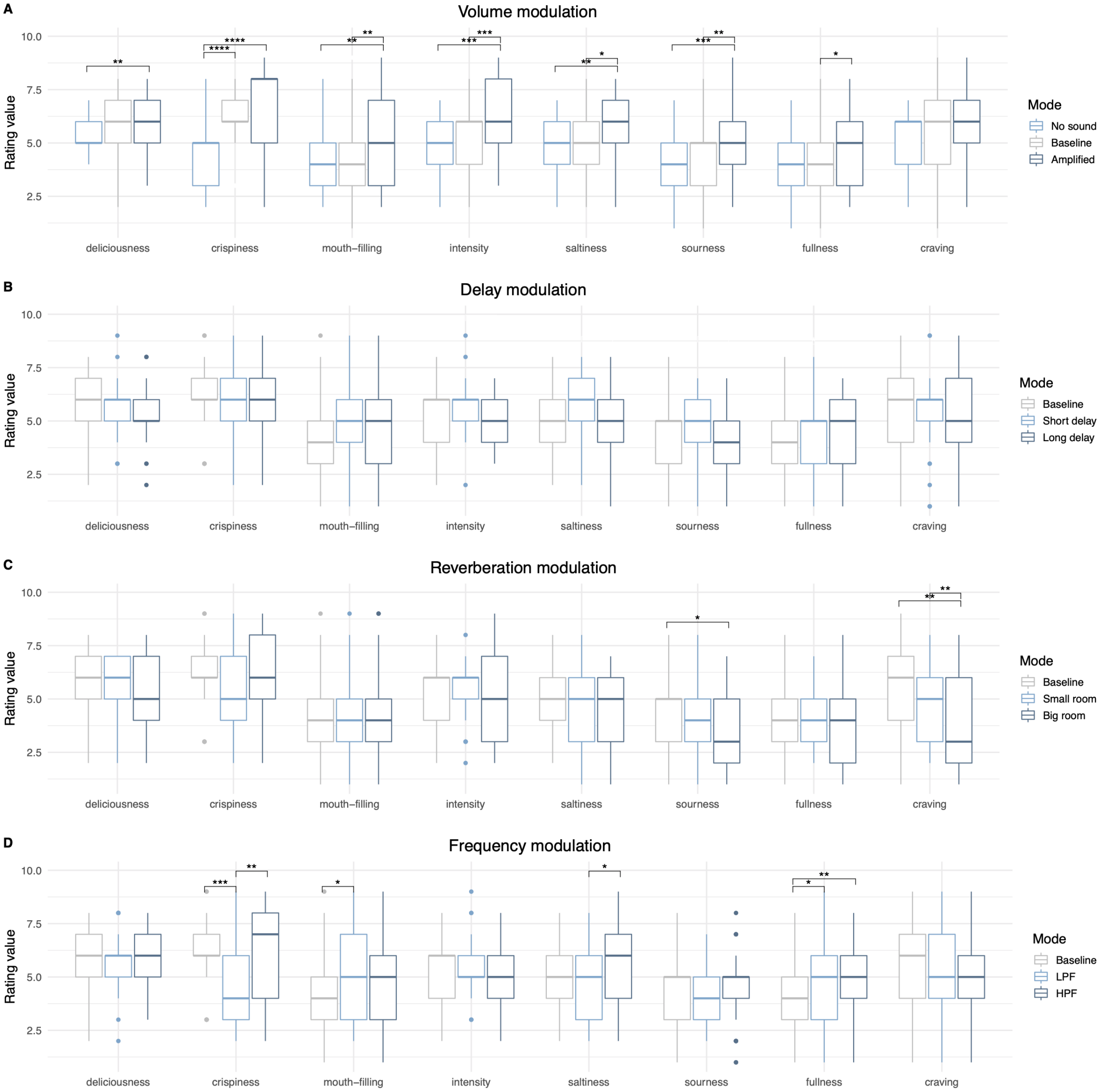


Figure 5: Pringles evaluation ratings as a result of audio feedback modulation by: A)volume, B) delay, C) reverberation, D) frequency range. Brackets indicate significant pairwise comparisons with Bonferroni correction (* p < .05, ** p < .01, *** p < .001, **** p < .0001).

participants reported no change in eating speed. Most other modes led to a majority of the reported decrease in eating speed with Long Delay being most reported as inducing slow down (17 people), followed by Amplified (16), Raw Voice (15), Large Room Reverb (14), 200ms Delay & Small Room Reverb (both 13).

6.7 Effect of auditory modes on attention to food

An analysis of auditory modes on the index of attentiveness via one-way rm-ANOVA revealed a significant effect of mode ($F(5, 197) = 9.34$, $p < .001$, partial $\eta^2 = 0.21$). As shown in Figure 6, follow-up post-hoc comparisons with Bonferroni correction illustrated that attentiveness to eating is significantly lower in the sound-dampened

Table 2: number of subjects who reported seeing no change, a slow down, or an increase in their eating speed in the open comment section

	(A) No sound	(J) No sound	Raw Voice	Amplify	Short delay	Long delay	Reverb small	Reverb large	LPF	HPF
no change	27	26	11	2	11	7	9	3	10	11
faster	1	4	5	5	7	5	8	6	4	6
slower	7	3	15	16	13	17	13	14	12	10

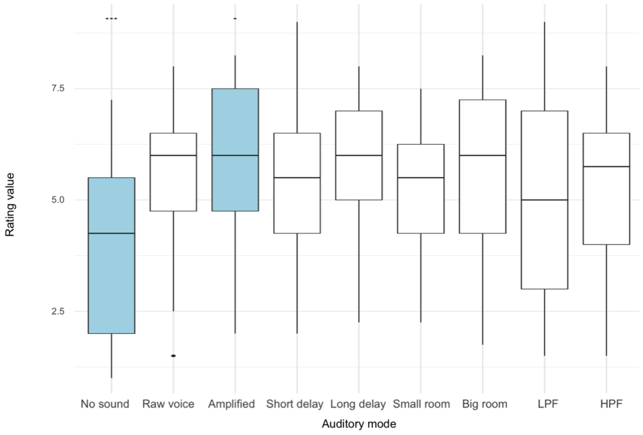


Figure 6: Attention to food (as rated by participants from 0 to 5) as a result of different modes of audio feedback modulation. * indicates significant differences compared to all other modes (i.e., comparison against the base mean), * $p < .05$, *** $p < .001$

mode compared to all other modes ($p < .001$); in contrast, attention in the amplified sound mode is significantly higher ($p = .03$).

6.8 Usability and Overall Experience

When asked to report on potential use cases of the tool in their everyday life, the theme most mentioned by participants was eating speed with 57% of subjects saying it could help them eat slower. The next most common theme was mindfulness with 46% of subjects suggesting it could increase awareness or attention to food, and reduce mindless eating. Another common thread regarded the quantity of food eaten and cravings with 22% of people saying it could help reduce cravings and eat less. 22% of people reported the tool could be used to make more healthy food choices (e.g. less sugar or salt). Finally, two subjects mentioned eating as a social experience and potential issues of eating with earphones. Another two subjects mentioned the potential of the tool for children to make vegetables taste less bitter or in gamifying eating. Table 3, summarizes subjects' comments according to four themes: eating speed, attention, craving, and healthy food choices.

Table 3: Subjects comments according to four themes: eating speed, attention, craving, and healthy food choices

Eating speed (57%)
"I could see myself using that to eat slower " (P30)
"I am pretty sure I was eater slower with some options" (P3)
Awareness and Attention (46%)
"It definitely made me more attentive and also appreciate the food more" (P2)
"the healthy distraction could bring the focus on different attention could help to eat less [...] not just mindlessly eating"
"It could be used to make yourself more aware of your chewing speed. So if you had an ambition to slow your chewing this app could be used to remind yourself." (P21)
"I could hear my saliva, and I am never aware of that and it was super interesting [...] I learned a lot about an everyday common experience and it made it more interesting" (P10)
Craving/eat less (22%)
"[it] slowed me down [...] With an addictive food like Pringles it would probably cause me to eat fewer chips" (P7)
"Maybe also eat less even, because it can be an intense experience" (P9)
Healthy foods/ less sugar & salt (22%)
"Some options really changed the taste that it pretty incredible" (P2)
"Overcoming craving for salty/savory food like crisps by 'neutralizing' their taste" (P6)
"Less sugar/salt while maximising flavour would be great. " (P8)
"the audio changed my perception of salt quite a lot [...], it could maybe help eat less salt because it is unhealthy" (P31)
"Could be fun try with different drinks/beers to make you drink slower and perhaps less" (P13)

When asked to report on the comfort of eating while wearing microphone and headphones, 7 out of 37 participants reported extremely comfortable, 14 reported somewhat comfortable, 11 reported neither comfortable nor uncomfortable, and 5 reported it was somewhat uncomfortable. No participants reported that it was extremely uncomfortable. 29 out of 37 participants thought the sounds varied "a great deal" or "a lot" between the different auditory modes, 7 thought the sounds varied "moderately", and only one thought the sounds did not vary at all between conditions. When asked to report their favorite mode, Large Room Reverb and High Pass Filter ranked equally highest each with 11 points. Small Room Reverb, Amplify and Delay all obtained 3 points while No Sound, Raw voice and Low Pass all obtained 2 points.

7 DISCUSSION

In this section, we discuss our findings, summarised in Table 4 and suggest possible implications for theory and design.

7.1 Interpretation of findings

From the empirical study, we found that **amplifying eating sounds** not only increases crispiness [118], but all flavor-related parameters including saltiness, sourness, mouth-filling, and overall flavor intensity. Louder eating sounds also led participants to feel more full. This is consistent with the theory of magnitude framework, since an increase in one sensory modality (sound) led to subsequent higher ratings in taste, texture, and flavor (smell+taste).

Participants' comments reflected this hypothesis of general amplification: "This sounded great like powerful, I felt like in an ad for chips", " I chewed slower because of the wider variety of the saltiness/sweetness/sourness sensations I could feel with my tongue", "it felt very surprising, as a new way to discover what I had in my mouth, I really liked it, it felt really alive and super exciting, as if someone was helping me to taste it better." Participants' reaction suggests major enthusiasm for this mode and positive user experience: "it's like those sounds were just for me and this chip was just for me to enjoy", "I feel like I wouldn't mind hearing and tasting this again, as if I wanted to hear myself bite in the chip again.", "This mode was awesome. I was smooshing the chip much more to enjoy the sounds of myself eating" Overall attention to food also followed volume levels, with participants in the no-sound condition reporting the lowest attention to food, and the amplified sound condition the highest. This has a two-fold implication: first, attention could be the mediating factor behind the observed effects, and secondly, auditory feedback could be utilized to modulate the diner's attention. Either way, this effect opens the door to various possible applications. Using carefully curated amplifications could raise awareness of food and potentially help reduce mindless eating. If the amplification were to happen gradually, it could also bring attention to overeating. Based on the time of day, a gain factor could be added to amplify sound when eating between meals making chewing sounds louder than when eating at regular meal times.

In terms of **delay**, we did not observe any effects on perceived texture or flavor. When asked to comment on potential changes in eating behaviors, participants' self-reported comments suggest a decrease in eating speed. Moreover, participants reported out-of-body experiences where they felt like they were eating with someone else, or that the chips tasted less flavourful. "It felt like it had more crunch, like I was eating with others, maybe I chew longer", "gave me the sensation that I was listening to someone else's eating. The chip became quite bland and it wasn't a very tasty experience", "it felt like I had two chips in my mouth". These results could be leveraged for specific applications. Future work on delayed auditory feedback could unveil possible thresholds for motor command disturbance of jaw motion similar to speech production. Such work would lead to the development of tools that make users unconsciously slow down their eating, but caution is required to ensure that motor disturbance won't lead to swallowing issues.

Large room reverberation made the chips taste less sour and led to less craving. This may have been due to the unfamiliar/ incongruent nature of the large room, which reduced appetite. Bregman's work on auditory scene analysis, mentioned in section 3, suggests that room reverberations contribute to our ability to make sense of our environment from an audio stream and affect our overall sense of space and safety [13]. By informing us about our environment in more or less congruent ways, mediated reverberations affect both perception and response behaviors. [18]. The results from our study may derive from one of two consequences of this. First, it may be linked with primitive reactions to environmental incongruence between sonic and visual/tactile feedback, leading to increased fight or flight reaction, and a more alert and attentive behavior less prone to eating. Some of the participants' comments lean toward this hypothesis "There was a disconnect between my mouth and my ears as if they did not belong to the same person." On the other hand, the decrease in craving could be linked with more culturally-informed behavior of restraint in large rooms when one can be observed and therefore is more self-conscious. This is also supported by some of the participant's comments regarding the large room reverb: "made the eating experience feel performative", and "I felt like everybody was looking at me! It made me feel very exposed [...] so it doesn't make me want another one.", or "I felt like I was in a ceremony of chips eating!" This can be compared to some comments regarding small room reverb: "I enjoyed this one, it felt a little intimate and personal, almost like if I were protected". These findings could be leveraged in designing curated eating experiences. Large room responses could be used similarly as sound amplification, to increase self-awareness when eating and reduce mindless eating. But by offering more parameters, this effect could also help transform people's relationship with their eating experience. For instance, by reducing sourness, large room reverberation could be used to habituate children and picky eaters to gradually broaden their food spectrum. Small room responses could help create a sense of comfort and safety and increase food intake for people with eating disorders, people with cognitive impairment, or elderly adults who might suffer from appetite loss. Varying room responses between courses could also create an architectural feeling of a journey throughout a gastronomic meal to enhance specific aspects of the food.

As hypothesized, the **low pass filter** led to less crispiness, more mouth-filling texture, and more fullness after eating. Interestingly, both high and low pass filters led to more fullness, which suggests that fullness may be mediated by attention to the food. Comments from participants supported this idea of increased attention to the food, both with the Low Pass Filter: "I ate slower, and felt compelled to close my eyes," "It really made me focus on the liquid sounds within my mouth, so [...] I let the chips remain in my mouth longer", "I was more aware of/listening more to the background noise" but also with the High Pass Filter: "the sound made me more conscious of the noises in my mouth", "It made me think more about the texture and sound of the chips. Probably made me eat slower". However, other comments suggest that increased fullness might be linked to different causes for high versus low pass filters. Indeed, participants reported their experience with High Pass similar to the Amplified mode as "crispier" and "more flavorful" suggesting a similar mechanism as the magnitude framework, whereby attention to sensory aspects of food led to greater fullness. In contrast, the low pass filter mode was described as more "liquid", "wet", and akin to a "paste" suggesting a more mouth-filling feeling. If these two fullness properties of low vs. high pass filters are validated, this could lead to specific applications. Mouth-filling-based satiety could be used to reduce the appetite for unhealthy food such as fried food. In contrast, high-pass filters could lead to more fulfillment from eating healthy food such as vegetables, apples, carrots, etc.

Put together, we found evidence that even simple auditory filters had complex effects on the eating experience spanning from the perception of flavor and texture to fullness and awareness. This suggests that modifying eating sounds can be used as a tool to influence complex eating behavior beyond just sensory perception, for example by altering satiety or promoting mindful eating. On the flip side, although we see a consistent and significant effect from

Table 4: Summary of literature-based hypotheses and actual significant findings from the present research, for each auditory mode tested. Significance based on comparison with baseline (raw voice) mode.

Label	Mode Name	Literature-based hypotheses	Significant findings
A, J, Off	No sound (dampened)	Less sensory intensity (taste, mouthfeel, flavor) [31, 118]	Tasting less crispy Paying less attention to food
B	Raw Voice	baseline	(Baseline)
C	Amplified	Greater sensory intensity (taste, mouthfeel, flavor) [22, 31, 118]	Tasting more mouth-filling, intense, salty, sour Feeling more full Paying more attention to food & eating slower
D	Short Delay	Slightly greater fullness and less craving [64]	No significant effects compared to baseline
E	Long Delay	Greater fullness and less craving [64]	Eating more slowly
F	Small Room Reverberation	More crispiness and saltiness [111]	No significant effects compared to baseline
G	Large Room Reverberation	Less crispiness and saltiness [111]	Tasting more sour Feeling less craving
H	Low Pass Filter (LPF)	Less sourness and crispiness, greater fullness [27, 84, 118]	Tasting less crispy, more mouth-filling Feeling more full
I	High Pass Filter (HPF)	More sourness and crispiness [27, 118]	Feeling more full

the auditory feedback on people's eating experience, our study suggests that there isn't an independent one-to-one straightforward mapping between simple audio parameters and food perception parameters. Audio feedback while eating contains a large mix of information that is not simply separable into independent variables. For instance, loudness influences both flavor intensity and saltiness. This justifies the approach of first testing simple filters, and then implementing richer modes targeting eating features. Future work can therefore target the creation of richer audio transformations by combining simple filters to tackle more targeted features of the eating experience (only satiety, only saltiness, etc.). Moreover, our dataset provides a jumping-off point for further exploratory analyses, for example uncovering individual differences in auditory seasoning effects, which can then be validated in subsequent studies.

7.2 Limitations

First, our experiment was conducted using only one specific type of food: sour cream and onion Pringles, therefore we are unable to generalize our findings to a broader variety of foods until further studies are conducted. Moreover, we did not precisely control how far our participants held the mic in front of their mouths, which has implications for the degree of experienced sound amplification and modulation. With advances in noise cancellation, source separation, and selective filtering, the need to hold the microphone will hopefully not be necessary for the future, which will make data collection more easily in the wild. Another drawback to our design of the tool is the need for wearing earphones. As eating can be a social experience, having to wear earphones while eating might prevent people from having a conversation with others during a meal. Therefore, the use case for an Auditory Seasoning app might be more appropriate in a solo eating setting, unless bone conduction headphones are used in combination with the app. Further work could also investigate the difference between using closed vs. open headphones with the app.

7.3 Design implications

In this section, we discuss design considerations for changing the self-perception of chewing sounds as a way to regulate the eating experience. We also propose some potential use cases where this approach can be used.

7.3.1 Volume, distraction, and misophonia. Although amplification appears to affect perception parameters most consistently, there is a balance to find. Audio levels should remain safe not to cause hearing issues, and increasing volume also increases the risk of feedback. Further work should evaluate the added value from various audio gains to establish the right balance. Chosen volume and wet/dry mix also influence the naturalness of the sounds and can cause added distraction. Future work will evaluate the effects of various time delays on eating speed, perception, and motor control. Beyond distraction and potential effects on muscle control, the experience might be unpleasant for people with misophonia who may have intense emotional reactions in response to specific sounds, particularly sounds of human origin such as oral or nasal noises made by other people. [100]

7.3.2 Social vs. Solo eating. As noted by some of the participants, eating is often a social activity where additional auditory stimuli might be detrimental to human interaction. Thus, our tool would be more suited in situations where people already eat alone and plan to use a device. According to several recent reports, about half of all adult meals or snacks are taken alone [1, 2]. In addition, people eating alone often tend to use a device while eating, which can keep them from feeling lonely but may distract them from satiety and increase caloric intake [65]. By bringing the attention back to the act of eating, our tool could alleviate some of the drawbacks of using a device when eating.

7.3.3 Just-in-time intervention. One common issue in behavior change in the context of eating is the "adherence problem" [74]. This is the case with any approach that requires implicit decision-making or even the use of a technological assistant. Other approaches use nudging with regular phone notifications to remind users about eating strategy regularly, but they may lack immediacy. One advantage of using ubiquitous hardware that people are often already using when eating is to trigger the experience as a just-in-time intervention. For instance, the phone could listen in the background for eating sounds and turn on the system when eating activity is detected [48]. Even though the trigger could be automatic, the purpose or requested experience modification could be done purposefully by the user prior to the intervention. This strategy is also in line with behavior change theory that supports planning ahead of time, such as through the use of smaller plates.

7.3.4 Interface Design. In order to pre-program the type of effects desired, the interface design could propose several dimensions for desired targets. One issue from our results is the interconnectedness of the modes and effects. Future work will assess the feasibility of decoupling effects by combining auditory modulations. The objective would be to put together combinations of auditory filters to build interfaces that act on three levels: 1) Muscle control: Regulating eating speed and chewing force. This could involve auditory modulations based on time delays. 2) Flavour augmentation: Modulating the level of taste/flavor intensity, and crispness. It could be done through layering filters and pitch shifts. 3) Appetite regulation: Changing the level of fullness and craving. This may be done by changing room reverberations while also modulating the frequency balance to normalize the effect on sourness and flavor.

7.3.5 Learning to listen. One major impact of the Auditory Seasoning approach is to bring attention to the sounds of eating by unveiling an existing experience that we do not generally pay attention to. Although not asking for conscious attention and self-regulation from users, the experience may offer a frame for introspection, that might help regulate eating in an implicit way. Regarding the role of technology in mindful eating when compared to low-tech mindfulness exercises, we believe that–at least in part–our system acts as a catalyst for mindfulness that could potentially be achieved by other means. However, the profoundly surprised and amazement expressed by subjects, even some older participants, suggested that only a few people appear to be exposed to such other approaches during their lives. The ubiquitous and accessible nature of our phone-based approach did appear, in contrast, to make an effect. Indeed, many subjects suggested that it could potentially affect their eating experience positively in the long term and several of them reported six months later that they never ate chips the same way again and are now much more prone to listen to the sound of their food, in a playful, candid, and ingenious way, thanks to this one-time exposure to our system. Longitudinal testing would be required for rigorous testing of such lasting effects.

8 CONCLUSION

Throughout this paper, we contributed a first exploration of how closed-loop auditory feedback can be used to manipulate flavor perception, eating behavior, and appetite. We extend prior work on sound-taste interactions, demonstrating on a theoretical level that enhanced chewing sounds can influence more than tactile evaluations, and from a design perspective, showcasing a mobile tool that is easily accessible and usable without specialist equipment. Our work therefore sits squarely in the recommended areas of future HFI development, specifically digital augmentation of eating experiences which can be used to highlight specific food features and influence appetite [108]. For future work, we plan to develop new interfaces with customizable functions, deploy them to a wider audience, and conduct full meal eating studies to validate the use of auditory seasoning in the real world.

REFERENCES

[1] [n. d.]. Power of Family Meals 2017 from the Food Industry Association. https://www.fmi.org/docs/default-source/familymeals/fmi-power-of-family-meals-whitepaper-for-web.pdf?sfvrsn=13d87f6e_2. Accessed: 2021-09-06.

[2] [n. d.]. The Sainsbury's Living Well Index. https://www.about.sainsburys.co.uk/~/media/Files/S/Sainsburys/living-well-index/sainsburys-living-well-index-sep-2018.pdf. Accessed: 2021-09-06.

[3] Niels Werner Adelman-Larsen, Eric R Thompson, and Anders C Gade. 2010. Suitable reverberation times for halls for rock and pop music. *The journal of the Acoustical Society of America* 127, 1 (2010), 247–255.

[4] Garrett E Alexander, Michael D Crutcher, and Mahlon R DeLong. 1991. Basal ganglia-thalamocortical circuits: parallel substrates for motor, oculomotor,"prefrontal" and "limbic" functions. *Progress in brain research* 85 (1991), 119–146.

[5] Ferran Altarriba Bertran, Samvid Jhaveri, Rosa Lutz, Katherine Isbister, and Danielle Wilde. 2019. Making sense of human-food interaction. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–13.

[6] Joy Armson and Michael Kiefte. 2008. The effect of SpeechEasy on stuttering frequency, speech rate, and speech naturalness. *Journal of Fluency Disorders* 33, 2 (2008), 120–134.

[7] Jean-Julien Aucouturier, Petter Johansson, Lars Hall, Rodrigo Segnini, Lolita Mercadié, and Katsumi Watanabe. 2016. Covert digital manipulation of vocal emotion alter speakers' emotional states in a congruent direction. *Proceedings of the National Academy of Sciences* 113, 4 (2016), 948–953.

[8] Yoram Baram and Ariel Miller. 2007. Auditory feedback control for improvement of gait in patients with Multiple Sclerosis. *Journal of the neurological sciences* 254, 1-2 (2007), 90–94.

[9] Raiyan Abdul Baten and Ehsan Hoque. 2021. Technology-Driven Alteration of Nonverbal Cues and its Effects on Negotiation. *Negotiation Journal* 37, 1 (2021), 35–47.

[10] Philippe Besnard. 2016. Lipids and obesity: Also a matter of taste? *Reviews in Endocrine and Metabolic Disorders* 17, 2 (2016), 159–170.

[11] Kailash P Bhatia and C David Marsden. 1994. The behavioural and motor consequences of focal lesions of the basal ganglia in man. *Brain* 117, 4 (1994), 859–876.

[12] Gavin M Bidelman, Mary Katherine Davis, and Megan H Pridgen. 2018. Brainstem-cortical functional connectivity for speech is differentially challenged by noise and reverberation. *Hearing research* 367 (2018), 149–160.

[13] Jens Blauert and Guy J Brown. 2020. Reflexive and reflective auditory feedback. In *The Technology of Binaural Understanding*. Springer, 3–31.

[14] Dieuwerke P Bolhuis and Russell SJ Keast. 2016. Assessment of eating rate and food intake in spoon versus fork users in a laboratory setting. *Food quality and preference* 49 (2016), 66–69.

[15] James R Booth, Lydia Wood, Dong Lu, James C Houk, and Tali Bitan. 2007. The role of the basal ganglia and cerebellum in language processing. *Brain research* 1133 (2007), 136–144.

[16] Ronald W Botto, Gary G Galbraith, and Robert M Stern. 1974. Effects of false heart rate feedback and sex-guilt upon attitudes toward sexual stimuli. *Psychological reports* 35, 1 (1974), 267–274.

[17] Albert S Bregman. 1994. *Auditory scene analysis: The perceptual organization of sound*. MIT press.

[18] Albert S Bregnian. 1993. Auditory scene analysis: Hearing in complex environments. (1993).

[19] Jas Brooks, Pedro Lopes, Judith Amores, Emanuela Maggioni, Haruka Matsukura, Marianna Obrist, Roshan Lalintha Peiris, and Nimesha Ranasinghe. 2021. Smell, Taste, and Temperature Interfaces. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–6.

[20] Norm Campbell, Ricardo Correa-Rotter, Bruce Neal, and FP Cappuccio. 2011. New evidence relating to the health impact of reducing salt intake. *Nutrition, Metabolism and Cardiovascular Diseases* 21, 9 (2011), 617–619.

[21] Felipe Reinoso Carvalho, Qian Janice Wang, Raymond Van Ee, and Charles Spence. 2016. The influence of soundscapes on the perception and evaluation of beers. *Food quality and preference* 52 (2016), 32–41.

[22] Carol M Christensen and Zata M Vickers. 1981. Relationships of chewing sounds to judgments of food crispness. *Journal of Food Science* 46, 2 (1981), 574–578.

[23] Tara Cornwell, Jane Woodward, Mengnan Wu, Brennan Jackson, Pamela Souza, Jonathan Siegel, Sumitrajit Dhar, Keith E Gordon, et al. 2020. Walking with Ears: Altered Auditory Feedback Impacts Gait Step Length in Older Adults. *Frontiers in Sports and Active Living* 2 (2020), 38.

[24] Jean Costa, Alexander T Adams, Malte F Jung, François Guimbretière, and Tanzeem Choudhury. 2016. EmotionCheck: leveraging bodily signals and false feedback to regulate our emotions. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 758–769.

[25] Jean Costa, Malte F Jung, Mary Czerwinski, François Guimbretière, Trinh Le, and Tanzeem Choudhury. 2018. Regulating feelings during interpersonal conflicts by changing voice self-perception. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–13.

[26] Anne-Sylvie Crisinel, Stefan Cosser, Scott King, Russ Jones, James Petrie, and Charles Spence. 2012. A bittersweet symphony: systematically modulating the taste of food by changing the sonic properties of the soundtrack playing in the background. *Food Quality and Preference* 24, 1 (2012), 201–204.

[27] Anne-Sylvie Crisinel and Charles Spence. 2009. Implicit association between basic tastes and pitch. *Neuroscience letters* 464, 1 (2009), 39–42.

[28] C Dacremont. 1995. Spectral composition of eating sounds generated by crispy, crunchy and crackly foods. *Journal of texture studies* 26, 1 (1995), 27–43.

[29] C Dacremont, B Colas, and F Sauvageot. 1991. Contribution of air-and bone-conduction to the creation of sounds perceived during sensory evaluation of foods. *Journal of Texture Studies* 22, 4 (1991), 443–456.

[30] Roelof AJ de Vries, Gijs HJ Keizers, Sterre R van Arum, Juliet AM Haarman, Randy Klaassen, Robby W van Delden, Bert-Jan F van Beijnum, and Janet HW van den Boer. 2020. Multimodal Interactive Dining with the Sensory Interactive Table: Two Use Cases. In *Companion Publication of the 2020 International Conference on Multimodal Interaction*. 332–340.

[31] M Luisa Demattè, Nicola Pojer, Isabella Endrizzi, Maria Laura Corollaro, Emanuela Betta, Eugenio Aprea, Mathilde Charles, Franco Biasioli, Massimiliano Zampini, and Flavia Gasperi. 2014. Effects of the sound of the bite on apple perceived crispness and hardness. *Food Quality and Preference* 38 (2014), 58–64.

[32] Jialin Deng, Yan Wang, Carlos Velasco, Ferran Altarriba Altarriba Bertran, Rob Comber, Marianna Obrist, Katherine Isbister, Charles Spence, and Florian 'Floyd'Mueller. 2021. The Future of Human-Food Interaction. In *Extended Abstracts of CHI Conference on Human Factors in Computing Systems*. 1–6.

[33] Rachel Doggett, Elizabeth J Sander, James Birt, Matthew Ottley, and Oliver Baumann. 2021. Using virtual reality to evaluate the impact of room acoustics on cognitive performance and well-being. *Frontiers in Virtual Reality* 2 (2021).

[34] Birger K Drake. 1963. Food crushing sounds. an introductory study a, b. *Journal of Food Science* 28, 2 (1963), 233–241.

[35] Lisa Duizer. 2001. A review of acoustic research for studying the sensory perception of crisp, crunchy and crackly textures. *Trends in food science & technology* 12, 1 (2001), 17–24.

[36] LM Duizer, OH Campanella, and GRG Barnes. 1998. Sensory, instrumental and acoustic characteristics of extruded snack food products. *Journal of texture studies* 29, 4 (1998), 397–411.

[37] Audrey Eertmans, Frank Baeyens, and Omer Van Den Bergh. 2001. Food likes and their relative importance in human eating behavior: review and preliminary suggestions for health promotion. *Health education research* 16, 4 (2001).

[38] Anke Ehlers, Jürgen Margraf, Walton T Roth, C Barr Taylor, and Niels Birbaumer. 1988. Anxiety induced by false heart rate feedback in patients with panic disorder. *Behaviour research and therapy* 26, 1 (1988), 1–11.

[39] Hiroshi Endo, Shuichi Ino, and Waka Fujisaki. 2016. The effect of a crunchy pseudo-chewing sound on perceived texture of softened foods. *Physiology & behavior* 167 (2016), 324–331.

[40] Grant Fairbanks. 1955. Selective vocal effects of delayed auditory feedback. *Journal of speech and hearing disorders* (1955).

[41] Grant Fairbanks and Newman Guttman. 1958. Effects of delayed auditory feedback upon articulation. *Journal of Speech and Hearing Research* (1958).

[42] Robert A Frank and Jennifer Byram. 1988. Taste–smell interactions are tastant and odorant dependent. *Chemical Senses* 13, 3 (1988), 445–455.

[43] Asma Ghandeharioun and Rosalind Picard. 2017. BrightBeat: effortlessly influencing breathing for cultivating calmness and focus. In *Proceedings of the CHI conference extended abstracts on human factors in computing systems*.

[44] Andrea Grimes and Richard Harper. 2008. Celebratory technology: new directions for food research in HCI. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 467–476.

[45] Nicolas Guéguen, Céline Jacob, Hélène Le Guellec, Thierry Morineau, and Marcel Lourel. 2008. Sound level of environmental music and drinking behavior: a field experiment with beer drinkers. *Alcoholism: clinical and experimental research* 32, 10 (2008), 1795–1798.

[46] Frank H Guenther, Satrajit S Ghosh, and Jason A Tourville. 2006. Neural modeling and imaging of the cortical interactions underlying syllable production. *Brain and language* 96, 3 (2006), 280–301.

[47] Juliet AM Haarman, Roelof AJ de Vries, Emiel C Harmsen, Hermie J Hermens, and Dirk KJ Heylen. 2020. Sensory Interactive Table (SIT)—Development of a Measurement Instrument to Support Healthy Eating in a Social Dining Setting. *Sensors* 20, 9 (2020), 2636.

[48] Simone Hantke, Felix Weninger, Richard Kurle, Fabien Ringeval, Anton Batliner, Amr El-Desoky Mousa, and Björn Schuller. 2016. I hear you eat and speak: Automatic recognition of eating condition and food type, use-cases, and impact on asr performance. *PloS one* 11, 5 (2016), e0154486.

[49] Louise C Hawkley and John T Cacioppo. 2010. Loneliness matters: A theoretical and empirical review of consequences and mechanisms. *Annals of behavioral medicine* 40, 2 (2010), 218–227.

[50] Sander Hermsen, Jeana H Frost, Eric Robinson, Suzanne Higgs, Monica Mars, and Roel CJ Hermans. 2016. Evaluation of a Smart Fork to Decelerate Eating Rate. *Journal of the Academy of Nutrition and Dietetics* 116, 7 (2016), 1066–1068.

[51] Peter Howell and Alexandra Archer. 1984. Susceptibility to the effects of delayed auditory feedback. *Perception & Psychophysics* 36, 3 (1984), 296–302.

[52] Hiroo Iwata, Hiroaki Yano, Takahiro Uemura, and Tetsuro Moriya. 2004. Food simulator: A haptic interface for biting. In *IEEE Virtual Reality 2004*. IEEE, 51–57.

[53] Richard J Johnson, Mark S Segal, Yuri Sautin, Takahiko Nakagawa, Daniel I Feig, Duk-Hee Kang, Michael S Gersch, Steven Benner, and Laura G Sánchez-Lozada. 2007. Potential role of sugar (fructose) in the epidemic of hypertension, obesity and the metabolic syndrome, diabetes, kidney disease, and cardiovascular disease. *The American journal of clinical nutrition* 86, 4 (2007), 899–906.

[54] Azusa Kadomura, Cheng-Yuan Li, Yen-Chang Chen, Koji Tsukada, Itiro Siio, and Hao-hua Chu. 2013. Sensing fork: Eating behavior detection utensil and mobile persuasive game. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems*. 1551–1556.

[55] Joseph Kalinowski, Andrew Stuart, Sarah Sark, and Joy Armson. 1996. Stuttering amelioration at various auditory feedback delays and speech rates. *International Journal of Language & Communication Disorders* 31, 3 (1996), 259–269.

[56] Dwayne Keough and Jeffery A Jones. 2009. The sensitivity of auditory-motor representations to subtle changes in auditory feedback while singing. *The Journal of the Acoustical Society of America* 126, 2 (2009), 837–846.

[57] Rohit Ashok Khot, Eshita Sri Arza, Harshitha Kurra, and Yan Wang. 2019. Fobo: Towards designing a robotic companion for solo dining. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–6.

[58] 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.

[59] Yui Kita and Jun Rekimoto. 2013. Spot-light: multimodal projection mapping on food. In *International Conference on Human-Computer Interaction*. Springer, 652–655.

[60] Rebecca Kleinberger. 2020. *Vocal connection rethinking the voice as a medium for personal, interpersonal, and interspecies understanding*. Ph. D. Dissertation. Massachusetts Institute of Technology.

[61] Rébecca Kleinberger, Stefanakis George, and Sebastian Franjou. 2019. Speech companions: Evaluating the effects of musically modulated auditory feedback on the voice. Georgia Institute of Technology.

[62] Naoya Koizumi, Hidekazu Tanaka, Yuji Uema, and Masahiko Inami. 2011. Chewing jockey: augmented food texture by using sound based on the cross-modal effect. In *Proceedings of the 8th international conference on advances in computer entertainment technology*. 1–4.

[63] Kazutaka Kurihara and Koji Tsukada. 2012. Speechjammer: A system utilizing artificial speech disturbance with delayed auditory feedback. *arXiv preprint arXiv:1202.6106* (2012).

[64] Marlou Lasschuijt, Monica Mars, Cees de Graaf, and Paul AM Smeets. 2020. How oro-sensory exposure and eating rate affect satiation and associated endocrine responses—a randomized trial. *The American journal of clinical nutrition* 111, 6 (2020), 1137–1149.

[65] Mailin Lemke and Hendrik NJ Schifferstein. 2021. The use of ICT devices as part of the solo eating experience. *Appetite* 165 (2021), 105297.

[66] Sook Ling Leong, Clara Madden, Andrew Gray, Debra Waters, and Caroline Horwath. 2011. Faster self-reported speed of eating is related to higher body mass index in a nationwide survey of middle-aged women. *Journal of the American Dietetic Association* 111, 8 (2011), 1192–1197.

[67] Jenni Leppanen, Valentina Cardi, Felicity Sedgewick, Janet Treasure, and Kate Tchanturia. 2020. Basal ganglia volume and shape in anorexia nervosa. *Appetite* 144 (2020), 104480.

[68] Benjamin J Li and Jeremy N Bailenson. 2018. Exploring the influence of haptic and olfactory cues of a virtual donut on satiation and eating behavior. *Presence* 26, 03 (2018), 337–354.

[69] Ying-Ju Lin, Parinya Punpongsanon, Xin Wen, Daisuke Iwai, Kosuke Sato, Marianna Obrist, and Stefanie Mueller. 2020. FoodFab: creating food perception illusions using food 3D printing. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–13.

[70] Maurizio Mancini, Radoslaw Niewiadomski, Gijs Huisman, Merijn Bruijnes, and Conor Patrick Gallagher. 2020. Room for one more?-Introducing artificial commensal companions. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–8.

[71] Signe Lund Mathiesen, Line Ahm Mielby, Derek Victor Byrne, and Qian Janice Wang. 2020. Music to eat by: A systematic investigation of the relative importance of tempo and articulation on eating time. *Appetite* 155 (2020), 104801.

[72] Yash Dhanpal Mehta, Rohit Ashok Khot, Rakesh Patibanda, and Florian'Floyd' Mueller. 2018. Arm-a-Dine: towards understanding the design of playful embodied eating experiences. In *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play*. 299–313.

[73] Bruno Mesz, Marcos A Trevisan, and Mariano Sigman. 2011. The taste of music. *Perception* 40, 2 (2011), 209–219.

[74] Kathryn R Middleton, Stephen D Anton, and Michal G Perri. 2013. Long-term adherence to health behavior change. *American journal of lifestyle medicine* 7, 6 (2013), 395–404.

[75] Michael T Motley. 1976. Stage fright manipulation by (false) heart rate feedback. *Communication Studies* 27, 3 (1976), 186–191.

[76] Takuji Narumi, Yuki Ban, Takashi Kajinami, Tomohiro Tanikawa, and Michitaka Hirose. 2012. Augmented perception of satiety: controlling food consumption by changing apparent size of food with augmented reality. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 109–118.

[77] Takuji Narumi, Shinya Nishizaka, Takashi Kajinami, Tomohiro Tanikawa, and Michitaka Hirose. 2011. Meta cookie+: an illusion-based gustatory display. In *International Conference on Virtual and Mixed Reality*. Springer, 260–269.

[78] Narihiro Nishimura, Asuka Ishi, Michi Sato, Shogo Fukushima, and Hiroyuki Kajimoto. 2012. Facilitation of affection by tactile feedback of false heratbeat. In *CHI'12 Extended Abstracts on Human Factors in Computing Systems*. 2321–2326.

[79] Masahiro Nishizawa, Wanting Jiang, and Katsunori Okajima. 2016. Projective-AR system for customizing the appearance and taste of food. In *Proceedings of the 2016 workshop on Multimodal Virtual and Augmented Reality*. 1–6.

[80] Taketoshi Ono, H Nishijo, and H Nishino. 2000. Functional role of the limbic system and basal ganglia in motivated behaviors. *Journal of neurology* 247, 5 (2000), V23–V32.

[81] Ye I PALTSEV. 1967. Change in the functional state of the segmental apparatus of the spinal cord under the influence of sound stimuli and its role in voluntary movement. *Biophysics* 12 (1967), 1219–1226.

[82] Danni Peng-Li, Signe L Mathiesen, Raymond CK Chan, Derek V Byrne, and Qian Janice Wang. 2021. Sounds Healthy: Modelling sound-evoked consumer food choice through visual attention. *Appetite* 164 (2021), 105264.

[83] Peter Q Pfordresher and Simone Dalla Bella. 2011. Delayed auditory feedback and movement. *Journal of Experimental Psychology: Human Perception and Performance* 37, 2 (2011), 566.

[84] Betina Piqueras-Fiszman and Charles Spence. 2012. The weight of the container influences expected satiety, perceived density, and subsequent expected fullness. *Appetite* 58, 2 (2012), 559–562.

[85] Nimesha Ranasinghe and Ellen Yi-Luen Do. 2016. Digital lollipop: Studying electrical stimulation on the human tongue to simulate taste sensations. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)* 13, 1 (2016), 1–22.

[86] Marianne Regard and Theodor Landis. 1997. "Gourmand syndrome": eating passion associated with right anterior lesions. *Neurology* 48, 5 (1997), 1185–1190.

[87] Eric Robinson, Suzanne Higgs, Amanda J Daley, Kate Jolly, Deborah Lycett, Amanda Lewis, and Paul Aveyard. 2013. Development and feasibility testing of a smart phone based attentive eating intervention. *BMC public health* 13, 1 (2013), 1–7.

[88] Gaëlle Roudaut, Catherine Dacremont, B Vallès Pàmies, B Colas, and Martine Le Meste. 2002. Crispness: a critical review on sensory and material science approaches. *Trends in food science & technology* 13, 6-7 (2002), 217–227.

[89] Charles Spence. 2015. Eating with our ears: assessing the importance of the sounds of consumption on our perception and enjoyment of multisensory flavour experiences. *Flavour* 4, 1 (2015), 1–14.

[90] Charles Spence. 2015. Just how much of what we taste derives from the sense of smell? *Flavour* 4, 1 (2015), 1–10.

[91] Charles Spence. 2015. On the psychological impact of food colour. *Flavour* 4, 1 (2015), 1–16.

[92] Charles Spence, Felipe Reinoso-Carvalho, Carlos Velasco, and Qian Janice Wang. 2019. Extrinsic auditory contributions to food perception & consumer behaviour: An interdisciplinary review. *Multisensory research* 32, 4-5 (2019), 275–318.

[93] Charles Spence and Maya U Shankar. 2010. The influence of auditory cues on the perception of, and responses to, food and drink. *Journal of Sensory Studies* 25, 3 (2010), 406–430.

[94] Charles Spence, Qian Janice Wang, Felipe Reinoso-Carvalho, and Steve Keller. 2021. Commercializing sonic seasoning in multisensory offline experiential events and online tasting experiences. *Frontiers in Psychology* (2021), 4150.

[95] Charles Spence and Massimiliano Zampini. 2006. Auditory contributions to multisensory product perception. *Acta acustica united with acustica* 92, 6 (2006), 1009–1025.

[96] Shrikant Srivastava, Lucy F Donaldson, Dheeraj Rai, Jan K Melichar, and John Potokar. 2013. Single bright light exposure decreases sweet taste threshold in healthy volunteers. *Journal of Psychopharmacology* 27, 10 (2013), 921–929.

[97] Ana Tajadura-Jiménez, Maria Basia, Ophelia Deroy, Merle Fairhurst, Nicolai Marquardt, and Nadia Bianchi-Berthouze. 2015. As light as your footsteps: altering walking sounds to change perceived body weight, emotional state and gait. In *Proceedings of the 33rd annual ACM conference on human factors in computing systems*. 2943–2952.

[98] Ana Tajadura-Jiménez, Pontus Larsson, Aleksander Väljamäe, Daniel Västfjäll, and Mendel Kleiner. 2010. When room size matters: acoustic influences on emotional responses to sounds. *Emotion* 10, 3 (2010), 416.

[99] Ana Tajadura-Jiménez, Aleksander Väljamäe, Iwaki Toshima, Toshitaka Kimura, Manos Tsakiris, and Norimichi Kitagawa. 2012. Action sounds recalibrate perceived tactile distance. *Current Biology* 22, 13 (2012), R516–R517.

[100] Steven Taylor. 2017. Misophonia: A new mental disorder? *Medical hypotheses* 103 (2017), 109–117.

[101] Ethan Teo, Daniel Goh, Kamalakannan M Vijayakumar, and Jean CJ Liu. 2018. To message or browse? Exploring the impact of phone use patterns on male adolescents' consumption of palatable snacks. *Frontiers in psychology* (2018).

[102] Junya Ueda and Katsunori Okajima. 2019. AR food changer using deep learning and cross-modal effects. In *2019 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR)*. IEEE, 110–1107.

[103] Vesa Valimaki, Julian D Parker, Lauri Savioja, Julius O Smith, and Jonathan S Abel. 2012. Fifty years of artificial reverberation. *IEEE Transactions on Audio, Speech, and Language Processing* 20, 5 (2012), 1421–1448.

[104] Stuart Valins. 1966. Cognitive effects of false heart-rate feedback. *Journal of personality and social psychology* 4, 4 (1966), 400.

[105] Kimberley van der Heijden, Anouk Festjens, and Caroline Goukens. 2021. On the bright side: The influence of brightness on overall taste intensity perception. *Food Quality and Preference* 88 (2021), 104099.

[106] VL Van Stokkom, AE Blok, O Van Kooten, C de Graaf, and M Stieger. 2018. The role of smell, taste, flavour and texture cues in the identification of vegetables. *Appetite* 121 (2018), 69–76.

[107] Carlos Velasco, Marianna Obrist, Olivia Petit, and Charles Spence. 2018. Multisensory technology for flavor augmentation: a mini review. *Frontiers in psychology* 9 (2018), 26.

[108] Carlos Velasco, Qian Janice Wang, Marianna Obrist, and Anton Nijholt. 2021. A Reflection on the State of Multisensory Human–Food Interaction Research. *Frontiers in Computer Science* (2021), 106.

[109] Peter Walker, Gabrielle Scallon, and Brian Francis. 2017. Cross-sensory correspondences: Heaviness is dark and low-pitched. *Perception* 46, 7 (2017), 772–792.

[110] Vincent Walsh. 2003. A theory of magnitude: common cortical metrics of time, space and quantity. *Trends in cognitive sciences* 7, 11 (2003), 483–488.

[111] Qian Janice Wang, Steve Keller, and Charles Spence. 2021. Metacognition and Crossmodal Correspondences Between Auditory Attributes and Saltiness in a Large Sample Study. *Multisensory Research* 1, aop (2021), 1–21.

[112] Qian Janice Wang and Charles Spence. 2019. Sonic packaging: how packaging sounds influence multisensory product evaluation. In *Multisensory Packaging*. Springer, 103–125.

[113] Qian Janice Wang, Charles Spence, and Klemens Knoeferle. 2020. Timing is everything: Onset timing moderates the crossmodal influence of background sound on taste perception. *Journal of Experimental Psychology: Human Perception and Performance* (2020).

[114] Qian Janice Wang, Sheila Wang, and Charles Spence. 2016. "Turn up the taste": Assessing the role of taste intensity and emotion in mediating crossmodal correspondences between basic tastes and pitch. *Chemical Senses* 41, 4 (2016), 345–356.

[115] Jun Wei, Xuan Wang, Roshan Lalintha Peiris, Yongsoon Choi, Xavier Roman Martinez, Remi Tache, Jeffrey Tzu Kwan Valino Koh, Veronica Halupka, and Adrian David Cheok. 2011. CoDine: an interactive multi-sensory system for remote dining. In *Proceedings of the 13th international conference on Ubiquitous computing*. 21–30.

[116] Aubrey J Yates. 1963. Delayed auditory feedback. *Psychological bulletin* 60, 3 (1963), 213.

[117] Martin R Yeomans and Sophie Boakes. 2016. That smells filling: Effects of pairings of odours with sweetness and thickness on odour perception and expected satiety. *Food Quality and Preference* 54 (2016), 128–136.

[118] Massimiliano Zampini and Charles Spence. 2004. The role of auditory cues in modulating the perceived crispness and staleness of potato chips. *Journal of sensory studies* 19, 5 (2004), 347–363.

[119] Massimiliano Zampini and Charles Spence. 2005. Modifying the multisensory perception of a carbonated beverage using auditory cues. *Food quality and preference* 16, 7 (2005), 632–641.

[120] Yong Zhu and James H Hollis. 2014. Increasing the number of chews before swallowing reduces meal size in normal-weight, overweight, and obese adults. *Journal of the Academy of Nutrition and Dietetics* 114, 6 (2014), 926–931.

[121] Nicolien Zijlstra, René 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.