Exploring Voice Input Opportunities in Multimodal Food Journaling

Lucas M. Silva
University of California, Irvine
Irvine, CA 92697, USA
silvald@ics.uci.edu

Yuqi Huai
University of California, Irvine
Irvine, CA 92697, USA
yhuai@ uci.edu

Daniel A. Epstein
University of California, Irvine
Irvine, CA 92697, USA
epstein@ics.uci.edu

CHI 2020 Extended Abstracts, April 25–30, 2020, Honolulu, HI, USA.

ACM ISBN 978-1-4503-6819-3/20/04.

Accepted to the CHI 2020 Workshop on Conversational Agents for Health and Wellbeing.

Abstract
Digital food journaling can support a range of personal goals, such as weight loss and healthy eating behavior. However, automatic tracking is limited, and manual tracking demands great effort, often leading to lapses or abandonment. Seeking to lessen tracking burden of food journaling, we explore opportunities for leveraging voice assistants on multiple devices and varying input modalities. We have developed technology probes for tracking through multiple devices, and report on participant feedback from a pilot study.

Author Keywords
Personal Informatics; food journaling; multimodality; voice input.

CSS Concepts
• **Human-centered computing~ Ubiquitous and mobile computing**; *Mobile Devices*; User studies.

Introduction
Tracking of food intake is increasingly common, such as calorie counting or being mindful of food choices. The practice of journaling can help people understand their food consumption practices, monitor progress towards behavior change goals like weight loss, and manage chronic conditions like diabetes and Irritable Bowel Syndrome [6,7,9,11,13]. Food journaling also has

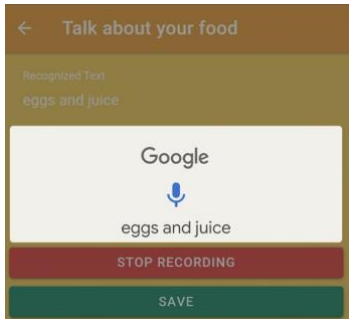

Figure 1: We integrate voice-based logging into our mobile app.

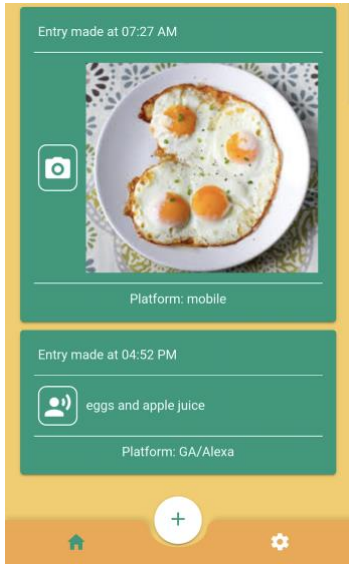

Figure 2: Integrating input from different modalities can provide additional detail.

clinical value, helping clinicians assess patients and provide treatment [12]. Although paper journals are traditional for assessment, digital journals have become increasingly popular and viable [12,19]. However, clinicians still typically lessen patient burden by having them journal for three days at most [18].

Despite promising opportunities for digital food journaling, the task of manually tracking one's every meal may feel daunting [5], and this burden can lead to fatigue or lapses in tracking [4]. Although there have been proposals of automating this process through crowdsourcing or computer vision (e.g., Platemate [14]), these methods can be inaccurate and lessen people's engagement with data [3,5].

Research prototypes and commercial apps have examined strategies to reduce the manual input burden by supporting lookup in food databases, barcode scanning, voice memo recording, and photo logging. Although these strategies help make food tracking more accessible, each of these strategies in isolation may be insufficient to reduce the burden to the point that people are willing and interested in sustaining the habit. Some entry methods may be preferable to others given social circumstances, types of food, and recognition of food types [6]. Digital food journaling is also typically limited to mobile phones. Few apps or platforms support journaling through digital voice assistants, such as those available on smartwatches, home assistants, and desktop/laptop computers.

We are examining the idea of tracking on multiple devices, using several entry styles and integrated with conversational agents. This has the potential to lower burden by aligning journaling to specific contexts. For instance, when a person's hands are occupied, voice assistants may be better suited than a mobile app, or a web app more adequate when eating at the office in front of the computer.

We therefore created technology probes [10] for food logging on smartphones, web browsers, Amazon Alexa, Google Home, and Apple Watch. We have deployed these probes in a pilot study with three participants and are preparing a field study with experienced food journalers or people enthusiastic to start journaling. The probes will help participants imagine how they would ideally track their food and integrate journaling in their daily lives, including how digital voice assistants might facilitate journaling and in what contexts.

Background

Mobile devices have been used extensively in food journaling research to promote self-awareness, eating mindfulness, and to replace traditional paper journaling in clinical contexts. For example, the MAHI system [13] aimed to help people with diabetes to manage their disease by collecting food intake and blood glucose level. PmEB [17] contributed a mobile application that calculates caloric balance based on logged food and physical activity to promote general wellbeing. Although some early systems integrated voice logging of food (e.g., Barcode Ed [16]), the majority of research and commercial apps have supported food logging via text database lookups on a mobile phone [16].

As mobile apps for food journaling have become more pervasive, studies have identified challenges to food journaling that might lead to abandonment or lapsing in use of the technology. For example, Cordeiro et al. suggest that people might have difficulty journaling due

Figure 3: Apple Watch app capturing voice description of food.

to barriers in identifying and quantifying foods eaten, non-reliable databases for food search, and eating context (e.g., some social situations, meals prepared by others, eating a variety of foods over time, etc.) [6].

To address these concerns, research has explored strategies for lowering journaling fatigue while still promoting self-reflection and healthy habits [5,6,8,12]. For example, Cordeiro et al. explored how photo capture can complement or improve on calorie and nutrient-based food tracking [5], suggesting that it eased collection and promoted reflection for certain food tracking goals. Considerable effort has also examined automatically identifying and classifying eating moments [1,2,15]. Accuracy and reliability continue to be active areas of study and can help people reflect on their habits in a lower-burden way.

In our research, we explore how conversational agents might be integrated into food journaling and how availability of multiple modalities of input and devices can lower tracking fatigue and increase engagement.

Our Technology Probes

Technology probes, or lightweight and flexible versions of technology, are often used to inspire ideas for new interactions in participatory design processes [10]. Technology probes are often deployed to understand people's needs and brainstorm new ideas in real-world settings, making the method useful for examining people's diverse interactions with devices. We have developed probes for a web app to be used on computers, a mobile app for Android and iOS (Figures 1 and 2), an Apple watch app (Figure 3), an Amazon Alexa skill and a Google Home Action (Figure 4).

The web and mobile probes have common features for food journaling: taking pictures, adding descriptions, barcode scanning, voice recording, and simulated database search. The smartwatch app and conversational agents only support voice recording.

The probes are intentionally flexible, providing open-ended fields, minimal conversation, and not offering suggestions for what or how to journal. This flexibility will enable participants to produce queries beyond current system's capabilities and allow us to understand the fidelity to which people practically want to journal their food (e.g., nutrients and amount of food). For instance, the voice assistant's probes are limited to journaling food in the form of a "journal X" command, speaking the latest food entry, or deleting the latest entry (interactions demonstrated in Figure 4).

Pilot Study

The goal of our study is to identify design opportunities to support food journaling over multiple devices and input modalities. Specifically, we aim to understand how a person's context influences their journaling choices, how people want to combine different modalities to facilitate journaling, and how to design underexplored modalities (e.g., conversational agents) to react to how people imagine journaling in practice.

Thus far, we have deployed the probes in a one-week pilot study with three participants, all researchers in our lab. P1 had no experience with food journaling and had no goal related to tracking, so his feedback was mostly technical and related to study management. P2 had journaled her food previously, while P3 was currently journaling her food with a mobile app.

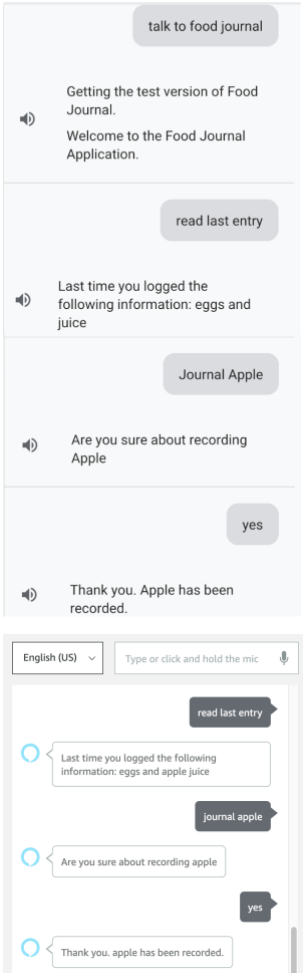

Figure 4: Interaction with voice assistant simulators to hear the last food entry and make a new entry.

Preliminary Results

Participants averaged between 2 and 3 inputs per day and mobile devices where the most used platform for food logging (~57% of entries), followed by digital voice assistants (~27%) and web application (~16%). P2 used the conversational agent exclusively on her phone, while P3 used it on her phone more times than on the Amazon Echo. P3 said that she would often journal her food in a rush, such as walking to some destination. In these situations, it was faster to speak with the conversational agent than to open the mobile app to journal via text or a simulated database search.

Both P2 and P3's modality choices revolved around convenience and if there was a need for a picture. P2 clarified that many of her meals where the same food. After the first time she recorded using picture and description, she started using only the conversational agent for that particular food: "*because it was convenient to record it fast and there was no more need for pictures*". P3 also mentioned some situations where the conversational agent felt more convenient: "*when I'm at home and mostly when I'm not near those devices [laptop and smartphone] especially for dinners, it was just easier for me to report through Echo. When you're cooking you can just record it or after meal you are cleaning dishes you can do it [record food]*".

P3 had defined a daily calorie limit and said it would be interesting if the voice assistant could remind her of how much of her calorie "budget" was still available. Both P2 and P3 wished that the conversational agent could summarize their food consumption for the week.

Despite these convenient uses, P3 stated she felt that conversational agents would not be appropriate in some social situations: "*So most of the time I entered things here [workplace], I mean either in front of a laptop or a smartphone so, like, [in] a social environment [you] cannot just go and use, like, digital assistant*".

Discussion

The pilot study has pointed to opportunities for future designs and integration between platforms. For instance, participants highlighted how conversational agents can support quick entries and to refer to previously tracked foods. Some participants also desired interaction with conversational agents to reflect on their eating patterns. We expect that the full study will surface further opportunity for conversational agents as well as its integration with other platforms and modalities.

Workshop Participation

We are excited to discuss potential of conversational agents and multimodality for journaling and tracking in other wellbeing and health domains with other researchers. We are currently recruiting for our field study, so we are eager to use this venue to discuss our findings and design considerations once it is complete.

Although digital voice assistants are becoming pervasive, there are still many challenges in creating conversational flows which match with people's expectations. We hope to learn from other researcher's experiences and insights on overcoming these challenges.

Acknowledgments

We thank our pilot participants, and Kimberly Flores for helping to develop the technology probes. This research was supported in part by the National Science Foundation under award IIS-1850389.

References

1. Abdelkareem Bedri, Gregory Abowd, Richard Li, Malcolm Haynes, Raj Prateek Kosaraju, Ishaan Grover, Temiloluwa Prioleau, Min Yan Beh, Mayank Goel, and Thad Starner. (2017). EarBit: Using Wearable Sensors to Detect Eating Episodes in Unconstrained Environments. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (IMWUT)*, *1*(3), 1–20. http://doi.org/10.1145/3130902
2. Joan-Isaac Biel, Nathalie Martin, David Labbe, and Daniel Gatica-Perez. (2018). Bites'n'Bits. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (IMWUT)*, *1*(4), 1–33. http://doi.org/10.1145/3161161
3. Eun Kyoung Choe, Saeed Abdullah, Mashfiqui Rabbi, Edison Thomaz, Daniel A. Epstein, Felicia Cordeiro, Matthew Kay, Gregory D. Abowd, Tanzeem Choudhury, James Fogarty, Bongshin Lee, Mark Matthews, and Julie A. Kientz. (2017). Semi-Automated Tracking: A Balanced Approach for Self-Monitoring Applications. *IEEE Pervasive Computing*, *16*(1), 74–84. http://doi.org/10.1109/MPRV.2017.18
4. Eun Kyoung Choe, Nicole B. Lee, Bongshin Lee, Wanda Pratt, and Julie A. Kientz. (2014). Understanding Quantified-Selfers' Practices in Collecting and Exploring Personal Data. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2014)*, 1143–1152. http://doi.org/10.1145/2556288.2557372
5. Felicia Cordeiro, Elizabeth Bales, Erin Cherry, and James Fogarty. (2015). Rethinking the Mobile Food Journal: Exploring Opportunities for Lightweight Photo-Based Capture. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2015)*, 3207–3216. http://doi.org/10.1145/2702123.2702154
6. Felicia Cordeiro, Daniel A. Epstein, Edison Thomaz, Elizabeth Bales, Arvind K. Jagannathan, Gregory D. Abowd, and James Fogarty. (2015). Barriers and Negative Nudges: Exploring Challenges in Food Journaling. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2015)*, 1159–1162. http://doi.org/10.1145/2702123.2702155
7. Alaina Darby, Matthew W Strum, Erin Holmes, and Justin Gatwood. (2016). A Review of Nutritional Tracking Mobile Applications for Diabetes Patient Use. *Diabetes Technology & Therapeutics. March 16, 2016. 200-212*. http://doi.org/10.1089/dia.2015.0299
8. Daniel A. Epstein, Felicia Cordeiro, James Fogarty, Gary Hsieh, and Sean A. Munson. (2016). Crumbs: Lightweight Daily Food Challenges to Promote Engagement and Mindfulness. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2016)*, 5632–5644. http://doi.org/10.1145/2858036.2858044
9. William D. Heizer, Susannah Southern, and Susan McGovern. (2009). The Role of Diet in Symptoms of Irritable Bowel Syndrome in Adults: A Narrative Review. *Journal of the American Dietetic Association*, *109*(7), 1204–1214. http://doi.org/10.1016/j.jada.2009.04.012
10. Hilary Hutchinson, Heiko Hansen, Nicolas Roussel, Björn Eiderbäck, Wendy Mackay, Bosse Westerlund, Benjamin B Bederson, Allison Druin, Catherine Plaisant, Michel Beaudouin-Lafon, Stéphane Conversy, and Helen Evans. (2003). Technology probes: Inspiring Design For and With Families. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2003)*, 17-24. http://doi.org/10.1145/642611.642616
11. Ravi Karkar, Jessica Schroeder, Daniel A. Epstein, Laura R. Pina, Jeffrey Scofield, James Fogarty, Julie A. Kientz, Sean A. Munson, Roger Vilardaga, and Jasmine Zia. (2017). TummyTrials: A Feasibility Study of Using Self-Experimentation to Detect Individualized Food Triggers. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2017)*, 6850–6863. http://doi.org/10.1145/3025453.3025480
12. Yuhan Luo, Peiyi Liu, and Eun Kyoung Choe. (2019). Co-Designing Food Trackers with Dietitians. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2019)*, 1–13. http://doi.org/10.1145/3290605.3300822

13. Lena Mamykina, Elizabeth Mynatt, Patricia Davidson, and Daniel Greenblatt. (2008). MAHI: Investigation of Social Scaffolding for Reflective Thinking in Diabetes Management. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2008), 477–486. http://doi.org/10.1145/1357054.1357131

14. Jon Noronha, Eric Hysen, Haoqi Zhang, and Krzysztof Z Gajos. (2011). Platemate: Crowdsourcing Nutritional Analysis From Food Photographs. *Proceedings of the 24th annual ACM symposium on User interface software and technology (UIST '11)*, 1-12. http://doi.org/10.1145/2047196.2047198

15. Keum San Chun, Sarnab Bhattacharya, and Edison Thomaz. (2018). Detecting Eating Episodes by Tracking Jawbone Movements with a Non-Contact Wearable Sensor. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (IMWUT)*, *2*(1), 1-21. http://doi.org/10.1145/3191736

16. Katie A. Siek, Kay H. Connelly, Yvonne Rogers, Paul Rohwer, Desiree Lambert, and Janet L. Welch. (2006). When Do We Eat? An Evaluation of Food Items Input into an Electronic Food Monitoring Application. Proceedings of *Pervasive Health Conference and Workshops (2006)*, 1–10. http://doi.org/10.1109/PCTHEALTH.2006.361684

17. Christopher C. Tsai, Gunny Lee, Fred Raab, Gregory J. Norman, Timothy Sohn, William G Griswold, and Kevin Patrick. (2007). Usability and Feasibility of PmEB: A Mobile Phone Application for Monitoring Real Time Caloric Balance. *Mobile Networks and Applications (2007)*, *12*(2–3), 173–184. http://doi.org/10.1007/s11036-007-0014-4

18. Jasmine K. Zia, Chia-Fang Chung, Jessica Schroeder, Sean A. Munson, Julie A. Kientz, James Fogarty, Elizabeth Bales, Jeanette M. Schenk, and Margaret M. Heitkemper. (2017). The Feasibility, Usability, and Clinical Utility of Traditional Paper Food and Symptom Journals for Patients with Irritable Bowel Syndrome. *Neurogastroenterology & Motility*, *29*(2), e12935. http://doi.org/10.1111/nmo.12935

19. Jasmine Zia, Jessica Schroeder, Sean Munson, James Fogarty, Linda Nguyen, Pamela Barney, Margaret Heitkemper, and Uri Ladabaum. (2016). Feasibility and Usability Pilot Study of a Novel Irritable Bowel Syndrome Food and Gastrointestinal Symptom Journal Smartphone App. *Clinical and Translational Gastroenterology*, *7*(3), e147. http://doi.org/10.1038/ctg.2016.9