Towards In-Home Deployments of Physically Assistive Robots: Insights from Robot-Assisted Feeding for People with Motor Impairments

Amal Nanavati

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington
2025

Reading Committee:
Siddhartha S. Srinivasa, Co-Chair
Maya Cakmak, Co-Chair
Anat Caspi

Program Authorized to Offer Degree:
Computer Science and Engineering

University of Washington

Abstract

Towards In-Home Deployments of Physically Assistive Robots: Insights from Robot-Assisted Feeding for People with Motor Impairments

Amal Nanavati

Co-chairs of the Supervisory Committee:

Siddhartha S. Srinivasa
Computer Science and Engineering

Maya Cakmak
Computer Science and Engineering

Over 1 billion people worldwide are estimated to experience significant disability, which impacts their ability to independently conduct activities of daily living (ADLs) such as eating, ambulating, and dressing. Physically assistive robots (PARs) have emerged as a promising technology to help people with disabilities conduct ADLs, thereby restoring independence and reducing caregiver burden. However, despite decades of research on PARs, deployments of them in end-users' homes are still few and far between.

This thesis focuses on robot-assisted feeding as a case study for how we can achieve in-home deployments of PARs. Our ultimate goal is to develop a robot-assisted feeding system that enables *any user*, in *any environment*, to feed themselves a *meal of their choice* in a way that *aligns with their preferences*. We collaborate closely with 2 community researchers with motor impairments to design, implement, and evaluate a robot-assisted feeding system that makes progress towards this ultimate goal. Specifically, this thesis presents the following work:

1. A systematic survey of research on PARs, identifying key themes and trends;
2. A formative study investigating the meal-related needs of people with motor impairments and their priorities regarding the design of robot-assisted feeding systems;

3. An action schema and unsupervised learning pipeline that uses human data to learn representative actions a robot can use acquire diverse bites of food; and
4. The key system design considerations, both software and hardware, that enabled us to develop a robot-assisted feeding system to deploy in users' homes.

We evaluate the system with two studies: (1) an out-of-lab study where 5 participants and 1 community researcher use the robot to feed themselves a meal of their choice in a cafeteria, conference room, or office; and (2) a 5-day, in-home deployment where 1 community researcher uses the robot to feed himself 10 meals across various spatial, social, and activity contexts. The studies reveal promising results in terms of the usability and functionality of the system, as well as key directions for future work that are necessary to achieve the aforementioned ultimate goal. We present key lessons learned regarding in-home deployments of PARs: (1) spatial contexts are numerous, customizability lets users adapt to them; (2) off-nominals will arise, variable autonomy lets users overcome them; (3) assistive robots' benefits depend on context; and (4) work with end-users and stakeholders.

Acknowledgements

They say it takes a village to raise a child. I would say it takes the same to get a PhD. This thesis would not have been possible without the support, encouragement, guidance, and collaboration of dozens of mentors, labmates, friends, and family. I'd like to take this space to offer them my deepest, heartfelt gratitude.

To my advisers, Sidd Srinivasa and Maya Cakmak: Our weekly meetings provided inspiration, learning, and growth. Your feedback and collaboration sharpened my research ideas. Your flexibility and mentorship allowed me to explore research directions of interest. Your support and encouragement helped me incubate socially impactful research ideas. So much of the researcher, engineer, writer, mentor, and mentee I am today is thanks to you.

To Leila Takayama: It was by complete happenstance that I started working on a grant you were co-PI on 6 years ago, but it was one of the best turns of fate of my PhD. You have been such an inspirational teacher and collaborator during our work together, both in academia and the industry. You have been one of the most supportive career mentors I could have asked for, and your advice has been invaluable as I continue to navigate human-robot interaction (HRI) in the industry. Let's continue to grow HRI's presence in the industry together.

To my thesis committee members, Anat Caspi, Henny Admoni, and Kim Ingraham: Thank you for your feedback, guidance, and encouragement as I proposed and worked on this thesis over the last 1.5 years. It has been truly invaluable to have such nuanced and diverse perspectives on my committee. I continue to learn and grow from your advice, and this thesis is stronger because of you.

To Tyler Schrenk and Johnathan Ko: I could not have asked for better community researchers to work on this research with. It has been an absolute privilege working with you and learning from you over the last 3 years. Much of the robot-assisted feeding system is the way it is because of your insights, and your impact will continue to reverberate amongst the broader research community for years to come.

To my labmates and collaborators: Ethan Gordon, Taylor Kessler-Faulkner, Patrícia Alves-Oliveira, Vy Nguyen, Vinitha Ranganeni, Bernie Zhu, Atharva Kashyap, Ray Song, Sriram Kutty, Haya Bolotski, Liander Rainbolt, Nick Walker, Chris Mavrogiannis, and Tapo Bhattacharjee, to name a few. Thank you for your collaboration and camaraderie over the years.

To the organizations who (partially) funded research in this thesis: UW Center for Research and Education on Accessible Technology and Experiences (CREATE); UW Allen School Postdoc Research Award; the National Science Foundation GRFP (DGE-1762114), NRI (#1925043 and #2132848), and CHS (#2007011); the National Institute of Biomedical Imaging and Bioengineering (#1R01EB034580-01); the Office of Naval Research (#N00014-24-S-B001 and #2022-016-01 UW); DARPA RACER (#HR0011-21-C-0171); Amazon; Collaborative Robotics; Cruise; and others. Thank you for generously supporting this work.

To Catherine Wyatt, who mailed nearly 200 pages of historic research papers on robot-assisted feeding [243, 12] to a random PhD student who cold-emailed her: Thanks to you, my research is informed by a deep historical perspective of where this field has come from. I hope that by archiving these papers digitally[1], I make it easier for others in our field to learn from the foundational research you and your colleagues conducted decades ago.

To Aaron Steinfeld, Takayuki Kanda, Zhi Tan, and Bernardine Dias: I could not have asked for a better set of formative research mentors to lay the foundation for the researcher I am today. Having now mentored several undergraduates myself, I have an even deeper appreciation for the amount of time and effort you put towards guiding me as a budding researcher, from as young as 18 years old.

To the teachers and mentors who have been deeply influential over my education: Mark Stehlik, Kunal Ghosh, Nico Slate, Katharina Reinecke, Judith Schachter, Aileen Owens, Mehbub Karmali, Jaime Richards, Sai Kumar, Dr. Ricks, Mrs. Ferreira, Ms. Warden, Ms. Cortez, and Sisirkana Dhar Choudhury. Thank you for nurturing my curiosity and creativity. The fact that I am a lifelong learner is thanks to you.

To the students I have mentored and taught: You have made me a better teacher and mentor, you have challenged me to continually deepen my knowledge, and you have made me proud of all the great things you have achieved. You are as much a part of my journey as I am of yours.

[1] `https://robotfeeding.io/publications/`

To my friends: There are far to many to name. The late night heart-to-heart conversations we have had, the frivolous (and sometimes fierce) board game nights we have had, and the many laughs we have shared, are what got me through this PhD.

To my partner, Anand: All the memes I sent you about the "PhD partner" are true. Thank you for cooking (and bearing with me) when I was working late, for offering an empathetic ear when I was stressed, for providing sprinkles of joy throughout the day, and for grounding and balancing me over the years.

To my parents, Chaya and Sameer, who nurtured in me the deep curiosity and perseverance that carried me through this PhD. To my brothers, Akshay and Anuj, and my honorary sibling Parul, who occasionally succeeded in pushing me to have a better worklife balance, and always succeeded in bringing out my inner goofiness. To my aunts, Priti and Priya, to whom I owe my human-centered perspective and my interest in writing. To my grandparents, who taught me the value of hard work and how to cook the meals that powered me through my PhD. I would not be who I am today without your love and support. Thank you.

I have almost certainly forgotten others who have been deeply influential along this journey. You know who you are. From the bottom of my heart, thank you.

DEDICATION

To my family, for their loving and unwavering support

Contents

List of Figures

List of Tables

Chapter 1

Introduction

"*[Physically assistive robots] would decrease the workload on family members, help with caregiver burnout, and maybe in the future help a disabled person [like me] have more independence.*" – Tyler Schrenk[1], 1985–2023

According to estimates by the World Health Organization (WHO), 1.3 billion people worldwide experience significant disability [313]. These disabilities hinder one's ability to independently perform activities of daily living (ADLs) such as eating, ambulating, and dressing [79]. In the United States alone, at least 6 million adults face challenges doing errands independently [281]. While most people with disabilities wish to live independently in their home [88, 119], these challenges can threaten their ability to do so, leaving them reliant on a caregiver for assistance with ADLs. Besides their impact on day-to-day activities, disabilities also take a psychological toll and can lead to mental health challenges [70].

The social model of disability argues that disability is a result of the mismatch between a person's abilities and their environment [188], and advocates to bridge the gap between our inaccessible world and diverse abilities. Universal design [228] has helped bridge the gap in accessing the digital world, allowing increasing proportions of people with disabilities to program computers and access the internet [2]. However, the gap in accessing the *physical* world remains. Physically assistive robots (PARs) have emerged as a promising way to help bridge that gap. PARs are robots that provides assistance to humans through physical interaction, for example by helping to feed users, dress users, and pick up and move objects for users[2].

[1] https://www.thetsf.org/. Quote from Tyler's keynote at the Inaugural Robot-Assisted Feeding Retreat on Oct 21, 2022

[2] This contrasts with socially assistive robots (SARs), which are robots that provide assistance to humans through social inter-

Figure 1.1: The robot-assisted feeding system developed in this thesis consists of a 6-degree-of-freedom (DoF) robot arm, that can attach to a power wheelchair or hospital table. Users interact with the system through a web app that they can interact with using assistive technologies of their choice. The images on the right show the meals fed as part of the multi-user out-of-lab study and the images on the left show meals fed as part of the single-user in-home deployment (Chapter 6).

One specific type of PAR that has been studied since at least the 1970s is robot-assisted feeding (RAF) systems [243]. RAF systems developed in research labs in the 1970s and 1980s were evaluated in clinical studies and home deployments [243]. This laid the groundwork for commercial RAF systems released in the 1990s and 2000s [127, 177, 11]. While these commercial systems have shown positive results in terms of promoting user independence [177, 164], they have struggled to achieve long-term adoption. One reason could be the technology's limitations in reliable bite acquisition [277, 159, 178] and inflexible bite transfer [220]. In response to these shortcomings, a contemporary push in RAF research began in the 2010s, to develop robot-feeding systems with more generalizable bite acquisition and more customized bite transfer [276, 93, 30, 236, 226, 218]. Despite considerable technical advances, there have been relatively few in-home deployments of these contemporary research systems, with notable exceptions including [109, 277, 218]. This mirrors a broader trend in PAR research over the last decade, where despite considerable technical advances there have been relatively few in-the-wild robot deployments [210].

This thesis focuses on robot-assisted feeding as a case study for achieving in-home deployments of physically assistive robots. RAF is a well-suited case study due to the established user need, the long history of technical research, and the interest from both academia and the industry.

action, e.g., helping to motivate users to exercise or providing autism therapy to children [190].

More specifically, this thesis is guided by the following research question:

RQ-Thesis **How can we develop a deployable robot-assisted feeding system that enables any user, in any environment, to feed themselves a meal of their choice, while aligning with their preferences?**

To answer RQ-Thesis, this thesis presents works that investigate the following research questions:

RQ0: What applications, methods, and themes underlie the last decade of research on physically assistive robots for people with disabilities?

RQ1: What challenges do people with motor impairments face during dining, and how should a robot-assisted feeding system be designed to address those challenges?

RQ2: How can a robot-assisted feeding system feed users the variety of food items they want to eat?

RQ3: How can we develop a robot-assisted feeding system to feed users in diverse out-of-lab and in-home contexts?

The *key challenge* of developing a deployable robot-assisted feeding system is enabling it to work reliably across the various contexts the robot will inevitably encounter.

- Users eat various different foods: sandwiches, noodles, salads, fried chicken, etc.
- Users eat in a various environments: in bed, in their living room, in their office, in a cafeteria, etc.
- Users eat in various social contexts: by themselves, with caregiver(s), with loved ones, etc.
- Users eat while doing various other activities: watching TV, working, conversing, etc.

Our *key observation* is that there is inherent structure within the broader system the robot embeds into.

- There is structure within the task of feeding. For example, bite acquisition motions tend to first approach the food, then grasp the food, and then extract the food from the plate.
- The people around the robot have structured relations with the robot. For example, the end-user is physically co-located with the robot, participates in the task synchronously with the robot (i.e., they have to chew after the robot delivers a bite), and has goal alignment with the robot (i.e., eating a meal).
- There is structure in the environment the robot is in. For example, wheeled, adjustable "hospital tables" are often used as part of the care routines of people with motor impairments.
- There are other tools that structure the user's interactions with the world. For example, users often have impairment-specific assistive technologies that enable them to use smartphones or laptops.

This informs the key insight and thesis statement of this work. *In order to develop deployable PARs, it is crucial to understand and leverage the structure within the broader system the robot embeds into. In the case of robot-assisted feeding, this involves: (1) understanding users' circumstances and priorities [RQ1]; (2) leveraging structure within bite acquisition motions to learn motion primitives that enable the robot to generalize across foods [RQ2]; and (3) leveraging structure in the user's relation with the robot to overcome off-nominal scenarios, by providing them with control and customizability over the system [RQ3].*

The outcome of this thesis is an open-source[3] robot-assisted feeding system that has demonstrated success feeding 6 users with motor impairments meals of their choice in diverse out-of-lab environments—a cafeteria, office, conference room, and end-user's home—for over 15 hours. Figure 1.1 shows the RAF system and several of the meals it fed.

1.1 Methodology

In the words of Ladner [169], "some of the best work comes when there are people with disabilities on the design and development team, contributing to all aspects of the design and implementation...the users of the technology are empowered to solve their own accessibility problems." Thus, the work presented in this thesis follows principles from community-based participatory research (CBPR). CBPR is a method where academics researchers work equitably with community members throughout all research stages, from ideation to dissemination [115, 301]. CBPR is rooted in the belief that community members and academic researchers each bring unique skills, expertise, and lived experiences to the team; for example, while academic researchers are familiar with rigorous research methodologies, community members are familiar with the nuances of their lived experience of disability. Thus, addressing a community need requires an equitable partnership between academics and the community, which involves sharing power, resources, credit, results, and knowledge [239, 198]. CBPR has been used in the health sciences for decades [138, 301], and is increasingly used in assistive technology research [188, 50, 167, 25].

More specifically, the research that investigated RQ1, RQ2, and RQ3 was done in collaboration with two community researchers (CRs), Tyler Schrenk[4] (CR1), and Jonathan Ko[5] (CR2).

[3] `https://robotfeeding.io/publications/hri25a/`
[4] https://www.thetsf.org/
[5] https://linkopllc.com/about/

For RQ1, Tyler worked as an equal participant in the research team. His involvement in this work included, but was not limited to, the following:

1. co-creating the study question (e.g., the fact that it focused on *social* aspects of dining was informed by conversations with Tyler);
2. co-creating study materials (e.g., the videos participants were shown);
3. co-designing the study procedure (e.g., study phases, interview questions, study length);
4. helping to recruit study participants, through the network he built up running his assistive technology non-profit[4];
5. conducting the studies (i.e., Tyler asked participants questions and led follow-up discussion, while academic researchers took notes);
6. co-analyzing the data (i.e., sharing insights on emergent and salient themes);
7. co-creating the diagrams that disseminated key insights from the data (e.g., Chapter 4's Figure 4.6 and Figure 4.7);
8. co-authoring the paper.

Throughout the collaboration, the research team (of whom Tyler was an equal member) met semi-weekly. Several of these meetings involved reflecting on the methodology we were following, to ensure we were creating a space where we all could equitably share power, share our expertise, and participate accessibly. Readers are encouraged to read Sec. 4.9, a section Tyler was heavily involved in writing, to better understand how CBPR was applied to that work.

Our approach to RQ2 involved far less collaboration with CRs. This was partly CR-requested; Tyler said he was not as interested in the technical details of how bite acquisition works, as long as it works reliably across various food items. To help us better understand the variety of food items he was interested in, Tyler tracked a list of the foods he ate over a week. This formed the basis for the list of food items we focused on in RQ2. We periodically provided Tyler updates about the research, but his primary involvement was in creating the list of foods we focused on.

RQ3 involved close collaboration with Tyler (for the first half) and Jonathan (for the second half). This is because Tyler passed away half-way through the research, and his friend, Jonathan, wanted to honor his legacy by continuing the work. Throughout all our work on RQ3, we had semi-weekly meetings with either

Tyler or Jonathan. Their involvement in this work included, but was not limited to:

1. co-creating the list of off-nominal scenarios that could arise (Table 6.2);
2. co-designing the system by sharing their insights on key system design decisions (e.g., the decision to use a web app as the interface to interact with the robot, the decision to put the web app in charge of system execution, the icons and colors used to communicate key concepts, etc.);
3. participating in pilot studies of system components (e.g., the bite selection pipeline) and the entire system;
4. co-designing the procedure for "Study 1: Multi-user, On-campus Study" (e.g., study length, phases, questionnaires, etc.);
5. co-designed the procedure for "Study 2: Single-user, In-home Deployment" (e.g., what Jonathan was doing for each meal, what he was eating, which meals he ate on which day, etc.);
6. sharing insights on how the system could better align with their relationships with caregivers, family, and friends; and
7. co-authoring the paper.

Throughout this work, we visited both Jonathan and Tyler's homes multiple times, including to eat meals together, which was a crucial part of building team cohesion and building a shared understanding of their meal contexts. Although academic researchers were creating the system hardware and software, Jonathan and Tyler provided the goals and features that academic researchers were working towards. Academic researchers provided Jonathan and Tyler the option to delve into the system at whatever depth they desired; in one context, this included directly sharing code with Jonathan, but more often this involved distilling key aspects of the system into an accessible presentation that we then shared with them. We taught them concepts that they were interested in learning, such as different approaches to object detection in computer vision. Throughout feature development, academic researchers iteratively met with CRs to get continuous user feedback. These meetings also sometimes involved reflection to ensure that this collaboration and research process was aligning with everyone's goals.

In summary, the work in this thesis aligns with multiple core tenets of CBPR [138]. It was based on a **long-term relationship** that spanned multiple research projects. We **equitably shared power**, for example through co-authoring papers and through monetary compensation at a rate that CR's felt was fair ($50/hr).

This equitable power-sharing **extended beyond the research**; for example, we connected Tyler with experts to help him appeal denied health insurance coverage. We **learned from each other**; Tyler and Jonathan taught us a great deal about their lived experiences with disability, and we taught them different technical aspects about how robots are developed (e.g., the sense-plan-act paradigm, different types of computer vision algorithms, etc.). We **reflected on the research process** we were following, the goals we each had for our collaboration, the accessibility of our collaboration, and more. We **paused the research process as community researchers needed/requested**, irrespective of whether that would align with academic timelines. The entire process was **cyclic and iterative**, involving semi-weekly meetings and repeatedly revisiting earlier topics of discussion. We worked to ensure research publications were **accessible to the community**, for example by presenting key insights through both textual and visual formats [208], ensuring screen-reader compatible documents [208, 212], creating a plain-language paper summary[6], and presenting the work at forums for people with motor impairments (e.g., the Northwest Regional Spinal Cord Injury System (NRSCIS)).

The work in this thesis also deviates from certain tenets of CBPR [138]. First, two individuals can never encompass the diversity of an entire community. While the 15 *participants* in RQ1 and RQ3 spanned multiple genders, ages, cultural backgrounds, and disability types, the 2 *community researchers* had multiple demographic similarities: technology enthusiasts, working professionals, middle-aged, males, with spinal cord injuries (SCIs). The community of people who can benefit from robot-assisted feeding is extremely diverse, and more fully encompassing that diversity within the group of community researchers would further strengthen the research and benefit the community. Second, we did not deeply engage with other community stakeholders such as friends, families, and more. While RQ3 did involve some collaboration with caregivers and occupational therapists, they were chiefly involved in the "Study 2", not throughout the whole research process. Third, a tenet of CBPR involves applying an ecological perspective to the work [138]; while RQ3 did focus on meal contexts, it focused little on how the system can integrate into care routines beyond just the isolated meal. Finally, although we took some actions to provide collective community benefit, such as presenting the work at the aforementioned NRSCIS forum, there is always room to deepen the work we do to benefit the community. For example, future research could approach RAF through a series of work-

[6] `https://robotfeeding.io/publications/hri25a/`

RQ-Thesis	How can we develop a deployable robot-assisted feeding system that enables any user, in any environment, to feed themselves a meal of their choice, while aligning with their preferences?
RQ0	What applications, methods, and themes underlie the last decade of research on physically assistive robots for people with disabilities?
RQ1	What challenges do people with motor impairments face during dining, and how should a robot-assisted feeding system be designed to address those challenges?
RQ2	How can a robot-assisted feeding system feed users the variety of food items they want to eat?
RQ3	How can we develop a robot-assisted feeding system to feed users in diverse out-of-lab and in-home contexts?

Figure 1.2: List of research questions that constitute this thesis.

shops, where community members learn how to develop hardware and softare assistive technologies while contributing to the research.

Overall, this thesis has been deeply enriched by the methodological inspiration we took from community-based participatory research (CBPR). We encourage future researchers to continue to follow tenets and best practice from CBPR, and address some of the deviations listed above, to more fully ground the research in community needs and priorities.

1.2 Roadmap

The remainder of this thesis approaches the research questions (Figure 1.2) as is detailed below. Note that where a chapter was originally published in another venue, we recommend the reader read that original paper, for two reasons. First, because the chapter was written for the formatting norms of that particular venue, the original paper may provide a better reading experience. Second, the original papers provide a more accessible reading experience, for example with detailed alternative text for figures.

Chapter 2 presents a systematic survey of research of mobile and manipulator physically assistive robots (PARs) for people with disabilities, originally published in [210]. This chapter seeks to answer RQ0, by studying the application domains, methods, and themes that underlie recent PAR research. A key result from this work is that PARs are primarily evaluated inside the lab, which motivates this thesis' focus on out-of-lab deployability. Another key result is that the themes of "levels of autonomy" and "adaptation" are relevant to PAR research across different application domains. This finding motivates our approach to system design and engineering in RQ3.

The thesis then moves on to Chapter 3, which motivates why robot-assisted feedings (RAFs) systems are a fitting case study for in-home deployability of PARs. This chapter presents the history of RAF research, from the 1970s till now, and highlights some of the key ways in which the research presented in this thesis aims to improve upon the state-of-the-art.

Chapter 4, originally published in [208], seeks to answer RQ1. Specifically, this chapter presents the results of a qualitative study with 10 participants with motor impairments, investigating their current dining experiences and how they would like robot-assisted feeding systems to be designed. This chapter particularly hones in on the social aspects of dining, a topic that has been under-studied in past works. Key insights from this work are that users desire control over their RAF system and want a highly customizable robot. These form the guiding design principles we follow when developing the system in RQ3.

Chapter 5, originally published in [104], focuses on RQ2. Specifically, this work presents a schema that represents a broad space of robot-executable single-utensil bite acquisition actions, and then learns which actions within that schema are representative strategies that people without motor impairments use to acquire food with a fork. This work generalize the robot feeding system's bite acquisition capabilities to be able to acquire a large variety of food items users may want to eat.

Chapter 6, originally published in [212], focuses on RQ3. Specifically, it presents the system design and integration work necessary to develop a robot-assisted feeding system that can be used out-of-lab to feed people meals of their choice. This work specifically focuses on the design principles of reliability, portability, safety, user control, and customizability. We evaluate this system with two studies (Figure 1.1). Study 1 is a multi-user, out-of-lab, quantitative study where 6 users with diverse motor impairments use the robot to eat a meal of their choice in a cafeteria, office, or conference room. Study 2 is a single-user, in-home deployment where the system feeds a user 10 meals over 5 consecutive days across various contexts: while watching TV, while talking to a caregiver, eating in-bed, etc. This chapter concludes with lessons learned regarding deploying PARs in users' homes.

Finally, Chapter 7 concludes the thesis with a discussion of limitations and future work.

Chapter 2

A Systematic Survey of Physically Assistive Robots for People with Disabilities

This chapter focuses on the overarching field of physically assistive robots (PARs), and seeks to answer the question:

RQ0 What applications, methods, and themes underlie the last decade of research on physically assistive robots for people with disabilities?

This chapter was originally published as "Physically Assistive Robots: A Systematic Review of Mobile and Manipulator Robots That Physically Assist People with Disabilities" in the *Annual Review of Control, Robotics, and Autonomous Systems* [210]. The literature review for this survey was conducted in spring 2023, and thus this chapter excludes papers published after that.

2.1 Introduction

The World Health Organization estimates that 1.3 billion people around the world experience significant disability [313]. Whether due to congenital conditions, injury, illness, or acquired with age, disabilities can impact people's ability to independently perform activities of daily living (ADLs) and therefore reduce their quality of life. According to the CDC, at least 6 million adults in the US have difficulty doing errands independently [281]. While most people with disabilities wish to live independently in their home [88, 119], such difficulties can threaten their ability to do so. Besides their impact on day-to-day activities, disabilities also take a psychological toll and can lead to mental health challenges [70].

The social model of disability argues that disability is a result of the mismatch between a person's abilities and their environment [188] , and advocates to bridge the gap between our inaccessible world and diverse abilities. Universal design has helped bridge the gap in accessing the digital world, allowing people of many abilities to program computers and access the internet. However, the ability gap in accessing the physical world remains.

Physically assistive robots (PARs) present a unique opportunity for enabling access to the physical world for people with disabilities. A PAR is a robot that provides assistance to humans through physical interaction. PARs include robots that help feed users, dress users, help users move, pick up and move objects for users, replace limbs (e.g., prosthetics), rehabilitate limbs, augment the body (e.g., exoskeletons and wearable robots), and more.[1]. Many activities of daily living that are difficult or impossible due to a person's impairment—such as independently feeding or ambulating—are physically possible for a robot to perform (Fig. 2.1A). However, developing robots that safely and robustly perform these tasks in diverse environments, with diverse user impairments and preferences, is challenging. Many open questions remain as to how robots should be designed, what user interfaces to use, what levels of autonomy they should have, and more. These questions have fueled research in PARs for decades.

Within the space of physically assistive robots intended for people with disabilities, this survey specifically focuses on *mobile and manipulator* robots. A mobile robot is a robot that can move its own base (e.g., a robotic vacuum cleaner). A manipulator is a robot that can manipulate objects; for instance, by picking them up and moving them around (e.g., a robotic arm). A mobile manipulator is a robot that can move do both (e.g., a humanoid robot). The reasons for focusing on mobile and manipulator robots are:

1. There are several recent surveys on prosthetics [15, 170], wearable robots [262, 314], and rehabilitation robots [182, 200, 182]. To the best of our knowledge, there has not been a comparable focus on surveying mobile and manipulator robots for physical assistance.
2. Over the past decade, the number of papers researching mobile and manipulator PARs has increased

[1]This contrasts with a socially assistive robot (SAR), a robot that provides assistance to humans through social interactions. Examples of SARs and that are not PARs include robots to: help provide autism therapy to children, serve as social companions to elderly people, and help motivate their users to exercise [190]

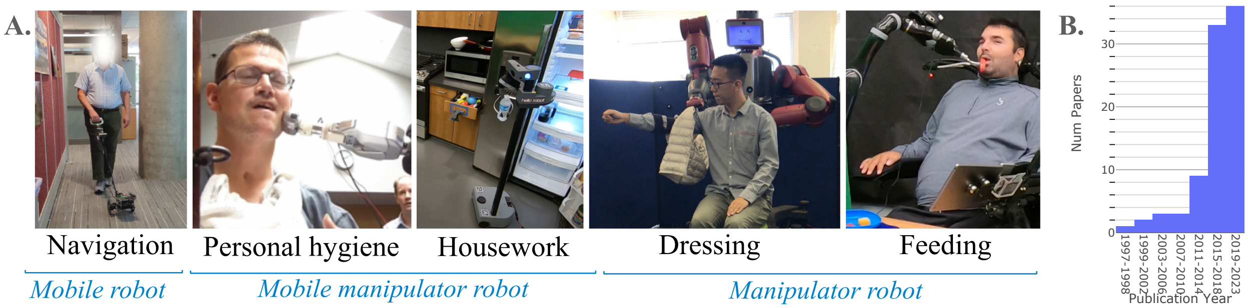


Figure 2.1: A. Common domains of assistance, exemplifying the different types of robots: mobile [251], mobile manipulator [321, 124, 40], and manipulator [194, 30]. (First, second, and fourth images: Reprinted from [251] (CC BY 4.0). ©2012 IEEE. Reprinted, with permission, from [124]. ©2019 IEEE. Reprinted, with permission, from [321].) **B.** Number of papers in this review by year published.

several-fold (Fig. 2.1B). Yet, this research has been siloed by domain of assistance, e.g., robot-assisted feeding and robot-assisted navigation, and there is little dialogue about takeaways that cut across these domains.

3. The formative studies that highlight the needs and preferences of people with disabilities tend to be published in venues focused more on human factors—e.g., CHI, ASSETS, PeTRA, RO-MAN—and do not always reach the roboticists capable of meeting those needs.
4. Mobile and manipulator PARs are increasingly being deployed in real-world settings [277, 155, 113, 219], which is a welcome advancement but makes it more important to have conversations within the field about safety, robustness, working with people with disabilities, and more.

Our goal with this survey is to fuel progress in mobile and manipulator PARs by: (1) highlighting existing research; (2) inspiring more roboticists to apply their skills towards PARs; and (3) systematizing methods so researchers can more easily work with people with disabilities.

2.1.1 Relation to Other Survey Papers

Newman et al. [215] present a survey of physically and socially assistive robotics in general. Our work differs from theirs by focusing on people with disabilities, who have specific needs and constraints that must be taken into account when developing assistive robots.

Matarić and Scassellati [190] focus on socially assistive robots. Although social assistance is beyond the scope of this survey, the reality is that disability is intersectional. As a result, robots that holistically support

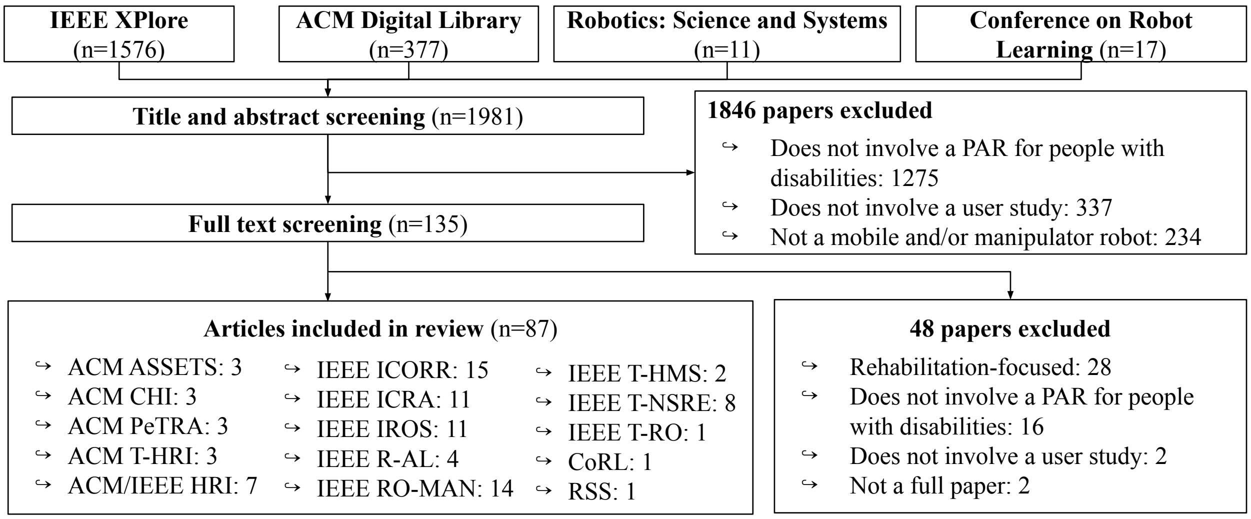


Figure 2.2: The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram for this paper. We screened 1981 papers and include 87 in this review.

people with disabilities will likely need to integrate both physical and social assistance.

Brose et al. [36], Chung et al. [62] focus on physically assistive robots. Although we report on some similar themes, such as user interfaces and shared control, their surveys were written before the last decade's drastic increase in PAR papers (Fig. 2.1B).

Mohebbi [200] reviews the human-robot interaction of physically assistive robots. While we have a section dedicated to interaction interfaces (Sec. 2.5.1), we also focus on other topics such as the methods used in user studies.

Some surveys focus on assistive robots for particular populations—people with quadriplegia [227], older adults [242], and people with visual impairments [152]. Our paper brings together work focused on multiple types of disabilities and domains of assistance, to facilitate meaningful dialogue *across* the field of physically assistive robots.

As mentioned above, there are several recent surveys on prosthetics [15, 170], wearable robots [262, 314], and rehabilitation robots [182, 200, 182]. Although beyond the the mobile/manipulator scope of this survey, such works no doubt have relevant insights for PAR research more broadly.

2.2 Survey Methodology

We began by curating a list of top conferences and journals in robotics and assistive technology (Fig. 2.2). From those venues, we searched for full papers whose title, abstract, or keywords had "robot" and either: "assistive," "accessibility," "disability," "impairment," or forms thereof. This resulted in 1981 papers. We then screened the title and abstract for the following inclusion criteria: The paper involves (1) a PAR for people with disabilities or older adults, (2) a user study, and (3) a mobile, manipulator, or mobile manipulator robot.

We aligned our interpretations of the above criteria by having a random selection of 60 papers tagged by two or three researchers and discussing any differences until we reached consensus. The rest of the papers were split amongst the three researchers for tagging. 135 papers remained after this title and abstract screening. We then conducted full-text screening. At this stage, we also removed works that had a rehabilitation focus, due to the existence of existing surveys devoted to recent trends in rehabilitation robotics [200, 182]. This resulted in 87 papers included in this review. Fig. 2.2 shows the entire pipeline.

While reading the papers, we iteratively met to converge upon dimensions along which the papers are similar/different that would be of interest to the PAR research community. These dimensions are: Descriptive Statistics (Sec 2.3), Types of User Studies (Sec 2.4), Interaction Interface (Sec 2.5.1), Levels of Autonomy (Sec 2.5.2), and Adaptation (Sec 2.5.3). For each dimension, we developed discrete codes by describing and clustering the works (bottom-up), and then identifying existing frameworks that the codes mapped to (top-down).

This systematic survey was conducted in spring 2023, and thus does not include papers published after that.

2.3 Descriptive Statistics About the Papers

2.3.1 Domain of Assistance

For every paper, we coded the domain(s) of assistance that the PAR helped the user with. This classification drew upon activities of daily living (ADLs) and instrumental activities of daily living (IADLs), a framework for classifying the skills and activities necessary to live independently. ADLs are the "skills required to

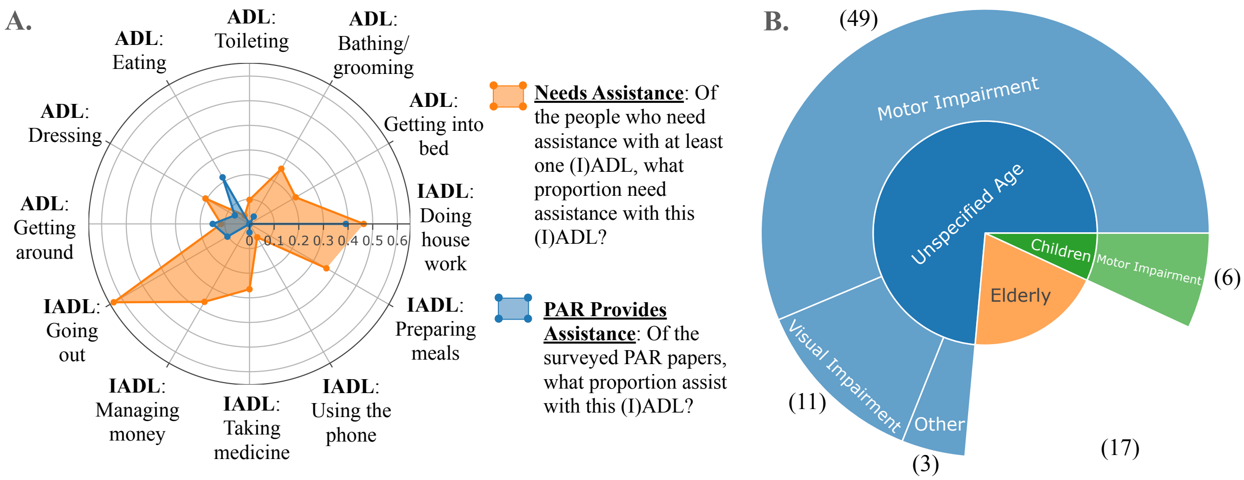


Figure 2.3: A. The proportion of people who need assistance with each (I)ADL versus the proportion of PAR papers that assist with that (I)ADL. **B.** Papers in this review by the target population's age (inner) and disability (outer).

manage one's basic physical needs" while IADLs are the "more complex activities related to the ability to live independently in the community" [79]. We then compared the proportion of PARs that focus on each (I)ADL to the proportion of people who need assistance with that (I)ADL [288], in Fig 2.3A.

There are three spikes amongst PAR research, for (I)ADLs focused on *navigation*, *feeding*, and *doing housework*. For the navigation domain, we characterized works that focused on navigating in *any environment* (e.g., fall prevention [97, 259], standing assistance [55, 61]) as "getting around," and works that focused on navigating in *environments outside the home* (e.g., guide robots for people who have visual impairments [251, 323, 155, 156]) as "going out." For the housework domain, we classified all "pick-and-place" works, that focused on assisting with the general manipulation of objects, as housework. However, such works can also help with going out (e.g., opening doors) and managing medication (e.g., bringing medication to a user). Note that even if the proportion of PAR research is similar to or greater than the proportion of people who need assistance, that does not mean out work is done; formative studies have found numerous ways in which PARs must be improved [208, 28, 5, 20, 40].

Some (I)ADLs have a high user need for assistance but proportionately little PAR research—dressing, bathing / grooming, and managing medications. Extending the existing research in these realms (Table 2.1) would be a fruitful direction for future work. There are also some (I)ADLs that have no papers from this survey. Some, such as difficulty toileting and difficulty getting out of bed, may require special hardware [114, 273] that go beyond the mobile/manipulator focus of this survey. Others, such as difficulty managing money or using the phone, are better served by either non-robotic solutions or SARs, as opposed to PARs [117, 260].

2.3.2 Target Population

We coded the target population age for each paper as one of: "children," "elderly," or "unspecified age" (which was typically adults across ages). We also coded the target population's disability (if any) as one of: "motor impairment," "visual impairment," or "other[2]." Fig 2.3B presents this data. The bulk of PAR research is motivated by three target populations: people with motor impairments, people who are blind or low-vision, and older adults. This drastically differs from the target populations of SAR research: people with autism, people with dementia, and older adults [190].

2.4 User Studies in PAR Research

For every work, we coded the type of study, number of participants with and without disabilities, what was being evaluated, and the methods used. We coded the type of study as either "formative," "summative," or "both." Formative studies take place in the early stages of research and help "form" the design for the system, while summative studies take place near the end of system development and help evaluate. or "sum up," the system [176, 123]. Fig 2.4 and Table 2.1 show the distribution of papers along these metrics. 14 papers (17%) included a formative study, with the rest including only summative studies[3] (Fig. 2.4A).

2.4.1 Involvement (or Lack Thereof) of Participants With Disabilities

Half of the papers involved no participants with disabilities, while the other half involved at least one[4] (Fig 2.4A). Notably, **nearly all formative works involved people with disabilities**. This is crucial to ensure that the early decisions that are made in a research area are informed by the needs of the target population. In

[2]Works with a target population of "other" either focused generally on "people with disabilities" [49, 84] or used a different form of categorization, e.g., "people in skilled nursing facilities" [320].

[3]All papers that collected data and trained a model were considered both formative—for the data collection and analysis—and summative—for the model evaluation.

[4]We determined whether people with disabilities were involved by reading the main text of the paper, which sometimes fails to mention whether a co-author has the target disability. Some papers where members of the research team have the target disability include [155, 149, 251, 14, 208]

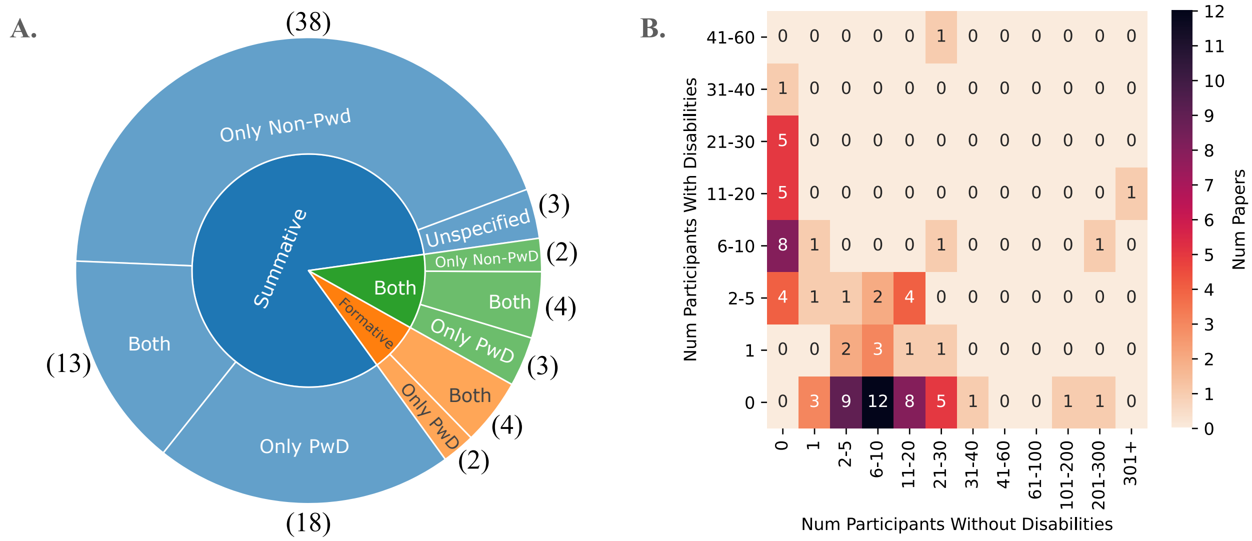


Figure 2.4: A. Papers included in this review by type of study (inner) and whether they included users with disabilities (outer). **B.** How many participants with(out) disabilities each paper had.

contrast, **the majority of summative evaluations involved only participants *without* disabilities**. Some works framed these evaluations as "preliminary," "pilot," or "proof-of-concept," [116, 248, 53, 3, 230, 64] giving the impression that an evaluation with participants with disabilities is forthcoming. We found a few instances amongst the reviewed papers with a followup evaluation with participants with disabilities, e.g., [40] followed up on [41], [30] followed up on [93]. In some works, researchers claimed to simulate disability amongst non-disabled participants through blindfolds [303, 3], braces [320, 321, 238], or intentional falls (e.g., to simulate older adults falling) [97, 259]. Although simulations can be a rapid way of testing capabilities of a robotic system, they are considered problematic in the disability studies literature and should always be complemented with studies involving the target population [38].

About a quarter of works involve participants with *and* without disabilities (Fig 2.4B). In some cases, participants without disabilities were caregivers [28, 203], occupational therapists [28], and other stakeholders [155, 249, 69]. In other cases, researchers ran a large-sample study with people *without* disabilities to collect statistical insights, followed by a small-sample study with people *with* disabilities to collect qualitative insights [9, 40, 234, 277, 113].

Table 2.1: Domain of assistance and type of study for all papers in this review

Domain of Assistance	Formative		Summative: What is being evaluated?				
	Formative	Dataset	Interaction Interface	Level of Autonomy	Specific Functionality	Whole System: In-Lab	Whole System: In-Context
Eating	[28, 208, 5, 10, 178]	[226]	[172, 30, 5]	[30, 179]	[226, 93, 87, 240, 234, 256, 47]	[101, 266, 277, 84]	[277]
Dressing	[]	[153]	[]	[]	[153, 82, 320, 83, 321, 47]	[]	[]
Bathing / Grooming	[]	[124]	[]	[]	[110]	[84]	[124]
Taking Medicine	[203, 40]	[]	[40]	[]	[]	[263]	[203]
Pick-and-Place / Housework	[20, 40, 5, 10]	[]	[40, 249, 270, 57, 17, 142, 245, 41, 53, 64, 293, 322, 139]	[134, 224, 126, 194, 248, 307, 306, 146, 103, 102]	[302, 9, 58]	[263, 270, 116, 230, 91]	[16, 69]
Playing	[35]	[]	[35]	[]	[]	[162]	[120, 68]
Working	[]	[]	[]	[]	[]	[]	[49]

2.4.2 Formative Studies

Involvement of target users in formative research is particularly critical to ensure that researchers: (a) work on problems that are actually important to the target users; and (b) are aware of user constraints and preferences that should be taken into account when developing assistive technologies. This was reflected in the proportion of formative research in our survey that involved people with disabilities. On the other hand, the proportion of formative research to summative research was small, with only five papers that involved solely formative studies [28, 20, 208, 14, 10] and five that included a formative study and summative study [35, 203, 155, 40, 5]. This is in contrast with other research focused on (non-robotic) technology for people with disabilities. For example, a recent survey of technology for people with visual impairments found more formative than summative research [37]. One reason for this finding could be the lack of familiarity with formative research methods in the robotics community and the emphasis on quantitative findings.

Dataset collection for training a model was rare in the PAR literature, with only four papers [124, 153, 226, 48], despite the popularity of the approach in the robotics community. In all cases, the data was collected to model a component of the system, e.g., for gait tracking [48], force prediction [124], failure prediction [153], and bite timing prediction [226]. None of the papers reported on generalizable formative insights based on the collected data.

A variety of formative research methods were exemplified in the papers: surveys [20, 10], interviews [208, 10, 203, 155, 16], group interviews [20, 35, 10], contextual inquiry [28, 10], participatory design [14, 10], observational studies [5, 40],workshops [16, 10], and ethnography [10]. Some papers combined methods. For example, Beer et al. conducted a written survey with older adults to assess the tasks they would like assistance with, and then followed up with a group interview to understand why they held those preferences [20].

Formative studies on PARs contribute insights that other researchers can use when designing, developing, and/or evaluating similar PARs. The findings from formative research can be presented as design constraints [203] or guidelines [208, 14, 237], evaluation frameworks [28], limitations of existing systems [40, 5], participants' concerns and potential opportunities [155, 35, 178], and directions for future work [20, 208]. Note that some works conducted a formative study to understand the users' needs and then a summative study to evaluate the resultant system [155, 203]. Further note that some summative studies can also yield formative insights such as users' preferences on the system's form factor [323].

2.4.3 Summative Studies

Summative studies either evaluate a specific component of the system (the middle three columns of Table 2.1) or the whole system (the last column of Table 2.1), and gather quantitative and/or qualitative data to conduct that evaluation.

2.4.3.1 What is being evaluated?

Studies evaluating a system component focused on the:

- **Interaction Interface**: how users send and receive information to/from the robot.
- **Level of Autonomy**: how much of the sensing, planning, and acting of the system is done by the robot versus the user.
- **Specific Functionality**: any robot functionality that does not fall into the above two categories, such as domain-specific functionality.

These studies typically compare the specific component of their system to one or more baselines, which are either state-of-the-art approaches [48, 146, 194, 226, 9, 55, 179, 248, 245, 302, 17, 57] or variants of their component with some subcomponents systematically removed, i.e., ablation studies [293, 83, 3, 320, 307, 303, 116, 35]. Most of these studies are within-subjects, where each participant experiences every condition, which is better when there is high variance across participants [98], such as with participants with disabilities.

Studies that evaluate the whole system sometimes move beyond the lab and into the user's context-of-use. Of these, some are **field studies**, which involve running a structured study in the context-of-use [68, 124, 120, 184, 49, 80], while others are **deployments**, which involve letting users freely interact with the robot in the context-of-use [69, 203, 277, 113, 156, 124, 16]. Note that most whole system evaluations are non-comparative. This may be due to the large amount of resources required to develop a whole other system.

2.4.3.2 What data is being collected?

Most summative studies in this review gathered **quantitative data**, which can further be divided into objective and subjective metrics. **Objective metrics** are often task-specific, such as task completion time [146, 40, 58, 245, 323, 134], the number of mode switches [126, 103], success rate [233, 302, 172], classification accuracy [110, 226, 153], among others. **Subjective metrics** often focus on user preferences regarding different versions of the robot. Many researchers create their own Likert-scale questions that focus on topics such as usability [245, 172, 134, 146, 179, 184], preference [134, 226, 323], satisfaction [277, 57, 126], feeling of control and safety [293, 9], and more. Others use standardized subjective metrics, such as the System Usability Scale [323, 263], NASA-TLX [40, 113, 251], and Psychosocial Impact of Assistive Devices [323]. Note that objective and subjective metrics have complementary benefits—objective metrics are not impacted by biases in self-reporting, but subjective metrics are more grounded in users' preferences [324]—resulting in many studies that use both [323, 251, 40, 172, 57, 134, 16]

Multiple summative studies paired quantitative data with **qualitative data**. Qualitative data can help to understand nuances of user preferences, gain insights into additional features users want, or contextualize quantitative results [274]. To gather qualitative data several summative studies held semi-structured interviews [35, 10, 203] or focus groups [155, 156] *after interacting* with the robot, while others had participants share thoughts, insights, and reactions *while interacting* with the robot [40].

2.4.4 Suggestions for Physically Assistive Robot (PAR) User Studies

First, **we caution PAR researchers to not over-generalize from evaluations involving people without disabilities**, as "there is not yet enough evidence supporting the generalization of findings from non-disabled subjects to the [target] population"[19]. Further, it is different to live with versus simulate an impairment: "putting on a blindfold for half an hour...can't give you the full experience of living with a visual impairment for...40 years" [294]. While we acknowledge the challenges in running large-sample in-person studies with people with disabilities, alternatives exist [188], including remote studies [40, 9, 28, 30], video studies [9, 208], and working with a community researcher [208, 10].

Second, **when using objective metrics (e.g., accuracy, efficiency) we call on PAR researchers to justify *why* those metrics align with user preferences**. There is often the implicit assumption that users want their assistive robot to optimize the metric that researchers are measuring, but prior work has shown that is not always the case [30, 194]. As opposed to assuming an objective metric aligns with user preferences, it is important to *work with users* to identify objective metrics that align with their preferences.

Third, **we recommend PAR researchers use standardized scales, such as the System Usability Scale [175] or NASA-TLX [122], for whole system evaluations.** Because most whole-system evaluations are non-comparative, it becomes difficult to compare research systems across different labs and papers. Standardized metrics can address this, since they are designed to work across a variety of technologies and have standard interpretations of their numeric scores [175, 111]. In addition to the above standardized subjective metrics, standardized objective metrics—that measure the user's performance on a benchmark task—can facilitate comparisons across works and create a universal interpretation of performance (e.g., the ARAT test used in [109]).

Fourth, **we call for more formative research involving PwDs** to inform the development of PARS. Formative research can be especially impactful if its findings are synthesized into open problems for robotics (e.g., [208]), allowing other researchers to work on important challenges even without direct involvement by PwDs. Frameworks for describing assistance tasks and user requirements in detailed, structured ways, like SPARCS [185], can further increase the impact of formative work. Another avenue for accelerating progress based on formative research is the creation of **robotics benchmarks and simulations** for physical assistance. Choi et al. created a list of household objects used by people with ALS [59], allowing researchers who work on pick-and-place tasks to focus on the objects most frequently needed by this user group. Ye et al. conducted formative research with motor-limited individuals, caregivers, and healthcare professionals to inform the design of RCareWorld [319]—a simulation environment with realistic human models representing different disabilities, home environments, and common assistance scenarios.

Finally, **we call for more in-context research and deployments**, as laboratory studies "reflect an oversimplified view of HRI [human-robot-interaction]" [150]. For the in-context deployments that do exist, there is an unfortunate trend of relegating findings to a small section within the paper [113, 124, 277]. However, there are several details of deployments that would be useful for the research community to know to help accelerate future deployments: (a) engineering details regarding how the system was made robust and safe; (b) qualitative and quantitative insights into the user's experience with the system; and (c) findings regarding where the robot fell short. Although some may argue that small-sample deployments lack the statistical power of large-sample studies, we note that there is a large body of work in the experimental design and statistical analysis of “n-of-1” studies that could add methodological rigor to PAR deployments [287].

2.5 Overarching Themes

2.5.1 Interaction Interface

One overarching theme across these works is the interaction interface that allows users to send and receive information from the robot. Some works explicitly focused on understanding the tradeoffs between different interfaces for different individuals [57] or in different contexts [30, 208]. Even those works that did not explicitly focus on interaction interfaces still made design decisions as to which interface(s) were best suited to their application. This section provides an overview of the interfaces that are commonly used and tradeoffs amongst them, based on the Senses and Sensors Taxonomy [13].

2.5.1.1 Input Interfaces

The Senses and Sensors Taxonomy [13] differentiates between *direct processing*, or sensors that directly measure electrical stimuli sent from the brain, and *indirect processing*, or sensors that measure the outcome of those stimuli.

A small number of works use **direct processing**, such as electromyography (EMG) or electroencephalogram (EEG), to convert the user’s neural signals into inputs to the robot. Most works used EMG or EEG to teleoperate a robot in the pick-and-place domain of assistance [293, 91]. Others combined EMG/EEG with another input device, such as muscle contraction [172, 17], brain signals [240] or eye gaze [307], to teleoperate the robot.

A larger set of papers involve indirect processing through modalities of vision, audition, touch, and kinesthetic inputs. The **vision** modality contain sensors that *see* user inputs and send them to the robot. One common application is detecting whether the user is ready for the robot to move towards their face in robot-assisted feeding [266, 208], or robot-assisted drinking [101]. Another common application for vision is detecting object that a user wants the robot to acquire, e.g., using a laser pointer [57] or a gaze tracker [266, 270, 64]. Yet another application is to have the users completely control the robot with vision inputs [116].

The **audition** modality contains sensors that *hear* user inputs and send them to the robot. This includes interfaces that allow the user to give vocal commands to teleoperate a robot arm [134, 245]. This also includes systems where the user uses voice to specify the object they want the robot to acquire, such as a specific bite of food [30]. While audition sensors have the benefit of not requiring any body motion on the part of users, they may not work well in noisy settings [113] or social settings [208, 30].

The **touch** modality contains sensors that *feel* user inputs through direct contact and send them to the robot. This includes traditional methods of interacting with technologies, such as a mouse and keyboard [263, 41, 40, 224], joystick [245, 302, 162, 277, 249], or a touchscreen [57, 30, 14]. This also includes custom force-torque sensors used for robot-assisted navigation [323] or robot-assisted drinking [101].

The **kinesthetic** modality contains sensors that *feel user motion* and send then to the robot. This includes using inertial measurement units (IMUs) to sense users' head [139, 172, 17, 139] and upper body movements [142, 53] for tele-operation, or using rotary sensors [148] or pressure sensors [55] for tele-operation. Ranganeni et al. [251] uses a force-torque sensor to detect when the user twists the robot's handle, and turns the robot accordingly.

2.5.1.2 Output Interfaces

Output interfaces are often used for the robot to communicate information to the user about its state, the state of the environment, or its feedback on how the user is completing the task. Relative to the number of PAR papers that incorporate input interfaces, comparatively few explicitly incorporate output interfaces. Papers that use the **vision** modality often display the robot's camera feed to the user for tele-operation [40, 91, 293] or interaction [224, 266]. Papers that use the **audition** modality use verbalization to greet the user [14, 263], provide feedback on what direction the user should move in [323], or give the user information on what the robot will be doing [156, 30, 3]. Those that use the **touch** modality use haptic vibrations to convey to the user what direction the robot will move in [251], the direction the user should move [3, 233], or the distance to obstacles [113]. Those that use the **kinesthetic** modality adjust the position of a walker to help users restore their balance [97], adjust the force profile of a walker to help users stand up [61], or guide a user's hand to their target [35, 270, 14]. Note that some works also incorporate multi-modality, such as using verbal instructions to instruct users who are blind on where to find the robot arm and then kinesthetically guiding their arm to the target [35].

2.5.1.3 Future Work on Interaction Interfaces

The observations above about interaction interfaces in prior PAR research point towards several opportunities for further research.

First, there has been comparatively less focus on output interaction interfaces than input interaction interfaces. This is despite the fact that research has shown that users' trust in robots, comfort around robots, and ability to help robots improves if the robot transparently communicates its current state and future intent to them [18, 286, 290]. Therefore, **we call on future research to investigate what output information users want to receive from their PAR, and how that information improves the user experience**. Note that robot motion is an implicit output interface that can expressively communicate the robot's intent [77, 286], but was not investigated by any works in this survey.

Second, we note that some input interaction interfaces require additional devices [270, 293, 172, 57]. However, past research has demonstrated that users want to limit the number of additional devices they have to work with in order to use an assistive technology [208]. Therefore, **we call for future research on how PARs can effectively integrate with assistive technology interfaces that users already use (e.g., sip-and-puff straws, button arrays, screen readers, etc.)**. PAR research that utilizes smartphones [156, 30, 113] or computers [41] as an interaction interface are one approach to this problem, as those devices already integrate with numerous assistive technologies.

Third, although some works focus on comparing interaction interfaces, they mostly evaluate preferences aggregated across all participants. Yet, the reality is that preferred interaction interface can vary drastically across individuals and contexts-of-use [57, 208, 30]. Further, users with different disabilities may need very different interfaces from one another. Therefore, **we call for future research to investigate in what ways users' interface preference vary with the individual and/or their context(s), and how we can provide a superset of interfaces—and a smooth experience of switching between them—to cover these various preferences and abilities**.

2.5.2 Levels of Autonomy

Another overarching research theme is levels of autonomy (LoA). Autonomy is "the extent to which a robot can sense its environment, plan based on that environment, and act upon that environment with the intent of reaching some task-specific goal (either given to or created by the robot)" [21]. This section provides an overview of LoA in PAR research, following the five guidelines in Beer et al.'s framework for levels of autonomy in HRI [21].

2.5.2.1 Determining autonomy: What task is the robot to perform?

Beer et al. state that a key consideration for determining the LoA of a robot is the impact of failures on its task [21]. With PARs, the impact of failures is often high; a failure in robot-assisted feeding can result in choking or cuts, and a failure in robot-assisted navigation can result in collisions or falls. Therefore, there have been multiple efforts to enable robots to detect, predict, and/or avoid failures. In the case of robot-assisted feeding or shaving, this includes: stopping as soon as an anomalous force is detected [30, 124, 234], as soon as the user winces or has other anomolous movements [110, 234], or as soon as an anomolous sound is detected [234]. In the case of robot-assisted navigation, this includes predicting other pedestrians' motion and avoiding them [149, 155, 156], or predicting when users are getting unbalanced and changing the robot's force profile to support them [259, 97, 61]. There are also standardized methods for hazard analysis, that have been applied to robot-assisted dressing [74].

Note, however, that automated ways of detecting and avoiding failure places accountability for system success on the robot, not the user. Users may not be comfortable with this. Studies have revealed that users want full control to stop their PAR at any time, e.g., by pressing an accessible emergency stop button in robot-assisted feeding [208, 28], or by letting go of or ceasing to push the robot in robot-assisted navigation [14, 251]. After stopping the robot, the user can teleoperate it and decide when it continues [147]. In addition to giving users control to stop the robot at any time, another approach is giving users sole control to move the robot when near safety-critical areas, e.g., the robot only moves towards the user's face if they continuously facing it or press on a force-torque sensor [101].

2.5.2.2 Determining autonomy: What aspects of the task should the robot perform?

Beer et al. divide tasks into three primitives: sensing, planning, and acting [21]. Within PARs, which primitives the robot should perform is heavily influenced by the target population's impairments. PARs for people with visual impairments assist with "sensing" the environment to account for the user's reduced ability to independently do so [14, 251, 323, 156]. PARs for elderly people who are sighted assist with "acting," adjusting their force profile to account for the user's reduced ability to independently maintain balance [259, 97, 61]. PARs for people with motor impairments assist with "acting," acquiring items and moving them to the user's face to account for their reduced ability to independently do so [93, 101, 124].

While the user's impairment can influence which aspects of the task they *need assistance with*, users also have preferences over which aspects of the task they *want control over*. Users often want control to set the robot's goal. For example, in robot-assisted feeding users often want to select the bite the robot will feed them [28, 208], and in robot-assisted navigation users often want to set the goal the robot is navigating them to [14, 156]. In addition, users sometimes want control over how the robot achieves the task. Works in robot-assisted feeding have shown that some users want control over when the robot feeds them [208, 28, 30], and a work in robot-assisted navigation found that some users want control over which direction the robot turns at a junction [251]. These works serve as important reminders that just because a PAR *can* do something autonomously does not mean that it *should*, a topic investigated in Bhattacharjee et al. [30].

2.5.2.3 Determining autonomy: To what extent can the robot perform those aspects of the task?

Researchers can aim to automate as much of the assistive task as users are willing to have robots perform. However, achieving robust, generalizable robot autonomy in unstructured human environments is extremely challenging. What is possible to automate heavily depends on robot hardware (sensors, actuators, compute) and the state-of-the-art algorithms of the day. When robust robot autonomy is not feasible, including the human-in-the-loop (e.g., giving users control to stop the robot [323, 101]) can enable the robot to reliably complete its assistive task. Alternatively, one can modify the user's environment to make tasks easier to automate [43]; e.g., attaching towels to make drawers easier to manipulate [216], or attaching fiducials to make light switches easier to perceive [217].

Although the three questions for determining LoA restrict the levels that are available in a given situation,

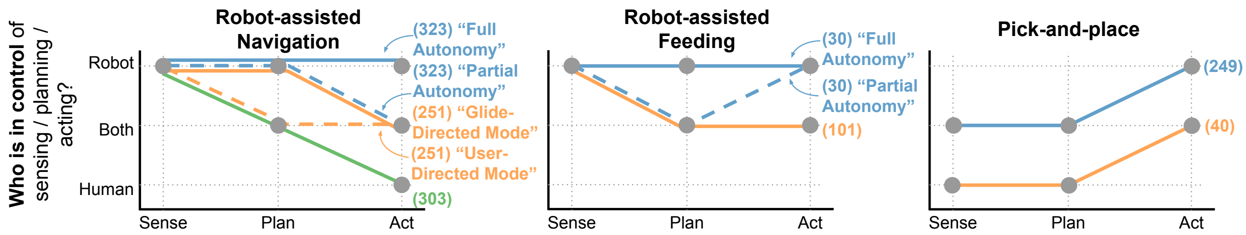


Figure 2.5: Case studies of the levels of autonomy used in three different domains of assistance: robot-assisted navigation [323, 251, 303], robot-assisted feeding [30, 101], and pick-and-place [249, 40].

there might still be multiple options. Making as many LoAs available on a robot is advisable, as it can allow for customizing based on user preferences, having different interfaces for different users (care recipient versus caregiver), and context-dependent LoA switching e.g., falling back on lower levels of autonomy when unexpected failures occur.

2.5.2.4 Categorizing automonomy

A variety of LoAs are exemplified in the PAR literature (Fig. 2.5). In robot-assisted navigation for people with visual impairments, although the robot has to be autonomous in sensing, there are a range of autonomy levels it can take on for planning and acting. Some robots autonomously plan and execute their route [323]. Others autonomously plan but share execution with the user, e.g., having the user push while the robot steers [323, 251]. Some yield part of the planning autonomy to users, letting them select the direction to turn [251]. Yet other robots fully yield execution to the user; the robot suggests a direction, but the user is the sole agent pushing and steering the robot [303].

In robot-assisted feeding for people with motor impairments, some robots acquire the bite and move it to the user's mouth autonomously [30]. Others let the user influence planning, by specifying high-level guidelines for how the robot should acquire the bite [30]. Others let the user influence acting, by controlling how much the robot tilts a drink glass [101].

In pick-and-place for people with motor impairments, some works have the user teleoperate the robot, by doing the sensing, planning, and controlling its base and arm motion [40]. Others have the robot and user sense the environment, have the robot present discrete grasping strategies to the user, and then have the robot autonomously grasp the item [249].

As indicated by this range in levels of autonomy for PARs, there is not one LoA that is strictly better than others. Multiple works have found that users preferences for LoA vary based on environmental and individual-level factors [251, 323, 30].

2.5.2.5 Influence of autonomy on human-robot interaction

The level of autonomy of a PAR affects users' feelings of comfort, trust, and safety. Some works found that users feel more comfortable when they have more control over their PAR [251, 323]. Others found that users have safety concerns regarding interacting with a fully autonomous robot [30, 208]. Another work found that users lose trust in a PAR that fails while operating autonomously, such as colliding into an obstacle [251]. Yet another work found that not just the level of autonomy, but also the level of transparency influences users' experience of the robot [224].

2.5.2.6 Future Work on Levels of Autonomy

Despite the finding that users value a variety of LoAs and will use them in different contexts [251, 323, 30], most PAR papers focus on just one LoA. Further, despite the finding that the LoA has an important impact on user experience (Sec. 2.5.2.5), most PAR papers do not justify why their LoA is a good match for the task and target population. Therefore, **we call for more PAR research that investigates the tradeoffs across different levels of autonomy, and provide guidelines on how to determine the most suitable level(s) of autonomy based on the PAR's domain of assistance, target population, and context(s) of use.**

2.5.3 Adaptation

Another overarching research theme is adaptation. We define *adaptation* as a process that changes the functionality, interface, or distinctiveness of a system to increase its relevance to an individual in a particular context [86]. Note that this process is also referred to as "personalization" or "customization" in the literature—we opt for "adaptation" as it is one of the recommended principles of ability-based design [312].

2.5.3.1 The Need for Adaptation

The need for adaptation is motivated by diversity in user's impairments, preferences, and contexts-of-use. Studies reveal that users want to customize their PAR's interaction interface, level of autonomy, and other specific functionality.

Regarding adaptation of interaction interfaces, one work found that users with greater mobility preferred a different interface for telling a robot to pick up an object than users with less mobility [57]. Other works found that users' preferred interface for interacting with a robot-assisted feeding system depended on whether they were in a social context [30, 208].

Regarding adaptation of levels of autonomy, some studies found that users' desired level of autonomy when using a robotic navigational aide was both context-dependant (e.g., is it a new environment or an unfamiliar one) [323, 251] and individual-dependant [251]. Bhattacharjee et al. [30] found that users with higher mobility impairment preferred higher levels of autonomy than users with lower levels of impairment. Yet another work found that age could impact users' preferred level of autonomy when interacting with PARs [224].

Regarding adaptation of specific functionality, Chugo et al. [61] found that the support profile users desired from a robotic walker differed based on their level of motor impairment. Choi et al. [58] found that how a robot should deliver items varies based on their posture and body type. Azenkot et al. [14] found that users with visual impairments had different preferred speeds for robot-assisted navigation systems. Works in robot-assisted feeding have found that users' preferred bite size, bite timing, bite transfer motion, bite transfer speed, and more varied based on their impairment(s), preferences, and social context [28, 208].

2.5.3.2 Adaptation in PARs

We draw upon the questions in Fan and Poole's [86] classification scheme to characterize adaptation in PARs.

What is adapted? There are several approaches to adapting interaction interfaces. Some studies found that, partly due to the large variance in ability levels for end-users, the sensors used in input interfaces need to be calibrated per user [17, 55, 266, 142]. Another study developed multiple interfaces: one for people with fine motor skills and the other for people without [249]. Yet another study leveraged existing adaptation in the user's assistive technology ecosystem, by allowing them to use their own screen-reading applications to customize hearing speed [156]. Note that companies such as Kinova[5] have for years provided the ability

[5]https://assistive.kinovarobotics.com/product/jaco-robotic-arm

to interact with their device through a variety of interfaces.

Regarding adapting levels of autonomy, Zhang et al. [323] let users of a robotic navigational aide choose whether the robot operates in full or partial autonomy, and found that users preferred less robot autonomy in environments that were less controlled (e.g., outdoor environments). These findings were mirrored by Ranganeni et al. [251].

Multiple works allowed users to adapt specific functionalities of the robot. One work enabled older adults to program custom skills on their robot, such as "raise the tray when the microwave is on" [263]. Another work allowed users to customize a parameter that controlling how much the robot followed its own policy versus the user's inputs [102]. Another study customized how close the robot brings an object to a user, based on the user's self-declared mobility level [9]. Yet another work allows the user to customize the robot's speed, speech, proximity to the user, and model of the user's movements [47].

Who does the adaptation? Works that allow the **user** to adapt the robot focus on providing the user knobs to tune the robot's functionality. In the case of Saunders et al. [263], those knobs consisted of an entire domain-specific language designed for customizing that PAR. In other cases, researchers designed multiple discrete modes and let the user select one [323, 249]. In yet another case, researchers exposed a continuous parameter to the user and let them adjust it [102].

Works that use **shared control** to customize the robot typically have the user provide some data during a calibration phase, and have the robot adapt its behavior based on that data. This includes calibrating the sensitivity of sensors [17, 55], asking users for self-reported mobility level [9], and to move through their full range of arm motion [320, 321].

Works where the **robot** adapts to the user have the robot observe or predict some attribute about the user and change its behavior accordingly. Erickson et al. [83] track the distance between the robot and the user's body in order to adjust the robot's motion to the user's contours. Ondras et al. [226] use information about when the user last took a bite and the gaze of co-diners to predict when to feed the user.

When does the adaptation take place? A variety of works adapt the robot **during its main execution**. This includes works that allow users to select one of multiple modes for robot behavior [323, 249], works where the user iteratively modifies a parameter [102], and works where the robot tracks and adapts to attributes of the user [83, 226]. On the other hand, all works that involve a calibration phase adapt the robot **outside of main execution** [17, 55, 9, 320, 321, 53, 266, 142]. Further, works that involve the user pre-programming robot actions also involve adapting outside of main execution [263].

2.5.3.3 Future Work on Adaptation

Although the broader field of assistive technology has had considerable focus on adaptation, summarized in Wobbrock et al. [312], it has been a smaller focus of PAR research. This presents several exciting directions for future work.

First, certain application domains tend to focus on specific types of adaptation. For example, research into interface adaptation was largely in the domain of pick-and-place [17, 55, 249], although the need has also been established in robot-assisted feeding [28, 208]. Similarly, research into LoA adaptation was largely in the application domain of robotic navigational aides for people with visual impairments [323, 251], although the need has also been established in robot-assisted feeding [28] and pick-and-place [224]. **We call for more cross-domain research in adaptation, particularly to investigate under what conditions insights on adaptation can be transferred across domains.**

Second, although there are works across the spectrum of "who does the adaptation," there are no works to the best of our knowledge that provide guidelines on how to decide who should do the adaptation for a particular robot, user, domain, or context. The same applies to guidelines regarding when the adaptation takes place. **We call for research into user perspectives regarding who should do the adaptation, when it should be done, and how that varies across the application domain, user, and context**.

2.6 Conclusion

In conclusion, this chapter presents the results of a systematic survey of PARs research, identifying key trends in domains of assistance, study methodology, involvement of users with disabilities, interaction interfaces, levels of autonomy, and adaptation. Key summary points and future issues are described below.

2.6.1 Summary Points

1. **Domains of Assistance (Sec. 2.3.1)**: There have been three main foci in PAR research: navigation, feeding, and general pick-and-place.
2. **Involvement of Participants with Disabilities (Sec. 2.4.1)**: Nearly all formative works involved people with disabilities, while about half of summative evaluations involved solely participants *without* disabilities.
3. **In-Context Deployments (Sec. 2.4.3.1)**: The few in-context deployments of PARs that have been done tend to be relegated to small sections within a paper, preventing the community from learning about and benefiting from the several research, engineering, and logistical decisions required to deploy a system.
4. **Quantitative Metrics (Sec. 2.4.3.2)**: Most summative evaluations gather task-specific objective data (e.g., completion time, number of mode switches, success rate), and/or subjective data based on custom questionnaires measuring usability, satisfaction, feelings of safety, etc.
5. **Interaction Interfaces (Sec. 2.5.1)**: PAR research covers a variety of input interfaces, from brain-computer interfaces to vision-based interfaces to touchscreens to kinesthetic interfaces. In contrast, comparatively less work has focused on *output* interfaces for the robot to communicate to the user.
6. **Levels of Autonomy (Sec 2.5.2)**: Most PAR research uses a single level of autonomy, despite the fact that past work has revealed that users' preferred level of autonomy varies with the individual and context.
7. **Adaptation (Sec. 2.5.3)**: Several studies have found that users want their interactions with PARs to be customized to their impairment, their preferences, and their context. Although some work has investigated adaptation, that work is segmented across application domains, and few works investigate tradeoffs across who is doing the adaptation and when it takes place.

2.6.2 Future Issues

1. **Domains of Assistance (Sec. 2.3.1)**: We call on researchers to study under-researched (I)ADLs such as dressing, bathing / grooming, and managing medication. This also includes conducting formative studies to ensure the design and development of PARs in these domains is rooted in user needs

(Sec. 2.4.2).

2. **Involvement of Participants with Disabilities (Sec. 2.4.4)**: We call on researchers to include more participants with disabilities in their works. In addition to in-person studies, other ways to do so can include remote studies, video studies, or working with a community researcher.
3. **In-Context Deployments (Sec. 2.4.4)**: We call on researchers to conduct and publish more in-context deployments. Experimental design theory for "n-of-1" studies can be used to add methodological and statistical rigor to PAR deployments [287].
4. **Quantitative Metrics (Sec. 2.4.4)**: We call on researchers to use standardized quantitative metrics such as the System Usability Scale and NASA-TLX when evaluating systems, to facilitate comparisons across PAR research. We also call on researchers to *work with users* to ensure that objective metrics they gather align with users' desires for system functionality.
5. **Interaction Interfaces (Sec. 2.5.1.3)**: We call on researchers to investigate the desired output information users want to receive from their PARs, as well as how PARs' input and output interfaces can integrate with users' existing assistive technology ecosystem.
6. **Levels of Autonomy (Sec 2.5.2.6)**: We call on researchers to be intentional about which level(s) of autonomy they use and justify why that is suitable for the task(s), user(s), and context(s). We further call for more research on the tradeoffs between levels of autonomy, in order to derive guidelines for how to determine the most suitable level(s) of autonomy for a PAR.
7. **Adaptation (Sec. 2.5.3.3)**: We call on researchers to investigate users' preferences regarding the different forms of adaptation—what is adapted, who does the adaptation, and when it takes place—and how that varies across domain of assistance, user, and context.
8. **PARs in Society**: Developing PARs that are widely used requires engaging with government regulations [44], ethics [297], and factors that influence technology adoption such as affordability and scale [202]. Therefore, we call for more research that places PARs within the context of the political, economic, and social systems that impact their usage.

Chapter 3

Why Robot-Assisted Feeding? A Brief History

The previous chapter delved into the applications, methods, and themes that underlie the last decade of research on physically assistive robots (PARs) for people with disabilities. It presented few (but notable) examples of in-context deployments, and advocated for more research involving in-context deployments of PARs.

This chapter delves into one specific application domain of PARs: robot-assisted feeding (RAF). It presents the 50 year history of RAF research, including formative user studies, technical advancements, summative clinical studies, and commercial products. *The well-established user need, the long history of technical research, and the interest from both academia and the industry makes RAF systems a good case study for achieving in-home deployability of PARs.* Portions of this chapter were originally published in [212, 208].

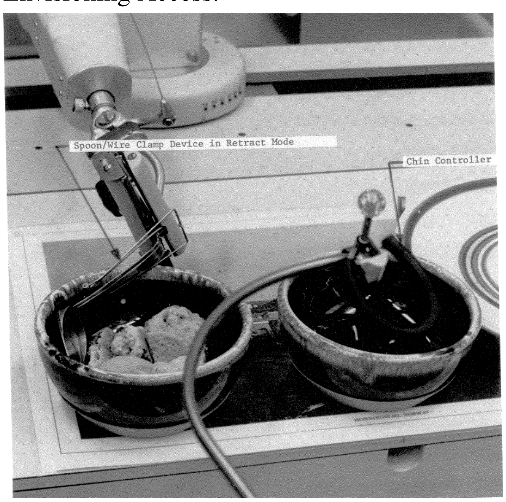
3.1 A History of Robot-assisted Feeding

Enabling people with motor impairments to eat independently has been a goal of research for around 50 years, and has been covered in the following surveys [243, 289, 165, 12, 36, 213].

One of the early research directions in the 1970s involved training capuchin monkeys as service animals

(a) A trained service monkey feeds a person with motor impairments, a research direction that began in the 1970s [181]. Reprinted from Envisioning Access.

(b) The Morewood Spoon Lifter, developed in 1974. Reprinted from Appendix C of [243].



(c) The Robot Arm Worktable, which began development in 1974 and was clinically evaluated by the US Department of Veterans Affairs (VA) starting in 1983. Reprinted from [267].

(d) The Handy 1 was developed in 1987 to help a boy with cerebral palsy to independently eat. Reprinted from [296].

Figure 3.1: Selected research foci (1970s-80s) to enable people with motor impairments to independently feed themselves.

to feed people with motor impairments [181] (Figure 3.1a), an effort that continued for decades [129]. Also in the 1970s, an early robot-assisted feeding system, the Morewood Spoon Lifter (Figure 3.1b), was developed. This portable system involved strapping a metal rod around the user's head, which they used to shovel food into the spoon and to press a switch that moved the spoon along a fixed trajectory from table-level to mouth-level. The VA clinically evaluated this device with 16 people with quadriplegia due to a SCI and studied it during a 3 year home deployment with one user. The clinical evaluation found that users were able to use the device to independently feed themselves, but felt it was "too strenuous" or "required too much training" [243]. Feedback from the clinical evaluation was incorporated into the device's design, and it was later manufactured and sold as a commercially as the "Winsford Feeder" from at least the 1990s to the 2010s [127, 213].

While the "Winsford Feeder" focusing on being portable, other systems focused on being multi-purpose. Developed in the 1980s, the "Robot Arm Work Table" (Figure 3.1c) consisted of a desk with a fixed robot arm and other tools mounted to it. The system was designed to enable people who only have neck mobility to drive their wheelchair up to it and use the robot arm for a variety of different tasks, including picking and hanging up the phone, typing on a computer, and eating from a bowl. This system was deployed and clinically evaluated by the VA with 20 people with quadriplegia, over up to a year, in environments as diverse as family home, a nursing home, and a hospital [267]. The clinical evaluation found that multiple users preferred the self-feeding functionality over the other functionalities, but problems with that system included food spilling and getting cold. The 1980s also brought with it other stationary, multi-purpose robot-assisted feeding systems such as the Handy 1 (Figure 3.1d), which was designed to help an 11 year old boy with cerebral palsy eat, drink, brush his teeth, and more [296].

By the 2000s, multiple commercial products for robot-assisted feeding were on the market. These included: the aforementioned Winsford Feeder (Figure 3.2a); Neater Eater, in development in the 1990s [197] and sold in the early 2000s [130] (Figure 3.2b); Bestic, first developed in 2004 and sold in 2012 [177] (Figure 3.2c); Obi, first developed in 2009 and sold in 2016 [11] (Figure 3.2d); My Spoon, in development in the 1990s and sold in 2002 [278]; and more [12, 213]. These products are all table-mounted and have the robot execute fixed trajectories to acquire food and move it to the user's mouth. They have undergone considerable user testing, including clinical evaluations, which have shown positive results in terms of enabling

(a) The Winsford Feeder has been sold commercially since at least the 1990s [127]. Reprinted from North Coast Medical & Rehabilitation Products.

(b) The Neater Eater is a commercial device that has been sold in the UK for around two decades. Reprinted from Neater Eater.

(c) Bestic, first sold in 2012, feeds its founder, Sten Hemmingsson. Reprinted from Bestic AB.

(d) The Obi is a commercial device that has been sold in the US since 2016. Reprinted from MeetObi.

Figure 3.2: Selected commercial systems (1990s-2010s) that were developed and sold to enable people with motor impairments to independently feed themselves.

users to eat a full plate of food [177], feel more independent and confident [164, 177], and improve their posture [164]. They have also been demonstrated to save caregiver time [92]. Despite the positive results, these devices have struggled to achieve long-term adoption; as of 2024, all but the Obi and Neater Eater were discontinued. This may be due to shortcomings including: being unable to acquire users' desired food items or acquiring too little food [277, 159]; dropping food [178]; and requiring precise positioning of the user's face, sometimes resulting in strained muscles [220]. These shortcomings can be traced to an inability of the robot to autonomously sense and react to the environment, e.g., adapting acquisition to perceived food properties, adapting transfer to the face pose.

(a) Bhattacharjee et al. observed caregivers feeding care recipients in assisted-living facilities. Reprinted from [28].

(b) Pascher et al. observed an end-user use a wheelchair-attached robot arm to feed herself. Reprinted from [237].

(c) Kim et al. conducted focus group interviews with care recipeints, caregivers, and doctors. Reprinted from [158].

Figure 3.3: Selected formative research (2010s-20s) into RAF systems.

Contemporary research in robot-assisted feeding has largely sought to address the aforementioned shortcomings. Specifically, the systems currently being used in research typically have more sensors than commercial systems—including RGB cameras, depth cameras, and/or force-torque sensors—which they use to be more adaptive to the food, user, and environment. Modern research systems have various forms, including wheelchair-mounted robot arms [93, 30, 226], table-mounted portable robot arms [276], table-mounted fixed robot arms [24], and mobile manipulator robots [236, 218]. Sec. 3.2, 3.3, and 3.4 focus on this contemporary wave of robot-assisted feeding research.

3.2 Formative Studies in Robot-assisted Feeding

Some research has conducted formative studies to understand participants needs and priorities when it comes to robot-assisted feeding. Bhattacharjee et al. [28] conducted a contextual inquiry in an assisted-living center, where they observed meals, showed care recipients and caregivers a video of robot-assisted feeding, and interviewed them (Figure 3.3a). They synthesized several evaluation indicators for robot-assisted feeding systems, including their technical function, technology robustness, information gaps, usability, social acceptance, and system integration. Pascher et al. [237] observed participants consuming a caregiver-fed or robot-fed meal in their homes, had them experience robot-assisted feeding through a virtual reality headset, and interviewed them (Figure 3.3b). They present several recommendations for the development of robot-assisted feeding systems, including that such systems should not involve lengthy familiarization periods, should be compact and unobtrusive, and should allow users to perform daily tasks being feeding. Kim et al. [158] conducted focus group interviews with care recipients, caregivers, and doctors to understand their priorities for robot-assisted feeding (Figure 3.3c). One of their key findings was that each group has different preferences; for example, doctors and caregivers want the robot to move twice as slow as care recipients do.

Ljungblad [178] and Nickelsen [220] conducted observations and interviews of people with motor impairments who used the Bestic robot to feed themselves. The former research found that the robot could not satisfy users' desires to eat food in traditional ways, such as peeling shrimp or taking crab meat out of the shell [178]. The latter research found that caregivers often needed to monitor and tinker with the setup and positioning of the robot and user throughout the meal, leading to longer meals and hindering long-term adoption. They also found that users internally rank feeding aides, irrespective of whether they are human or robot; for one user, Tonni, the preference ranking was her caregiver June, followed by Bestic, followed by her caregivers Nete or Helge [220].

The work presented in Chapter 4 (RQ1) [208] adds to this space of formative studies by focusing on user's needs and priorities when it comes to robot-assisted feeding in *social settings*.

3.3 Technical Advances in Robot-assisted Feeding

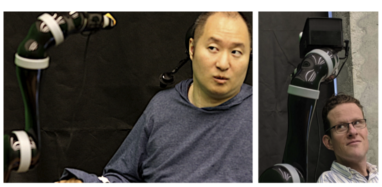
Technical advances in robot-assisted feeding research often focus on one of two sub-tasks: bite acquisition (Figure 3.4a) and bite transfer (Figure 3.4b).

Bite acquisition is the process of acquiring a bite-sized amount of food. Prior works have focused on the robot's ability to acquire food with a fork [105, 125, 285, 284], spoon [255, 221, 136], chopsticks [317, 222], or multiple tools [277, 236, 108]. Some of these works focus on the motion primitives a robot should use to acquire food [125, 29, 255, 93, 284], others focus on online learning techniques to select which motion primitives to use on novel food items [105, 106], and yet others focus on chaining together motion primitives for more complex acquisition (e.g., pushing food together before scooping it) [285, 145]. The work presented in Chapter 5 (RQ2) [104] adds to this by presenting a pipeline to learn, in an unsupervised manner, motion primitives from human bite acquisition data.

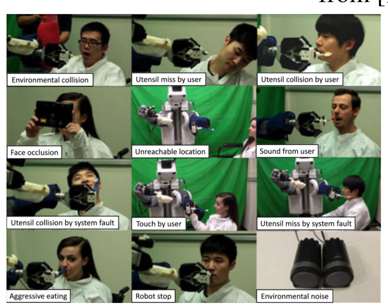
Bite transfer is the process of moving the acquired bite to the user's mouth. Prior works have: presented motion-planning techniques that account for user comfort [23]; developed ways for caregivers to teach to a

(a) Bite acquisition is the process of acquiring a bite of food. Reprinted from [89].

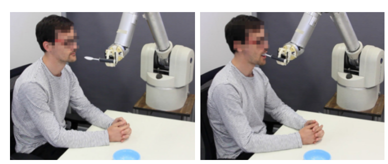
(b) Bite transfer is the process of moving a bite of food to the user's mouth. Reprinted from [93].



(c) Bhattacharjee et al. studied levels of **user control** and how that varies across individual and technical factors. Reprinted from [30]



(d) Park et al. studied **anomaly detection** for safe and robust robot-assisted feeding. Reprinted from [235]



(e) Canal et al. studied **customizability** in robot speed and distance-to-mouth. Reprinted from [47]

Figure 3.4: Selected technical research (2010s-20s) into RAF systems.

robot how to properly transfer bites [46]; and revealed the coupling between how a bite is acquired and how it can be transferred to the mouth [93]. Although most works have the robot arm stop a few centimeters in front of the user's mouth, some recent works have investigated in-mouth bite transfer [271, 144] (Figure 3.5b).

Finally, some works have extended beyond the dichotomy of bite acquisition and transfer. Some works investigated way to ensure robot-assisted feeding systems are robust to errors, through **user control** (Figure 3.4c) [30] and **anomaly detection** (Figure 3.4d) [235]. Others studied ways to enable users to **customize** the robot's speed and delivery position (Figure 3.4e) [47, 236]. Other areas of focus include: **bite timing**, or predicting when users want a bite [226, 125]; **bite sequencing**, or predicting what bite the user wants [143, 145]; **food detection** [125, 89, 145]; **mouth detection** [240, 144]; and **user interfaces** for robot-assisted feeding [30, 232]. Other works have focused on feeding-adjacent tasks such as fetching food [231, 141, 252], robot-assisted drinking [101, 166], and robot-assisted napkin-wiping [231, 109, 236]. Recent works have also developed simulation environments for caregiving tasks [319, 185, 84].

The work presented in Chapter 6, (RQ3) [212] incorporates several of the aforementioned technical components into an end-to-end system for out-of-lab robot-assisted feeding. It additionally improves upon the state-of-the-art with: (1) a user interface that provides substantial user customizability and control, (2) a food detection implementation that incorporates users-in-the-loop to generalize across food items, and (3) portable hardware that facilitates system use in diverse environments without inhibiting user mobility.

3.4 Deployments of Robot-Assisted Feeding Systems

As with the broader field of physically assistive robots (PARs) (Chapter 2), out-of-lab deployments of robot-assisted feeding systems are few and notable. Park et al. [236] deployed a mobile manipulator robot in the home of a person with motor impairments for 3 consecutive days (Figure 3.5a). The user used a web-based interface to have the robot feed him 6 meals of yogurt, and "found the system to be effective, safe, and easy to use" [236]. Multiple works have deployed a teleoperated mobile manipulator robot in a user's home to enable him to successfully feed himself food items like bananas and nut butter [223, 231, 252, 218], in addition to performing other activities of daily living (Figure 3.5c). One study deployed a table-mounted feeding robot in a user's home for 5 days [277], whereas another had a wheelchair-mounted robotic system feed a participant one meal in their home [144] (Figure 3.5b). Notably, there is space for longer deployments

(a) A mobile manipulator robot-assisted feeding system feeding a user in their home. Reprinted from [236].

(b) A wheelchair-attached robot-assisted feeding system feeding a user in their home. Reprinted from [144].

(c) A user teleoperates a mobile manipulator to feed himself in his home. Reprinted from [252].

Figure 3.5: Selected research (2010s-20s) involving in-home use of RAF systems.

and deployments where users freely use their wheelchair-mounted robot-assisted feeding system across various meal contexts. This motivated the in-home deployment we present in Chapter 6 (RQ3) [212], where a user uses the robot-assisted feeding system to feed themselves 10 meals across 5 consecutive days while watching TV, working, socializing, and more.

3.5 Conclusion: Why Robot-assisted Feeding?

As has been established above, robot-assisted feeding (RAF) has been a research focus for at least 50 years. Multiple research and commercial RAF systems have been developed and studied through clinical studies and in-home evaluations. This has established a wealth of cumulative knowledge around users' needs and technical ways to meet those needs. However, existing commercial RAF systems still face technical challenges that have hindered their long-term adoption at scale. Contemporary research has made technical contributions to overcome these shortcomings, and we as a field are moving closer to achieving long-term deployments of robot-assisted feeding systems. *This well-established user need, the long history of technical research, and the interest from both academia and the industry makes robot-assisted feeding (RAF) systems a good case study for achieving in-home deployments of PARs*. The remaining sections of this thesis focus specifically on robot-assisted feeding (RAF) systems as a case study for achieving in-home deployability of physically assistive robots (PARs).

Chapter 4

Investigating Users' Needs and Priorities for Robot-Assisted Feeding Systems

This chapter presents our investigation into end-users' current dining routines and how robot-assisted feeding systems should be designed to better support them during meals. Specifically, this formative study seeks to answer the question:

RQ1 What challenges do people with motor impairments face during dining, and how should a robot-assisted feeding system be designed to address those challenges?

Because of the existence of past formative studies that focus on robot-assisted feeding in general [28, 237, 158], this work focuses on the under-studied *social* aspects of dining.

This chapter was originally published as "Design Principles for Robot-Assisted Feeding in Social Contexts" at the *ACM/IEEE Conference on Human-Robot Interaction* in 2023 and won the **Best Design Paper** award [208].

4.1 Introduction

Take a moment to recall the last time you shared a meal. What made it meaningful? The company, the food, the ambiance? The stories that were told, relationships that were strengthened, milestones that were celebrated? If you were asked, "How does it feel to eat socially?", you might say it is a pleasant experience.

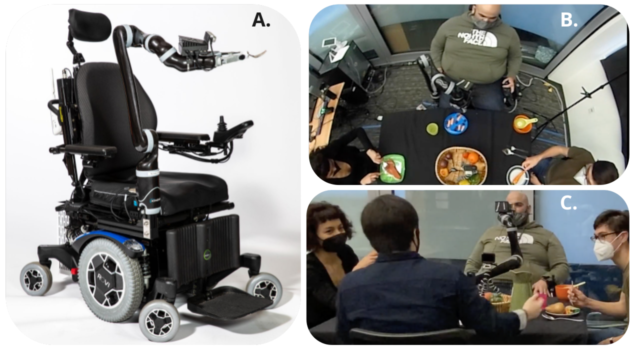


Figure 4.1: A. robot-assisted feeding system used in this work, with **B.** top-view and **C.** side-view of its social use.

Now consider this response from a participant in our study: *"Sometimes I wait longer to ask [my caregiver] for a bite or a drink because it might mess up a conversation. It's something that's always in the back of my mind when eating socially... Sometimes I'm not eating, or I'm barely eating, because I'm self-conscious of interrupting a conversation."* This participant is paralyzed from the neck down. For him and at least 1.8 million Americans who need assistance eating [292], social dining may be the opposite of pleasant.

Eating is not only a *functional* experience, but also a *meaningful* one. Specifically, social dining introduces nuances, such as synchronizing of eating pace [121], avoiding a bite while being addressed [201], and making special efforts to eat in a socially appropriate manner [310]. For those with motor impairments, robot-assisted feeding (Fig 4.1) has emerged as a promising technology to alleviate some of the challenges faced during dining. However, much prior work in this area focuses on the functional tasks of picking up food and moving it to a person's mouth [105, 221, 23, 236, 317]. These tasks are indeed technically challenging, and prior work significantly improves the state-of-the-art. Nonetheless, there is an open design space to create meaningful social dining experiences for people with motor impairments.

4.2 Contributions

We conducted design explorations (Fig 4.2) of robot-assisted feeding in social contexts, driven by the following questions:

1. **What *challenges* do people with motor impairments face during social dining, and how can robot-assisted feeding address them?** Participants' challenges include divided attention, caregiver

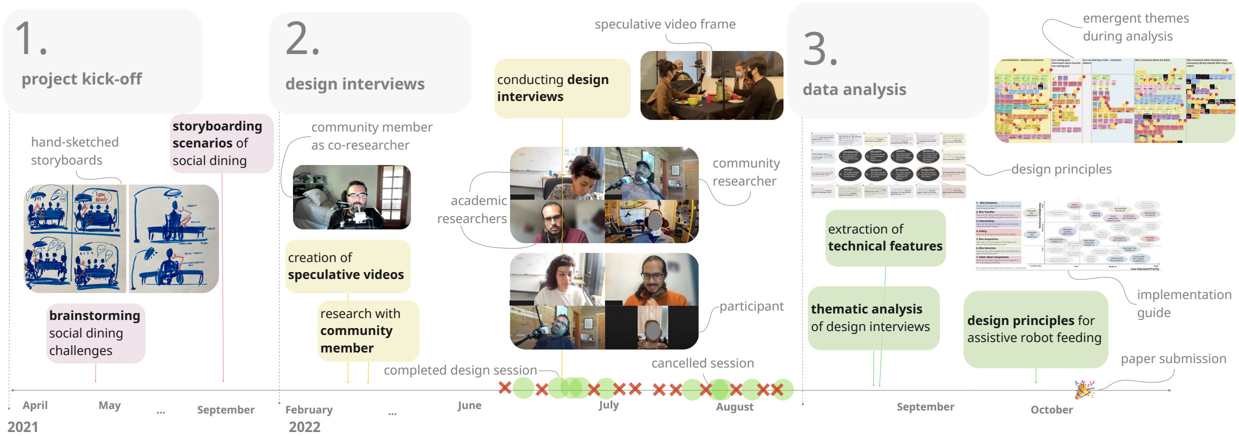


Figure 4.2: Timeline of our design project on robot-assisted social dining.

variability, and more (Sec 4.6).

2. **What *principles* should guide the design of robot-assisted feeding systems for social dining?** The robot should be subtle, customizable, reliable, and more (Sec 4.7).

3. **How can these insights *guide the implementation* of robot-assisted feeding systems for social dining?** Our interview data reveals key features, such as unobtrusive bite transfers, feeding others (e.g., kids), and more (Sec 4.8).

While some social dining challenges cannot (and should not) be solved using a robot, **our key insight is that robots can be designed with assistive qualities that address some challenges people with motor impairments face during social dining.** Specifically, they can promote *empowerment* by enabling users to eat without human assistance and *belonging* by increasing user's opportunities for meaningful social interactions.

4.3 Related Work

4.3.1 The Power of Social Dining

Social dining has biological, psychological, and cultural benefits [174, 195]. Food is a "vehicle" to establish social linkages, has symbolic functions, and is a medium for aesthetic expression [204, 257]. Families that eat together build stronger relationships that improve well-being and lower the rates of risk-taking behavior [298] and depression [299]. People with Alzheimer's who share meals show an increased sense of autonomy [52]. Unfortunately, for people with motor impairments who rely on caregivers to eat, shared meals tend to be less about *socialization* and more about *functionality* (e.g., meal prep, food intake) [206, 183]. This excludes them from social dining benefits [100, 268]. Our work surfaces priorities voiced by people with motor impairments, towards a robot that enables meaningful social dining experiences.

4.3.2 Robot-Assisted Feeding

Robot-assisted feeding systems have existed since at least the 80's [267, 296]. Since then, over a dozen such systems have been developed, both commercially [191, 192, 205, 193, 254, 22, 27, 214] and for research. Most research has focused on functional aspects of eating, not the nuances inherent to social dining. This includes the robot's ability to *acquire* food items with a fork [105, 125], spoon [255, 221, 136], or chopsticks [317, 222]; also, the robot's ability to *transfer* food items to the user's mouth by accounting for user comfort [23], learning from demonstrations [46], and adjusting based on how the food was acquired [93]. Needs assessments used interviews and ethnographic observations to develop evaluation indicators for robot-assisted feeding systems [28, 158] but did not directly examine social dining nuances.

Some research has included social dining. One work compared three robots and found that users preferred the one that enabled more socialization [127]. Another work found that users preferred a slower robot and a non-voice interface in social contexts [30]. Other studies modeled when a robot should automatically feed a user during social dining [125, 226]. To the best of our knowledge, there is no thorough investigation of user needs and priorities for robot-assisted feeding in social contexts. Our work fills this gap.

4.3.3 Design Principles for Assistive Technology

Universal design (UD) originated in the 90's from the disability rights movement and consists of 7 principles, such as equitable and flexible use [157]. Critics of UD maintain that it tries to accommodate two contradictory goals: designing for mass marketing and designing for specialized communities, such as people with disabilities [246, 132, 67]. Previous work tried to reconcile these goals with the EMFASIS framework [244].

Our work unearths design principles that are distinctive to social contexts and tailored to robot-assisted feeding. The design principles we propose differ from others, [157, 261, 72], as they are actionable within the realm of robot-assisted feeding. They differ from work in robot-assisted feeding [158, 28], because the

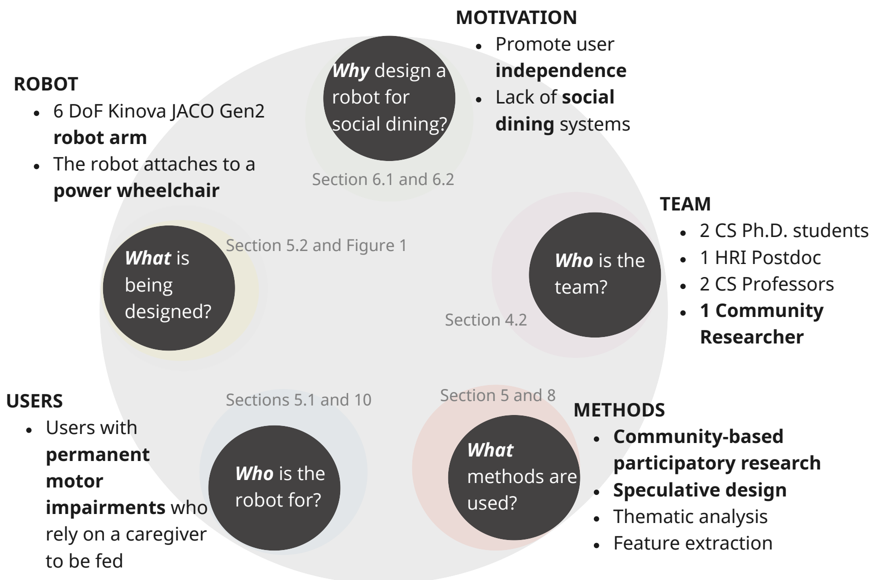


Figure 4.3: Applied framework for Inclusive Design [132].

principles are meant to guide designers *throughout technology creation*, not just during evaluation.

4.4 Design Framework

4.4.1 Framework for Inclusive Design

Our work is based on Kat Holmes' framework for inclusive design [132]. This framework provides 5 questions to help teams avoid perpetuating exclusion: (1) *Why* make this artifact? (2) *Who* makes it? (3) *How* do we make it? (4) *Who* will use it? (5) *What* do we make? Fig 4.3 shows the framework applied to our research.

4.4.2 Community-Based Participatory Research

To wholly include people with permanent motor impairments in this work, we followed CBPR, which harnesses community wisdom in equal partnership with academic methodological rigor throughout the research process [115, 301]. According to CBPR, an equitable partnership between research and community requires sharing power, resources, credit, results, and knowledge [239, 198]. This method is increasingly common in assistive technology research [51, 50, 71, 25, 291, 167, 295].

Table 4.1: Study participant demographics. No participant used robotic assistance for feeding.

ID	Age	Gender	Living at	Self-described impairment	Impairment time	Eating assistance providers[3]
CR1[4]	37	Male	Home	Paralyzed from the neck down	> 5 years	Formal caregivers, parents, friends
P1	28	Female	Home	Unable to move arms	> 5 years	Parents, formal caregiver
P2	40	Male	Home	Paralyzed from the neck down	> 5 years	Formal caregiver, girlfriend
P3	42	Male	Home	Muscular dystrophy	> 5 years	Formal caregiver, sister, mother
P4	40	Male	Home	Paralyzed from the neck down	> 5 years	Formal caregiver, family, friends
P5[5]	18	Male	Home	Cerebral palsy	> 5 years	Parent, friend, formal caregiver
P6	58	Male	Care facility	Quadriplegia due to multiple sclerosis	> 5 years	Formal caregiver, family, friends
P7	49	Male	Care facility	Quadriplegia due to multiple sclerosis	> 5 years	Formal caregiver, family, friends
P8	30	Male	Home	Spinal Muscular Atrophy	> 5 years	Wife, family, friends
P9	18	Male	Home	Almost paralyzed from neck down	3-5 years	Parent, formal caregiver
P10	34	Female	Home	Spinal Muscular Atrophy	> 5 years	Parents

Near the beginning of this project, the academic researchers[1] engaged a person with permanent motor impairments as a community researcher. This individual has been a recurring participant in our lab's user studies since 2019, which gives him familiarity with robot-assisted feeding. A C1 quadriplegic[2], he was injured in 2012. He runs a non-profit organization that connects people with motor impairments to assistive technologies (ATs), is on advisory boards related to AT, and runs a business focused on smart homes and AT, making him a valuable community researcher for our project. Throughout the project, he has been involved in creating design materials, co-running design interviews, analyzing data, and co-authoring this paper, spending an average of 1 hour per week since joining the team in Feb 2022.

[1]Throughout the paper, "we" refers to the entire research team, including the community researcher. To differentiate, we use "academic researchers."

[2]A person diagnosed with a C1 quadriplegic injury will probably lose function from the neck down and be ventilator-dependent. For more information, see https://www.spinalinjury101.org/details/levels-of-injury.

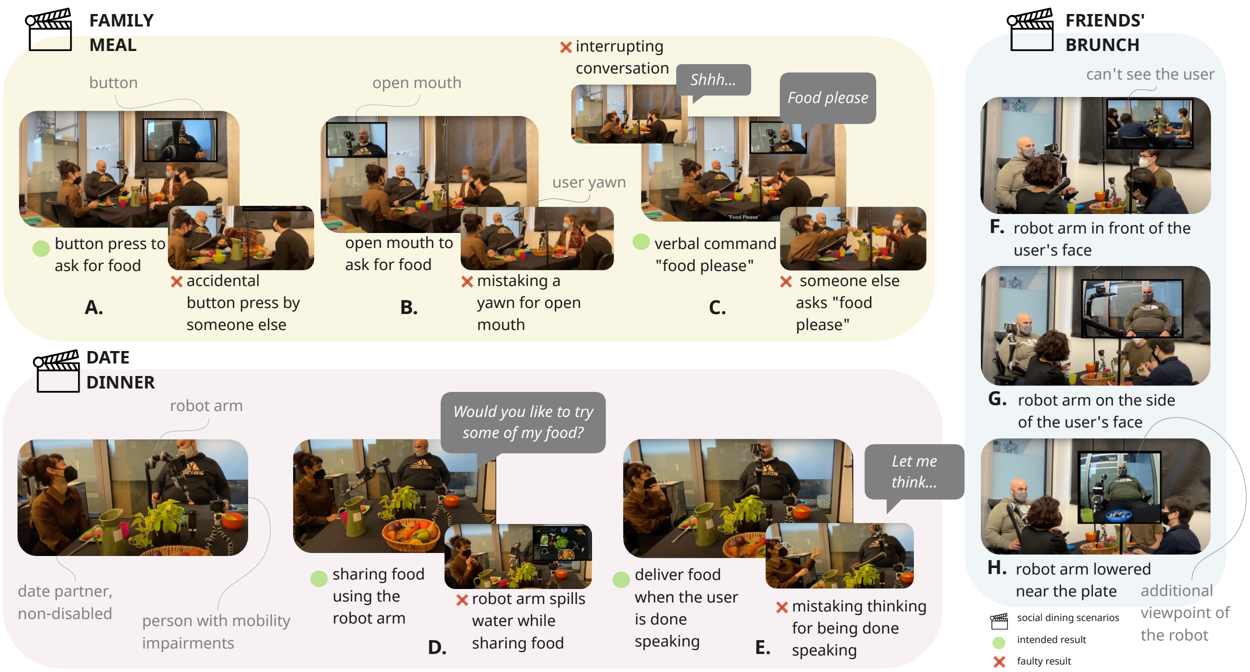


Figure 4.4: Speculative videos on robot-assisted social dining used during our design interviews.

4.5 Design Method

We interviewed participants using speculative videos of robot-assisted social dining and applied qualitative analysis methods to identify key themes. Fig 4.2 shows an overview of our method.

4.5.1 Participants

We recruited participants primarily from the community researcher's connections. The inclusion criteria were to have a permanent motor impairment and to rely on a caregiver to be fed. Table 4.1 shows demographic information about our 10 participants.

[3]Formal caregivers are paid and trained professionals.
[4]This individual is the community researcher (Sec 4.4.2).
[5]Since P5had difficulty speaking, their parent sometimes clarified what they said.

4.5.2 Design Materials

4.5.2.1 Speculative Videos

We showed participants speculative videos of how robot-assisted feeding might be used in social settings[6]. These videos were intended to familiarize them with robot-assisted feeding and invite them to share their views on robot design. We created the videos following speculative design guidelines [199]. These videos include three common social dining scenarios: family meal, dinner date, and brunch with friends. They feature a person who acted as someone with motor impairments using the robot and other social dining partners without motor impairments.

We themed the videos around areas where the design direction was unclear: (1) how should a user ask the robot for food? (2) should a robot share food with dining partners, and how? (3) where should the robot rest its arm while delivering food? We recorded polarized versions of each robot behavior, showing intended robot performance vs faux pas the robot could cause. The community researcher helped design the videos, ensuring they would be understandable to participants with little experience with assistive robots. In total, we created a playlist of nine 1 minute videos (Fig 4.4). More details can be found in our HRI'23 Video submission[7] [209].

4.5.2.2 Robot System

We used a 6 degree-of-freedom Kinova JACO Gen2 robot arm attached to a power wheelchair base (Fig 4.1). The robot arm has an RGB-D mounted on-board, which it uses to perceive food and the user's face. It uses a custom 3D-printed fork to pick up food, and a force-torque sensor to know when it has skewered food and to guarantee user safety. In the videos, the robot autonomously acquires pieces of fruits and vegetables and feeds them to the user.

4.5.3 At-Home Interviews

Design interviews were held virtually and consisted of the following steps: (1) introduction of the research team and participant, (2) questions about participants' current social dining routines, (3) watching and dis-

[6]https://youtube.com/playlist?list=PLv0SEVdRS7GqvB1eWGUrEvMwfNgdcbuMt
[7]https://youtu.be/BInhARANKaU

cussing videos, and (4) session wrap up. The decision to hold interviews virtually was made in consultation with the community researcher in order to promote accessibility for participants. The community researcher led the interviews, while other team members took notes and participated in the discussion.

4.5.4 Thematic Analysis

We employed qualitative methods since it can surface understanding around particular people's nuanced experiences, emotions, needs, and motivations [274, 247, 95]. Specifically, we used thematic analysis [274] to analyze the data, which consisted of video recordings from the design interviews. To develop the codes and themes that emerged from the data, two researchers independently coded each interview recording and performed calibration exercises to ensure consistency [161]. Overall, the two researchers met 10 times, with the community researcher participating in 5 meetings to reconcile divergence in the coding [135, 94]. The thematic analysis took over 70 hours across all researchers to transcribe and code the over 500 sentences from design interviews.

4.5.5 Synthesis as Visual Knowledge

From this thematic analysis, we extracted the **three key outcomes of this work: interview results (Sec 4.6), design principles (Sec 4.7), and an implementation guide for robot-assisted social dining (Sec 4.8)**. For each outcome, we synthesized key insights as standalone visual knowledge, presented in Figs 4.5-4.7. This builds upon the growing awareness that visuals have unique strengths compared to text [272, 78], particularly for "creating and articulating knowledge about interactivity" [33], and are becoming prevalent in interdisciplinary fields including HRI [6, 7, 131, 173, 168].

4.6 Interview Results

Participants engaged in social dining in restaurants, breweries, sports games, picnics, road trips, theaters, and remote socialization. We now elaborate on their challenges during social dining (Sec 4.6.1), thoughts on how robot-assisted feeding could address them (Sec 4.6.2), and preferences with respect to robot behaviors (Sec 4.6.3). Several of the quotes referred to in-text are presented in Fig 4.5 and Fig 4.6; we recommend readers interactively read through the figure as they go through the text.

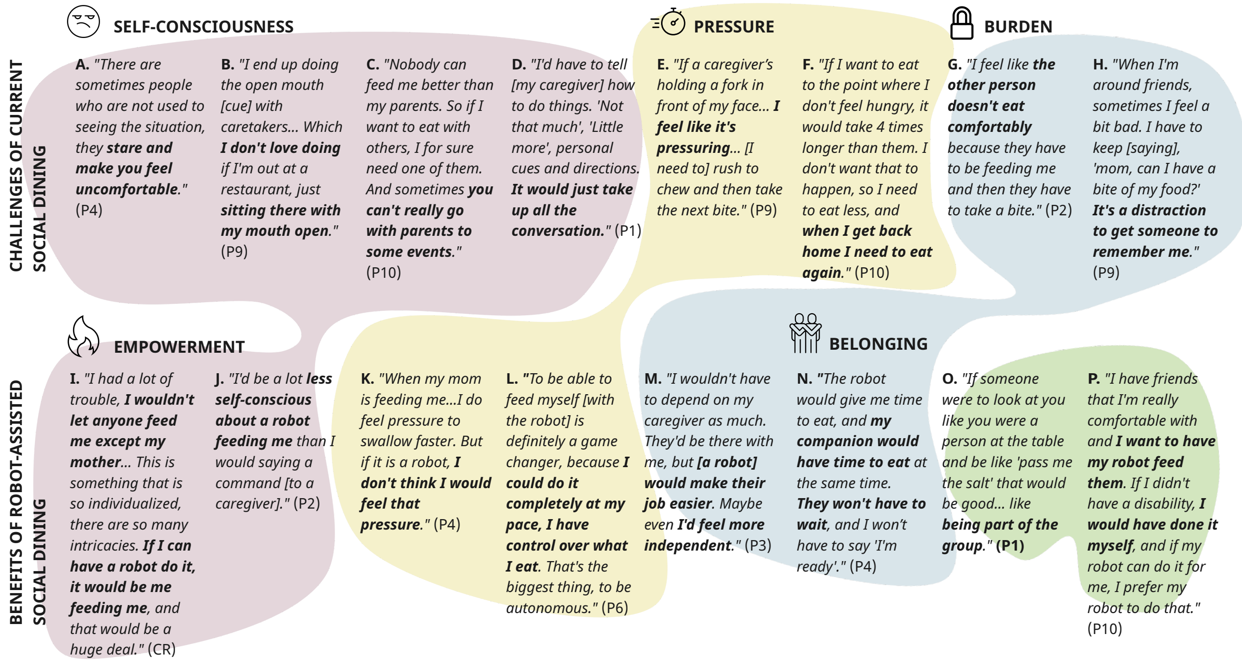


Figure 4.5: The top row shows the negative emotions participants felt during social dining, while the bottom shows benefits they perceived from robot-assisted social dining. Background shapes connect related quotes.

4.6.1 Social Dining Challenges

Some participants enjoyed social dining: *"They take a bite, I take a bite, it becomes part of the interaction. We enjoy the meal together."* (P7). However, most **preferred not to eat socially** due to **repeated challenging experiences**. *"I don't like it. I'll arrive and be like 'nope, I'm good, I already ate'... A lot of people eat out for enjoyment. For me, it's not like that. Eating is a necessity, I don't do it for fun."* (P1). Fig 4.5A-H provides an overview of these challenging experiences.

4.6.1.1 Divided Attention

A challenge participants discussed is that caregiver attention is divided among feeding, eating, and social interactions. Therefore, participants have to verbally remind caregivers to feed them (Fig 4.5H), interrupting conversations and making them feel **self-conscious** (Fig 4.5D) or **burdensome** (Fig 4.5G, 4.5H).

4.6.1.2 Caregiver Variability

Participants noted caregivers' lack of consistency in meeting their needs because **different caregivers feed them differently**. Some feed too fast, causing **pressure** (Fig 4.5E), while others feed too slowly, causing **frustration**. Some offer bites that are too large, a **choking hazard**, while others' are too small: *"One day my dad's shoving half a chicken down my throat, the next a nurse is cutting the tiniest pieces; I'm like: 'I'm gonna be here for centuries!"'* (CR1). To cope, some participants rely on a few consistent caregivers to feed them but felt that it can be **inappropriate to bring specific caregivers to some social events** (Fig 4.5C).

4.6.1.3 Undesired Attention

Participants raised the challenge of bystanders staring or pointing at them, causing **discomfort** (Fig 4.5A). Such discomfort could be sufficiently powerful to influence preferred dining venues: *"I prefer outside dining because it's more laid back, more distractions. It's not me in a big chair in this tiny diner. I have less people watching me."* (P9).

4.6.1.4 Mismatch Between Participant Needs and Social Dining Norms

Some participants felt that signaling readiness for a bite with an open mouth, a common non-verbal way to communicate with caregivers [28], was **awkward** in social settings (Fig 4.5B). Another said she needs to eat slower than is typical in social meals to avoid choking, resulting in feeling **hungry** (Fig 4.5F). Others wanted to avoid food spilling on them but felt **embarrassed** wearing a bib.

4.6.1.5 Mismatch with Environmental Factors

Participants faced challenges due to a mismatch between environmental factors, e.g., **a lack of space and too much noise**, and their needs. One needed to tilt his wheelchair to regulate blood pressure and was constantly concerned: *"Am I going to... tilt back and crash into a waiter?"* (P9). Others adjust how they sit, making it **difficult to interact with others**: *"My chair is oversized, so I don't fit going straight into a table. I have to sit sideways."* (P4). Noise is a concern, too: *"Communicating when it's loud is difficult. I don't have as strong a voice[8]"* (P2). This makes it **difficult to communicate** with dining partners and caregivers.

[8] Not having a strong voice was one of the impacts of this participant's disability.

4.6.2 Trade-Offs of Robot-Assisted Social Dining

4.6.2.1 Benefits

Participants felt that robot-assisted feeding systems could address the challenges they face with caregiver variability and divided attention. They felt that a robot could provide **customization and consistency** (Fig 4.5I), which is difficult to achieve due to caregiver variability. By using the robot to feed themselves, participants envisioned feeling **empowered** (Fig 4.5L), less self-conscious (Fig 4.5J), and less pressured (Fig 4.5K). Participants also felt that using a robot could promote a sense of **belonging**. Specifically, a robot would free up their caregiver's time (Fig 4.5M), enabling them to eat at the same time as the participant (Fig 4.5N). A robot could also open a new realm of belonging by enabling participants to share food with others (Sec 4.6.3.2). Notably, participants did not want the robot to replace their caregiver but rather sought **caregiver-robot teaming**: *"It wouldn't be a problem if my mom just cut the steak and put it in front of me, and the robot could then feed me"* (P4).

4.6.2.2 Lingering Challenges

Participants recognized that a robot-assisted feeding system could not address all challenges. For example, it was still likely to draw **unwanted attention**, and for some that would be a deal breaker: *"If it is going to cause more attention on me, then I probably wouldn't want it"* (P1). They also recognized that a robot **could not address mismatches with social dining norms**, such as taking longer to eat than a 'typical' social meal (Fig 4.5F). Further, they recognized that the robot arm was **unlikely to address mismatches with environmental factors**, too little space and too much noise, but they proposed desired features for the feeding system that could avoid worsening those experiences (see Sec 4.6.3).

4.6.3 User Preferences about Robot Behaviors

After watching the videos, participants shared their preferences and ideas for behaviors of the robot-assisted feeding system.

4.6.3.1 Initiating a Bite

Participants saw 4 ways to instruct the robot to initiate a bite: (1) button-based[9] (Fig 4.4A); (2) open mouth (Fig 4.4B); (3) voice command (Fig 4.4C), and (4) automatic (Fig 4.4E).

Button This was the most desired option for bite initiation. Participants liked that a button is subtle (Fig 4.6P) and *"fewer things [can] go wrong"* (P7). Yet, some felt it would not work well if they have to press it frequently, like when they *"eat lots of popcorn"* (P8). Some participants wanted the button to be part of a phone app.

Open Mouth Participants liked open-mouth bite initiation for its inclusivity (Fig 4.6D) and because it aligns with how they currently interact with caregivers (Fig 4.5B). Yet, some said they would feel awkward opening their mouth socially and were also concerned about face detection failures or robot misinterpretations: *"What if you're talking and the robot thinks you want food?"* (P5).

Voice Participants had concerns about voice detection failing in loud social settings (Fig 4.6A). They also felt it would require them to interrupt conversations (Fig 4.5D), and that the robot may not understand them due to speech impediments. Yet, participants saw the value of voice commands for quieter social settings (Fig 4.6A).

Automatic In automatic bite initiation the robot waits until participants stop speaking before feeding them. Participants were very concerned about no longer having control of the robot with this option (Fig 4.6E) and about potential misunderstandings, e.g., the robot feeds them while they are listening to someone else.

Customizable Bite Initiation Participants saw the different bite initiation mechanism options as complementary. They repeatedly mentioned wanting to decide which option to use based on factors like noise (Fig 4.6A) or lighting. Others wanted multiple options as backup: *"If I'm having a bad day where I can't press [a button], then [I'd like] voice commands"* (P1).

[9] This includes a micro-switch users can mount anywhere, e.g., a tongue switch.

4.6.3.2 Sharing Food

Some participants felt that using the robot to pass food to others would help them feel like an equal participant during a meal (Fig 4.5O). Others did not consider this a priority: *"[due to a] lifetime of being disabled, people don't expect that"* (P8). Some felt that feeding romantic partners was not part of their dynamic, whereas others were excited about letting good friends taste their food (Fig 4.5P) and feeding children (Fig 4.6I) or a pet.

4.6.3.3 Arm Resting Pose

Participants saw 2 aspects of the robot's arm resting pose: (1) *before delivering* a bite, the arm rests in front (Fig 4.4F) or to the side (Fig 4.4G) of them, and (2) *between* bites, the arm rests above the plate or is lowered (Fig 4.4H).

Before Delivering a Bite Participants did not like the robot arm in front of their face since that would obstruct their interactions with others (Fig 4.6O). This was consistent even for participants who could not eat from the side: *"I can't turn my head, so I'd need the food to come directly from the front, but that's just the into-mouth motion. I think to the side is better [for the resting pose]"* (P8). This was an important finding since multiple current robot-assisted feeding systems have the arm rest in front of a user [30, 93, 236].

Between Bites Participants had mixed preferences about where the arm should rest between bites. Some felt it should go above the plate to make the next bite faster: *"It has to reach over the plate to pick the food up. So if it rested in that position, it wouldn't have to make the extra motion"* (P6). Others felt it should be lowered since that *"is less obtrusive, and down out of the line of sight"* (P7). Yet others felt it should be configurable: *'If you want to eat quickly and have it over your plate, that could be one mode... But say you're letting the food settle, it would be nice to have it rest [lowered]."* (P8).

4.7 Design Principles

Using the thematic analysis method described in Sec 4.5.4, we synthesized 8 design principles from participant input to guide the development of robot-assisted social dining systems (see Fig 4.6).

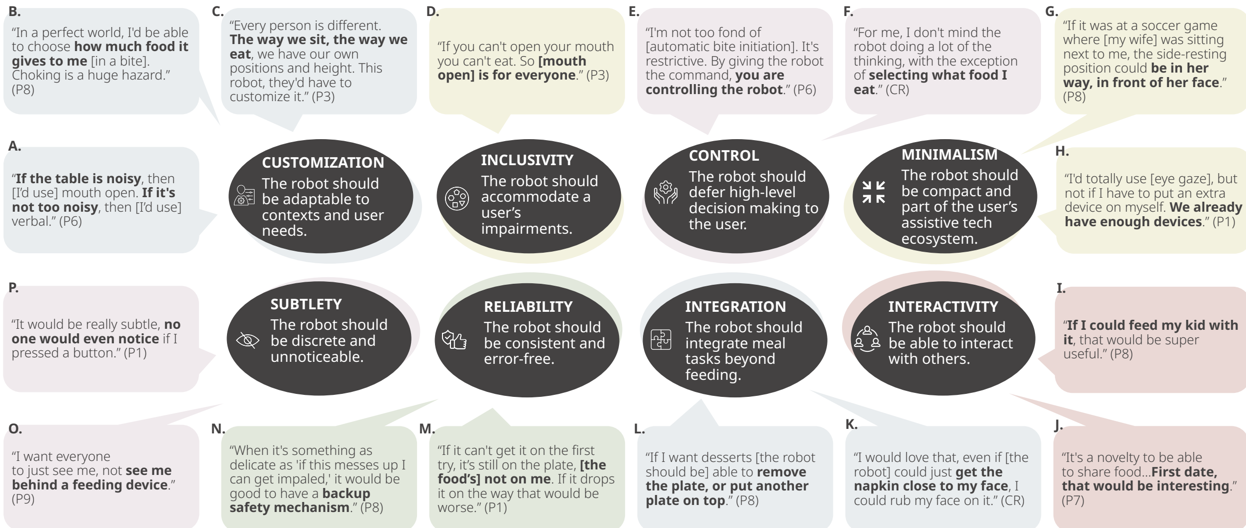


Figure 4.6: Design principles for robot-assisted feeding.

Participants wanted to **customize** their robot so it could work in a variety of environments (Fig 4.6A), be tailored to impairment-specific needs (Fig 4.6B), and work with other individual preferences (Fig 4.6C). This relates to participants' desire for an **inclusive** robot that works across a spectrum of disabilities (Fig 4.6D).

Participants also wanted the robot to be **subtle**. They wanted to communicate with it in a way that would not be noticeable (Fig 4.6P) and not have the robot get in-between them and others (Fig 4.6O). They also wanted it to be **minimalist**, by not adding extra devices to their current assistive technology ecosystem (Fig 4.6H) and not interfering in others' personal space (Fig 4.6G).

Participants wanted a **reliable** robot. They did not want the robot to make errors that have social repercussions, such as spilling food (Fig 4.6M). They also wanted access to an emergency stop in case of robot errors (Fig 4.6N). Further, they wanted to be in **control** of the robot; most participants reacted negatively to the proposal of a robot automatically deciding when to initiate a bite (Fig 4.6E) or what food item they should eat (Fig 4.6F).

Finally, participants wanted a robot that **interacts** with social partners and **integrates** with other meal components, e.g., using the robot to feed others (Fig 4.6I, 4.6J), move plates between courses (Fig 4.6L), wipe their face (Fig 4.6K), or team up with caregivers to achieve tasks it cannot do by itself.

Participants differed in how much, and in what realm, they prioritized each principle. For example, consider **control**. Some participants wanted to control the robot's pace of feeding (Fig 4.5L), while others

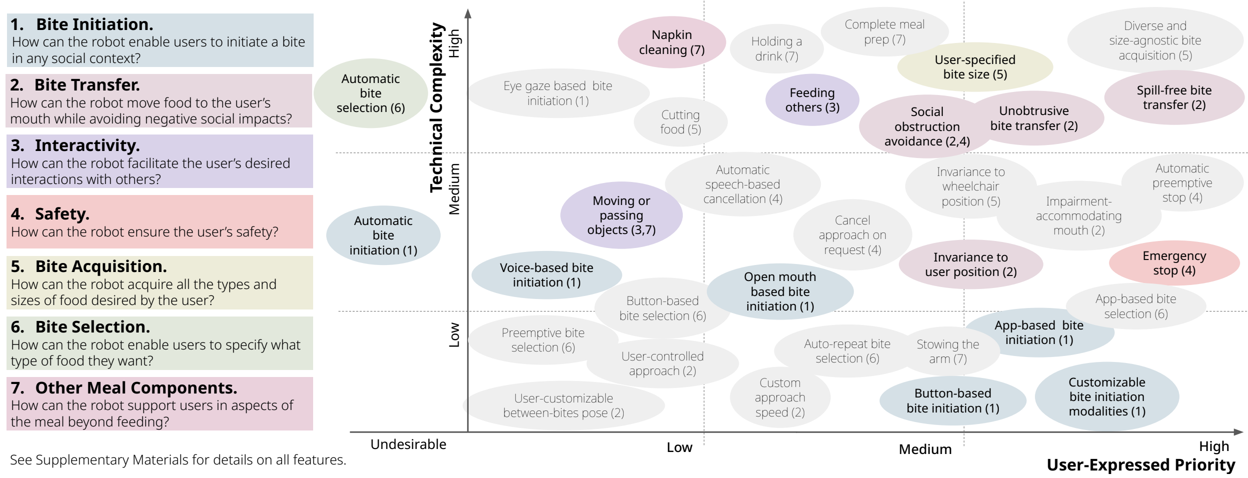


Figure 4.7: This implementation guide contains features users discussed, organized by user priority and technical complexity. Highlighted features are mentioned in the main paper; those in grey appear in Supplementary Materials (Sec 4.11).

were *"willing to let the robot decide pace...because it's a thing I already deal with with caregivers"* (CR1). Or consider **reliability**. For some participants, unreliability in bite size would render the robot unusable because *"I can only open my mouth so far because of atrophy"* (P8). For others, the behavior that needed to be reliable was face detection because *"I have limited movement, so if it doesn't detect my open mouth that would be the most frustrating"* (P2). This diversity of user preferences is a reminder that design principles are only guides and cannot replace user studies for in-depth identification of specific users' priorities.

We recommend that researchers use these principles when making design decisions about robot-assisted feeding. For example, when designing the robot behavior of passing food, the principles of **control** might lead a researcher to not have the robot directly indulge a request from someone else, but rather wait for the user to instruct it to pass food. When designing the before-bite resting pose, the principle of **subtlety** might lead a researcher to have the arm rest on the side, not the front, of the user's face. However, if approaching from the side reduces the accuracy of face detection and impacts the robot's **reliability**, then design principles conflict and should be resolved via a user study.

4.8 Implementation Guide

Integrating the preceding findings, we present a guide for implementing robot-assisted social dining. This guide, Fig 4.7, is intended to help researchers identify and prioritize technical features to work on when developing a robot-assisted feeding system.

To develop this guide, two researchers with experience developing robot-assisted feeding systems analyzed all user quotes related to (un)desirable robot behaviors. For each quote, they identified the **technical features** that would be needed to implement that behavior and grouped similar features together. For example, participants' desired robot behavior when transferring food to their mouth contains multiple underlying technical features: "unobtrusive bite transfer" (Fig 4.6O) and "social obstruction avoidance" (Fig 4.6G).

The researchers then labeled each feature with a **technical complexity** (y-axis, Fig 4.7) of high, medium, or low. 'High' was assigned to features that require novel research to implement; 'medium' to features in prior work that require adaptation to implement; and 'low' to features implementable with out-of-the-box code. For example, "open mouth bite initiation" was assigned medium since it can use out-of-the-box face detection but requires camera calibration and accounting for obstructions (e.g., utensils blocking the mouth).

The researchers then analyzed the quotes associated with each feature and assigned the feature a **user-expressed priority** of high, medium, low, or undesirable (x-axis, Fig 4.7). For example, the feature "open mouth bite initiation" was given a priority of medium because some participants liked it but others had concerns about it failing or being socially awkward to use.

There are several ways to use this guide. A PhD student looking for a dissertation topic might focus on multiple features in the same group. A first-time researcher might focus on a feature with low complexity. A startup developing a minimum viable product might focus on features with high priority. In general, this guide serves to facilitate future work in robot-assisted feeding.

4.9 Reflections on Community-Based Participatory Research (CBPR)

In this section, we present reflections on benefits of and best practices for CBPR, to facilitate the use of CBPR in HRI research.

Shared Experiences. During interviews, the community researcher and participants discussed shared experiences living with motor impairments, which the academic researchers did not have. These moments of empathetic support created trust that enabled deeper insights from the conversation, which would have been impossible without the community researcher.

Building Community. During interviews, participants sometimes raised challenges they faced with assistive technologies, and the community researcher offered advice. This occasionally extended further, with the community researcher sharing resources and meeting participants post-interviews to offer further support.

Demystifying Research. Research can be confusing for newcomers. For example, there are methods for asking questions without biasing participants, procedures for running studies, and terms like "semi-structured interview" that can be obscure. Thus, we integrated explanations of the research processes throughout our collaboration. Having common terminology and expectations enabled the community researcher to make informed decisions when co-creating timelines, protocols, and action items.

Accessible Collaboration. One topic we frequently discussed was how to collaborate accessibly. This included holding all meetings virtually, at a time that accommodated the community researcher's disability-related needs, and having preparatory meetings before design interviews. Holding weekly team meetings was also essential to counteract the knowledge imbalance between the academic team (more familiar with research) and the community researcher (more familiar with living with motor impairments).

Research Time. Throughout our collaboration, the community researcher and participants experienced challenges such as illness, insurance challenges, and technical problems. At those times, the academic team also paused, progressing only when the full team reunited (see Fig 4.2 for canceled interviews). Where possible, they provided support, such as by connecting the community researcher with a resource to appeal denied health insurance coverage. Accommodating delays and supporting the community researcher beyond the project are crucial to an equitable and sustainable partnership.

4.10 Limitations and Future Work

Our sample does not represent all stakeholders in a few dimensions: (a) only $2/10$ participants were women; (b) we did not interview caregivers or social dining partners who indirectly use the system; (c) and all participants had permanent impairments (as opposed to temporary, e.g., a broken arm). Future work involves diversifying participants, particularly including informal caregivers to understand how they think a robot might alter social dining dynamics.

Participant preferences were derived from discussing speculative videos. However, interacting with a physical robot involves nuances that cannot be captured in videos. An important future step is to implement the features in Fig 4.7 and have a long-term deployment. Participants may then evaluate the features in social settings and provide further insights into future directions for development.

Yet another interesting direction is investigating features that can facilitate the caregiver-robot teaming discussed in Sec 4.6.2.1.

4.11 Supplementary Materials

Supplementary materials are hosted on the Open Science Foundation at [8]. They include the study protocol, codebook, tagged quotes, details on Fig 4.7 features, attribution for icons, and more.

Chapter 5

Generalizing Bite Acquisition With Human-Informed Actions

This chapter investigates how we can enable a robot-assisted feeding system to feed users a meal of their choice. Specifically, this chapter focuses on the following question:

RQ2 How can a robot-assisted feeding system feed users the variety of food items they want to eat?

This chapter was originally published as "Towards General Single-Utensil Food Acquisition with Human-Informed Actions" at the *Conference on Robot Learning* in 2023 [104].

5.1 Introduction

Eating is a fundamental part of the human experience, and robots can play an important role in facilitating the transfer of food from farm to kitchen to plate to mouth. But the prerequisite process of actually picking food up can be tricky. Food is fragile and can fall apart with excessive forces or stick to both the robot and the environment. Food is visually diverse and hard to simulate. In industrial transport and packaging settings, specialized end-effectors can utilize suction [229] or soft enveloping links [308] to safely and reliably grasp a variety of foods. However, for in-home food manipulation, such hardware can be unavailable, impractical due to limited space, or uncomfortable for humans to interact with. In particular, this work is motivated by

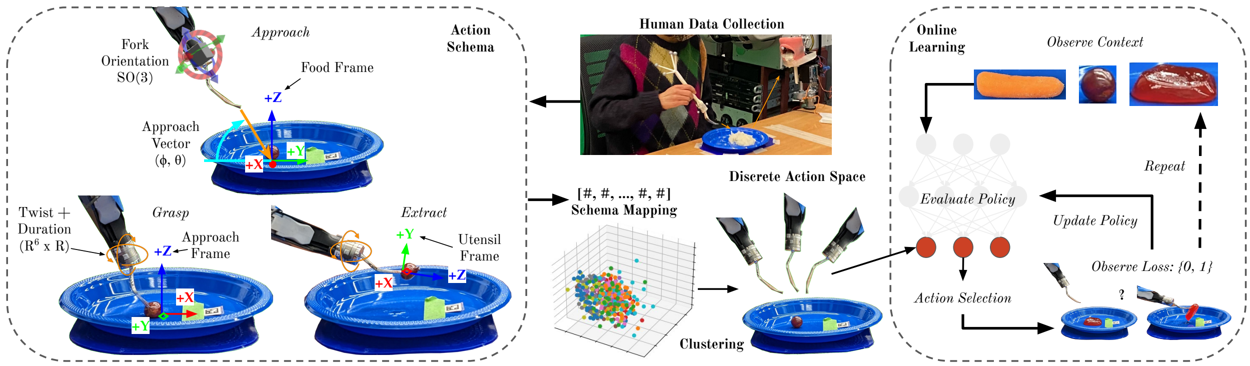


Figure 5.1: *(Left)*: Visual description of the action schema. Robot motions are in orange. Reference frames are represented as three-color axes with X in red, Y in green, and Z in blue. *(Right)*: General food acquisition pipeline. Human data is collected, mapped into the schema, and clustered into a discrete action space. This space is small enough to treat food acquisition as a contextual bandit to learn online the optimal action for new food items.

the application of robot-assisted feeding for those who cannot eat on their own (1.8 million people in the United States alone [292]). In this setting, there is strong coupling between the problems of food acquisition and mouth transfer [93]. Therefore, we focus on the problem of acquiring food with common, single-piece utensils (i.e. a fork or spoon) with which users are already familiar.

Formally, our goal is to learn a map from food context (e.g., RGBD images of the food item) to a sufficiently good utensil trajectory and control strategy to acquire the food. In the assisted feeding context, previous work [31] suggests that that "sufficiently good" is on the order of an 80% success rate, depending on the level of mobility impairment.

Humans are generally experts at food acquisition, so a natural approach to this problem is to use a human dataset collected on a variety of food items to learn this map directly. This approach runs into sample complexity issues, as food is hard to simulate, and it is difficult to capture the quantity of real world data necessary to cover the diversity of food and the diversity of possible ways to acquire it. Previous work [29, 89] got around this issue by using human data as a qualitative starting point for the manual design of a small set of actions with some success on a limited set of food items.

Supported by this work and other work exploiting the haptic similarities between food items [106], our key insight is that a very small subset of the space of possible acquisition actions is sufficient to acquire almost all food items that a human is capable of picking up with a fork. We can capture this subset in a principled, data-driven way by observing a relatively small number of human acquisition trajectories on arbitrary food. By parameterizing these trajectories within an interpretable metric space, we can use off-the-shelf clustering to create an even smaller representative discrete space. We empirically show that this discrete space is simultaneously large enough to effectively cover the space of food contexts and small enough that online learning can identify the best action for multiple new food contexts over the course of a 30 minute meal.

In summary, this work presents three contributions in the field of robotic food acquisition. (1) In Section 5.3 we present the schema that defines an over-parameterized action space to capture human food acquisition techniques with a common single-piece utensil (i.e. fork or spoon). (2) In Section 5.4, we present a dataset of human food acquisition trajectories on a user-motivated variety of food items. (3) In Section 5.4.3, we describe both a method for distilling discrete actions from human data and a set of 11 actions constructed from our dataset. This method is tailored to in-home robot-assisted feeding, but we believe a similar structure could be used for other applications where the key insight holds.

In Section 5.5 we demonstrate that the discrete action set is sufficient to acquire a variety of food items that are visually dissimilar from those used in the human study. We demonstrate in Section 5.6 that we can quickly use off-the-shelf online learning techniques to determine a sufficiently optimal action for previously-unseen food items within 13 trials per item. Finally, Section 5.7 discusses avenues for future work in both general food manipulation research and the application of robot-assisted feeding.

5.2 Related Work

5.2.1 Robot-Assisted Feeding: Food Manipulation

Robot-assisted feeding has been explored in industry and academia, yet still contains many unsolved problems. Commercial table-mounted systems [191, 192, 205, 193, 254, 22, 27, 214] are available, but work with fixed trajectories for food acquisition. This can result in issues such as food being pushed off of the plate, as found in a study of the Bestic system [220]. There are also teleoperated options available, such as [45], but users can have difficulties using a fully teleoperated system [63].

Food acquisition is a fundamental part of robotic feeding devices [60]. Some prior work creates specialized tools for food manipulation [229, 309, 81, 308]. We focus instead on food acquisition with a fork, as it is a common household utensil with which users are familiar. Other research that has focused on picking up food with forks uses a limited set of skewering actions, which works well for a small set of food items but has difficulties on more varied items [105, 125, 89]. Other work added variety with utensil swapping [236], which adds hardware complexity, but kept the set of food items and action trajectories small.

Further work in food acquisition uses vision or haptic data to improve the choice of acquisition action. Vision can be used to classify visually different food items, and haptic data can assist in identifying foods that look different but require similar actions (e.g. grapes and cherry tomatoes). Some prior works require additional "probing" actions for every food item [264, 96, 29] or specialized sensors beyond force torque sensors [315]. Other work uses haptic data to improve food acquisition during the feeding process, but the expert-designed action space does not cover the variety of food necessary for in-home applications [106, 283]. This paper leverages the learning approaches of past work with a human-informed action space that is likely to cover a wider variety of food items that users might want to eat.

5.2.2 Learning Grasps from Human Demonstrations

As humans are generally expert tool users, a plethora of work has gone into transferring those skills to robots [34, 305, 76]. Some work focuses on higher-level task planning [304, 241]. Others learn more granular motions by generating dynamic motion primitives that a model learns to stitch together [189, 4]. Still others investigate a hand-design restricted action space for use with end-to-end models [54]. In contrast, this work looks at leveraging application-specific structure and human data to systematically restrict the action space prior to learning a model.

Other works [275, 318, 187, 316, 85] utilize simulations or extensive human and environment data to learn offline RL policies that yield good generalization performance at test time. In contrast, this work exploits application-specific structure and the simpler contextual-bandit setting to learn a simpler model that can be refined online without the need for simulation or large datasets.

5.3 Acquisition Action Schema

Previous work [29] qualitatively captured a taxonomy of human food acquisition techniques with a fork. This taxonomy included skewering and scooping and highlighted the importance of the approach angle and in-food manipulation strategies (e.g., wiggling for greater pressure or rotation for greater contact area). Yet, that work does not provide a quantitative way to represent those motions so a robot can execute them.

Filling this gap, the following schema describes an acquisition action space that is narrow enough to distill these taxonomic elements but flexible enough to capture variants of those elements (e.g., additional wiggling, or a different approach angle). The action is defined by 26 continuous parameters divided into three phases: approach, grasp, and extraction (Figure 5.1, left).

5.3.1 Approach (Pre-Grasp)

This phase captures the fork tilt and and approach angle elements of the qualitative grasp taxonomy.

Frame Definitions: Define the world frame with an arbitrary origin and orientation such that -Z is the direction of gravity. We assume the existence of a food manipulation target bounded by an ellipsoid which may possibly intersect a flat plane defined by a table/plate/other surface parallel to the X-Y plane (from here just referred to as the plate). The projection of this ellipsoid onto the plate is the *bounding ellipse* of the food. The *food frame* is the default reference frame in which all parameters are defined unless otherwise specified. The origin is defined as the center of the bounding ellipse. +Z is aligned with that of the world frame and the X-axis is aligned with the major axis of the bounding ellipse.

Parameters: The approach consists of the following 9 parameters: Fork orientation ($SO(3)$), Approach polar and azimuthal angle ($[0, \frac{\pi}{2}] \times [0, 2\pi)$), Target approach point within the food ($\mathbb{R}^3$), Force threshold ($+\mathbb{R}$).

Implementation: During implementation, the utensil begins an arbitrary distance from the food and moves in a straight line towards the target approach point until either that point is reached or the force on the utensil exceeds the threshold. Note that the definition of the food frame introduces a π-rotation symmetry depending on which direction along the X-axis is +X. In this work, this symmetry is broken during the on-robot experiments based on which approach direction is within the robot's workspace and easiest for the on-board planning algorithm.

5.3.2 Grasp

This phase captures the wiggling, twirling, and in-food scooping motions of the qualitative taxonomy.

Frame Definitions: Define the *approach frame* as the food frame rotated by the azimuthal angle of the approach direction. Define the *utensil frame* with an origin at the very tip of of the utensil (e.g., between the middle two tines on a fork) such that +Z points along the handle of the utensil and the X-axis goes across the face of the utensil. In other words, the Euler angles in this frame correspond with roll (Z), pitch (X), and yaw (Y). Using this frame instead of the food frame allows the approach and grasp to be parameterized independently. For example, approaching from the side instead of the front should not result in a grasp rotation yawing instead of pitching the fork.

Parameters: The grasp consists of the following 9 parameters: Angular velocity in utensil frame ($\mathbb{R}^3$), Linear velocity in approach frame ($\mathbb{R}^3$), Duration ($+\mathbb{R}$), Force and torque thresholds ($+\mathbb{R} \times +\mathbb{R}$).

Implementation: During implementation, the utensil will execute the provided velocities for the provided duration, or cut short if the force or torque thresholds are reached.

5.3.3 Extraction

This phase captures any stabilizing rotations that take place after the food is on the fork.

Parameters: Similarly to grasp, the extraction consists of the following 7 parameters: Angular velocity in utensil frame ($\mathbb{R}^2 \times +\mathbb{R}$), Linear velocity in approach frame ($\mathbb{R}^3$), Duration ($+\mathbb{R}$). While a force and torque threshold can be introduced, it is rendered unnecessary in this work by requiring the extraction motion to move against gravity away from the plate.

Implementation: Extraction is implemented the same way “grasp“ is.

5.4 Human Bite Acquisition Strategies

Although the 26 dimensional action schema is large, we hypothesized that only specific points (acquisition actions) within this schema will actually be commonly used to acquire food items. To identify those points, we had able-bodied participants acquire a variety of food items and feed them to an actuated mouth.

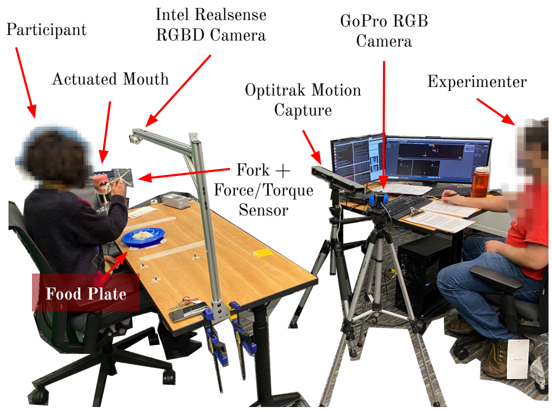
5.4.1 Study Design

The study involved participants acquiring bite-sized pieces of a variety of food items with a fork and feeding them to an actuated mouth. The choice of food items was informed by ongoing collaboration with an end-user with C1 quadripledia[1], who had his caregiver take pictures of all the meals he ate in a week. One researcher then grouped similar food items (e.g., bread bun and bagel), resulting in a final set of 13 diverse food items: bagel chunks, mini sub sandwiches, pizza, chicken tenders, fries, broccoli, glazed doughnut holes, mashed potatoes, lettuce, spinach mix, whole jello, instant ramen noodles, and brown rice with beans. The bagels, sub sandwich, and pizza were pre-cut into bite-sized chunks, building off of past research that found that users are okay with caregivers cutting their food into bites before the robot feeds them [208]. The same brand ingredients and preparation procedure were followed for every food item.

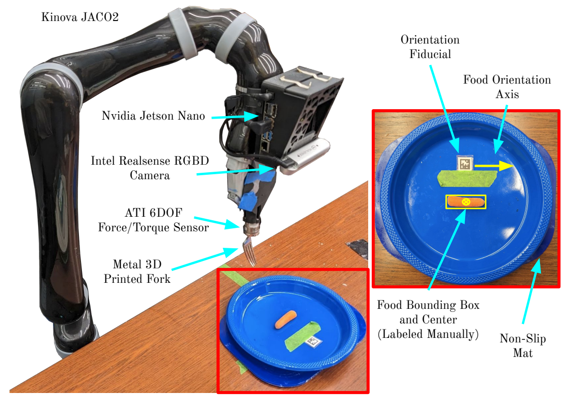
The study space consisted of a table with a plate of food on it, a fork near the plate, a chair for the participant to sit in, and an actuated mouth to the left of the chair[2]. An RGB-D camera[3] above the plate captured visual aspects of the food. The fork, table, RGB-D camera, and actuated mouth all had motion capture markers on them, which were tracked by a motion capture system[4] in front of the table. The fork was also actuated with a force-torque sensor[5] to measure haptic aspects of the participant's acquisition action. An experimenter sitting behind the motion capture system, in full view of the participant, oversaw data collection. Figure 5.2a shows the study setup.

When each participant arrived, they were first briefed on the study and given time to read and fill out a consent form. They were then given a chance to familiarize themselves with the fork[6] by feeding baby carrots to the actuated mouth. When participants were ready, the actual data collection began, where they were provided a plate with one of the 13 food items, in randomized order, and were asked to feed one bite at a time to the actuated mouth. For each bite, participants were first asked to hold the fork in a comfortable "ready" position above the plate. When the experimented said "start," they lowered the fork to acquire the food item, moved it to the actuated mouth, and held it there until the experimenter said "stop." For each plate of food, participants were asked to feed at least 5 bites, and possibly more if the motion capture system lost tracking of the fork during the bite. In total, each study session took one hour and participants were compensated with a $10 gift card. Three researchers ran the study, and the study procedure was approved by our university's IRB. We had 9 participants, who all happened to be right-handed.

[1]C1 quadriplegia refers to paralysis of all four limbs as a result of an injury to the first, or top-most, cervical vertebrate.
[2]Because all participants happened to be right-handed, this positioning allowed them easy access to feed the actuated mouth.
[3]Intel RealSense D415
[4]OptiTrak V120 Trio
[5]ATI Industrial Automation 6DOF Nano25 with Net F/T Interface
[6]Due to the force-torque sensor and motion capture dots, the fork was a different shape and weight from regular forks



(a) Human food acquisition data collection setup.



(b) On-robot experiment setup.

Figure 5.2: Food acquisition trials. For each trial, a single food item was acquired. Food perception (both center-of-mass and orientation) was performed with classical computer vision through a fiducial and color-based background rejection.

5.4.2 Dataset and Qualitative Observations

For each bite the participant acquired, our time series data consisted of the fork pose, force-torque sensor readings, and RGB-D images of the plate. We first cleaned this data by removing mistrials (i.e., trials with missing or corrupted data) and transforming all poses to a uniform frame of reference. We then published this dataset [207] to facilitate future research in food acquisition strategies. This dataset consists of 496 trials, totaling over 1.25 hours of food acquisition data across 13 food items and 9 participants.

Similarly to previous work [29], we observed some patterns in user interaction during data collection.

- Different participants held the fork differently (thumb in front of or behind the fork), which gave rise to different acquisition actions;
- Some food items could be acquired with multiple types of actions, e.g., some users scooped noodles whereas others twirled them;
- Users often tilted the fork while putting downwards pressure on it, in order to get it to pierce the food (e.g., broccoli).

Some of these lead to emergent behaviors identified in Section 5.4.3.2.

5.4.3 Human Data Analysis

5.4.3.1 Extracting an Action Schema Point

For each bite's acquisition data, we extracted a point within the acquisition action schema that was close to that motion. We developed the procedure by iteratively extracting action schema points from bites, and then visualizing the actual participant's motion superimposed with the extracted action on a random subset of bites to determine how to improve the extraction procedure. The complete procedure is detailed below and in our code repository[7].

Exclusion Criteria We excluded any trial where tracking of the fork tip was lost for more than 0.5 seconds, or where the motion capture system did not read the stationary object poses (e.g., table, mouth) for the entire trial. After exclusions, we had a total of 410 trials.

Pre-processing To remove noise from the motion capture system, we smoothed all the fork tip poses by applying a median filter of 0.33 seconds separately to the x, y, z, roll, pitch, and yaw of the pose.

Significant Timestamps We first extracted the significant timestamps of the user's acquisition action. Specifically, we defined the *contact time* as the first timestamp when the distance from the fork tip to the camera exceeded the distance from the pixel corresponding to the fork tip to the camera in the initial depth image of the plate (i.e., the first time the fork pierced the surface of food on the plate). We then worked backwards from contact time, defining *start time* as the end of the 0.5 sec interval where the fork tip was consistently within a sphere of radius 5 cm, was more than 5 cm from its lowest point, and was more then 35 cm from the mouth. This criteria was based on our experimental design, where we asked the participant hold the fork stationary at a "ready" position before they acquired the food. In the event of multiple such periods where the fork was held stationary, we chose the one where the fork was the highest. We then defined *end time* to be the last timestamp when the fork was within 7 cm from its lowest point. And finally, we worked backwards from end time, defining *extraction time* to be the latest time when the fork was 1 cm away from its lowest point, within 2 sec of the end time.

[7] https://github.com/personalrobotics/corl23_towards_general_food_acquisition

All-in-all, the motion the participant took between *start time* and *contact time* corresponds to their **pre-grasp** motion, the motion they took between *contact time* and *extraction time* corresponds to their **grasp** motion, and the motion they took between *extraction time* and *end time* corresponds to their **extraction** motion.

Food Reference Frame Although a robot will know the food it is targeting before beginning its motion, with human data we have to extract the food item they were targeting from the data. We did so by segmenting the visually separate food items in the first RGB image of the plate of food. First, we detected the plate by: (a) using inpainting to remove glare; (b) using k-means clustering (k=3) to simplify the colors; and (c) finding the largest contour of the image that has at least 50% blue pixels. In practice, this reliably detected the plate for every food item in our study. We then masked out all the non-plate pixels from the de-glared image, and detected the food bounding box by: (a) running k-means (with k=2 for most food items, and k=3 for broccoli since its colors were closest to blue) to simplify colors; (b) masking out all the blue colors; (c) narrowing the mask to separate touching food items; (d) computing contours; and (e) fitting rotated rectangles to every contour with an area between within a hardcoded range. In practice, this reliably segmented separate food items, like bagel pieces or chicken tenders. For food items with a lot of overlap, like fries, this approach sometimes segmented multiple pieces of fries as the same. However, since participants often also acquired multiple overlapping pieces of those food items, we accepted those slight errors in food detection. For foods that weren't separated into bites like noodles or mashed potatoes, this algorithm rightly segmented it as one contiguous chunk of food.

Once we segmented separate bites of food, we defined the food reference frame to be centered at the center of the bounding box the fork tip was in at contact time, rotated to align with its major axis (i.e the center of the bite the user selected.)

Pre-Grasp The above preliminaries enable straightforward extraction of the pre-grasp, grasp, and extraction components of the action schema. For pre-grasp, we computed the target offset as the the target offset as the fork position at contact time in the food reference frame. We computed the initial utensil transform by taking the fork's linear velocity during a 0.5 second window before contact time and extrapolating that backwards 0.1 m, with a fixed orientation. And we took the force threshold to be 50% of the max force between start time and contact time.

Grasp We defined the in-food twist to be the transformation between the fork pose at extraction time and contact time, and the duration of the twist to be the duration between extraction time and contact time. We defined the force and torque thresholds to be 50% of the max force and torque between contact time and extraction time.

Extraction We defined the out-of-food twist to be the transformation between the fork pose at end time and extraction time, and the duration of the twist to be the duration between end time and extraction time.

Cleaning Three trials resulted in values of NaN or inf for at least one of the dimensions of the action schema. We eliminated these, resulting in 407 actions used for clustering.

5.4.3.2 Clustering Actions

We ran k-medoids on the entire dataset of 407 extracted actions, where each action has 26 dimensions. Notably, the clustering did not consider aspects of the action that would be outputted by the perception system, such as food reference frame, and only included the aspects of the acquisition motion that might generalize across food items. This resulted in $k = 11$ actions (corresponding to the within-cluster-sum-of-square-distances elbow point). A full quantitative parameterization and video of each action are provided in the supplementary materials. Qualitatively, we observed emergent behavior consistent with findings in previous work [29], such as in-food wiggling, tilted extraction, and the use of vertical tines for high force.

5.4.3.3 From Human to Robot Actions

Although this procedure outputs representative actions for the motions that participants took when acquiring food items, it also learns some aspects of motion that are particular to the morphology of a human arm. For example, participants' motions tended to approach food from the right, since they were right-handed and feeding a person to their left. However, since the robot approaches food from above and feeds someone behind it (sitting in the wheelchair), it no longer needs to be constrained to right-to-left motion. As described in Section 5.3, the definition of the food frame is ambiguous with respect to a π rotation about the axis of gravity. Since the food location was fixed, we manually broke this symmetry by choosing the orientation that was easiest for the on-board planning algorithms. Further, some in-food grasp motions that humans executed, specifically tiny rotations of the fork (3° or less), produced negligible motion of the robot with significant planning and collision checking time; so we truncated those rotations to 0°.

5.5 Experiment 1: Action Evaluation

This experiment was designed to test the utility of the discrete 11-action space on a variety of food items. Previous work in assisted feeding suggests that, depending on level of mobility impairment, users can generally tolerate up to 20% failure rate in food acquisition [31]. Our hypothesis was two-fold: (1) *Coverage:* for each food item, at least one action would meet or outperform baseline performance and meet the 80% user threshold. (2) *Minimal Bad Actions:* each action would have acceptable performance on at least one type of food. If the action set lacks coverage, it is likely too small to adequately acquire new food items in the home, while if there are many bad actions, it is likely too big and will be difficult to use for online learning.

5.5.1 Experiment Setup Details

Our experimental setup is shown in Figure 5.2b and was performed with a 6 DoF JACO2 robotic arm [140] with a 3D-printed handle and fork-shaped end-effector. To implement the force/torque thresholding, we instrumented the fork with a 6-axis ATI Nano25 Force-Torque sensor [269]. The center of the food (and the bounding box used for visual context in the online learning experiment described in Section 5.6) were annotated manually from the robot's eye-in-hand vision system. This was done in order to run a controlled experiment specifically focused on food acquisition, as opposed to introducing additional variance with a (possibly imperfect) food perception system. This system includes the Intel RealSense D415 RGBD camera and the NVidia Jetson Nano for wireless image transmission. Food was placed on a plate equipped with an AprilTag [225] for camera calibration and mounted on an anti-slip mat commonly found in assisted living facilities [28].

In each trial, the food placed in the vicinity of the AprilTag on the plate and oriented such that the major axis of the bounding ellipse is parallel to the bottom edge of the fiducial. The end-effector moved to a fixed position above the plate, and the location of the center of the food is annotated manually in the RGBD camera image. After action execution, we wait at least 3s before recording success or failure. For most food items, success is defined as the entire item being removed from the plate. If a homogeneous food item breaks (a common occurrence with banana slices), at least half of the item needs to end up on the fork. For the sandwich, success required that all layers (both pieces of bread, the lettuce, and the cheese) make it off the plate. Finally, for multi-piece and continuous items (i.e. potatoes, rice, noodles), a conservative success metric was set at 200mg ($\sim$15 grains of rice, or 1 full noodle).

5.5.2 Experiment Design

We evaluated our action space on 14 diverse food items. Some food items were identical to those used during the human acquisition data collection (Section 5.4): fries, broccoli, mashed potatoes, spinach mix, and jello (cut into ~ 1.5cm slices to obviate the need for cutting). Some food items had similar properties to those in the human data collection with different visual characteristics: powdered doughnut holes, white rice, white bread sandwich, and flat noodles. Finally, some food items were new: baby carrots, grapes, half-strawberries, banana slices, and kiwi slices.

The baseline action set consisted of 3 skewering techniques pulled from past fork-acquisition work [89]. *Vertical skewer* (VS) orients the handle of the fork to be orthogonal to the table and moves straight down applying up to 15N of force before moving straight back up. *Tines vertical* (TV) orients the tines of the fork to be orthogonal to the table and again applies 15N of force straight down before moving back up. Finally, *tilted angle* (TA) orients the handle of the fork 45 degrees off the table normal with the fork flat facing up and approaches the food at that same angle, moving straight upwards after skewering.

For each food item, we perform 10 trials with each of the 3 baseline and 11 human-informed actions for a total of 14 actions $\times$ 14 food types $\times$ 10 $=$ 1960 trials. At about 1 minute per trial, data collection took about 33 hours.

5.5.3 Results

These results are summarized in Figure 5.3(Left). All error bars represent Wilson Binomial Proportion 95% Confidence Intervals ($n = 140$ in aggregate, $n = 10$ per food item). Overall, the best action for each food item from the human-informed set significantly outperforms both the best action from the baseline set and the the user-defined benchmark with a success rate of 94.6% ($p < 0.05$ necessarily by non-overlapping confidence intervals).

Coverage All food items except for spinach exhibited a success rate of 90% or higher within the new action space, exceeding the 80% user benchmark. Single-leaf spinach, difficult to acquire due to its thinness, came close with a 70% success rate with the best human-informed action. The nearly complete coverage suggests that this action space is large enough to handle the variety of food items necessary for in-home deployment. Additionally, reducing k maintained a subset of the $k = 11$ medioids down to at least $k = 5$, and so coverage is achieved for $k \geq 8$, as below that, the action space does not include the only action that covers jello.

Bad Actions Almost all human-informed actions exhibited good performance on at least one food item. Actions 0, 1, 2, 3, 6, 8, and 10 were the optimal action for spinach, carrot, banana, strawberry, potato, jello, and sandwich respectively. While not an optimal action for any food item, actions 4, 5, and 7 exhibited $\geq 70\%$ success on at least one food item. The only exception was action 9. This action captured the "cutting" motion that humans used on the full, undivided jello cups. Therefore, only the side of the fork comes into contact with the food, making success less likely. That 10/11 actions exhibited good performance suggests that this action space is not excessively large.

Baseline Comparison Carrots, grapes, strawberries, bananas, and broccoli exhibited good ($\geq 80\%$) performance with the optimal baseline actions that were designed for them in previous work [93], with the human-informed actions performing as well or slightly better. Sandwich, fries, noodles, and rice exhibited $60 - 70\%$ baseline performance, with insufficient in-food contact (e.g. not enough of the fork present inside of every sandwich layer) as the primary failure mode. These failures were remedied by the increased in-food grasp motion of the human-informed actions. Finally, jello and spinach were completely impossible for the baseline actions to acquire. Jello needed a significant rotation during extraction to prevent the heavy chunk from slipping off the fork. Spinach needed a significant lateral force during the grasp phase to wedge the fork between the flat leaf and the plate. Finally, the human-informed actions were generally able to acquire a greater mass of rice (258mg vs. 212mg) and potato (890mg vs 5930mg) than TA, the only baseline action with any form of scooping-like motion.

5.6 Experiment 2: Online Action Selection

As described in Section 5.1, we assert that online learning is a necessary component of any in-home robot food acquisition system to handle previously-unseen food items. This is supported by the extensive in-lab evaluation time required for the previous experiment. In this experiment, we evaluate the ability of an off-the-shelf online learning procedure to identify sufficiently good human-informed actions for acquiring previously unseen food items. Our hypothesis is that, given that there are only 11 actions (and as few as 4 can cover the space), such a system should be able to reach the user benchmark on the order of 11 trials for each new type of food. At about 1 minute per trial, in an assistive feeding context, this would happen well within the bounds of a 20-30min meal.

5.6.1 Learning System

As in Experiment 1, this experiment constituted a series of trials. We cycled through all 14 food items, with 1 trial per item. A group of 14 trials, one per food item, constitutes a *round*. Data collection ceased once the success rate over the course of a round exceeded 90% (approaching the 94% optimal). For each trial, food orientation, manual perception, action execution, and success definition, are identical to what is described in Section 5.5.2. The setup for this experiment was identical to the one for Experiment 1(Sec. 5.5.1).

Our online learning procedure models food acquisition as a contextual bandit [32, 105] with visual context and augmented with haptic post hoc context based on related work [106]. Specifically, each trial prior to action selection, we manually annotate a bounding box around the food item in the RGBD image. The cropped image is run through the SPANet [89] model to create a feature vector that constitutes the visual context. During the approach phase of action execution, we collect raw 6D force-torque data, isolate the period immediately following food contact using a z-force threshold, and run the result through a HapticNet multi-layer perceptron model [29] to generate a feature vector that constitutes the haptic context.

We assume a linear relationship holds between both forms of context and the expected reward (i.e. the success rate) of each of the 11 actions. Therefore, we utilize LinUCB [311], which uses the data collected

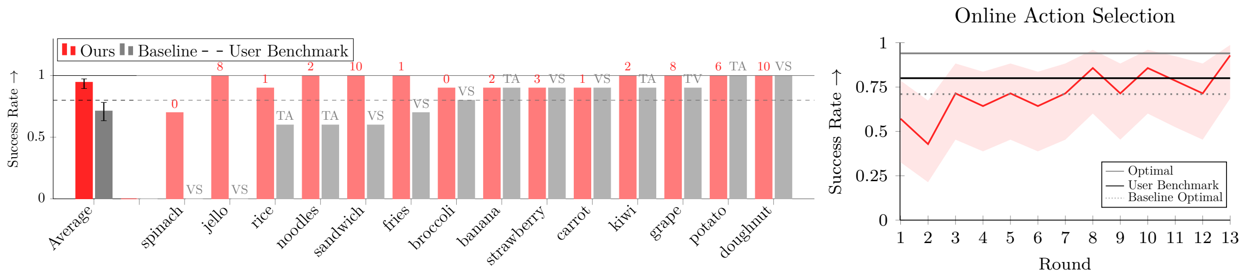


Figure 5.3: *(Left)* Best action for each food item from both the baseline and our action spaces. The specific action is labeled above each bar. *(Right)* Acquisition success rate for each round of 14 trials (1 per food item) using LinUCB. Error bars represent the 95% confidence interval.

so far to assign an upper confidence bound to the reward of each action and optimistically selects the action that has the highest upper confidence bound. Previous work [106] suggests that, despite being collected after action selection, optimizing the linear model with the haptic context can decrease the time needed to converge to the optimal action.

5.6.2 Results

These results are summarized in Figure 5.3(Right), where we plot the success rate across all 14 food items in each round (with Wilson 95% Confidence Interval, $n = 14$). Since most human-informed actions perform well on most food items, we see that we get a 50% success rate even in the first round. By round 8, the action performance is on par with the user benchmark. And by round 13, we have approached the expected optimal performance with this action space (i.e. the Average shown in Figure 5.3). At one minute per trial, this suggests that this system can successfully learn an acceptable acquisition action for 2-4 new types food within the span of at 30min meal, assuming that all foods of a given type have a similar success rate for each action. And that number is likely higher for foods with similar haptic properties. Most food items converged to a sufficiently good action (e.g. Action 1) with 5 rounds. The food items that took the longest to learn were the haptically "unusual" food items like jello, rice, noodles, and mashed potato, which exhibited particularly poor performance on the skewering actions that worked well on the firmer food items.

5.7 Limitations and Discussion

In this work, we present a methodology to use human trajectory data to identify a subset of food acquisition actions that can acquire a wide variety of food items for robot-assisted feeding applications. The 11 actions we distill from our publicly available dataset [207] are sufficient to pick up 14 food items including hard carrots, soft bananas, slippery jello, compound sandwiches, and continuous mashed potatoes. And the set is so small that we can reasonably expect to determine the optimal action for 2-4 food items over the course of a 30-minute meal.

A major avenue for future work is evaluation in a real in-home context. The foods selected for both the human data collection and on-robot experiments were motivated by surveying a participant and co-designer with mobility impairments about their eating habits. We believe they cover a wide variety of rheological contexts, but there may still exist food types that are sufficiently distinct as to not be covered by the actions presented here. Future work in the home can help identify such foods. These can possibly be addressed with online action space expansion. For example, the caregiver can provide some kinesthetic demonstrations that can be mapped into the action schema and averaged. Additional work can investigate the coupling between these actions and the bite transfer process [93].

One significant hurdle to any future in-home experiments is food perception, a prerequisite for any long-term in-home deployment. In this work, food localization was done manually by clicking on the food center. Autonomous food perception is currently being investigated utilizing general segmentation models [160] and could introduce errors that lower the overall success rate.

In general, the interaction between sources of error (e.g. perception, calibration, handling arbitrary food orientation, environmental hazards like obstacles or a non-level table) and optimal action selection is a key realm for future investigation. Since many foods were well-covered by multiple actions, it is possible that the current action space will be robust to such disturbances, but further study is needed. Overall, this work represents a step towards in-home general food manipulation in both feeding and food preparation contexts, and we hope that the provided method, data, and actions can help enable in-home experimentation in the near future.

Chapter 6

A System for Out-of-Lab Robot-Assisted Feeding

This chapter presents the system development and evaluation for the robot-assisted feeding system. Specifically, this chapter focuses on the following question:

RQ3 How can we develop a robot-assisted feeding system to feed users in diverse out-of-lab and in-home contexts?

This chapter was originally published as "Lessons Learned from Designing and Evaluating a Robot-assisted Feeding System for Out-of-lab Use" at the *ACM/IEEE Conference on Human-Robot Interaction* in 2025 [212].

6.1 Introduction

Eating is a basic ADL, one of the "fundamental skills required to independently care for oneself" [79]. Satisfaction with food-related matters is positively correlated with physical and mental health [112, 299, 52]. Unfortunately, for the millions of people who need caregiver assistance to eat,[1] mealtimes can lead to feelings of self-consciousness, pressure, and being burdensome to others. [208, 28].

Robot-assisted feeding is emerging as a promising way to alleviate these challenges [208, 28, 237]. Research in this area often focuses on specific technical components of eating, including bite acquisition [107, 284, 108, 222, 145], transfer [24, 144, 271], and timing [226, 125]. These contributions are

[1]In 2010, 1.8M Americans needed assistance eating due to a disability [292].

Figure 6.1: We evaluate the robot feeding system with: (Bottom) an $n = 5$ study across 3 out-of-lab locations; (Top) a 5-day, $n = 1$ in-home deployment.

evaluated via targeted studies that control for other aspects of eating, e.g., being in a controlled lab environment [24], limiting food positions and types [30, 107], and limiting the number of bites per user [144]. Such limitations are necessary to isolate the component under investigation from other meal-related factors. However, this leaves a gap in developing and evaluating an end-to-end system for robot-assisted feeding.

This paper addresses that gap. *Our goal is to develop an end-to-end robot feeding system that users can independently use to feed themselves meals of their choice outside the lab*. Except for system setup, onboarding, and pre-cut bite-sized meal preparation, users should be able to use the system to independently feed themselves entire meals.

The *key challenge* of developing a system for out-of-lab use is the wide variety of off-nominal scenarios that can arise: e.g., the user may cough; the robot may not acquire a food bite; or the plate may shift. Our *key insight* is that users can overcome many off-nominals, provided acceptable levels of customizability and control over the system.

We worked with two CRs with motor impairments to co-design and evaluate the feeding system. In Study 1, 5 participants and one CR[2] used the robot to eat a meal of their choice in a cafeteria, office, or conference room. In Study 2, one CR used the robot *in his home* over 5 days to feed himself 10 meals in

[2]All CRs and participants need caregiver assistance to eat.

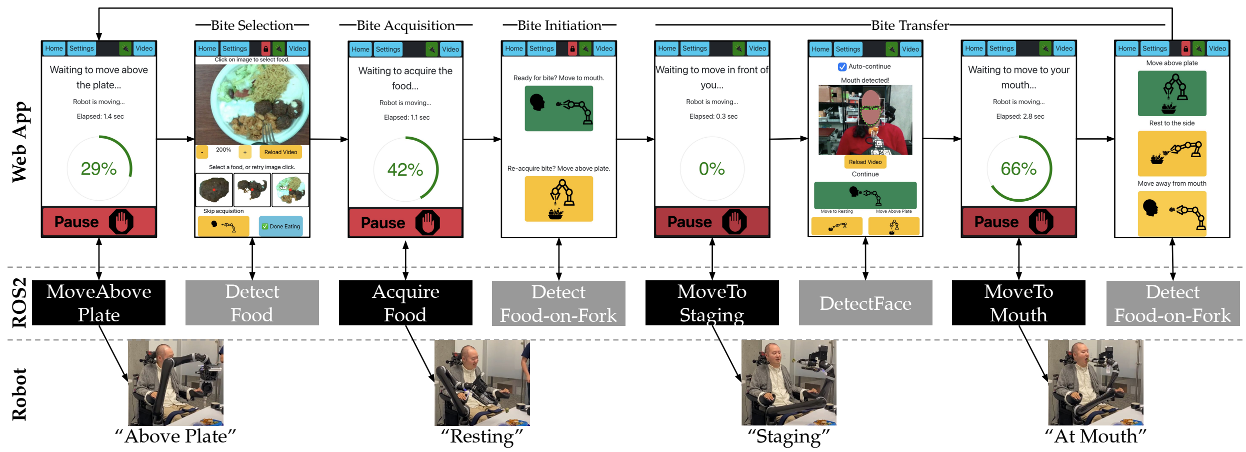


Figure 6.2: An overview of the robot-assisted feeding system's software. The user interacts with the system using the web app (top), which invokes ROS2 interfaces (middle) that move the robot to key configurations (black) and perceive the environment (grey). The default configurations are shown at the bottom.

diverse contexts. Both studies were approved by our institution's ethical review board.[3]

Our *key contribution* is a novel, open-source[4] robot-assisted feeding system (Figure 6.2) that has demonstrated success feeding real meals to real users in real environments for over 15 hours. To our best knowledge, this has not been previously demonstrated by modern research system [276, 236, 218, 144]. A majority of users rated the system as being average-or-above in usability and outperforming caregivers in user independence and control. We attribute the system's success to three key improvements over the state-of-the-art:

1. A unique *web app* that is the seat of system logic and provides substantial customizability and control to users.
2. A novel *bite selection implementation*, with user-in-the-loop input, to accommodate diverse foods.
3. *Portable, flexible hardware* (Figure 6.3), facilitating system use in diverse environments without hindering user mobility.

We also contribute 3 *key lessons learned* from developing and deploying this system: (1) spatial contexts are numerous, customizability lets users adapt to them; (2) off-nominals will arise, variable autonomy lets users overcome them; and (3) assistive robots' benefits depend on context.

[3]UW IRB: STUDY00005607, STUDY00020357
[4]https://robotfeeding.io/publications/hri25a/

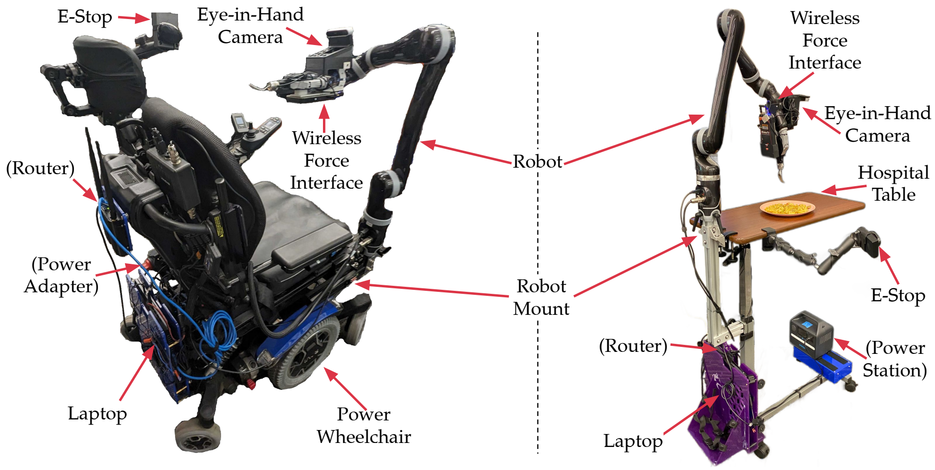


Figure 6.3: The robot feeding system's hardware, mounted on a wheelchair (left) or hospital table (right). Power and compute is self-contained; no wires leave.

6.2 Related Work

History of Robot-assisted Feeding. Enabling people with motor impairments to eat independently has been a long-standing research goal [243, 12, 213]. Research in the 1970s included trained capuchin monkeys [181] and the Morewood Spoon Lifter, where a user shovels food into a pedal-controlled spoon using a head-mounted rod. This robot was sold as the "Winsford Feeder" [213] and was clinically evaluated in labs and homes [243]. In the 1980s, multi-purpose systems emerged that allowed feeding, brushing teeth, and more [296, 267].

Contemporary Robot-assisted Feeding. The last two decades have seen many commercial robotic feeding systems: Bestic [177], Obi [11], Neater Eater [197, 130], and more [12, 213, 278]. These table-mounted robots improved users' mental and physical well-being [177, 164]; however, all but Obi and Neater Eater were discontinued. This may be due to an overreliance on fixed bite acquisition and transfer motions that led to acquisition failures, dropped food, or neck strain [277, 159, 178, 220].

Researchers address these limitations by integrating perception—e.g., cameras and force-torque (F/T) sensors—onto their robotic feeding systems. These robots can be mounted on wheelchairs [93, 30, 226] or tables (fixed [24] or portable [276]), or be mobile manipulators [236, 218].

Research on these systems often focuses on bite acquisition and bite transfer. *Bite acquisition* involves using a fork [105, 125, 285, 284], spoon [255, 221, 136], chopsticks [317, 222], or other tool [277, 236,

Robot	Approx-imate Cost	Mounting	Auto-nomous Motion	(General) Food De-tection[5]	Face Detec-tion	Collision Detection / Avoidance	Portable & Self-contained	User Can Stop / Restart Motion	Custom-izable Robot Motion	Multiple UI Modalities
Obi [11]	$8,625	Table	✓	(✗) ✗	✗	✓/ ✗	✓	✓/ ✓	✓	✓
Neater Eater [197, 130]	$6,500	Table	✓	(✗) ✗	✗	✗/ ✗	✓	✓/ ✓	✓	✓
Song et al. [276, 277]	–	Table	✓	(✗) ✗	✗	✗/ ✗	✓	✓/ ✓	✗	✓
Park et al. [236]	$400,000	Mobile base	✓	(✗) ✓	✓	✓/ ✓	✗	✓/ ✓	✓	✓
Nguyen [218]	$17,950	Mobile base	✗	(✗) ✗	✓	✗/ ✗	✗	✓/ ✓	✗	✓
Bhattacharjee et al. [30]	$50,000	Wheelchair	✓	(✗) ✓	✓	✓/ ✓	✗[6]	✓/ ✗	✗	✓
Jenamani et al. [144, 145]	$50,000	Wheelchair	✓	(✗) ✓	✓	✓/ ✓	✗[6]	✓/ ✗	✗	✗
This paper	$50,000	Wheelchair or Table	✓	(✓) ✓	✓	✓/ ✓	✓	✓/ ✓	✓	✓

Table 6.1: Comparison between this and other robot-assisted feeding systems (Top: commercial, Bottom: research).

108] to grasp food. Utensils often follow motion primitives [125, 29, 93, 284] derivable from human data [107] and chained for complex motions [285, 145]. Online learning can improve primitive selection over time [105, 106]. *Bite transfer* involves handing off an acquired bite to the user's mouth, using trajectories created with heuristic-based planning [24] or learned from demonstrations [46]. Recent work studied in-mouth bite transfer for users without neck mobility [144, 271].

Other research includes: studying the coupling of acquisition and transfer [93], predicting food preference [143], detecting food [125, 89] or the mouth [240], predicting bite timing [226, 125], detecting anomalies [234], and studying natural language interfaces [232]. Simulation environments for robotic caregiving have also been developed [319, 185].

Table 6.1 compares technical capabilities across contemporary robot feeding systems. Ours differs from others in its general food detection capability and from wheelchair-mounted systems in its portability, customizability, and user control.

Out-of-lab Deployments. There is growing interest in out-of-lab deployments of PARs [210]. This includes a robotic guide for blind museum visitors [156], a teleoperated mobile manipulator deployed over weeks that also fed its user [231, 252, 223], and a table-mounted robot-assisted feeding system [277]. These deployments, where users freely use systems in-the-wild, yield valuable insights about task nuances, user

[5]This refers to detecting food masks, irrespective of semantic labels.

[6]These systems have wires connecting the robot's end-effector to external power or compute. This restricts robot motion and poses a trip hazard [30, 144].

preferences, and needed system improvements.

6.3 System

6.3.1 System Development

6.3.1.1 CBPR

We conducted this research with two community researchers (CRs) following CBPR principles [115, 301]. CBPR maintains that community members and academic researchers have unique expertise and experiences, so addressing a community need requires sharing power, resources, and knowledge [239, 198]. CBPR is used in health sciences [138, 301] and increasingly in assistive technology (AT) research [208, 188, 50, 167, 25].

We met CR1 in 2018 through our network. He was passionate about assistive robots. "For a long time, I would only let my mom feed me. I wondered, why am I so uncomfortable with others feeding me that I'll just not eat? I realized that eating is so individualized, with so many intricacies. If I can have a robot do it, I can learn to adapt to it, but it would be *me feeding me*, and that would be huge" (CR1). He participated in pilot studies and more, and in 2021 we began working with him as a CR.[7] The first multi-day deployment was planned in his home, but he passed away months before. His friend, CR2, wanted to honor his legacy by continuing the work. We began working with CR2, culminating in a 5-day deployment in his home. Both CRs have quadriplegia due to a SCI and are paper co-authors.

6.3.1.2 An Analysis of Off-Nominals

Past work surfaced the importance of anomaly monitoring, detection, and correction in robot-assisted feeding [28, 234, 209]. We worked with CR1 to compile a list of such off-nominal scenarios[8] (Table 6.2). Despite the diversity of off-nominals, CR1 observed that users could resolve many of them if provided control (e.g., to retry robot motions, to teleoperate the robot) and customizability (e.g., to adapt the robot to their environment).

[7] Semi-weekly meetings with CRs involved learning about their meal experiences, teaching them technical concepts, and iterating upon the system.

[8] In an off-nominal scenario, something involved in system execution—the user, robot, or environment—does not proceed "according to plan" [90, 196].

User	Robot	Environment
User no longer wants the bite	Robot collides with object	Food falls off of the fork
User gets pulled into a conversation	Robot fails to perceive bite	Unexpected relative configuration of user/robot/plate
User cannot eat (e.g., is coughing)	Robot fails to acquire bite	Local area network fails
User takes a partial bite	Robot fails to perceive face	Device running the web app fails
User clicks an unintended button	Robot stops too far from face	Voice-based assistive technology fails (e.g., due to noise)

Table 6.2: Off-nominal scenarios that can arise during robot-assisted feeding, co-created with CR1.

6.3.1.3 Guiding Design Principles

Informed by formative research in robot-assisted feeding [208, 28, 237], we worked with CR1 to develop the following guiding design principles:

- **Portability**. The system must not hinder the user's or others' mobility. It must be easy to transport and set up.
- **Safety**. The system must not harm the user, other people, or objects in the environment.
- **Reliability**. The system must be able to reliably acquire and transfer the food items a user regularly eats.
- **Customizability**. The user should be able to customize the system to their contexts and preferences.
- **User Control**. The user should have fallback control and enough transparency into the system to utilize it.

6.3.2 System Overview

This section provides a system overview, including how we implement the "portability" and "safety" design principles. All software and custom hardware is available open-source.[4]

6.3.2.1 Hardware

Figure 6.3 shows system hardware.

Robot. A 6 DoF Kinova JACO arm.

Camera. An eye-in-hand system with an Intel Realsense D415 RGBD camera attached to an Nvidia Jetson Nano for wireless image transport. It accesses the robot's internal power through a hole drilled above the last joint. Its position was designed to maintain continuous wrist rotation.

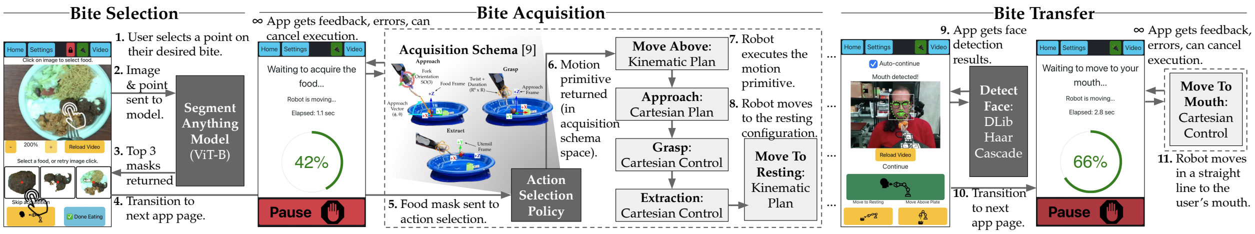


Figure 6.4: A system diagram of bite selection, acquisition, and transfer, showing how the web app communicates with machine learned models (dark grey) and robot motion code (light grey). Components surrounded in a dashed line are represented as a BT. Acquisition schema visualization adapted from [107].

Fork. A custom 3D-printed fork assembly, held in the robot's two-finger gripper. The fork has a 6 DoF ATI Nano25 F/T sensor attached to a battery-powered transmitter that charges with a magnetic connection to the eye-in-hand system.

Compute. A Lenovo Legion 5 laptop (RTX 3060 6GB GPU) that connects to the robot over USB and to a standard accessibility button for emergency stop (e-stop) over 3.5mm aux. The e-stop is mounted in user-accessible location.

Network. A local area network that enables system component communication. Users can either use a home router or the Cradlepoint IBR900 that travels with the system

Mount. A portable mount for the system. Components can be mounted to a wheelchair or hospital table (Figure 6.3), or the robot can be on a tripod with other components in a backpack.

Power. A 24V DC power supply. This can be provided by a power wheelchair's internal battery, a portable power station, or a wall outlet. For the former two, no wires leave the mount.

6.3.2.2 Software

Figure 6.2 shows system software.

Hardware Interface and Controllers. The software stack is built on ROS2 and ros2-control. All controllers are "force-gated," so execution is aborted if measured force or torque exceeds configurable thresholds.[9] All Cartesian control uses a selectively damped pseudo-inverse Jacobian [39, 265].

Planning. We use MoveIt2 for planning, kinematics (`pick_ik` [280]), and collision detection. We use RRT-Connect [163] with shortcutting and hybridization [180] for planning due to its successful prior

[9]Thresholds are 1N when approaching the face, $\leq$50N when acquiring foods, else 4N, far below max force standards for collaborative robots [137].

use [24]. The planning scene has a hull around the user and wheelchair, tight workspace walls, and an Octomap [133] for user-specific obstacles (e.g., ATs). We reject plans whose joint rotations exceed a threshold.

Robot Behaviors. The robot exposes modular behaviors through ROS2 actions and services. For example, “SegmentFromPoint” takes in a user-specified seed pixel in the robot’s image and returns contender masks of that food item. “AcquireFood” computes the food reference frame and executes an acquisition action. `py_trees` represents each motion-related action as a behavior tree (BT) [66], which encapsulates complex robot actions while re-using constituent behaviors.[10]

User Interface. Users interact with the robot via a React web app, accessed from any device with a browser. Thus, they can use *their own ATs* to interact with the system. The web app controls system execution, letting users navigate the state machine by invoking robot actions (Figure 6.2). Unlike prior systems that have the robot control system execution [30, 144], our architecture increases robustness to off-nominals; users can at any time pause, go back, or redirect the system.

Safety Watchdog. A 60Hz watchdog verifies invariants, e.g., the F/T sensor and e-stop are connected and the e-stop has not been clicked. Robot motion stops if an “all clear” watchdog message is not received for 0.5 secs, simplifying verification of safety-critical code by centralizing it.

6.3.3 Robot-Assisted Feeding Procedure

This section focuses on the “reliability” design principle. Figure 6.4 shows key components of the feeding procedure. Figure 6.2 shows quoted robot arm configurations (below).

6.3.3.1 Bite Selection

Users specify their bite preference through a UI that was informed by a pilot study with CR1. At the “above plate” configuration, the robot sends the live RGB view to the web app via WebRTC. The user then selects a pixel on the image. SegmentAnything (Vit-B pre-trained) [160] generates 3 candidate masks, which are rendered on the web app with a dot showing roughly where the robot will skewer. The user can then select a candidate or re-select a pixel.

While waiting for bite selection, the system runs table detection to generalize across table heights. It

[10] We chose BTs over other models (e.g., finite state machines) for their readability and availability of documented open-source software [66, 99].

uses OpenCV's Hough Transform to detect the plate, takes the depth within a 50px ring around the plate, removes outliers, and fits a plane.

6.3.3.2 Bite Acquisition

The user-selected mask is sent to a policy that selects an acquisition action for the arm to execute.

Acquisition Action. This action is based on the specifications and schema defined in [107]. Each action consists of 3 linear Cartesian motions: *approach* (pre-contact), *grasp* (in-food manipulation), and *extraction*. The approach is defined by the initial fork orientation, the approach vector, and a target contact location on the food. The grasp and extraction are defined as a Cartesian twist (angular and linear velocity) and duration. Each motion has end-effector F/T thresholds that, when exceeded, abort the current motion and move to the next. All 3 motions are defined with respect to the food center (+x: the major axis of the bounding ellipse, -z: gravity).

We use 7 actions, based on those learned from human data in [107].[11] We manually adjust the actions to improve stability: we scale twists to a constant angular velocity, remove angular rotations $< 5°$, and define the food frame to be the top instead of bottom of the bite (to align with the perceived depth).

Action Selection Policy. We implement [105]'s online learning system. The policy linearly maps the bite's visual features (last layer of a custom-trained RetinaNet [93]) to the 7 primitives. Map parameters are learned online via LinUCB [171].

Post-Acquisition. The robot moves to a "resting" configuration. The user can initiate bite transfer when ready or have the robot move back "above plate" if acquisition failed.

6.3.3.3 Bite Transfer

The robot moves to a "staging" configuration with a view of the user's face, detects it with the Haar Cascade classifier [300], uses Cartesian control to move the fork to the mouth, and then returns to "staging" and "above plate."

While by default the preceding transitions await user input, each has an optional auto-continue setting.

[11] Specifically, from [107], we use the 3 baseline actions, variants of those three with tilt-back extraction, and the human-informed action #3.

1. **Post-Acquisition**. The robot uses a food-on-fork detector to predict if acquisition succeeded and moves forward to "staging" configuration or back to "above plate."
2. **Moving to Mouth**. The robot auto-continues once face detection perceives a face within the expected distance.
3. **Moving from Mouth**. The robot moves away if the food-on-fork detector (see Appendix) perceives no food.

6.3.4 Implementing the Remaining Design Principles

This section focuses on the remaining two design principles.

6.3.4.1 Customizability

Customizability is useful for ATs [312], PARs [210], and robot feeding [30, 47]. We provide it through:

Arm Configurations. Users have full control of the "above plate," "resting," and "staging" configurations, which are used as waypoints in all robot motions.

Bite Transfer. Users can customize how far from their mouth the robot stops and its speed near their mouth.

Auto-Continue. Users can customize whether the web app waits for their input or uses perception to transition states.

Planning Scenes. User can choose from pre-defined planning scenes: (1) the user and robot are on a wheelchair or (2) the user is in bed, and the robot is on a hospital table.

Customization is done via a web app settings menu. Given prior findings that users and caregivers tinker with assistive robots [220], we design the settings menus using the "Designing for Tinkerability" framework [253]. We provide "fluid experimentation" to users through direct access to the parameter space and "immediate feedback" by allowing them to try out the robot motions that result from their customizations.

6.3.4.2 User Control

Past work showed the importance of *variable autonomy* for PARs [323, 251, 252, 231]. Thus, we provide users multiple levels of control (LoCs), as defined in [21]. When the robot is moving, users have "supervi-

ID	Age	Gender	Impairment	Eating assistance providers	Feeds Self?	Study Device	Device interaction	Study location(s)	Selected meal(s) items[12]
P11	49	M	C3 SCI[13]	Parent(s)	Never	Phone	Voice control	Conference room	Pizza, broccoli
P12	42	F	C5 SCI	FCs, parent(s)	Never	Phone	Stylus	Office	Chicken, salad
P13	45	M	Arthrogryposis	FCs, spouse	Sometimes	Phone	Stylus	Conference room	Sandwich, brownies
P14	62	M	C3 SCI	FCs	Never	Phone	Touch	Office	Chicken, potatoes
P15	61	F	C5-6 SCI	FCs, spouse	Sometimes	Tablet	Touch	Office	Salmon, brussels
CR2	43	M	C2 SCI	FCs	Never	Phone	Mouth joystick	Cafeteria	Stir-fry beef, tofu

Table 6.3: Participant demographics and details for Study 1. FC refers to formal caregivers, i.e., paid and trained professionals.

sory control" to pause it. Doing so drops the LoC to "decision support," where the robot provides multiple options for what to do next. At any time, users can drop the LoC to "teleoperation," where the web app gives them direct Cartesian and joint control.

6.4 Study 1: Multi-user, On-campus Study

Study 1 quantitatively investigates: **How does the system perform across *different users* in out-of-lab settings?** To answer, we invited 5 participants and CR2 (Table 6.3) to eat a meal of their choice in a campus cafeteria, conference room, or office.[14] Following the CR insight that users focus on core AT features at first, we introduced users to all features but customizability and teleoperation.

After obtaining informed consent, we asked participants pre-meal questions while cutting the food, mounted the e-stop, loaded the web app on their device, and walked them through a bite. We explained system features but did not prescribe how participants should behave. Participants then began their meal. One researcher ate with them as a social partner, another took notes, and a third monitored system software to terminate code if necessary. After the meal, we asked post-meal questions. We conducted system patches between studies (see Appendix).

We collected data using evaluation indicators from [28], including *objective metrics* (e.g., meal time profile; acquisition and transfer success rate; system errors) and *subjective metrics* (pre/post ratings of caregiver and robot feeding, the NASA-TLX [122] for cognitive workload, and the System Usability Scale

[12]Meals were bought from restaurants or made by the user's caregivers.
[13]SCI severity is classified by the injured vertebra; C1 is nearest the neck.
[14]Though not as controlled as labs, these are semi-controlled settings, e.g., with standardized lighting, less clutter, etc. Locations were chosen by availability (conference room, office) and user willingness (cafeteria).

(SUS) [175]). As widely used metrics, the TLX and SUS have baselines from meta-analyses; the TLX's is 37 ± 11 [122], and the SUS's is a standardized grade where C is average [175].

6.4.1 Results

Table 6.4 highlights results from Study 1.

6.4.1.1 Bite Duration

Bite duration[15] ranged from 1:00–2:26 minutes.[16] In contrast, people without disabilities take 18–30s [279, 118]. Figure 6.5 shows the time profile for P14's meal. Most time was spent in bite acquisition and moving to his mouth, which have the most interleaved perception, planning, and execution. The largest differences across participants depended on whether they used mouth-based (e.g., voice control, mouth joystick) or touch-based assistive technologies since users could not use the former while talking or chewing.

6.4.1.2 Bite Acquisition and Transfer

Prior work found that 80% acquisition success[17] was sufficient for practical use [30]. For all users, the system neared or exceeded that for the most successful food items, and for all but two users it did so throughout the entire meal. The transfer success rate[18] was 94% for P11 and 100% for the others.

6.4.1.3 Off-nominal Scenarios

Each meal had off-nominal scenarios. Many were user recoverable (e.g., 16% for P11, 88% for P12) via the web app; these included acquisition and transfer failures, robot action errors, mistaken app clicks, and browser interruptions. For some off-nominals, researchers intervened physically (e.g., moving the plate, re-aligning the fork in the gripper) or digitally (e.g., restarting code).

[15] "Bite duration" excludes time between bites (e.g., conversations).
[16] System Patch 1 increased speed by 66%; P11 used the slower robot.
[17] A bite acquisition success is recorded if the food is on the fork at the end of acquisition; else, failure.
[18] A bite transfer success is recorded if the robot stops where the user can eat the bite; else, failure.
[19] Full-screen pop-ups appeared often, hindering P14's web app use.

ID	Meal Time	Bites Eaten	Median Bite Time (IQR)	User-resolved Off-nominals	Researcher Interventions (Physical, Software)	Acquisition Success Rate	Most Successful Food	Cognitive Workload (Baseline: 37 [128])	Usability Grade (Baseline: C [175])
P11	52:37	15	2:26 (0:54)	8	1, 0	0.79 (15/19)	Pizza: 0.78 (14/18)	17.50	D
P12	54:53	24	1:10 (0:12)	2	5, 5	0.65 (24/37)	Chicken: 0.85 (11/13)	29.17	C
P13	54:06	31	1:00 (0:09)	7	6, 1	0.69 (31/45)	Sandwich: 0.94 (16/17)	38.33	F
P14	56:52	30	1:10 (0:21)	22[19]	2, 1	0.88 (30/34)	Chicken 1.0 (13/13)	20.00	A+
P15	51:05	23	1:15 (0:20)	5	0, 1	0.79 (23/29)	Brussels: 0.86 (6/7)	19.17	B+
CR2	28:44	14	1:41 (0:24)	3	1, 1	0.78 (14/18)	Tofu: 1.0 (3/3)	19.17	A

Table 6.4: Study 1: Per-participant time profile (mins:secs), number of interventions, acquisition results, and subjective results.

6.4.1.4 Caregiver Feeding Comparison

Figure 6.6 shows user ratings of caregiver vs. robot feeding. Robot feeding outperforms in users' sense of control (Q1-2) and independence (Q3). 4/5 participants and CR2 agreed that "When I ate with the robot, I was confident that I would remain safe" (Q6).

6.4.1.5 Cognitive Workload

All users but P13 reported experiencing a cognitive workload below the baseline. This indicates that the cognitive workload required to use the system was relatively low despite its many user-in-the-loop components.

6.4.1.6 System Usability

Three of five participants and CR2 rated the system as average-or-above average usability. There was wide variability, from P13's `F` to P14's `A+`. This variability is to be expected given the diversity of users. For example, P13's current self-feeding technique provides many of the robot's functionalities, but he imagined it would help "others who can't use a [self-feeding] system like me."

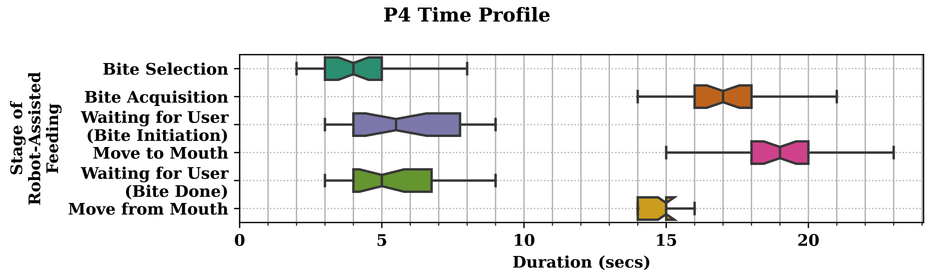


Figure 6.5: How long each stage of feeding took across P14's 30 successful bites[20].

6.5 Study 2: Single-user, In-home Deployment

Study 2 qualitatively investigates: **How does the system perform across the *diverse contexts* that arise when eating in the home?** To answer, we deployed the robot in CR2's home for 5 consecutive days to help him eat 2 meals/day. Pre-deployment, we worked with CR2 and an occupational therapist (OT) to identify CR2's meal-related contexts and goals. Context is "any information that can...characterize the situation of entities...considered relevant to the interaction between a user and an application" [75]. These included:

- **Spatial Context**. CR2 cannot sit up for consecutive days and so alternates between bed and wheelchair days.
- **Social Context**. CR2 has three caregivers, C1-C3 (see Appendix), who typically feed him.
- **Temporal Context**. Mornings are busy with CR2's care routine and daytimes with work, but evenings are relaxed.
- **Activity Context**. CR2's deployment goals were to (1) feed himself dinner while watching television, (2) spend time with a caregiver while both eat dinner, (3) feed himself while a caregiver does other care work, (4) feed himself breakfast while working, and (5) feed himself a mid-day snack while working.
- **Food Context**. CR2 is a flexible eater, enjoying ramen, pizza, chicken teriyaki, fruits/vegetables, and more.

On Mon, Wed, and Fri (wheelchair days), CR2 used the robot for breakfast and dinner. On Tues and Thurs (bed days) he used it for snack and dinner. CR2 selected meal locations and times. Before meals, one researcher set up the robot and acquired test bites, while another cut the food. We then brought the

[20] Box: 25th, 50th, 75th percentile. Whiskers: 1.5·IQR. Outliers excluded.

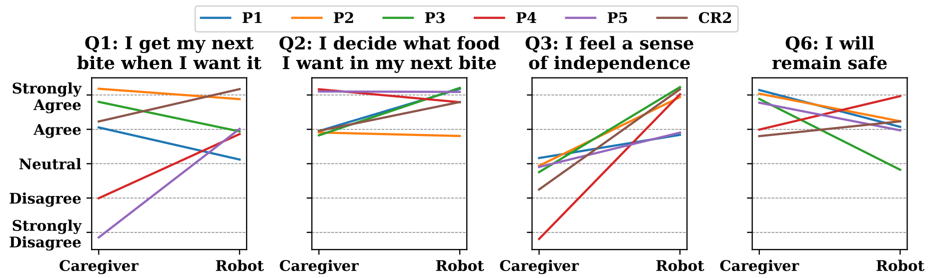


Figure 6.6: Users self-reported comparison: eating with caregivers vs. the robot.

robot to CR2 and positioned the e-stop near him. He then customized the system and began his meal. One researcher monitored software; the other 1–2 took notes. After each meal, we had a semi-structured interview with CR2 and/or his caregivers; all gave informed consent before the study.

6.5.1 Results and Lessons Learned

CR2 used the robot to fed himself all 10 meals, including store-bought foods (e.g., fruits) and caregiver-prepared ones (e.g., avocado toast); he ate various cuisines (e.g., pizza, chicken teriyaki, charcuterie). Using the Medicare Section GG scale [1], the OT assessed that due to system use, CR2's *level of independence during meals increased* from "dependent" (his baseline) to "supervision," where the caregiver is on standby to provide intermittent assistance. We now present qualitative findings grouped into key lessons learned.

6.5.1.1 Spatial Contexts are Numerous, Customizability Lets Users Adapt to Them

The home setting's spatial contexts differed from campus settings. There were *many environmental objects:* CR2 had a mouth joystick near his face and a laptop or phone in front, often on a face-height hospital table. These objects and the e-stop constrained the robot's motion enough that CR2 sometimes said it was "threading a needle." *Spatial configurations varied* between the robot, user, and plate: the bed's tilt, height, and the user's lateral position varied; on a wheelchair, the plate's position, height, and chair's tilt varied. *Lighting conditions varied*: sources of light included windows, lamps, and ceiling lights, and many surfaces were white or reflective, creating backlight, reflections, and shadows.

To enable the robot to work given these varied spatial contexts, CR2 started all meals by checking the previously customized configurations relative to the current meal's context. First, he customized the configurations to account for context, e.g., changing the "above plate" configuration to be above the plate. Second, he customized for preferences, e.g., trying to adjust "staging" to approach him below his eyeline.

Transparency into the downstream impacts of changes was crucial to this process. CR2 iteratively tuned the "resting" configuration and tried motions to/from it until finding one that gave his computer a wide berth. He iteratively tuned "staging" and checked face detection's precision until he found a configuration with reliable face detection in that context.

This process also involved environmental modifications. CR2 sometimes asked a caregiver or us to adjust his laptop or mouth joystick to give the robot arm more room. Once, after realizing that forehead reflections were causing false positive face detections, he had a caregiver place a cap on his head.

Having access to the right level of customization was vital: "This was the sweet spot. I don't want to have to type in code." As we demonstrated the planning scene for CR2 and discussed ways for him to customize it, he said, "Totally automate that. Just the thought of it makes my head hurt."

These observations led to the following lesson learned. *Tinkering is vital for assistive robots to work in users' contexts [220]. Systems should be customizable to foster ease of tinkering. This requires intuitive control over parameters and transparency into the downstream impacts of parameters.*

6.5.1.2 Off-nominals Will Arise, Variable Autonomy Lets Users Overcome Them

Although customizability enabled the user to adapt the robot to spatial contexts, some contexts remained challenging for autonomous behaviors. First, some lighting conditions lowered face detection's precision so much that CR2 did not want the robot to autonomously approach him: "[I don't want it] nearing my eyes." Second, some spatial configurations between the robot and plate hindered bite acquisition. At times, this was due to planning failures: all plans to move above the bite were rejected due to a joint rotation that exceeded thresholds. At other times, this was due to off-centering: a camera enclosure screwhole was damaged in transit, changing the camera's extrinsics.[21] In both cases, variable autonomy let CR2 overcome these off-nominals.

To address face detection failures, if customization did not work CR2 bypassed face detection altogether by teleoperating the robot from "resting" to his mouth. This could be mentally taxing due to many Cartesian

[21] Plate-to-camera distances varied $\geq$ 10cm across meals. So even a 0.05rad extrinsics error could move the fork 5mm off-center, making foods roll away.

and joint motions. Thus, CR2 devised a novel way to customize the "resting" configuration so a single joint-1 rotation moved the fork directly to his mouth, reducing teleoperation to a single button press. Thus, *variable autonomy helped him overcome the off-nominal, but customization let him lower the cognitive workload involved*.

To address bite acquisition failures, CR2 interspersed teleoperation with autonomous robot behaviors. He occasionally teleoperated before autonomous acquisition, changing the robot's starting configuration to overcome planning failures. At other times he used autonomous acquisition to move the fork above the food and then paused it, teleoperating the remainder. Yet other times he teleoperated the entire acquisition. These multiple levels of autonomy helped him avoid frustration: "I would [get frustrated] if it wasn't working, and it just kept on doing it and doing it. I'd be like, 'Oh, stop. Just give me regular control.' But [with this system], it is within my control."

All-in-all, for 6/10 meals the robot successfully autonomously transferred $\geq 80\%$ of bites; for 5/10 meals it successfully autonomously acquired $\geq 80\%$ of bites. Other meals combined autonomy with teleoperation. Despite needing to occasionally teleoperate, CR2 still found the robot empowering to use. "Sometimes people feed me, and I don't like how they're doing it. It's weirdly empowering, as someone who's been paralyzed as long as I have, to say, 'I'm going to eat this. It'll take me 3 times as long, but I'm not going to be frustrated while I eat."'

These observations led to the following lesson learned. *The challenging contexts present in home environments can hinder a robot's autonomous behaviors. Users can help the robot navigate through these challenges provided varied ways to control it. The robot's benefit may outweigh the users' cognitive workload required to control it when autonomy fails.*

The above two lessons validate our key insight about the importance of customizability and control. During system development, we under-appreciated *how often* customizability or control would be needed. However, because they were provided, CR2 leveraged them to resolve scenarios we had not anticipated and to develop novel strategies for system use.

6.5.1.3 Assistive Robots' Benefits Depend on Context

Contexts beyond the spatial affected CR2's meal experiences.

Activity Context. CR2 attempted to perform each of his 5 goals at least once during the deployment. He felt he achieved the first 3. “Eating while watching TV; it’s totally possible. [Using the robot] is not distracting to the point where you can’t do it. I’ve also accomplished eating while my caregiver did something else, because C1 did laundry.” He also ate dinner alongside his caregiver (Figure 6.1). However, CR2 could not achieve his latter 2 goals. “[I couldn’t] eat while working [because of] my over-expectation to not pay attention to the robot. And I had to pay attention to it.” One reason is overlapping demands on his faculties: CR2 types using dictation and so cannot type while chewing; he reads visually and so cannot do it while looking at the robot.[22] CR2’s perspective on these goals shifted over the deployment. “I realized, ‘Food is important. You need to eat more than you need to finish work. And doing that is worth your attention.’”

Social Context. When present, caregivers participated in the meals. “I was involved, but he was doing everything by himself. I was [checking if] the robot dropped something, [giving him] a napkin, refilling the plate. When he chewed, I watched the movie.” (C1). Caregivers were also involved in food preparation. After seeing the robot acquire her carrots but not her zucchini, C3 said, “[Today] it was [too] soft or small. That could be improved if I prepared [meals] several times with the robot." Importantly, CR2’s envisioned future robot use was conditioned on caregiver effort: “[If it takes too long], they will say, ‘Let me just feed you and not set up [and tinker with] the robot.’ And that would be reasonable.”

Researchers were also part of the social context. “If a plate were there the whole day—if you guys weren’t here—I would’ve gotten the work done. I would’ve taken one bite, waited 30 minutes, finished a task, and taken another bite. I would do it guilt-free [because no one has to wait for me].”

Food Context. Caregivers had concerns about the robot’s acquisition limits. “We need to choose foods that CR2 likes and the robot can pick up.” (C3). “[The robot currently] has too many limits with his diet and what he likes” (C1). CR2 envisioned food-dependent robot use. “I wouldn’t eat all my meals with it. Some foods I like [e.g., ramen] can be difficult for it. [But] I like pizza a lot; it did fine with pizza.”

Temporal Context. CR2 enjoyed the robot more during dinner. “[When I’m] eating for enjoyment, during dinner, [using the robot] is great. For breakfast and snack, where I feel I should be working, things are rushed.” His caregivers agreed: “Sometimes, CR2’s in rush. So we don’t have time to set [the robot] up. So we have to feed him.” (C2). Despite contextual differences, CR2 found the robot rewarding to use.

[22] In contrast, CR2 *can* look at the robot while *listening to* a TV show.

"That Wednesday morning, there was a flow state. I was succeeding at such a rate that it felt good. I was like, 'We're getting into it, no matter how long it takes.' At that point, my satisfaction levels are really high."

These observations led to the following lesson learned. *Assistive robots integrate into a user's life. They provide benefits in some contexts but not others. Such contextual benefits may still be sufficient to make them a valuable addition to the tools users and caregivers adopt for ADLs.*

6.6 Limitations and Future Work

In this work, we collaborated with community researchers (CRs) to develop a robot-assisted feeding system that people with motor impairments can use to independently feed themselves outside of lab settings. We evaluated the system quantitatively with 5 users and CR2 in 3 locations (Sec. 6.4) and qualitatively in CR2's home for 5 days (Sec. 6.5). Although this work made progress towards our goal (Sec. 6.1), results reveal system limitations to address to fully reach the goal.

For **bite acquisition**, a limitation was missed bites. An important future direction is incorporating closed-loop feedback into action primitives, e.g., adjusting the motion if the bite starts tilting or the fork fails to pierce it. Another important direction is expanding the food types the robot can acquire to include e.g., ramen. For **bite transfer**, a limitation was that no matter how much customization CR2 tried, the robot approached at his eyeline; future work should orchestrate transfer motions to approach from below. For **customizability**, users must be able to customize the planning scene; in our system, users had to choose among hard-coded scenes. For **user control**, the system must provide ways for users and caregivers to debug system problems, such as transparently explaining what error the robot encountered and how they could resolve it. For **user comfort**, approaches like compliant control from physical human-robot-interaction could improve the system [258, 73, 144]. For **commercial viability**, future work should focus on reducing system cost. Finally, co-designing setup and maintenance procedures with caregivers could improve **integration into care routines**.

This in-home deployment is only the beginning. An open-source, deployable system lays the foundation for: (1) follow-up deployments with CR2 as the system matures and (2) in-home deployments with other users to study additional meal contexts and address potential CR biases. Long-term, researchers should not be present since our presence influences system use (Sec. 6.5.1.3).

6.7 Supplementary Materials

Supplementary materials, hosted on Open Science Foundation at [211], include per-meal event annotations, user quotes, and appendices with system and study details.

Chapter 7

Conclusion and Future Work

This thesis presented research focused on developing and deploying a robot-assisted feeding (RAF) system for people with motor impairments. More specifically, it focused on the following research questions:

Figure 7.1: List of research questions that constitute this thesis.

RQ-Thesis	How can we develop a deployable robot-assisted feeding system that enables any user, in any environment, to feed themselves a meal of their choice, while aligning with their preferences?
RQ0	What applications, methods, and themes underlie the last decade of research on physically assistive robots for people with disabilities?
RQ1	What challenges do people with motor impairments face during dining, and how should a robot-assisted feeding system be designed to address those challenges?
RQ2	How can a robot-assisted feeding system feed users the variety of food items they want to eat?
RQ3	How can we develop a robot-assisted feeding system to feed users in diverse out-of-lab and in-home contexts?

RQ0's findings, presented in Chapter 2 and originally published in [210], motivated the overarching thesis. Specifically, we found that physically assistive robots (PARs) are primarily evaluated inside the lab, which motivated the thesis's overall focus on achieving out-of-lab deployments of a PAR. That work also found that the themes of "levels of autonomy" and "adaptation" are relevant to PAR research across different application domains, which motivated our approach to system design and engineering in RQ3.

Chapter 3 then argued for why RAF systems are a good case study for achieving in-home deployments of PARs. Specifically, it presented the nearly 50 year history of RAF research, the multiple commercial endeavors that have emerged from that research, and the key contributions and gaps of state-of-the-art contemporary RAF systems. RAF is a well-suited case study due to this established user need, long history of technical research, and interest from both academia and the industry.

RQ1, presented in Chapter 4 and originally published in [208], delved into users' current meal experiences and their needs and priorities when it comes to the design of RAF systems. Key insights from this work were that users desire control over their RAF system and want a highly customizable robot, which formed the design principles that guided system development in Chapter 6.

RQ2, presented in Chapter 5 and originally published in [104], focused on enabling a robot-assisted feeding system to acquire the large variety of foods users may want to eat. This work presented a structured schema to represent bite acquisition actions, and used unsupervised learning to extract key actions within that schema that people without disabilities use to acquire food. It then demonstrated that this set of learnt actions provides better coverage and learnability than a previously state-of-the-art set of handcrafted actions.

Finally, RQ3, presented in Chapter 6 and originally published in [212], delved into the system design and development necessary to create a RAF system that can be deployed outside the lab. It presented results from two studies: (1) a quantitative study where 6 people with motor impairments used the robot to feed themselves meals of their choice in a cafeteria, conference room, or office; and (2) a qualitative study where 1 community researcher (CR) used the robot to feed himself 10 meals over 5 consecutive days in his home, across diverse contexts such as eating in bed, watching TV while eating, eating while working, and more.

7.1 Lessons Learned

This thesis focused on robot-assisted feeding as a case study for achieving in-home deployments of physically assistive robots (PARs). What have we learned, over the course of this work, to help future researchers deploy physically assistive robots (PARs)?

The first three lessons were presented in Chapter 6. We encourage readers to revisit the respective sections to delve into those lessons.

1. *spatial contexts are numerous, customizability lets users adapt to them* (Sec. 6.5.1.1);
2. *off-nominals will arise, variable autonomy lets users overcome them* (Sec. 6.5.1.2); and
3. *assistive robots' benefits depend on context* (Sec. 6.5.1.3).

The final, and perhaps most important, lesson is to *work with end-users and stakeholders*.

- **Community Researchers**. Tyler (CR1) and Jonathan's (CR2) insights were foundational to the system design, study design, and more. Examples of this are described in-detail in Sec. 1.1 and Sec. 4.9.
- **Caregivers**. Jonathan's caregivers provided invaluable feedback on how the robot should acquire and feed food, and ways in which it can integrate into care routines. In this work, we only involved caregivers during the evaluation, but not the design, of the system. Future works should involve caregivers *throughout* the research process, as was done in Ranganeni et al. [252].
- **Occupational Therapists**. Vy Nguyen, the occupational therapist we worked with in Chapter 6, provided deep expertise in how clinical outcomes of assistive technologies can be measured and quantified. She also provided creative insights into how robots can better integrate into care routines, including through environment and tool modifications. For example, she suggested adding a removable plastic cover to Jonathan's mouth joystick, so food flavor does not linger even after he's done eating. In this work, we only involved occupational therapy expertise during the evaluation, but not the design, of the system. Future works should involve occupational therapists *throughout* the research process, as was done in Olatunji et al. [223]. Ranganeni [250] presents results from a study investigating occupational therapists' perspectives on assistive robots, which should serve as a foundation for future collaborations between roboticists and occupational therapists.

We hope these four lessons will accelerate progress in this field by making it easier for future researchers to develop and deploy physically assistive robots (PARs).

7.2 Limitations and Future Work

The work presented in this thesis has limitations that create opportunities for future work.

Bite Acquisition

Although the bite acquisition system did acquire many food items with near or above 80% accuracy[1], Chapter 6 revealed multiple limitations in bite acquisition that should be improved upon.

The bite acquisition system is very vulnerable to robot calibration errors. Small angular errors in the

[1] 80% is the geometric mean of users' self-declared tolerance for bite acquisition errors from [30].

camera's extrinsics calibration, or in the robot's internal model of its own form factor, can be enough to off-center the fork my multiple millimeters, which can be enough to misacquire the food. This is particularly problematic when one considers that those small miscalibrations are more likely to occur during deployments: the camera is more likely to get jostled when the robot is being moved; and the robot's wrist is more likely to slightly droop over an extended period of time holding the fork and the force-torque (F/T) transducer. This motivates two system improvements for future work. First, the system should have an auto-calibration procedure, so that both the camera's extrinsics and the robot's internal model can be kept up-to-date. One way to do this could involve placing fiducial markers [151] in known locations on the robot and its fork, and then having the robot move to known arm configurations to auto-calibrate[2]. Second, the robot should use visual servoing during bite acquisition, tracking the food and the fork to ensure contact is made at the desired position. That way, even if there are small errors in the calibration, the robot can overcome them by using its live sensor readings to constantly update its estimate of the food pose relative to the fork.

Expanding upon the theme of robustness to errors, the bite acquisition action schema presented in Chapter 5 is very prescriptive about where the fork should be before acquisition; specifically, it specifies a single end-effector pose for the fork to be in. As a result, if the planning system is unable to compute a plan to that fork pose in the allowed time, bite acquisition fails without the robot even moving. However, based on qualitative observations of the participants from Chapter 5 acquiring food, it is clear that motions exist to acquire the same bite of food from a variety of different starting fork poses. Thus, instead of only supplying a single start fork pose, the acquisition action schema should be expanded to specify a range of start poses, for example with a task-space region [26], to improve the planning success rate. This will likely necessitate other extensions of the schema as well, to enable the approach, grasp, and extract phases of an acquisition to still work effectively given a range of start fork poses.

Another important improvement to the bite acquisition system is allowing real-time error recovery during execution of a motion primitive. Currently, once the online learning system selects a motion primitive, the robot executes that motion primitive till the end. However, errors sometimes arise during the execution. In addition to the aforementioned off-centering error, other errors we saw included the robot tilting and/or

[2]This is the procedure used by Hello Robot's Stretch 3: `https://docs.hello-robot.com/0.2/stretch-ros/stretch_calibration/`

flipping the food, and the robot applying force but not actually piercing the food deep enough. In all these cases, if the error gets detected early, it can be resolved, for example by re-centering the fork on the food, by lessening the fork's downward force to let the bite un-tilt, or by applying more downwards force to fully pierce the food, respectively. Further, it should be possible for the robot to detect these errors early, because users in both studies often detected these errors while they were still recoverable, and then got frustrated when the robot continued executing the primitive without adjusting to resolve the error. Further, people without disabilities often do this real-time error detection and resolution during bite acquisition, for example by adjusting the angle and downwards pressure of the fork tines if they feel a grape start to roll away. Thus, an important direction for future work is expanding the bite acquisition system and the action schema to incorporate real-time error detection and recovery. This will result in higher bite acquisition success rate and lower bite duration (since the robot no longer has to finish its original bite acquisition attempt before trying again).

Users found it jarring when the robot arm made large motions, particularly swivels, as part of bite acquisition; this made the robot's motion less legible [77] and could make users more uncomfortable around the robot. Thus, future work should focus on reducing the joint-space length of trajectories the robot takes during bite acquisition. One approach to this is further tuning the planning algorithms to achieve shorter trajectories. Another approach is adding joints on the fork, as Jenamani et al. [145] did with their articulated fork, so that moving the fork tip does not require large motions of the joints on the robot arm.

Once CR2 became familiar with the different bite acquisition motion primitives, he would sometimes request we have the robot only use one particular motion primitive, based on his knowledge of what motion would work best for that food item. This points towards future work that allows users to customize toggle on/off online learning, and to select which action primitive to use if online learning is turned off. This also indicates the need for future work that speeds up the rate of online learning, since CR2 sometimes felt that a manual decision of what action to take would work better than running the online learning system.

Finally, expanding the robot's bite acquisition capabilities to additional types of food, additional arrangements of food, and multi-step bite acquisitions are all exciting directions for future work. One food CR2 really wanted to eat was ramen, which the system cannot currently feed due to its inability to use a spoon and its inability to perceive noodles that are covered by a liquid. CR2 also observed that the robot currently requires pre-cut bite sized pieces of food that are visually separated on a plate, which does not correspond to many realistic arrangements of food. P14 asked if the robot can dip his potatoes in ketchup before feeding him, revealing the importance of multi-step bite acquisitions. Jenamani et al. [145]'s work represents a promising improvement in the latter two directions.

Bite Transfer

For bite transfer, a big limitations was that face detection did not generalize well to different lighting conditions and different relative spatial configurations between the camera and user's face. Thus, an important direction for future work is making face detection robust to the types of images of the user's face that the robot will realistically see during in-home feeding. Jenamani et al. [144] presents one such approach—a technique to make face detection robust to occlusions that could arise from the fork or food—although it requires adding a second camera to the robot arm.

As was mentioned in Chapter 6, CR2 really wanted the robot to move to his mouth from below his eye level. During several meals, he tried to customize the "staging configuration" to achieve that goal. However, he was unable to, partly because the camera has to see his entire face from the "staging configuration," and partly because of the length of the robot arm's links that have to move as the robot approaches the user's mouth. Regardless, prior work should provide users a more granular and intuitive way to choreograph bite transfer motion, to allow them to specify certain goals like not crossing their eye-line.

The robot arm often had to traverse multiple dozen centimeters of distance to move from the staging configuration to the user's face. Not only does this increase bite duration, but it is also increases the time the food has to fall or drip onto the user. In contrast, when caregivers feed people with motor impairments, they often keep the plate or bowl right under the user's mouth [28], reducing the transfer distance to a few centimeters. Future work should experiment with the placement of the plate, possibly in tandem with the form factor of the robot arm, to reduce the distance the robot has to traverse to transfer bites to the user's mouth.

Although in theory users could toggle on auto-continue to use food-on-fork detection to automatically move the fork away from their mouth, in practice during some of CR2's meals it was working unreliably. Like with bite acquisition, the culprit was that the food-on-fork detection algorithm (Appendix 2.1.4) was very vulnerable to errors in robot calibration (camera extrinsics calibration error, or the fact that the fork tines sometimes got bent over the course of a meal). Prior work should develop food-on-fork detection approaches that are more robust to calibration errors. One such approach, which uses vision-language models (VLMs) to detect whether there is food on the fork, was presented in Jenamani et al. [145]'s open-sourced code.

Customization

A limitation of the current system is that it provides users limited control over customizing the planning scene. However, as was mentioned in Chapter 6, every one of CR2's meals involved a slightly different spatial configuration and obstacle configuration. Although CR2 stated that the robot should be able to autonomously detect and adapt to the planning scene, P13 expressed interest in being able to manually modify the planning scene, for example to ensure the robot does not move towards people sitting next to him. Thus, an exciting direction for future work is developing tractable, functional, and usable approaches for inferring and customizing the robot's planning scene.

The system's bite selection currently requires users to specify specific bites. However, multiple users expressed an interest in being able to specify a semantic category for a bite instead, such as "pizza," and having the system pick up any bite of pizza. Providing users with that option, as was done in Jenamani et al. [145], and studying in which scenario they use which option, is an exciting direction for future work.

User Control

A crucial limitation of the system is that debugging and resolving some issues still requires researchers in-the-loop. First, the user is not given any control to restart the system; when the e-stop is pressed, researchers must manually re-start the system. Secondly, although the web app tells users when the robot encounters an error during an action, it does not give the user details about the error or guidance on how to resolve the errors. Sometimes during the in-home deployment, this information had to be verbally communicated from a researcher to CR2, to account for this system limitation. Thus, an exciting direction for future work is investigating ways to give users intuitive transparency into system errors and the control and guidance necessary to resolve the errors, including by restarting system components.

Integration into Care Routines

A limitation of the current system is that setting it up for feeding and maintaining it (e.g., unscrewing the fork to clean it) takes technical expertise that caregivers may not have. An exciting direction of future work is co-designing the setup and maintenance of the robot-assisted feeding system with caregivers and end-users, because easy setup and maintainence is crucial to long-term adoption of assistive technologies.

Eating food is not an isolated task, and is often also interspersed with feeding-adjacent activities such as drinking fluids and wiping one's face with a napkin. In addition, other activities need to be completed before and after the meal, e.g., preparing and warming the food, cleaning dishes, brushing teeth, etc. Thus, an exciting opportunity for future work is either extending a RAF system to provide those additional functionalities, or ensuring it integrates with other ways the user completes those activities (e.g., it can work alongside other assistive technologies, or alongside the caregiver who is completing feeding-adjacent activities). Note that multiple feeding-adjacent activities recommended by users are documented in Chapter 4's Fig 4.7.

Physically Assistive Robots (PARs)

For both RAF systems specifically and PARs generally, an important and exciting direction for future work is conducting more in-home robot deployments. Every users' contexts, needs, preferences, and perspectives vary, so every new deployment will yield valuable information about how best to develop PARs to address the diverse needs of their target users. In the long term, it would be particularly interesting and valuable to conduct PAR deployments where researchers are not present during robot usage, as is sometimes done in socially assistive robot (SAR) deployments [42, 65].

As was revealed in Chapter 2, multiple ADLs have been insufficiently studies amongst PAR researchers, including dressing and bathing/grooming. Further, some tasks that do not cleanly fall under an ADL would still provide high-value to users' and caregivers. For example, CR2 repositioning a care recipient's body in bed as a task that needs to be done multiple times every day and night, can be physically difficult for caregivers, and can have serious health repercussions (pressure sores) for care recipients if not done right. As another example, CR1 mentioned the task of transferring himself between his wheelchair and bed as one that he would like to be able to do independently, with the assistance of a robot. These tasks have a few works dedicated to them (e.g., Madan et al. [186] for bathing and Cheng et al. [56] for bed-transfer), but could benefit from the concerted, decades-long research focus that an activity like robot-assisted feeding has had. The fact that these tasks have not been investigated much means that various technical challenges will reveal themselves as researchers try to develop PARs for these tasks, making them particularly exciting directions for future work.

Finally, research into PARs is only the beginning; for users to benefit, it is crucial to be able to sustainably develop and distribute PARs and to have users and caregivers adopt them over the long-term. This requires concerted collaborations across multiple stakeholders: researchers, businesses, insurance and regulation agencies, occupational therapists, and more. This also requires reducing the cost of PARs. Thus, an exciting direction for future work is studying, understanding, and developing sustainable pathways for PARs to reach end-users, to scale the positive impact of PARs research to the millions of people who can benefit from it.

Bibliography

[1] Coding Section GG Self-Care & Mobility Activities Included on the Post-Acute Care Item Sets: Key Questions to Consider When Coding. *U.S. Center for Medicare and Medicaid Services*.

[2] Disability and the digital divide: Internet subscriptions, internet use and employment outcomes. *U.S. Department of Labor*, Jun 2022. URL `https://www.dol.gov/sites/dolgov/files/ODEP/pdf/disability-digital-divide-brief.pdf`.

[3] Shivendra Agrawal, Mary Etta West, and Bradley Hayes. A novel perceptive robotic cane with haptic navigation for enabling vision-independent participation in the social dynamics of seat choice. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 9156–9163. IEEE, 2022.

[4] Baris Akgun, Maya Cakmak, Karl Jiang, and Andrea L Thomaz. Keyframe-based learning from demonstration: Method and evaluation. *International Journal of Social Robotics*, 4:343–355, 2012.

[5] Reem K Al-Halimi and Medhat Moussa. Performing complex tasks by users with upper-extremity disabilities using a 6-dof robotic arm: a study. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 25(6):686–693, 2016.

[6] Patrícia Alves-Oliveira, Maria Luce Lupetti, Michal Luria, Diana Löffler, Mafalda Gamboa, Lea Albaugh, Waki Kamino, Anastasia K. Ostrowski, David Puljiz, Pedro Reynolds-Cuéllar, et al. Collection of metaphors for human-robot interaction. In *Designing Interactive Systems Conference 2021*, pages 1366–1379, 2021.

[7] Patrícia Alves-Oliveira, Matthew Bavier, Samrudha Malandkar, Ryan Eldridge, Julie Sayigh, Elin A

Björling, and Maya Cakmak. Flexi: A robust and flexible social robot embodiment kit. In *Designing Interactive Systems Conference*, pages 1177–1191, 2022.

[8] Patrícia Alves-Oliveira, Amal Nanavati, Tyler Schrenk, Ethan Gordon, Maya Cakmak, and Siddhartha S Srinivasa. Design principles for robot-assisted feeding in social contexts (supplementary materials), Dec 2022. URL `osf.io/392hp`.

[9] Paola Ardón, Maria E Cabrera, Eric Pairet, Ronald PA Petrick, Subramanian Ramamoorthy, Katrin S Lohan, and Maya Cakmak. Affordance-aware handovers with human arm mobility constraints. *IEEE Robotics and Automation Letters*, 6(2):3136–3143, 2021.

[10] Stephanie Arevalo Arboleda, Max Pascher, Annalies Baumeister, Barbara Klein, and Jens Gerken. Reflecting upon participatory design in human-robot collaboration for people with motor disabilities: Challenges and lessons learned from three multiyear projects. In *PErvasive Technologies Related to Assistive Environments Conference*, pages 147–155. ACM, 2021.

[11] Dar Artman, Mary Parkinson, and Pat Adanti-Joy. New obi robotic dining device a breakthrough for people living with physical challenges, Jul 2016. URL `https://www.multivu.com/players/English/7850651-obi-robotic-dining-device/`.

[12] Assistive. Assistive dining technology 1970 through 2005. *Mealtime Partners, Inc.*, page 1–37, 2005. URL `https://robotfeeding.io/publications/archive/mealtime2005assistive.pdf`.

[13] Mirjam Augstein and Thomas Neumayr. A human-centered taxonomy of interaction modalities and devices. *Interacting with Computers*, 31(1):27–58, 2019.

[14] Shiri Azenkot, Catherine Feng, and Maya Cakmak. Enabling building service robots to guide blind people a participatory design approach. In *ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 3–10. ACM, 2016.

[15] Neil M Bajaj, Adam J Spiers, and Aaron M Dollar. State of the art in prosthetic wrists: Commercial and research devices. In *IEEE International Conference on Rehabilitation Robotics (ICORR)*, pages 331–338. IEEE, 2015.

[16] Markus Bajones, David Fischinger, Astrid Weiss, Paloma De La Puente, Daniel Wolf, Markus Vincze, Tobias Körtner, Markus Weninger, Konstantinos Papoutsakis, Damien Michel, et al. Results of field trials with a mobile service robot for older adults in 16 private households. *ACM Transactions on Human-Robot Interaction (THRI)*, 9(2):1–27, 2019.

[17] Tommaso Lisini Baldi, Giovanni Spagnoletti, Mihai Dragusanu, and Domenico Prattichizzo. Design of a wearable interface for lightweight robotic arm for people with mobility impairments. In *International Conference on Rehabilitation Robotics (ICORR)*, pages 1567–1573. IEEE, 2017.

[18] Kim Baraka and Manuela M Veloso. Mobile service robot state revealing through expressive lights: formalism, design, and evaluation. *International Journal of Social Robotics*, 10:65–92, 2018.

[19] Jan Bedrosian. Limitations in the use of nondisabled subjects in aac research. *Augmentative and Alternative Communication*, 11(1):6–10, 1995.

[20] Jenay M Beer, Cory-Ann Smarr, Tiffany L Chen, Akanksha Prakash, Tracy L Mitzner, Charles C Kemp, and Wendy A Rogers. The domesticated robot: design guidelines for assisting older adults to age in place. In *ACM/IEEE international conference on Human-Robot Interaction (HRI)*, pages 335–342. ACM, 2012.

[21] Jenay M. Beer, Arthur D. Fisk, and Wendy A. Rogers. Toward a framework for levels of robot autonomy in human-robot interaction. *J. Hum.-Robot Interact.*, 3(2):74–99, jul 2014. doi: 10.5898/JHRI.3.2.Beer. URL `https://doi-org.offcampus.lib.washington.edu/10.5898/JHRI.3.2.Beer`.

[22] Beeson. Beeson automaddak feeder, Feb 2004. URL `https://larimer.co.networkofcare.org/aging/assistive/detail.aspx?id=18902&cid=797&cn=Feeders&org=`.

[23] S. Belkhale, E.K. Gordon, Y. Chen, S. S. Srinivasa, T. Bhattacharjee, and D. Sadigh. Balancing efficiency and comfort in robot-assisted bite transfer. In *IEEE International Conference on Robotics and Automation*, 2022.

[24] Suneel Belkhale, Ethan K Gordon, Yuxiao Chen, Siddhartha Srinivasa, Tapomayukh Bhattacharjee, and Dorsa Sadigh. Balancing efficiency and comfort in robot-assisted bite transfer. In *2022 International Conference on Robotics and Automation (ICRA)*, pages 4757–4763. IEEE, 2022.

[25] Jeanette Bell and Tuck Wah Leong. Collaborative futures: Co-designing research methods for younger people living with dementia. In *Proceedings of the 2019 chi conference on human factors in computing systems*, pages 1–13, 2019.

[26] Dmitry Berenson, Siddhartha Srinivasa, and James Kuffner. Task space regions: A framework for pose-constrained manipulation planning. *The International Journal of Robotics Research*, 30(12): 1435–1460, 2011.

[27] Bestic. Bestic, May 2021. URL `https://web.archive.org/web/20220120094337/https://www.camanio.com/us/products/bestic/`.

[28] Tapomayukh Bhattacharjee, Maria E Cabrera, Anat Caspi, Maya Cakmak, and Siddhartha S Srinivasa. A community-centered design framework for robot-assisted feeding systems. In *The 21st international ACM SIGACCESS conference on computers and accessibility*, pages 482–494. ACM, 2019.

[29] Tapomayukh Bhattacharjee, Gilwoo Lee, Hanjun Song, and Siddhartha S Srinivasa. Towards Robotic Feeding: Role of Haptics in Fork-based Food Manipulation. *IEEE Robotics and Automation Letters*, 2019.

[30] Tapomayukh Bhattacharjee, Ethan K Gordon, Rosario Scalise, Maria E Cabrera, Anat Caspi, Maya Cakmak, and Siddhartha S Srinivasa. Is more autonomy always better? exploring preferences of users with mobility impairments in robot-assisted feeding. In *2020 15th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 181–190. IEEE, ACM, 2020.

[31] Tapomayukh Bhattacharjee, Ethan K Gordon, Rosario Scalise, Maria E Cabrera, Anat Caspi, Maya Cakmak, and Siddhartha S Srinivasa. Is more autonomy always better? exploring preferences of users with mobility impairments in robot-assisted feeding. In *2020 15th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 181–190. IEEE, 2020.

[32] Alberto Bietti, Alekh Agarwal, and John Langford. A Contextual Bandit Bake-off. *arXiv preprint 1802.04064*, February 2018.

[33] Eli Blevis, Sabrina Hauser, and William Odom. Sharing the hidden treasure in pictorials. *interactions*, 22(3):32–43, 2015.

[34] Jeannette Bohg, Antonio Morales, Tamim Asfour, and Danica Kragic. Data-driven grasp synthesis—a survey. *IEEE Transactions on robotics*, 30(2):289–309, 2013.

[35] Mayara Bonani, Raquel Oliveira, Filipa Correia, André Rodrigues, Tiago Guerreiro, and Ana Paiva. What my eyes can't see, a robot can show me: Exploring the collaboration between blind people and robots. In *International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS)*, pages 15–27. ACM, 2018.

[36] Steven W Brose, Douglas J Weber, Ben A Salatin, Garret G Grindle, Hongwu Wang, Juan J Vazquez, and Rory A Cooper. The role of assistive robotics in the lives of persons with disability. *American Journal of Physical Medicine & Rehabilitation*, 89(6):509–521, 2010.

[37] Emeline Brulé, Brianna J Tomlinson, Oussama Metatla, Christophe Jouffrais, and Marcos Serrano. Review of quantitative empirical evaluations of technology for people with visual impairments. In *Conference on Human Factors in Computing Systems (CHI)*, pages 1–14. ACM, 2020.

[38] Sheryl Burgstahler, Tanis Doe, et al. Disability-related simulations: If, when, and how to use them in professional development. *Review of Disability Studies: An International Journal*, 1(2), 2004.

[39] Samuel R Buss and Jin-Su Kim. Selectively damped least squares for inverse kinematics. *Journal of Graphics tools*, 10(3):37–49, 2005.

[40] Maria E Cabrera, Tapomayukh Bhattacharjee, Kavi Dey, and Maya Cakmak. An exploration of accessible remote tele-operation for assistive mobile manipulators in the home. In *IEEE International Conference on Robot & Human Interactive Communication (RO-MAN)*, pages 1202–1209. IEEE, 2021.

[41] Maria E Cabrera, Kavi Dey, Kavita Krishnaswamy, Tapomayukh Bhattacharjee, and Maya Cakmak. Cursor-based robot tele-manipulation through 2d-to-se2 interfaces. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 4230–4237. IEEE, 2021.

[42] Bengisu Cagiltay, Nathan Thomas White, Rabia Ibtasar, Bilge Mutlu, and Joseph Michaelis. Understanding factors that shape children's long term engagement with an in-home learning companion robot. In *Proceedings of the 21st Annual ACM Interaction Design and Children Conference*, pages 362–373, 2022.

[43] Maya Cakmak. What if your robot designed your next home? In *Human Computer Interaction Consortium (HCIC)*. HCIC, 2017.

[44] Praminda Caleb-Solly, Chris Harper, and Sanja Dogramadzi. Standards and regulations for physically assistive robots. In *2021 IEEE international conference on intelligence and safety for robotics (ISR)*, pages 259–263. IEEE, 2021.

[45] Alexandre Campeau-Lecours, Véronique Maheu, Sébastien Lepage, Hugo Lamontagne, Simon Latour, Laurie Paquet, and Neil Hardie. Jaco assistive robotic device: Empowering people with disabilities through innovative algorithms. *Rehabilitation Engineering and Assistive Technology Society of North America (RESNA)*, 2016.

[46] Gerard Canal, Guillem Alenyà, and Carme Torras. Personalization framework for adaptive robotic feeding assistance. In *International conference on social robotics*, pages 22–31. Springer, 2016.

[47] Gerard Canal, Carme Torras, and Guillem Alenyà. Are preferences useful for better assistance? a physically assistive robotics user study. *ACM Transactions on Human-Robot Interaction (THRI)*, 10 (4):1–19, 2021.

[48] Georgia Chalvatzaki, Xanthi S Papageorgiou, Costas S Tzafestas, and Petros Maragos. Augmented human state estimation using interacting multiple model particle filters with probabilistic data association. *IEEE Robotics and Automation Letters*, 3(3):1872–1879, 2018.

[49] Pyung Hun Chang, Sang Rae Park, Gun Rae Cho, Je Hyung Jung, and Sang Hyun Park. Development

of a robot arm assisting people with disabilities at working place using task-oriented design. In *International Conference on Rehabilitation Robotics (ICORR)*, pages 482–487. IEEE, 2005.

[50] Dorota Chapko, Pino Frumiento, Nalini Edwards, Lizzie Emeh, Donald Kennedy, David McNicholas, Michaela Overton, Mark Snead, Robyn Steward, Jenny M Sutton, et al. " we have been magnified for years-now you are under the microscope!": Co-researchers with learning disabilities created an online survey to challenge public understanding of learning disabilities. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–17, 2020.

[51] Dorota Chapko, Pedro Andrés Andrés Pérez Rothstein, Lizzie Emeh, Pino Frumiento, Donald Kennedy, David McNicholas, Ifeoma Orjiekwe, Michaela Overton, Mark Snead, Robyn Steward, et al. Supporting remote survey data analysis by co-researchers with learning disabilities through inclusive and creative practices and data science approaches. In *Designing Interactive Systems Conference 2021*, pages 1668–1681, 2021.

[52] Kevin Charras and Michèle Frémontier. Sharing meals with institutionalized people with dementia: a natural experiment. *Journal of Gerontological Social Work*, 53(5):436–448, 2010.

[53] Sheryl Chau, Sanders Aspelund, Ranjan Mukherjee, Mei-Hua Lee, Rajiv Ranganathan, and Florian Kagerer. A five degree-of-freedom body-machine interface for children with severe motor impairments. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 3877–3882. IEEE, 2017.

[54] Lawrence Yunliang Chen, Huang Huang, Ellen Novoseller, Daniel Seita, Jeffrey Ichnowski, Michael Laskey, Richard Cheng, Thomas Kollar, and Ken Goldberg. Efficiently learning single-arm fling motions to smooth garments, 2022.

[55] Yang Chen, Diego Paez-Granados, Hideki Kadone, and Kenji Suzuki. Control interface for hands-free navigation of standing mobility vehicles based on upper-body natural movements. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 11322–11329. IEEE, 2020.

[56] Gong Cheng, Yanjiang Huang, Xianmin Zhang, Haoyang Chen, and Jun Ota. An overview of transfer

nursing robot: Classification, key technology, and trend. *Robotics and Autonomous Systems*, 174: 104653, 2024.

[57] Young Sang Choi, Cressel D Anderson, Jonathan D Glass, and Charles C Kemp. Laser pointers and a touch screen: intuitive interfaces for autonomous mobile manipulation for the motor impaired. In *International ACM SIGACCESS conference on Computers and accessibility (ASSETS)*, pages 225–232. ACM, 2008.

[58] Young Sang Choi, Tiffany Chen, Advait Jain, Cressel Anderson, Jonathan D Glass, and Charles C Kemp. Hand it over or set it down: A user study of object delivery with an assistive mobile manipulator. In *IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pages 736–743. IEEE, 2009.

[59] Young Sang Choi, Travis Deyle, Tiffany Chen, Jonathan D Glass, and Charles C Kemp. A list of household objects for robotic retrieval prioritized by people with als. In *2009 IEEE international conference on rehabilitation robotics*, pages 510–517. IEEE, 2009.

[60] PY Chua, T Ilschner, and DG Caldwell. Robotic manipulation of food products–a review. *Industrial Robot: An International Journal*, 30(4):345–354, 2003.

[61] Daisuke Chugo, Shohei Kawazoe, Sho Yokota, Hiroshi Hashimoto, Takahiro Katayama, Yasuhide Mizuta, and Atsushi Koujina. Pattern based standing assistance adapted to individual subjects on a robotic walker. In *IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pages 1216–1221. IEEE, 2017.

[62] Cheng-Shiu Chung, Hongwu Wang, and Rory A Cooper. Functional assessment and performance evaluation for assistive robotic manipulators: Literature review. *The journal of spinal cord medicine*, 36(4):273–289, 2013.

[63] Cheng-Shiu Chung, Hongwu Wang, Dan Ding, and Rory A Cooper. Feasibility analysis of daily activities using assistive robotic manipulators. In *Proc. RESNA Annu. Conf.*, pages 1–9, 2016.

[64] Yann-Seing Law-Kam Cio, Maxime Raison, Cedric Leblond Menard, and Sofiane Achiche. Proof

of concept of an assistive robotic arm control using artificial stereovision and eye-tracking. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 27(12):2344–2352, 2019.

[65] Caitlyn Clabaugh, Kartik Mahajan, Shomik Jain, Roxanna Pakkar, David Becerra, Zhonghao Shi, Eric Deng, Rhianna Lee, Gisele Ragusa, and Maja Matarić. Long-term personalization of an in-home socially assistive robot for children with autism spectrum disorders. *Frontiers in Robotics and AI*, 6: 110, 2019.

[66] Michele Colledanchise and Petter Ögren. *Behavior trees in robotics and AI: An introduction*. CRC Press, 2018.

[67] Albert M Cook and Janice Miller Polgar. *Assistive technologies-e-book: principles and practice*. Elsevier Health Sciences, 2014.

[68] Albert M Cook, MQ-H Meng, Jason J Gu, and Kathy Howery. Development of a robotic device for facilitating learning by children who have severe disabilities. *IEEE transactions on neural systems and rehabilitation engineering*, 10(3):178–187, 2002.

[69] Albert M Cook, Brenda Bentz, Norma Harbottle, Cheryl Lynch, and Brad Miller. School-based use of a robotic arm system by children with disabilities. *IEEE transactions on neural systems and rehabilitation engineering*, 13(4):452–460, 2005.

[70] Robyn A Cree, Catherine A Okoro, Matthew M Zack, and Eric Carbone. Frequent mental distress among adults, by disability status, disability type, and selected characteristics—united states, 2018. *Morbidity and Mortality Weekly Report*, 69(36):1238, 2020.

[71] Lilian de Greef, Dominik Moritz, and Cynthia Bennett. Interdependent variables: Remotely designing tactile graphics for an accessible workflow. In *The 23rd International ACM SIGACCESS Conference on Computers and Accessibility*, pages 1–6, 2021.

[72] Cee W de Jong, Klaus Klemp, Erik Mattie, and Donald Goodwin. *Ten Principles for Good Design: Dieter Rams: The Jorrit Maan Collection*. Munich: Prestel,[2017]., 2017.

[73] Agostino De Santis, Bruno Siciliano, Alessandro De Luca, and Antonio Bicchi. An atlas of physical human–robot interaction. *Mechanism and Machine Theory*, 43(3):253–270, 2008.

[74] Daniel Delgado Bellamy, Gregory Chance, Praminda Caleb-Solly, and Sanja Dogramadzi. Safety assessment review of a dressing assistance robot. *Frontiers in Robotics and AI*, 8:667316, 2021.

[75] Anind K Dey, Gregory D Abowd, and Daniel Salber. A conceptual framework and a toolkit for supporting the rapid prototyping of context-aware applications. *Human–Computer Interaction*, 16 (2-4):97–166, 2001.

[76] Praveen Kumar Donepudi. Reinforcement learning for robotic grasping and manipulation: a review. *Asia Pacific Journal of Energy and Environment*, 7(2):69–78, 2020.

[77] Anca D Dragan, Kenton CT Lee, and Siddhartha S Srinivasa. Legibility and predictability of robot motion. In *ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 301–308. ACM, 2013.

[78] Johanna Drucker. *Graphesis: Visual forms of knowledge production*. Harvard University Press Cambridge, MA, 2014.

[79] Peter F Edemekong, Deb Bomgaars, Sukesh Sukumaran, and Shoshana B Levy. Activities of daily living. 2019.

[80] Eleni Efthimiou, Stavroula-Evita Fotinea, Anna Vacalopoulou, Xanthi S Papageorgiou, Alexandra Karavasili, and Theodore Goulas. User centered design in practice: adapting hri to real user needs. In *ACM International Conference on PErvasive Technologies Related to Assistive Environments*, pages 425–429, 2019.

[81] Gen Endo and Nobuhiro Otomo. Development of a food handling gripper considering an appetizing presentation. In *2016 IEEE International Conference on Robotics and Automation (ICRA)*, pages 4901–4906. IEEE, 2016.

[82] Zackory Erickson, Henry M Clever, Greg Turk, C Karen Liu, and Charles C Kemp. Deep haptic

model predictive control for robot-assisted dressing. In *IEEE international conference on robotics and automation (ICRA)*, pages 4437–4444. IEEE, 2018.

[83] Zackory Erickson, Maggie Collier, Ariel Kapusta, and Charles C Kemp. Tracking human pose during robot-assisted dressing using single-axis capacitive proximity sensing. *IEEE Robotics and Automation Letters*, 3(3):2245–2252, 2018.

[84] Zackory Erickson, Yijun Gu, and Charles C Kemp. Assistive vr gym: Interactions with real people to improve virtual assistive robots. In *IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, pages 299–306. IEEE, 2020.

[85] Nabil Ettehadi and Aman Behal. Implementation of feeding task via learning from demonstration. In *2018 Second IEEE International Conference on Robotic Computing (IRC)*, pages 274–277. IEEE, 2018.

[86] Haiyan Fan and Marshall Scott Poole. What is personalization? perspectives on the design and implementation of personalization in information systems. *Journal of Organizational Computing and Electronic Commerce*, 16(3-4):179–202, 2006.

[87] Qinyuan Fang, Maria Kyrarini, Danijela Ristic-Durrant, and Axel Gräser. Rgb-d camera based 3d human mouth detection and tracking towards robotic feeding assistance. In *PErvasive Technologies Related to Assistive Environments Conference*, pages 391–396. ACM, 2018.

[88] Cara Bailey Fausset, Andrew J Kelly, Wendy A Rogers, and Arthur D Fisk. Challenges to aging in place: Understanding home maintenance difficulties. *Journal of Housing for the Elderly*, 25(2): 125–141, 2011.

[89] Ryan Feng, Youngsun Kim, Gilwoo Lee, Ethan K Gordon, Matt Schmittle, Shivaum Kumar, Tapomayukh Bhattacharjee, and Siddhartha S Srinivasa. Robot-Assisted Feeding: Generalizing Skewering Strategies Across Food Items on a Plate. In *The International Symposium of Robotics Research*, pages 427–442. Springer, 2019.

[90] Donald Firesmith. The need to specify requirements for off-nominal behavior. *Software Engineering*

Institute Blog, Jan 2012. URL https://insights.sei.cmu.edu/blog/the-need-to-specify-requirements-for-off-nominal-behavior/.

[91] Osamu Fukuda, Toshio Tsuji, and Makoto Kaneko. An emg controlled robotic manipulator using neural networks. In *IEEE International Workshop on Robot and Human Communication (RO-MAN)*, pages 442–447. IEEE, 1997.

[92] Lone Gaedt. Summary of the report "dining assistive technology for people with physical disabilities", Nov 2012. URL https://www.neater.co.uk/_files/ugd/7274b5_8e8c40148f064b49bae377e9b2d60c5d.pdf.

[93] Daniel Gallenberger, Tapomayukh Bhattacharjee, Youngsun Kim, and Siddhartha S Srinivasa. Transfer depends on acquisition: Analyzing manipulation strategies for robotic feeding. In *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 267–276. IEEE, IEEE, 2019.

[94] D Randy Garrison, Martha Cleveland-Innes, Marguerite Koole, and James Kappelman. Revisiting methodological issues in transcript analysis: Negotiated coding and reliability. *The internet and higher education*, 9(1):1–8, 2006.

[95] Clifford Geertz et al. *The interpretation of cultures*, volume 5019. Basic books, 1973.

[96] M C Gemici and A Saxena. Learning haptic representation for manipulating deformable food objects. In *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 638–645, September 2014.

[97] Milad Geravand, Wolfgang Rampeltshammer, and Angelika Peer. Control of mobility assistive robot for human fall prevention. In *IEEE International Conference on Rehabilitation Robotics (ICORR)*, pages 882–887. IEEE, 2015.

[98] Darren Gergle and Desney S Tan. Experimental research in hci. *Ways of Knowing in HCI*, pages 191–227, 2014.

[99] Razan Ghzouli, Thorsten Berger, Einar Broch Johnsen, Swaib Dragule, and Andrzej Wasowski. Behavior trees in action: a study of robotics applications. In *Proceedings of the 13th ACM SIGPLAN International Conference on Software Language Engineering*, pages 196–209, 2020.

[100] Karen Glanz, Jessica J Metcalfe, Sara C Folta, Alison Brown, and Barbara Fiese. Diet and health benefits associated with in-home eating and sharing meals at home: A systematic review. *International journal of environmental research and public health*, 18(4):1577, 2021.

[101] Felix Ferdinand Goldau, Tejas Kumar Shastha, Maria Kyrarini, and Axel Gräser. Autonomous multi-sensory robotic assistant for a drinking task. In *IEEE International Conference on Rehabilitation Robotics (ICORR)*, pages 210–216. IEEE, IEEE, 2019.

[102] Deepak Gopinath, Siddarth Jain, and Brenna D Argall. Human-in-the-loop optimization of shared autonomy in assistive robotics. *IEEE Robotics and Automation Letters*, 2(1):247–254, 2016.

[103] Deepak E Gopinath and Brenna D Argall. Active intent disambiguation for shared control robots. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 28(6):1497–1506, 2020.

[104] E.K. Gordon*, A. Nanavati*, R. Challa, Bernie H. Zhu, Taylor A. Kessler Faulkner, and S. S. Srinivasa. Towards general single-utensil food acquisition with human-informed actions. In *Conference on Robot Learning*, 2023.

[105] Ethan K Gordon, Xiang Meng, Tapomayukh Bhattacharjee, Matt Barnes, and Siddhartha S Srinivasa. Adaptive robot-assisted feeding: An online learning framework for acquiring previously unseen food items. In *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 9659–9666. IEEE, 2020. doi: 10.1109/IROS45743.2020.9341359.

[106] Ethan K Gordon, Sumegh Roychowdhury, Tapomayukh Bhattacharjee, Kevin Jamieson, and Siddhartha S Srinivasa. Leveraging post hoc context for faster learning in bandit settings with applications in robot-assisted feeding. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pages 10528–10535. IEEE, 2021. doi: 10.1109/ICRA48506.2021.9561520.

[107] Ethan Kroll Gordon, Amal Nanavati, Ramya Challa, Bernie Hao Zhu, Taylor Annette Kessler

Faulkner, and Siddhartha Srinivasa. Towards general single-utensil food acquisition with human-informed actions. In *Conference on Robot Learning*, pages 2414–2428. PMLR, 2023.

[108] Jennifer Grannen, Yilin Wu, Suneel Belkhale, and Dorsa Sadigh. Learning bimanual scooping policies for food acquisition. *arXiv preprint arXiv:2211.14652*, 2022.

[109] Phillip M Grice and Charles C Kemp. In-home and remote use of robotic body surrogates by people with profound motor deficits. *PloS one*, 14(3):e0212904, 2019.

[110] Phillip M Grice, Andy Lee, Henry Evans, and Charles C Kemp. The wouse: A wearable wince detector to stop assistive robots. In *IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pages 165–172. IEEE, 2012.

[111] Rebecca A Grier. How high is high? a meta-analysis of nasa-tlx global workload scores. In *Human Factors and Ergonomics Society Annual Meeting*, volume 59, pages 1727–1731. SAGE Publications Sage CA: Los Angeles, CA, 2015.

[112] Klaus G Grunert, Moira Dean, Monique M Raats, Niels Asger Nielsen, Margaret Lumbers, et al. A measure of satisfaction with food-related life. *Appetite*, 49(2):486–493, 2007.

[113] Fabien Grzeskowiak, Louise Devigne, François Pasteau, Gustavo Souza Vieira Dutra, Marie Babel, and Sylvain Guégan. Swalkit: A generic augmented walker kit to provide haptic feedback navigation assistance to people with both visual and motor impairments. In *2022 International Conference on Rehabilitation Robotics (ICORR)*, pages 1–6. IEEE, IEEE, 2022.

[114] Florian Güldenpfennig, Peter Mayer, Paul Panek, and Geraldine Fitzpatrick. An autonomy-perspective on the design of assistive technology experiences of people with multiple sclerosis. In *CHI Conference on Human Factors in Computing Systems*, pages 1–14. ACM, 2019.

[115] Karen Hacker. *Community-based participatory research*. Sage publications, 2013.

[116] Hairong Jiang, Juan P. Wachs, and Bradley S. Duerstock. Integrated vision-based robotic arm interface for operators with upper limb mobility impairments. In *IEEE International Conference on*

Rehabilitation Robotics (ICORR), pages 1–6, Seattle, WA, June 2013. IEEE. ISBN 978-1-4673-6024-1 978-1-4673-6022-7. doi: 10.1109/ICORR.2013.6650447. URL http://ieeexplore.ieee.org/document/6650447/.

[117] Till Halbach, Ivar Solheim, Siri Ytrehus, and Trenton Schulz. A mobile application for supporting dementia relatives: a case study. In *Transforming our World Through Design, Diversity and Education*, pages 839–846. IOS Press, 2018.

[118] Yuka Hamada, Akane Miyaji, Yoko Hayashi, Naoyuki Matsumoto, Masato NISHIWAkI, and Naoyuki Hayashi. Objective and subjective eating speeds are related to body composition and shape in female college students. *Journal of nutritional science and vitaminology*, 63(3):174–179, 2017.

[119] T Hammarberg. The right of people with disabilities to live independently and be included in the community. *Issue Paper, Council of Europe Commissioner for Human Rights*, 2012.

[120] Søren Tranberg Hansen, Thomas Bak, and Claus Risager. An adaptive game algorithm for an autonomous, mobile robot-a real world study with elderly users. In *IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pages 892–897. IEEE, 2012.

[121] Amanda W Harrist and Ralph M Waugh. Dyadic synchrony: Its structure and function in children's development. *Developmental review*, 22(4):555–592, 2002.

[122] Sandra G Hart and Lowell E Staveland. Development of nasa-tlx (task load index): Results of empirical and theoretical research. In *Advances in psychology*, volume 52, pages 139–183. Elsevier, 1988.

[123] Rex Hartson and Pardha S Pyla. *The UX Book: Process and guidelines for ensuring a quality user experience*. Elsevier, 2012.

[124] Kelsey P Hawkins, Chih-Hung King, Tiffany L Chen, and Charles C Kemp. Informing assistive robots with models of contact forces from able-bodied face wiping and shaving. In *IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pages 251–258. IEEE, 2012.

[125] Laura V Herlant. Algorithms, implementation, and studies on eating with a shared control robot arm. March 2018. doi: 10.1184/R1/6714647.v1. URL https://kilthub.cmu.edu/articles/thesis/Algorithms_Implementation_and_Studies_on_Eating_with_a_Shared_Control_Robot_Arm/6714647.

[126] Laura V Herlant, Rachel M Holladay, and Siddhartha S Srinivasa. Assistive teleoperation of robot arms via automatic time-optimal mode switching. In *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 35–42. IEEE, ACM, 2016.

[127] Richard P Hermann, Anna C Phalangas, Richard M Mahoney, and Micheala Alexander. Powered feeding devices: an evaluation of three models. *Archives of physical medicine and rehabilitation*, 80 (10):1237–1242, 1999.

[128] Morten Hertzum. Reference values and subscale patterns for the task load index (tlx): a meta-analytic review. *Ergonomics*, 64(7):869–878, 2021.

[129] Emmanuelle Hien and Bertrand L Deputte. Influence of a capuchin monkey companion on the social life of a person with quadriplegia: an experimental study. *Anthrozoös*, 10(2-3):101–107, 1997.

[130] Michael Hillman. 2 rehabilitation robotics from past to present–a historical perspective. *Advances in rehabilitation robotics: human-friendly technologies on movement assistance and restoration for people with disabilities*, pages 25–44, 2004.

[131] Marius Hoggenmüller, Wen-Ying Lee, Luke Hespanhol, Malte Jung, and Martin Tomitsch. Eliciting new perspectives in rtd studies through annotated portfolios: A case study of robotic artefacts. In *Designing Interactive Systems Conference 2021*, pages 1875–1886, 2021.

[132] Kat Holmes. *Mismatch: How inclusion shapes design*. Mit Press, 2020.

[133] Armin Hornung, Kai M Wurm, Maren Bennewitz, Cyrill Stachniss, and Wolfram Burgard. Octomap: An efficient probabilistic 3d mapping framework based on octrees. *Autonomous robots*, 34:189–206, 2013.

[134] Brandi House, Jonathan Malkin, and Jeff Bilmes. The VoiceBot: a voice controlled robot arm. In *SIGCHI Conference on Human Factors in Computing Systems*, pages 183–192, Boston MA USA, April 2009. ACM. ISBN 978-1-60558-246-7. doi: 10.1145/1518701.1518731. URL `https://dl.acm.org/doi/10.1145/1518701.1518731`.

[135] Rick H Hoyle, Monica J Harris, and Charles M Judd. Research methods in social relations. 2002.

[136] Sumio Ishii, Shinji Tanaka, and Fumiaki Hiramatsu. Meal assistance robot for severely handicapped people. In *Proceedings of 1995 IEEE International Conference on Robotics and Automation*, volume 2, pages 1308–1313. IEEE, 1995.

[137] Standard ISO. Robots and robotic devices—collaborative robots (iso-15066: 2016). *International Organization for Standardization*, 2016.

[138] Barbara A Israel, Amy J Schulz, Edith A Parker, and Adam B Becker. Review of community-based research: assessing partnership approaches to improve public health. *Annual review of public health*, 19(1):173–202, 1998.

[139] Anja Jackowski, Marion Gebhard, and Roland Thietje. Head motion and head gesture-based robot control: A usability study. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 26 (1):161–170, 2017.

[140] Kinova JACO, 2018. `https://www.kinovarobotics.com/en/products/robotic-armseries`.

[141] Siddarth Jain and Brenna Argall. Probabilistic human intent recognition for shared autonomy in assistive robotics. *ACM Transactions on Human-Robot Interaction (THRI)*, 9(1):1–23, 2019.

[142] Siddarth Jain, Ali Farshchiansadegh, Alexander Broad, Farnaz Abdollahi, Ferdinando Mussa-Ivaldi, and Brenna Argall. Assistive robotic manipulation through shared autonomy and a body-machine interface. In *IEEE international conference on rehabilitation robotics (ICORR)*, pages 526–531. IEEE, 2015.

[143] Shervin Javdani, Henny Admoni, Stefania Pellegrinelli, Siddhartha S Srinivasa, and J Andrew Bagnell. Shared autonomy via hindsight optimization for teleoperation and teaming. *The International Journal of Robotics Research*, 37(7):717–742, 2018.

[144] Rajat K Jenamani, ..., and Tapomayukh Bhattacharjee. Robot-assisted inside-mouth bite transfer using robust mouth perception and physical interaction-aware control. In *2024 19th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 313–322, 2024.

[145] Rajat Kumar Jenamani, Priya Sundaresan, Maram Sakr, Tapomayukh Bhattacharjee, and Dorsa Sadigh. Flair: Feeding via long-horizon acquisition of realistic dishes. *arXiv preprint arXiv:2407.07561*, 2024.

[146] Hong Jun Jeon, Dylan P Losey, and Dorsa Sadigh. Shared autonomy with learned latent actions. *Robotics: Science and Systems (RSS)*, 2020.

[147] Aleksandar Jevtić, Andrés Flores Valle, Guillem Alenyà, Greg Chance, Praminda Caleb-Solly, Sanja Dogramadzi, and Carme Torras. Personalized robot assistant for support in dressing. *IEEE transactions on cognitive and developmental systems*, 11(3):363–374, 2018.

[148] N Jin, J Kang, and SK Agrawal. Design of a novel assist interface where toddlers walk with a mobile robot supported at the waist. In *IEEE International Conference on Rehabilitation Robotics (ICORR)*, pages 577–582. IEEE, 2015.

[149] Pengju Jin, Eshed Ohn-Bar, Kris Kitani, and Chieko Asakawa. A-exp4: Online social policy learning for adaptive robot-pedestrian interaction. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 5086–5093. IEEE, 2019.

[150] Malte Jung and Pamela Hinds. Robots in the wild: A time for more robust theories of human-robot interaction, 2018.

[151] Michail Kalaitzakis, Brennan Cain, Sabrina Carroll, Anand Ambrosi, Camden Whitehead, and Nikolaos Vitzilaios. Fiducial markers for pose estimation: Overview, applications and experimental comparison of the artag, apriltag, aruco and stag markers. *Journal of Intelligent & Robotic Systems*, 101: 1–26, 2021.

[152] Roya Norouzi Kandalan and Kamesh Namuduri. Techniques for constructing indoor navigation systems for the visually impaired: A review. *IEEE Transactions on Human-Machine Systems*, 50(6): 492–506, 2020.

[153] Ariel Kapusta, Wenhao Yu, Tapomayukh Bhattacharjee, C Karen Liu, Greg Turk, and Charles C Kemp. Data-driven haptic perception for robot-assisted dressing. In *IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pages 451–458. IEEE, 2016.

[154] Sertac Karaman and Emilio Frazzoli. Incremental sampling-based algorithms for optimal motion planning. *Robotics Science and Systems VI*, 104(2):267–274, 2010.

[155] Seita Kayukawa, Daisuke Sato, Masayuki Murata, Tatsuya Ishihara, Akihiro Kosugi, Hironobu Takagi, Shigeo Morishima, and Chieko Asakawa. How users, facility managers, and bystanders perceive and accept a navigation robot for visually impaired people in public buildings. In *IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, pages 546–553. IEEE, 2022.

[156] Seita Kayukawa, Daisuke Sato, Masayuki Murata, Tatsuya Ishihara, Hironobu Takagi, Shigeo Morishima, and Chieko Asakawa. Enhancing blind visitor's autonomy in a science museum using an autonomous navigation robot. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, pages 1–14. ACM, 2023.

[157] Simeon Keates, P John Clarkson, Lee-Anne Harrison, and Peter Robinson. Towards a practical inclusive design approach. In *Proceedings on the 2000 conference on Universal Usability*, pages 45–52, 2000.

[158] Hyun K Kim, Heejin Jeong, Jangwoon Park, Jaehyun Park, Won-Seok Kim, Nahyeong Kim, Subin Park, and Nam-Jong Paik. Development of a comprehensive design guideline to evaluate the user experiences of meal-assistance robots considering human-machine social interactions. *International Journal of Human–Computer Interaction*, 38(17):1–14, 2022.

[159] JS Kim, W-K Song, SR Han, JE Heo, and KJ Lee. A study on the scooping and serving korean food using meal assistance robot. *Gerontechnology*, 21, 2022.

[160] Alexander Kirillov, Eric Mintun, Nikhila Ravi, Hanzi Mao, Chloe Rolland, Laura Gustafson, Tete Xiao, Spencer Whitehead, Alexander C Berg, Wan-Yen Lo, et al. Segment anything. *arXiv preprint arXiv:2304.02643*, 2023.

[161] Klaus Krippendorff. *The content analysis reader*. Sage, 2009.

[162] Gernot Kronreif, Barbara Prazak, Stefan Mina, Martin Kornfeld, Michael Meindl, and M Furst. Playrob-robot-assisted playing for children with severe physical disabilities. In *International Conference on Rehabilitation Robotics (ICORR)*, pages 193–196. IEEE, 2005.

[163] James J Kuffner and Steven M LaValle. Rrt-connect: An efficient approach to single-query path planning. In *Proceedings 2000 ICRA. Millennium Conference. IEEE International Conference on Robotics and Automation. Symposia Proceedings (Cat. No. 00CH37065)*, volume 2, pages 995–1001. IEEE, 2000.

[164] Arun Kumar and Margaret Frances Phillips. Use of powered mobile arm supports by people with neuromuscular conditions. *Journal of Rehabilitation Research & Development*, 50(1), 2013.

[165] Vijay Kumar, Tariq Rahman, and Venkat Krovi. Assistive devices for people with motor disabilities. *Wiley Encyclopedia of Electrical and Electronics Engineering*, 1997.

[166] Tejas Kumar Shastha, Maria Kyrarini, and Axel Gräser. Application of reinforcement learning to a robotic drinking assistant. *Robotics*, 9(1):1, 2019.

[167] Raja S Kushalnagar and Christian Vogler. Teleconference accessibility and guidelines for deaf and hard of hearing users. In *Proceedings of the 22nd International ACM SIGACCESS Conference on Computers and Accessibility*, pages 1–6, 2020.

[168] Joseph La Delfa, Mehmet Aydin Baytas, Rakesh Patibanda, Hazel Ngari, Rohit Ashok Khot, and Florian'Floyd' Mueller. Drone chi: Somaesthetic human-drone interaction. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–13, 2020.

[169] Richard E Ladner. Design for user empowerment. *interactions*, 22(2):24–29, 2015.

[170] Carlos M Lara-Barrios, Andrés Blanco-Ortega, Cesar H Guzmán-Valdivia, and Karla D Bustamante Valles. Literature review and current trends on transfemoral powered prosthetics. *Advanced Robotics*, 32(2):51–62, 2018.

[171] Tor Lattimore and Csaba Szepesvári. *Bandit algorithms*. Cambridge University Press, 2020.

[172] Clemente Lauretti, Francesca Cordella, Francesco Scotto Di Luzio, Stefano Saccucci, Angelo Davalli, Rinaldo Sacchetti, and Loredana Zollo. Comparative performance analysis of M-IMU/EMG and voice user interfaces for assistive robots. In *International Conference on Rehabilitation Robotics (ICORR)*, pages 1001–1006, London, July 2017. IEEE. ISBN 978-1-5386-2296-4. doi: 10.1109/ICORR.2017.8009380. URL `https://ieeexplore.ieee.org/document/8009380/`.

[173] Matthew V Law, JiHyun Jeong, Amritansh Kwatra, Malte F Jung, and Guy Hoffman. Negotiating the creative space in human-robot collaborative design. In *Proceedings of the 2019 on Designing Interactive Systems Conference*, pages 645–657, 2019.

[174] David Lee. *The Origins of an Everyday Behavior: Why do People Share Meals?* PhD thesis, The University of Mississippi, 2014.

[175] James R Lewis. The system usability scale: past, present, and future. *International Journal of Human–Computer Interaction*, 34(7):577–590, 2018.

[176] Jessica Lindblom, Beatrice Alenljung, and Erik Billing. Evaluating the user experience of human–robot interaction. *Human-Robot Interaction: Evaluation Methods and Their Standardization*, pages 231–256, 2020.

[177] Ann-Louise Lindborg and Maria Lindén. Development of an eating aid–from the user needs to a product. In *pHealth 2015*, pages 191–198. IOS Press, 2015.

[178] Sara Ljungblad. Applying "designerly framing" to understand assisted feeding as social aesthetic bodily experiences. *ACM Transactions on Human-Robot Interaction*, 12(2):1–23, 2023.

[179] Dylan P Losey, Krishnan Srinivasan, Ajay Mandlekar, Animesh Garg, and Dorsa Sadigh. Control-

ling assistive robots with learned latent actions. In *IEEE International Conference on Robotics and Automation (ICRA)*, pages 378–384. IEEE, 2020.

[180] Ryan Luna, Ioan A Şucan, Mark Moll, and Lydia E Kavraki. Anytime solution optimization for sampling-based motion planning. In *2013 IEEE international conference on robotics and automation*, pages 5068–5074. IEEE, 2013.

[181] J. Tevere MacFayden. Educated monkeys help the disabled to help themselves. *Smithsonian*, 17: 125–132, 1986.

[182] Paweł Maciejasz, Jörg Eschweiler, Kurt Gerlach-Hahn, Arne Jansen-Troy, and Steffen Leonhardt. A survey on robotic devices for upper limb rehabilitation. *Journal of neuroengineering and rehabilitation*, 11(1):1–29, 2014.

[183] Dennis Maciuszek, Johan Aberg, and Nahid Shahmehri. What help do older people need? constructing a functional design space of electronic assistive technology applications. In *Proceedings of the 7th international ACM SIGACCESS conference on Computers and accessibility*, pages 4–11, 2005.

[184] Shane MacNamara and Gerard Lacey. A smart walker for the frail visually impaired. In *IEEE International Conference on Robotics and Automation.*, volume 2, pages 1354–1359. IEEE, 2000.

[185] Rishabh Madan, Rajat Kumar Jenamani, Vy Thuy Nguyen, Ahmed Moustafa, Xuefeng Hu, Katherine Dimitropoulou, and Tapomayukh Bhattacharjee. Sparcs: Structuring physically assistive robotics for caregiving with stakeholders-in-the-loop. In *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 641–648. IEEE, 2022.

[186] Rishabh Madan, Skyler Valdez, David Kim, Sujie Fang, Luoyan Zhong, Diego T Virtue, and Tapomayukh Bhattacharjee. Rabbit: A robot-assisted bed bathing system with multimodal perception and integrated compliance. In *Proceedings of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*, pages 472–481, 2024.

[187] Priyanka Mandikal and Kristen Grauman. DexVIP: Learning Dexterous Grasping with Human Hand Pose Priors from Video. In Aleksandra Faust, David Hsu, and Gerhard Neumann, editors, *Proceed-*

ings of the 5th Conference on Robot Learning, volume 164 of *Proceedings of Machine Learning Research*, pages 651–661. PMLR, 08–11 Nov 2022.

[188] Jennifer Mankoff, Gillian R Hayes, and Devva Kasnitz. Disability studies as a source of critical inquiry for the field of assistive technology. In *International ACM SIGACCESS conference on Computers and accessibility (ASSETS)*, pages 3–10. ACM, ACM, 2010.

[189] Ren Mao, Yezhou Yang, Cornelia Fermüller, Yiannis Aloimonos, and John S Baras. Learning hand movements from markerless demonstrations for humanoid tasks. In *2014 IEEE-RAS International Conference on Humanoid Robots*, pages 938–943. IEEE, 2014.

[190] Maja J Matarić and Brian Scassellati. Socially assistive robotics. *Springer handbook of robotics*, pages 1973–1994, 2016.

[191] Meal Buddy. Meal buddy, 2022. URL `https://www.performancehealth.com/meal-buddy-systems`. `https://www.performancehealth.com/meal-buddy-system`,[Online; Retrieved on 25th January, 2018].

[192] Mealtime Partner. The mealtime partner dining system description, 2020. URL `https://www.mealtimepartners.com/dining/mealtime-partner-dining-device.htm`.

[193] Meet Obi. Meet obi, Aug 2022. URL `https://meetobi.com/`.

[194] Shaunak A Mehta, Sagar Parekh, and Dylan P Losey. Learning latent actions without human demonstrations. In *International Conference on Robotics and Automation (ICRA)*, pages 7437–7443. IEEE, 2022.

[195] Herbert L Meiselman. Dimensions of the meal: Science, culture, business, art, 2000.

[196] Merriam-Webster. Nominal. In *Merriam-Webster.com dictionary*. 2023. URL `https://www.merriam-webster.com/dictionary/nominal`.

[197] J Michaelis. Mechanical methods of controlling ataxia. *Bailliere's Clinical Neurology*, 2(1):121–139, 1993.

[198] Meredith Minkler and Nina Wallerstein. *Community-based participatory research for health: From process to outcomes*. John Wiley & Sons, 2011.

[199] Ivica Mitrović. An introduction to speculative design practice. *An Introduction to Speculative Design-Eutropia, a Case Study Practice, Croatian Designers Association, Department for Visual Communications Design, Arts Academy, University of Split*, pages 8–23, 2015.

[200] Abolfazl Mohebbi. Human-robot interaction in rehabilitation and assistance: a review. *Current Robotics Reports*, 1:131–144, 2020.

[201] Lorenza Mondada. The methodical organization of talking and eating: Assessments in dinner conversations. *Food quality and preference*, 20(8):558–571, 2009.

[202] Niels Christian Mossfeldt Nickelsen. Imagining and tinkering with assistive robotics in care for the disabled. *Paladyn, Journal of Behavioral Robotics*, 10(1):128–139, 2019.

[203] Tamanna Motahar, Md Fahim Farden, Dibya Prokash Sarkar, Md Atiqul Islam, Maria E Cabrera, and Maya Cakmak. Sheba: A low-cost assistive robot for older adults in the developing world. In *IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, pages 1–8. IEEE, 2019.

[204] Anne Murcott. The cultural significance of food and eating. *Proceedings of the Nutrition Society*, 41 (2):203–210, 1982.

[205] MySpoon. My spoon, 2022. URL `https://web.archive.org/web/20210510063115/http://www.secom.co.jp/english/myspoon/`.

[206] Maia Naftali and Leah Findlater. Accessibility in context: understanding the truly mobile experience of smartphone users with motor impairments. In *Proceedings of the 16th international ACM SIGACCESS conference on Computers & accessibility*, pages 209–216, 2014.

[207] Amal Nanavati, Ramya Challa, Ethan K. Gordon, and Siddhartha S. Srinivasa. A Dataset of Food Manipulation Strategies for Diverse Foods, 2022. URL `https://doi.org/10.7910/DVN/C8SI1D`.

[208] Amal Nanavati, Patricia Alves-Oliveira, Tyler Schrenk, Ethan K. Gordon, Maya Cakmak, and Siddhartha S. Srinivasa. Design principles for robot-assisted feeding in social contexts. In *ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, HRI '23, page 24–33, New York, NY, USA, 2023. Association for Computing Machinery. ISBN 9781450399647. doi: 10.1145/3568162.3576988. URL `https://doi.org/10.1145/3568162.3576988`.

[209] Amal Nanavati, Patrícia Alves-Oliveira, Tyler Schrenk, Ethan Gordon, Maya Cakmak, and Siddhartha S Srinivasa. Unintended failures of robot-assisted feeding in social contexts. In *Proceedings of the Companion of the 2023 ACM/IEEE International Conference on Human-Robot Interaction*, HRI '23, pages 884–886, New York, NY, USA, 2023. Association for Computing Machinery.

[210] Amal Nanavati, Vinitha Ranganeni, and Maya Cakmak. Physically assistive robots: A systematic review of mobile and manipulator robots that physically assist people with disabilities. *Annual Review of Control, Robotics, and Autonomous Systems*, 7, 2023.

[211] Amal Nanavati, Ethan Gordon, Taylor A K Faulkner, Yuxin Song, Jonathan Ko, Tyler Schrenk, Vy Nguyen, Bernie H Zhu, Haya Bolotski, Atharva Kashyap, and et al. Lessons learned from designing and evaluating a robot-assisted feeding system for out-of-lab use (supplementary materials), Dec 2024. URL `https://doi.org/10.17605/OSF.IO/PRSXT`.

[212] Amal Nanavati, Ethan K Gordon, Taylor A Kessler Faulkner, Yuxin (Ray) Song, Johnathan Ko, Tyler Schrenk, Vy Nguyen, Bernie Hao Zhu, Haya Bolotski, Atharva Kashyap, Sriram Kutty, Raida Karim, Liander Rainbolt, Rosario Scalise, Hanjun Song, Ramon Qu, Maya Cakmak, and Siddhartha S Srinivasa. Lessons learned from designing and evaluating a robot-assisted feeding system for out-of-lab use. In *Proceedings of the 2025 ACM/IEEE International Conference on Human-Robot Interaction*, 2025.

[213] Isira Naotunna, Chamika Janith Perera, Chameera Sandaruwan, RARC Gopura, and Thilina Dulantha Lalitharatne. Meal assistance robots: A review on current status, challenges and future directions. In *2015 IEEE/SICE international symposium on system integration (SII)*, pages 211–216. IEEE, 2015.

[214] Neater Eater. Neater eater robotic, 2022. URL `https://www.neater.co.uk/neater-eater-robotic`.

[215] Benjamin A Newman, Reuben M Aronson, Kris Kitani, and Henny Admoni. Helping people through space and time: Assistance as a perspective on human-robot interaction. *Frontiers in Robotics and AI*, page 410, 2022.

[216] Hai Nguyen and Charles C Kemp. Bio-inspired assistive robotics: Service dogs as a model for human-robot interaction and mobile manipulation. In *2008 2nd IEEE RAS & EMBS International Conference on Biomedical Robotics and Biomechatronics*, pages 542–549. IEEE, 2008.

[217] Hai Nguyen, Matei Ciocarlie, Kaijen Hsiao, and Charles C Kemp. Ros commander (rosco): Behavior creation for home robots. In *2013 IEEE International Conference on Robotics and Automation*, pages 467–474. IEEE, 2013.

[218] Vy Nguyen. Increasing independence with stretch: A mobile robot enabling functional performance in daily activities. 2021.

[219] Vy Nguyen. Increasing independence with stretch: A mobile robot enabling functional performance in daily activities, 2021.

[220] Niels Christian Mossfeldt Nickelsen. Imagining and tinkering with assistive robotics in care for the disabled. *Paladyn, Journal of Behavioral Robotics*, 10(1):128–139, 2019.

[221] Yutaro Ohshima, Yuichi Kobayashi, Toru Kaneko, Atsushi Yamashita, and Hajime Asama. Meal support system with spoon using laser range finder and manipulator. In *2013 IEEE workshop on robot vision (WORV)*, pages 82–87. IEEE, 2013.

[222] Tomohiro Oka, Jorge Solis, Ann-Louise Lindborg, Daisuke Matsuura, Yusuke Sugahara, and Yukio Takeda. Kineto-elasto-static design of underactuated chopstick-type gripper mechanism for meal-assistance robot. *Robotics*, 9(3):50, 2020.

[223] Samuel A Olatunji, Vy Nguyen, Maya Cakmak, Aaron Edsinger, Charles C Kemp, Wendy A Rogers,

and Harshal P Mahajan. Immersive participatory design of assistive robots to support older adults. *Ergonomics*, pages 1–15, 2024.

[224] Samuel Adeolu Olatunji, Tal Oron-Gilad, Noa Markfeld, Dana Gutman, Vardit Sarne-Fleischmann, and Yael Edan. Levels of automation and transparency: interaction design considerations in assistive robots for older adults. *IEEE Transactions on Human-Machine Systems*, 51(6):673–683, 2021.

[225] Edwin Olson. AprilTag: A robust and flexible visual fiducial system. In *2011 IEEE International Conference on Robotics and Automation*, pages 3400–3407, 2011. doi: 10.1109/ICRA.2011.5979561.

[226] Jan Ondras, Abrar Anwar, Tong Wu, Fanjun Bu, Malte Jung, Jorge Jose Ortiz, and Tapomayukh Bhattacharjee. Human-robot commensality: Bite timing prediction for robot-assisted feeding in groups. *arXiv preprint arXiv:2207.03348*, 2022.

[227] Juan F Orejuela-Zapata, Sarita Rodríguez, and Gonzalo Llano Ramírez. Self-help devices for quadriplegic population: a systematic literature review. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 27(4):692–701, 2019.

[228] Elaine Ostroff. Universal design: an evolving paradigm. *Universal design handbook*, 2:34–42, 2011.

[229] Alex Owen-Hill. How Robots With Vacuum Grippers Improve Food Safety, 2019. URL `https://blog.robotiq.com/how-robots-with-vacuum-grippers-improve-food-safety`.

[230] Eimei Oyama, Woo-Keun Yoon, Yujin Wakita, Hideyuki Tanaka, Masahiro Yoshikawa, Takeshi Sakaguchi, and Yoshio Matsumoto. Development of evaluation indexes for assistive robots based on icf. In *IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pages 221–227. IEEE, 2012.

[231] Akhil Padmanabha, Janavi Gupta, Chen Chen, Jehan Yang, Vy Nguyen, Douglas J Weber, Carmel Majidi, and Zackory Erickson. Independence in the home: A wearable interface for a person with quadriplegia to teleoperate a mobile manipulator. In *Proceedings of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*, pages 542–551, 2024.

[232] Akhil Padmanabha, Jessie Yuan, Janavi Gupta, Zulekha Karachiwalla, Carmel Majidi, Henny Admoni, and Zackory Erickson. Voicepilot: Harnessing llms as speech interfaces for physically assistive robots. In *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology*, pages 1–18, 2024.

[233] Yi-Tsen Pan, Chin-Cheng Shih, Christian DeBuys, and Pilwon Hur. Design of a sensory augmentation walker with a skin stretch feedback handle. In *IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pages 832–837. IEEE, 2018.

[234] Daehyung Park, Hokeun Kim, Yuuna Hoshi, Zackory Erickson, Ariel Kapusta, and Charles C Kemp. A multimodal execution monitor with anomaly classification for robot-assisted feeding. In *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 5406–5413. IEEE, IEEE, 2017.

[235] Daehyung Park, Yuuna Hoshi, and Charles C Kemp. A multimodal anomaly detector for robot-assisted feeding using an lstm-based variational autoencoder. *IEEE Robotics and Automation Letters*, 3(3):1544–1551, 2018.

[236] Daehyung Park, Yuuna Hoshi, Harshal P Mahajan, Ho Keun Kim, Zackory Erickson, Wendy A Rogers, and Charles C Kemp. Active robot-assisted feeding with a general-purpose mobile manipulator: Design, evaluation, and lessons learned. *Robotics and Autonomous Systems*, 124:103344, 2020.

[237] Max Pascher, Annalies Baumeister, Stefan Schneegass, Barbara Klein, and Jens Gerken. Recommendations for the development of a robotic drinking and eating aid-an ethnographic study. In *Human-Computer Interaction–INTERACT 2021: 18th IFIP TC 13 International Conference, Bari, Italy, August 30–September 3, 2021, Proceedings, Part I 18*, pages 331–351. Springer, 2021.

[238] Joao Paulo, Paulo Peixoto, and Urbano J Nunes. Isr-aiwalker: Robotic walker for intuitive and safe mobility assistance and gait analysis. *IEEE Transactions on Human-Machine Systems*, 47(6):1110–1122, 2017.

[239] Karie Jo Peralta. Politics of knowledge in community-based work. In *Dimensions of Community-Based Projects in Health Care*, pages 67–78. Springer, 2018.

[240] Chamika Janith Perera, Thilina Dulantha Lalitharatne, and Kazuo Kiguchi. Eeg-controlled meal assistance robot with camera-based automatic mouth position tracking and mouth open detection. In *2017 IEEE international conference on robotics and automation (ICRA)*, pages 1760–1765. IEEE, IEEE, 2017.

[241] Claudia Pérez-D'Arpino and Julie A Shah. C-learn: Learning geometric constraints from demonstrations for multi-step manipulation in shared autonomy. In *2017 IEEE International Conference on Robotics and Automation (ICRA)*, pages 4058–4065. IEEE, 2017.

[242] Helen Petrie and Jenny Darzentas. Older people and robotic technologies in the home: perspectives from recent research literature. In *International conference on pervasive technologies related to assistive environments*, pages 29–36. ACM, 2017.

[243] Guy N Phillips. Feasibility study for assistive feeder. *Southwest Research Institute*, page 1–130, 1987. URL `https://robotfeeding.io/publications/archive/phillips1987feasibility.pdf`.

[244] Ornella Plos, Stéphanie Buisine, Améziane Aoussat, Fabrice Mantelet, and Claude Dumas. A universalist strategy for the design of assistive technology. *International Journal of Industrial Ergonomics*, 42(6):533–541, 2012.

[245] Samuel Poirier, Francois Routhier, and Alexandre Campeau-Lecours. Voice Control Interface Prototype for Assistive Robots for People Living with Upper Limb Disabilities. In *IEEE International Conference on Rehabilitation Robotics (ICORR)*, pages 46–52, Toronto, ON, Canada, June 2019. IEEE. ISBN 978-1-72812-755-2. doi: 10.1109/ICORR.2019.8779524. URL `https://ieeexplore.ieee.org/document/8779524/`.

[246] Graham Pullin. *Design meets disability*. MIT press, 2009.

[247] André Queirós, Daniel Faria, and Fernando Almeida. Strengths and limitations of qualitative and quantitative research methods. *European Journal of Education Studies*, 2017.

[248] Gabriel Quere, Annette Hagengruber, Maged Iskandar, Samuel Bustamante, Daniel Leidner, Freek Stulp, and Jörn Vogel. Shared control templates for assistive robotics. In *IEEE International Conference on Robotics and Automation (ICRA)*, pages 1956–1962. IEEE, 2020.

[249] Camilo Perez Quintero, Oscar Ramirez, and Martin Jägersand. Vibi: Assistive vision-based interface for robot manipulation. In *IEEE International Conference on Robotics and Automation*, pages 4458–4463. IEEE, 2015.

[250] Vinitha Ranganeni. *Developing Accessible Teleoperation Interfaces for Assistive Robots with Stakeholders*. PhD thesis, University of Washington, 2024.

[251] Vinitha Ranganeni, Mike Sinclair, Eyal Ofek, Amos Miller, Jonathan Campbell, Andrey Kolobov, and Edward Cutrell. Exploring levels of control for a navigation assistant for blind travelers. *arXiv preprint arXiv:2301.02336*, pages 4–12, 2023.

[252] Vinitha Ranganeni, Vy Nguyen, Henry Evans, Jane Evans, Julian Mehu, Samuel Olatunji, Aaron Rogers, Wendy age Edsinger, Charlie Kemp, and Maya Cakmak. Robots for humanity: In-home deployment of stretch re2. In *Companion of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*, pages 1299–1301, 2024.

[253] Mitchel Resnick and Eric Rosenbaum. Designing for tinkerability. *Design, make, play: Growing the next generation of STEM innovators*, pages 163–181, 2013.

[254] Rehabmart Reviews. Winsford self-feeder, Apr 2013. URL `https://www.youtube.com/watch?v=vYG2dVuPmos&ab_channel=RehabmartReviews`.

[255] Travers Rhodes and Manuela Veloso. Robot-driven trajectory improvement for feeding tasks. In *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 2991–2996. IEEE, 2018.

[256] Gustavo Alfonso Garcia Ricardez, Jun Takamatsu, Tsukasa Ogasawara, and Jorge Solis Alfaro. Quantitative comfort evaluation of eating assistive devices based on interaction forces estimation using an accelerometer. In *IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pages 909–914. IEEE, 2018.

[257] Paul Rozin. The meaning of food in our lives: a cross-cultural perspective on eating and well-being. *Journal of nutrition education and behavior*, 37:S107–S112, 2005.

[258] Matteo Rubagotti, Inara Tusseyeva, Sara Baltabayeva, Danna Summers, and Anara Sandygulova. Perceived safety in physical human–robot interaction—a survey. *Robotics and Autonomous Systems*, 151:104047, 2022.

[259] Francisco J Ruiz-Ruiz, Alberto Giammarino, Marta Lorenzini, Juan M Gandarias, Jesús H Gómez-De-Gabriel, and Arash Ajoudani. Improving standing balance performance through the assistance of a mobile collaborative robot. In *International Conference on Robotics and Automation (ICRA)*, pages 10017–10023. IEEE, 2022.

[260] Kirsti Sääskilahti, Riitta Kangaskorte, Sakari Pieskä, Juhana Jauhiainen, and Mika Luimula. Needs and user acceptance of older adults for mobile service robot. In *IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pages 559–564. IEEE, 2012.

[261] Jon A Sanford. *Universal design as a rehabilitation strategy: Design for the ages*. Springer Publishing Company, 2012.

[262] Mohammadhadi Sarajchi, Mohamad Kenan Al-Hares, and Konstantinos Sirlantzis. Wearable lower-limb exoskeleton for children with cerebral palsy: A systematic review of mechanical design, actuation type, control strategy, and clinical evaluation. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 29:2695–2720, 2021.

[263] Joe Saunders, Dag Sverre Syrdal, Kheng Lee Koay, Nathan Burke, and Kerstin Dautenhahn. "teach me–show me"—end-user personalization of a smart home and companion robot. *IEEE Transactions on Human-Machine Systems*, 46(1):27–40, 2015.

[264] Amrita Sawhney, Steven Lee, Kevin Zhang, Manuela Veloso, and Oliver Kroemer. Playing with food: Learning food item representations through interactive exploration. In *Experimental Robotics: The 17th International Symposium*, pages 309–322. Springer, 2021.

[265] S. Scherzinger, A. Roennau, and R. Dillmann. Forward dynamics compliance control (fdcc): A new

approach to cartesian compliance for robotic manipulators. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 4568–4575, 2017. doi: 10.1109/IROS.2017.8206325.

[266] John R Schultz, Andrew B Slifkin, Hongkai Yu, and Eric M Schearer. Proof-of-concept: A hands-free interface for robot-assisted self-feeding. In *International Conference on Rehabilitation Robotics (ICORR)*, pages 1–6. IEEE, 2022.

[267] W Seamone and G Schmeisser. Early clinical evaluation of a robot arm/worktable system for spinal-cord-injured persons. *Journal of rehabilitation research and development*, 22(1):38–57, 1985. URL `https://www.rehab.research.va.gov/jour/85/22/1/pdf/seamone.pdf`.

[268] Teresa E Seeman. Social ties and health: The benefits of social integration. *Annals of epidemiology*, 6(5):442–451, 1996.

[269] ATI-IA Force-Torque Sensor, 2018. `https://www.ati-ia.com/products/ft/ft_models.aspx?id=Nano25`.

[270] Ali Shafti, Pavel Orlov, and A Aldo Faisal. Gaze-based, context-aware robotic system for assisted reaching and grasping. In *International Conference on Robotics and Automation (ICRA)*, pages 863–869. IEEE, 2019.

[271] Lorenzo Shaikewitz, Yilin Wu, Suneel Belkhale, Jennifer Grannen, Priya Sundaresan, and Dorsa Sadigh. In-mouth robotic bite transfer with visual and haptic sensing. In *2023 IEEE International Conference on Robotics and Automation (ICRA)*, pages 9885–9895. IEEE, 2023.

[272] Richard K Sherwin, Neal Feigenson, and Christina Spiesel. What is visual knowledge, and what is it good for? potential ethnographic lessons from the field of legal practice. *Visual Anthropology*, 20 (2-3):143–178, 2007.

[273] Sivashankar Sivakanthan, Eline Blaauw, Mark Greenhalgh, Alicia M Koontz, Riemer Vegter, and Rory A Cooper. Person transfer assist systems: a literature review. *Disability and Rehabilitation: Assistive Technology*, 16(3):270–279, 2021.

[274] Shoshanna Sofaer. Qualitative methods: what are they and why use them? *Health services research*, 34(5 Pt 2):1101, 1999.

[275] Shuran Song, Andy Zeng, Johnny Lee, and Thomas Funkhouser. Grasping in the Wild: Learning 6DoF Closed-Loop Grasping From Low-Cost Demonstrations. *IEEE Robotics and Automation Letters*, 5(3):4978–4985, 2020. doi: 10.1109/LRA.2020.3004787.

[276] Won-Kyung Song and Jongbae Kim. Novel assistive robot for self-feeding. *Robotic Systems-Applications, Control and Programming*, pages 43–60, 2012.

[277] Won-Kyung Song, Won-Jin Song, Yale Kim, and Jongbae Kim. Usability test of knrc self-feeding robot. In *2013 IEEE 13th International Conference on Rehabilitation Robotics (ICORR)*, pages 1–5. IEEE, IEEE, 2013.

[278] Ryoji Soyama, Sumio Ishii, and Azuma Fukase. The development of meal-assistance robot'my spoon'. In *Proceedings of the 8th International Conference on Rehabilitation Robotics, 2003*, pages 88–91, 2003.

[279] Theresa A Spiegel, Thomas A Wadden, and Gary D Foster. Objective measurement of eating rate during behavioral treatment of obesity. *Behavior Therapy*, 22(1):61–67, 1991.

[280] Sebastian Starke, Norman Hendrich, and Jianwei Zhang. Memetic evolution for generic full-body inverse kinematics in robotics and animation. *IEEE Transactions on Evolutionary Computation*, 23 (3):406–420, 2018.

[281] Alissa C Stevens, Dianna D Carroll, Elizabeth A Courtney-Long, Qing C Zhang, Michelle L Sloan, Shannon Griffin-Blake, and Georgina Peacock. Adults with one or more functional disabilities—united states, 2011–2014. *Morbidity and Mortality Weekly Report*, 65(38):1021–1025, 2016.

[282] Ioan A Sucan, Mark Moll, and Lydia E Kavraki. The open motion planning library. *IEEE Robotics & Automation Magazine*, 19(4):72–82, 2012.

[283] Priya Sundaresan, Suneel Belkhale, and Dorsa Sadigh. Learning Visuo-Haptic Skewering Strategies for Robot-Assisted Feeding, November 2022. arXiv:2211.14648 [cs].

[284] Priya Sundaresan, Suneel Belkhale, and Dorsa Sadigh. Learning visuo-haptic skewering strategies for robot-assisted feeding. In *Conference on Robot Learning*, pages 332–341. PMLR, 2023.

[285] Priya Sundaresan, Jiajun Wu, and Dorsa Sadigh. Learning sequential acquisition policies for robot-assisted feeding. *arXiv preprint arXiv:2309.05197*, 2023.

[286] Daniel Szafir, Bilge Mutlu, and Terrence Fong. Communication of intent in assistive free flyers. In *ACM/IEEE international conference on Human-robot interaction (HRI)*, pages 358–365. ACM, 2014.

[287] Robyn L Tate and Michael Perdices. N-of-1 trials in the behavioral sciences. *The essential guide to N-of-1 trials in health*, pages 19–41, 2015.

[288] Danielle M Taylor. Americans with disabilities: 2014. *US Census Bureau*, pages 1–32, 2018.

[289] Noriyuki Tejima. Evaluation of rehabilitation robots for eating. In *Proceedings 5th IEEE International Workshop on Robot and Human Communication. RO-MAN'96 TSUKUBA*, pages 118–120. IEEE, 1996.

[290] Stefanie Tellex, Ross Knepper, Adrian Li, Daniela Rus, and Nicholas Roy. Asking for help using inverse semantics. 2014.

[291] Tiffany Thang, Alice Liang, Yechan Choi, Adrian Parrales, Sara H Kuang, Sri Kurniawan, and Heather Perez. Providing and accessing support during the covid-19 pandemic: Experiences of mental health professionals, community and vocational support providers, and adults with asd. In *The 23rd International ACM SIGACCESS Conference on Computers and Accessibility*, pages 1–6, 2021.

[292] Kristina A Theis, Amy Steinweg, Charles G Helmick, Elizabeth Courtney-Long, Julie A Bolen, and Robin Lee. Which one? what kind? how many? types, causes, and prevalence of disability among us adults. *Disability and health journal*, 12(3):411–421, 2019.

[293] Emmanuele Tidoni, Pierre Gergondet, Gabriele Fusco, Abderrahmane Kheddar, and Salvatore M Aglioti. The role of audio-visual feedback in a thought-based control of a humanoid robot: a bci study in healthy and spinal cord injured people. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 25(6):772–781, 2016.

[294] Garreth W Tigwell. Nuanced perspectives toward disability simulations from digital designers, blind, low vision, and color blind people. In *Conference on Human Factors in Computing Systems (CHI)*, pages 1–15. ACM, 2021.

[295] Garreth W Tigwell, Roshan L Peiris, Stacey Watson, Gerald M Garavuso, and Heather Miller. Student and teacher perspectives of learning asl in an online setting. In *The 22nd International ACM SIGACCESS Conference on Computers and Accessibility*, pages 1–6, 2020.

[296] Mike Topping. An overview of the development of handy 1, a rehabilitation robot to assist the severely disabled. *Artificial Life and Robotics*, 4(4):188–192, 2000.

[297] Carme Torras. Assistive robotics: Research challenges and ethics education initiatives. *Dilemata*, (30):63–77, 2019.

[298] Jennifer Utter, Simon Denny, Elizabeth Robinson, Theresa Fleming, Shanthi Ameratunga, and Sue Grant. Family meals and the well-being of adolescents. *Journal of Paediatrics and Child Health*, 49 (11):906–911, 2013.

[299] Jennifer Utter, Simon Denny, Roshini Peiris-John, Emma Moselen, Ben Dyson, and Terryann Clark. Family meals and adolescent emotional well-being: findings from a national study. *Journal of nutrition education and behavior*, 49(1):67–72, 2017.

[300] Paul Viola and Michael Jones. Rapid object detection using a boosted cascade of simple features. In *Proceedings of the 2001 IEEE computer society conference on computer vision and pattern recognition. CVPR 2001*, volume 1, pages I–I. Ieee, 2001.

[301] Meera Viswanathan, Alice Ammerman, Eugenia Eng, Gerald Garlehner, Kathleen N Lohr, Derek Griffith, Scott Rhodes, Carmen Samuel-Hodge, Siobhan Maty, Linda Lux, et al. Community-based participatory research: Assessing the evidence: Summary. *AHRQ evidence report summaries*, 2004.

[302] Dinh-Son Vu, Ulysse Côté Allard, Clément Gosselin, François Routhier, Benoit Gosselin, and Alexandre Campeau-Lecours. Intuitive adaptive orientation control of assistive robots for people living with upper limb disabilities. In *International Conference on Rehabilitation Robotics (ICORR)*, pages 795–800. IEEE, 2017.

[303] Andreas Wachaja, Pratik Agarwal, Mathias Zink, Miguel Reyes Adame, Knut Möller, and Wolfram Burgard. Navigating blind people with a smart walker. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 6014–6019. IEEE, 2015.

[304] Naoki Wake, Riku Arakawa, Iori Yanokura, Takuya Kiyokawa, Kazuhiro Sasabuchi, Jun Takamatsu, and Katsushi Ikeuchi. A Learning-from-Observation Framework: One-Shot Robot Teaching for Grasp-Manipulation-Release Household Operations. In *2021 IEEE/SICE International Symposium on System Integration (SII)*, pages 461–466, 2021. doi: 10.1109/IEEECONF49454.2021.9382750.

[305] Chao Wang, Xuehe Zhang, Xizhe Zang, Yubin Liu, Guanwen Ding, Wenxin Yin, and Jie Zhao. Feature sensing and robotic grasping of objects with uncertain information: A review. *Sensors*, 20 (13):3707, 2020.

[306] Ming-Yao Wang, Alexandros A Kogkas, Ara Darzi, and George P Mylonas. Free-view, 3d gaze-guided, assistive robotic system for activities of daily living. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 2355–2361. IEEE, 2018.

[307] Yanxin Wang, Guozheng Xu, Aiguo Song, Baoguo Xu, Huijun Li, Cong Hu, and Hong Zeng. Continuous shared control for robotic arm reaching driven by a hybrid gaze-brain machine interface. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 4462–4467. IEEE, 2018.

[308] Zhongkui Wang, Yuuki Torigoe, and Shinichi Hirai. A Prestressed Soft Gripper: Design, Modeling, Fabrication, and Tests for Food Handling. *IEEE Robotics and Automation Letters*, 2(4):1909–1916, October 2017. ISSN 2377-3766. doi: 10.1109/LRA.2017.2714141. Conference Name: IEEE Robotics and Automation Letters.

[309] Zhongkui Wang, Ryo Kanegae, and Shinichi Hirai. Circular shell gripper for handling food products. *Soft robotics*, 8(5):542–554, 2021.

[310] Alan Warde and Lydia Martens. *Eating out: Social differentiation, consumption and pleasure*. Cambridge University Press, 2000.

[311] Chu Wei, Lihong Li, Lev Reyzin, and Robert E Schapire. Contextual bandits with linear payoff functions. In *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, pages 208–214, 2011.

[312] Jacob O Wobbrock, Shaun K Kane, Krzysztof Z Gajos, Susumu Harada, and Jon Froehlich. Ability-based design: Concept, principles and examples. *ACM Transactions on Accessible Computing (TACCESS)*, 3(3):1–27, 2011.

[313] *Disability*. World Health Organization, Mar 2023. URL `https://www.who.int/news-room/fact-sheets/detail/disability-and-health`.

[314] Michele Xiloyannis, Ryan Alicea, Anna-Maria Georgarakis, Florian L Haufe, Peter Wolf, Lorenzo Masia, and Robert Riener. Soft robotic suits: State of the art, core technologies, and open challenges. *IEEE Transactions on Robotics*, 38(3):1343–1362, 2021.

[315] Akihiko Yamaguchi and Christopher G Atkeson. Combining finger vision and optical tactile sensing: Reducing and handling errors while cutting vegetables. In *Humanoid Robots (Humanoids), 2016 IEEE-RAS 16th International Conference on*, pages 1045–1051. IEEE, 2016.

[316] Akihiko Yamaguchi, Christopher G Atkeson, Scott Niekum, and Tsukasa Ogasawara. Learning pouring skills from demonstration and practice. In *2014 IEEE-RAS International Conference on Humanoid Robots*, pages 908–915. IEEE, 2014.

[317] Akira Yamazaki and Ryosuke Masuda. Autonomous foods handling by chopsticks for meal assistant robot. In *ROBOTIK 2012; 7th German Conference on Robotics*, pages 1–6. VDE, 2012.

[318] Yezhou Yang, Yi Li, Cornelia Fermuller, and Yiannis Aloimonos. Robot learning manipulation action plans by “watching” unconstrained videos from the world wide web. In *Proceedings of the AAAI conference on artificial intelligence*, volume 29(1), 2015.

[319] Ruolin Ye, Wenqiang Xu, Haoyuan Fu, Rajat Kumar Jenamani, Vy Nguyen, Cewu Lu, Katherine Dimitropoulou, and Tapomayukh Bhattacharjee. Rcare world: A human-centric simulation world for caregiving robots. In *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 33–40. IEEE, 2022.

[320] Fan Zhang, Antoine Cully, and Yiannis Demiris. Personalized robot-assisted dressing using user modeling in latent spaces. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 3603–3610. IEEE, 2017.

[321] Fan Zhang, Antoine Cully, and Yiannis Demiris. Probabilistic real-time user posture tracking for personalized robot-assisted dressing. *IEEE Transactions on Robotics*, 35(4):873–888, 2019.

[322] Jingzhe Zhang, Lishuo Zhuang, Yang Wang, Yameng Zhou, Yan Meng, and Gang Hua. An egocentric vision based assistive co-robot. In *IEEE International Conference on Rehabilitation Robotics (ICORR)*, pages 1–7. IEEE, 2013.

[323] Yan Zhang, Ziang Li, Haole Guo, Luyao Wang, Qihe Chen, Wenjie Jiang, Mingming Fan, Guyue Zhou, and Jiangtao Gong. " i am the follower, also the boss": Exploring different levels of autonomy and machine forms of guiding robots for the visually impaired. In *CHI Conference on Human Factors in Computing Systems*, pages 1–22. ACM, 2023.

[324] Megan Zimmerman, Shelly Bagchi, Jeremy Marvel, and Vinh Nguyen. An analysis of metrics and methods in research from human-robot interaction conferences, 2015–2021. In *ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 644–648. ACM, 2022.

Chapter A

Appendix for Chapter 5

Below is the appendix for Chapter 5 "Generalizing Bite Acquisition With Human-Informed Actions." As with Chapter 5, the original version of this appendix was published as part of "Towards General Single-Utensil Food Acquisition with Human-Informed Actions" at the *Conference on Robot Learning* in 2023 [104].

A.1 Human Data Collection

We ran a user study to identify which actions within this 26 dimensional schema are effective for acquiring a diverse food items.

1.1.1 Study Design

The study involved participants acquiring bite-sized pieces of a variety of food items with a fork and feeding them to an actuated mouth. The choice of food items was informed by ongoing collaboration with an end-user with C1 quadripledia[1], who had his caregiver take pictures of all the meals he ate in a week. One researcher then grouped similar food items (e.g., bread bun and bagel), resulting in a final set of 13 diverse food items: bagel chunks, mini sub sandwiches, pizza, chicken tenders, fries, broccoli, glazed doughnut holes, mashed potatoes, lettuce, spinach mix, whole jello, instant ramen noodles, and brown rice with beans. The bagels, sub sandwich, and pizza were pre-cut into bite-sized chunks, building off of past research that

[1]C1 quadriplegia refers to paralysis of all four limbs as a result of an injury to the first, or top-most, cervical vertebrate.

found that users are okay with caregivers cutting their food into bites before the robot feeds them [208]. The same brand ingredients and preparation procedure were followed for every food item.

The study space consisted of a table with a plate of food on it, a fork near the plate, a chair for the participant to sit in, and an actuated mouth to the left of the chair[2]. An RGB-D camera[3] above the plate captured visual aspects of the food. The fork, table, RGB-D camera, and actuated mouth all had motion capture markers on them, which were tracked by a motion capture system[4] in front of the table. The fork was also actuated with a force-torque sensor[5] to measure haptic aspects of the participant's acquisition action. An experimenter sitting behind the motion capture system, in full view of the participant, oversaw data collection. Figure 5.2a shows the study setup.

When each participant arrived, they were first briefed on the study and given time to read and fill out a consent form. They were then given a chance to familiarize themselves with the fork[6] by feeding baby carrots to the actuated mouth. When participants were ready, the actual data collection began, where they were provided a plate with one of the 13 food items, in randomized order, and were asked to feed one bite at a time to the actuated mouth. For each bite, participants were first asked to hold the fork in a comfortable "ready" position above the plate. When the experimented said "start," they lowered the fork to acquire the food item, moved it to the actuated mouth, and held it there until the experimenter said "stop." For each plate of food, participants were asked to feed at least 5 bites, and possibly more if the motion capture system lost tracking of the fork during the bite. In total, each study session took one hour and participants were compensated with a $10 gift card. Three researchers ran the study, and the study procedure was approved by our university's IRB. We had 9 participants, who all happened to be right-handed.

1.1.2 Dataset and Qualitative Observations

For each bite the participant acquired, our time series data consisted of the fork pose, force-torque sensor readings, and RGB-D images of the plate. We first cleaned this data by removing mistrials (i.e., trials with missing or corrupted data) and transforming all poses to a uniform frame of reference. We then published

[2]Because all participants happened to be right-handed, this positioning allowed them easy access to feed the actuated mouth.
[3]Intel RealSense D415
[4]OptiTrak V120 Trio
[5]ATI Industrial Automation 6DOF Nano25 with Net F/T Interface
[6]Due to the force-torque sensor and motion capture dots, the fork was a different shape and weight from regular forks

this dataset [207] to facilitate future research in food acquisition strategies. This dataset consists of 496 trials, totaling over 1.25 hours of food acquisition data across 13 food items and 9 participants.

Similarly to previous work [29], we observed some patterns in user interaction during data collection.

- Different participants held the fork differently (thumb in front of or behind the fork), which gave rise to different acquisition actions;
- Some food items could be acquired with multiple types of actions, e.g., some users scooped noodles whereas others twirled them;
- Users often tilted the fork while putting downwards pressure on it, in order to get it to pierce the food (e.g., broccoli).

Some of these lead to emergent behaviors identified in Section 5.4.3.2

A.2 Action Schema Point Extraction from Human Data

Exclusion Criteria We excluded any trial where tracking of the fork tip was lost for more than 0.5 seconds, or where the motion capture system did not read the stationary object poses (e.g., table, mouth) for the entire trial. After exclusions, we had a total of 410 trials.

Pre-processing To remove noise from the motion capture system, we smoothed all the fork tip poses by applying a median filter of 0.33 seconds separately to the x, y, z, roll, pitch, and yaw of the pose.

Significant Timestamps We first extracted the significant timestamps of the user's acquisition action. Specifically, we defined the *contact time* as the first timestamp when the distance from the fork tip to the camera exceeded the distance from the pixel corresponding to the fork tip to the camera in the initial depth image of the plate (i.e., the first time the fork pierced the surface of food on the plate). We then worked backwards from contact time, defining *start time* as the end of the 0.5 sec interval where the fork tip was consistently within a sphere of radius 5 cm, was more than 5 cm from its lowest point, and was more then 35 cm from the mouth. This criteria was based on our experimental design, where we asked the participant hold the fork stationary at a "ready" position before they acquired the food. In the event of multiple such periods where the fork was held stationary, we chose the one where the fork was the highest. We then defined *end* *time* to be the last timestamp when the fork was within 7 cm from its lowest point. And finally, we worked backwards from end time, defining *extraction time* to be the latest time when the fork was 1 cm away from its lowest point, within 2 sec of the end time.

All-in-all, the motion the participant took between *start time* and *contact time* corresponds to their **pre-grasp** motion, the motion they took between *contact time* and *extraction time* corresponds to their **grasp** motion, and the motion they took between *extraction time* and *end time* corresponds to their **extraction** motion.

Food Reference Frame Although a robot will know the food it is targeting before beginning its motion, with human data we have to extract the food item they were targeting from the data. We did so by segmenting the visually separate food items in the first RGB image of the plate of food. First, we detected the plate by: (a) using inpainting to remove glare; (b) using k-means clustering (k=3) to simplify the colors; and (c) finding the largest contour of the image that has at least 50% blue pixels. In practice, this reliably detected the plate for every food item in our study. We then masked out all the non-plate pixels from the de-glared image, and detected the food bounding box by: (a) running k-means (with k=2 for most food items, and k=3 for broccoli since its colors were closest to blue) to simplify colors; (b) masking out all the blue colors; (c) narrowing the mask to separate touching food items; (d) computing contours; and (e) fitting rotated rectangles to every contour with an area between within a hardcoded range. In practice, this reliably segmented separate food items, like bagel pieces or chicken tenders. For food items with a lot of overlap, like fries, this approach sometimes segmented multiple pieces of fries as the same. However, since participants often also acquired multiple overlapping pieces of those food items, we accepted those slight errors in food detection. For foods that weren't separated into bites like noodles or mashed potatoes, this algorithm rightly segmented it as one contiguous chunk of food.

Once we segmented separate bites of food, we defined the food reference frame to be centered at the center of the bounding box the fork tip was in at contact time, rotated to align with its major axis (i.e the center of the bite the user selected.)

Pre-Grasp The above preliminaries enable straightforward extraction of the pre-grasp, grasp, and extraction components of the action schema. For pre-grasp, we computed the target offset as the the target offset as the fork position at contact time in the food reference frame. We computed the initial utensil transform by taking the fork's linear velocity during a 0.5 second window before contact time and extrapolating that backwards 0.1 m, with a fixed orientation. And we took the force threshold to be 50% of the max force between start time and contact time.

Grasp We defined the in-food twist to be the transformation between the fork pose at extraction time and contact time, and the duration of the twist to be the duration between extraction time and contact time. We defined the force and torque thresholds to be 50% of the max force and torque between contact time and extraction time.

Extraction We defined the out-of-food twist to be the transformation between the fork pose at end time and extraction time, and the duration of the twist to be the duration between end time and extraction time.

Cleaning Three trials resulted in values of NaN or inf for at least one of the dimensions of the action schema. We eliminated these, resulting in 407 actions used for clustering.

A.3 Experiment Setup Details

Our experimental setup is shown in Figure 5.2b and was performed with a 6 DoF JACO2 robotic arm [140] with a 3D-printed handle and fork-shaped end-effector. To implement the force/torque thresholding, we instrumented the fork with a 6-axis ATI Nano25 Force-Torque sensor [269]. The center of the food (and the bounding box used for visual context in the online learning experiment described in Section 5.6) were annotated manually from the robot's eye-in-hand vision system. This was done in order to run a controlled experiment specifically focused on food acquisition, as opposed to introducing additional variance with a (possibly imperfect) food perception system. This system includes the Intel RealSense D415 RGBD camera and the NVidia Jetson Nano for wireless image transmission. Food was placed on a plate equipped with an AprilTag [225] for camera calibration and mounted on an anti-slip mat commonly found in assisted living facilities [28].

In each trial, the food placed in the vicinity of the AprilTag on the plate and oriented such that the major axis of the bounding ellipse is parallel to the bottom edge of the fiducial. The end-effector moved to a fixed position above the plate, and the location of the center of the food is annotated manually in the RGBD camera image. After action execution, we wait at least 3s before recording success or failure. For most food items, success is defined as the entire item being removed from the plate. If a homogeneous food item breaks (a common occurrence with banana slices), at least half of the item needs to end up on the fork. For the sandwich, success required that all layers (both pieces of bread, the lettuce, and the cheese) make it off the plate. Finally, for multi-piece and continuous items (i.e. potatoes, rice, noodles), a conservative success metric was set at 200mg ($\sim$15 grains of rice, or 1 full noodle).

Chapter B

Appendix for Chapter 6

Below is the appendix for Chapter 6 "A System for Out-of-Lab Robot-Assisted Feeding." As with Chapter 6, the original version of this appendix was published as part of "Lessons Learned from Designing and Evaluating a Robot-assisted Feeding System for Out-of-lab Use" at the *ACM/IEEE Conference on Human-Robot Interaction* in 2025 [212].

B.1 System

This section contains additional details of the system, beyond the core details presented in Sec. 6.3.

2.1.1 System Hardware

Figure B.1 shows a close-up of the elements mounted to the robot's wrist and end-effector: the eye-in-hand system, a fork assembly with a force-torque sensor, and a wireless force-torque transmitter.

2.1.2 Co-Design Sessions with the Community Researcher

Many components of the app were co-designed with community researchers. This includes the following.

We workshopped the app state-machine with a community researcher, discussing what the robot and user would do at each step. A key insight from this was the value of including "auto-continue" options so interested users can reduce the number of steps they have to take; this led to the "auto-continue" button before the robot moves to the user's mouth.

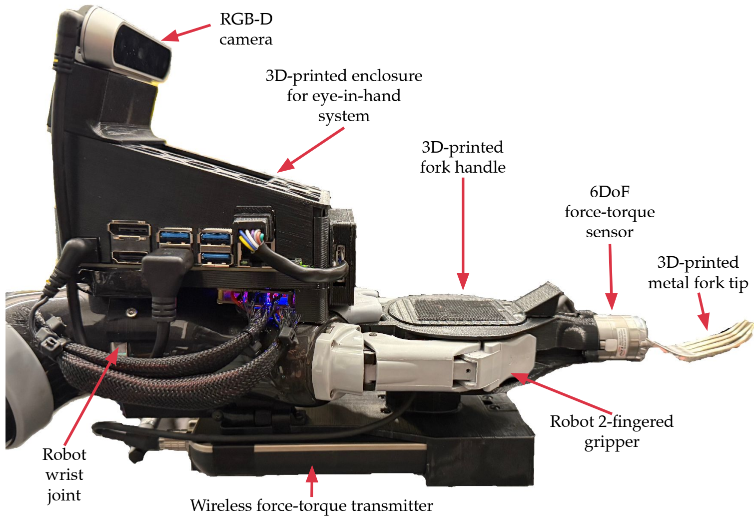


Figure B.1: A close-up view of the system's end-effector.

We co-designed the pause, back, resume, and retry options by having the community researcher verbally describe how he would like to control those aspects of robot motion, and creating a mock-up of that in real-time using a picture-editing application. He then gave us feedback, we discussed the pros and cons of the design, and continued until we converged on a user interface.

In a similar fashion, we co-designed the types of transparency the robot should provide the user as it is or is not moving. This involved teaching the community researcher the high-level concepts of robot planning versus motion. This co-design process led us to converge to the robot displaying elapsed time while it is planning but not yet moving, displaying the percent of motion it has completed while it is moving, and displaying a "lock" icon if it is on a screen where it will not move unless the user presses a button.

Finally, we ran a pilot study with a community researcher to investigate how to design the bite selection interface. We first introduced the community researcher to the concepts of object detection, segmentation, and classification. We then had him load a URL on his phone with 3 mock-ups of different bite selection interfaces: (a) the one described in Sec. 6.3.3.1; (b) an interface where the food segmentation algorithm segments several possible bites without a seed point, and the app renders them all to the user as buttons; and (c) an interface where the app renders semantic labels of the food type as buttons. This revealed several pros and cons about each. The first provides the most user control to select a bite, but some assistive technologies make it difficult to select an arbitrary point on an image. The second is more accessible since it has buttons, but can easily become cluttered. The third can be good for users who don't want to be so involved that they are selecting individual bites, but it provides users little recourse if the robot regularly misdetects a food. Through the discussion, we all converged to starting with the first option, and eventually adding the third option as a choice for users who want less control over their feeding process.

2.1.3 Planning

We use MoveIt2 with the Open Motion Planning Library (OMPL) [282] for path planning. This section contains additional details on our use of planning algorithms.

Across both studies, we use the default path length optimization objective, which minimizes the trajectory length in configuration space. Every path planning request has a corresponding time budget: 0.5 seconds for all motions but bite acquisition, which has 2.0 seconds. We generate 5 plans in parallel and hybridize between them [180]. Then, any time remaining in the budget is spent shortcutting the resulting trajectory [180].

The two studies differed in which planning algorithm we used. For Study 1, we use RRT* [154] as the planning algorithm, with all default parameters except `range`, which we set to 3.0 after some informal tuning that sought to balance between planning time and path length within a fixed time budget. For Study 2, in an attempt to speed up the planning times, we switched to RRT-Connect [163] with all default parameters (OMPL by default sets `range` to 20% of the maximum extent of the state space).

Most motions are kinematic plans that use the above pipeline. Of those motions, the ones to hard-coded configurations have goals specified as 6DoF joint goals. For bite acquisition's motion above the food, the goal is specified as a pose goal for the fork tip, and MoveIt2 samples multiple 6DoF joint goals from the inverse kinematics solver.

A few motions do not involve kinematic plans. For bite acquisition's motion into the food (approach), we use MoveIt2's default cartesian planner, which interpolates between the fork tip pose at the start and end, divides it into small intervals, and uses inverse kinematics to get joint configurations for each of those poses. For bite acquisition's grasp and extract, we use cartesian control to directly execute twists (linear and angular velocities) on the fork tip, using a selectively damped pseudo-inverse Jacobian [39, 265]. For the motions between the "staging" configuration and the user's mouth, we also use the aforementioned cartesian control approach to move the fork in a straight line to their mouth.

To avoid collisions, the robot's static planning scene includes meshes for the user's head, their body, the furniture they are sitting in (e.g., wheelchair or bed), and the table the food is on. The "body" mesh is a large hull intended to cover diverse body types. The "head" mesh moves and the "body" mesh scales based on the results of face detection. To prevent the robot from generating unnecessarily large motions, tight workspace walls are computed and statically placed in the planning scene to contain the user, the robot (in all of the hard-coded configurations), and some or all of their furniture. To account for un-modeled obstacles, such as user-specific assistive technology (AT) or the user's laptop, we use the depth image to populate an Octomap [133] at a resolution of 2cm. For plans involving bite acquisition, the robot is allowed to collide with the Octomap and table, because it is intended to come into contact with the food and sometimes the bottom of the plate.

Psychological safety and user comfort is crucial when there is a robot arm moving in close proximity to the user. We promote feelings of user safety with respect to robot motion in a few ways:

1. We reject any plan with joint rotations greater than a specified threshold[1], to avoid plans with large swivels that users might find scary or unpredictable.
2. For kinematic plans that may move near the face (i.e., the motions to the "staging" and "resting" configurations), we add a wall to the planning scene roughly 0.3m in front of the user's face. This is intended to keep motions away from the user's face.
3. As mentioned above, the fork moves in a straight line to and from the user's mouth, to promote interpretability of the robot's trajectory.

In terms of constraints beyond goal constraints, the system allows for orientation path constraints to be placed on the fork to ensure it remains face-up after acquiring foods. However, since all user-selected foods were skewerable, food falling off due to fork rotations was less of an issue. Thus, we did not use orientation path constraints during either study. Empirically, we found that adding orientation path constraints increased planning time roughly fourfold.

This work optimized planning times to the minimum amount necessary for the system to be deployable out-of-lab. Thus, we do believe the system's planning can be sped up considerably through, e.g., better

[1]For all motions, the plan was rejected if the sum of joint motion across all joints exceeded 10.0 radians. For acquisition, plans were additionally rejected if joint 1 exceeded $5\pi/6$ or joint 2 exceeded $\pi/2$ radians.

tuning of planner parameters, faster inverse kinematics computations, sending multiple goals to the planner, and using MoveIt's Python bindings as opposed to the ROS2 interface.

2.1.4 Food-on-Fork Detection Algorithm

Challenges: Detecting whether there is food on the fork in an RGB image is difficult due to reflections off of the metallic fork and due to the large diversity of food colors. Detecting whether there is food on the fork in a depth image is difficult because the fork is reflective and most of the fork tip is empty space; thus, whether or not the fork is even perceived in the depth image varies depending on lighting condition and objects behind the fork. Both are difficult due to the large variety of food shapes. Finally, detecting food on the fork with the F/T sensor is difficult due to hysterisis errors; after the sensor experiences high force during acquisition, it takes time to regain the level of sensitivity required to detect whether there is food on the fork.

Key Insights: Our approach hinges on two key insights:

1. Although food comes in a variety of shapes, *the fork has only one shape*.
2. Although the fork may or may not be perceived by the depth camera when there is no food on the fork, *food is always perceived when there is food on the fork*.

Algorithm Overview: Our algorithm uses depth images to memorize the shape of the fork without food. It then learns to predict the likelihood of food on the fork based on the deviation of an input depth image from that memorized shape.

Algorithm Details: Our algorithm operates on de-noised depth images[2], converted to pointclouds. During train time, it stores a representative set of points from the "no food on fork" pointclouds—essentially, it memorizes the shape of the fork[3]. Then, for each pointcloud, it computes the distance between each of its points and the closest point in the stored set, and then takes the 90^{th} percentile of those distances—essentially, this measures how far the farther points in the pointcloud are from the memorized fork shape. Finally, it trains a logistic regression classifier on those distances (x) and whether there is food on the fork (y). During test time, it passes a depth image through the same preprocessing steps, computes the 90^{th} per-

[2] During pre-processing, the algorithm crops the depth images to a rectangle around the fork, passes it through a temporal filter that only keeps depth points that are perceived across 5 consecutive images, and passes it through the morphological "opening" operation to remove isolated, noisy points.

[3] To reduce redundant points, it only stores a point if it is $\geq$ 1mm away from all other stored points

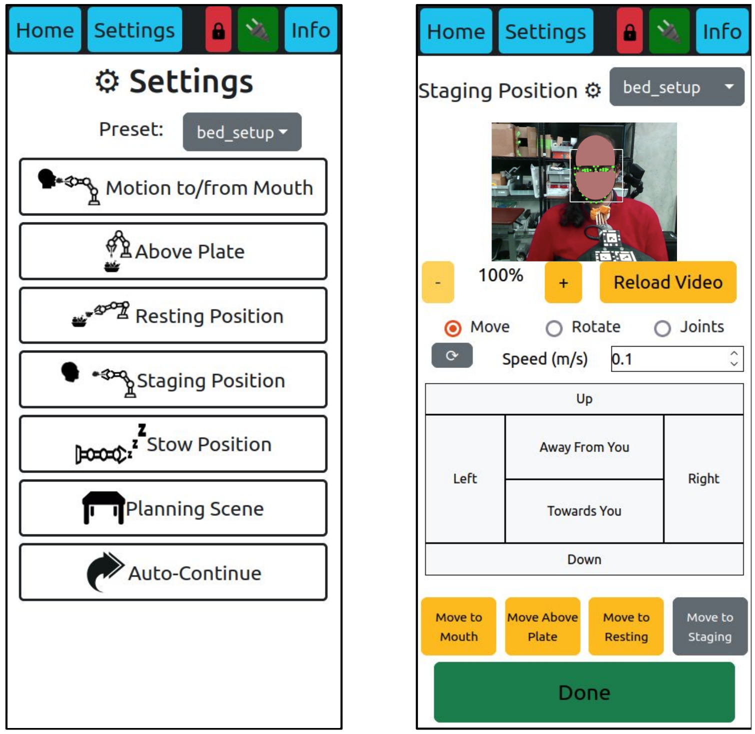


Figure B.2: (Left) The settings menu. (Right) The screen to customize the staging configuration.

centile difference between that pointcloud and the stored points, passes that through the logistic regression model, and uses the output as its confidence. If there are < 100 points in the depth image, it outputs `nan` as its confidence.

Usage: After bite acquisition, if the user has enabled "auto-continue," the web app toggles on food-on-fork detection and subscribes to its output. If, over the last 3 seconds, the output of food-on-fork is consistently `nan` of ≤ 0.25, the web app invokes the action to move the robot above the plate (i.e., acquisition failed). If it is consistently ≥ 0.75, the web app invokes the action to move the robot to the staging configuration. Else, the web app waits for user input. Similarly, when the robot is at the user's mouth, if the user has enabled "auto-continue," the web app toggles on food-on-fork detection and subscribes to its output. If, over the last 3 seconds, the output is consistency ≤ 0.25, the web app invokes the action to move the robot from the user's mouth and above the plate (i.e., the user ate the bite from the fork). Else, the web app waits for user input. Note that in the latter case, the web app waits for input on `nan` predictions, because it is possible to get too few points in the pointcloud while the user's mouth is on the fork.

Limitations: Although this model works, it is susceptible to slight changes in the pose and shape of the fork in the camera frame, which can occur if the fork bends, the camera moves, or the utensil changes. Another approach, which may be more robust, involves using a VLM to assess whether there is food on the fork [145][4].

2.1.5 User Interfaces for Customization

As mentioned in Sec. 6.3.4.1, the user is able to customize system parameters through a settings menu in the web app. Figure B.2 Left shows this settings menu. It allows them to customize: properties of the bite transfer motion (distance to mouth, speeds when approaching the mouth); key robot arm configurations that all motions start or end from; choose which planning scene to use; and change at which points the web app auto-continues. Figure B.2 Right shows the screen the app goes to if the user wants to customize the staging configuration. This illustrates several of the principles from "Designing for Tinkerability" [253] mentioned in Sec. 6.3.4.1. For example, the user is given "immediate feedback" by being able to transparently see the impacts of their parameter changes. This happens in two places: (a) seeing how face detection performs in the new configuration, at the top of the screen; and (b) allowing them to invoke actions to/from this configuration to see how the customization impacted robot motion, via the buttons at the bottom of the screen. As another example, the user is given "fluid experimentation" by having access to the full teleoperation interface, seen in the middle of the screen. The user can switch between "move," "rotate," and "joints" mode. The former two have 6 buttons, to move the robot in the positive and negative directions of each of the three cartesian motions in that mode. The latter mode has 12 buttons, to move each of the 6 joints in the positive and negative directions.

2.1.6 Users' Multiple Levels of Control

Figure B.3 graphically shows the multiple levels of control users have access to (Sec. 6.3.4.2), using the levels of control described in Beer et al. [21]'s framework. The solid line shows the nominal variation of level of control across a single bite. During bite selection, the level of control is "decision support," since the system presents the user with several masks, which they choose from. During all robot motion, the nominal level of control is "supervisory control," but the user can drop it down into "decision support" by pausing

[4] https://github.com/empriselab/FLAIR/blob/main/bite_acquisition/scripts/food_on_fork.py

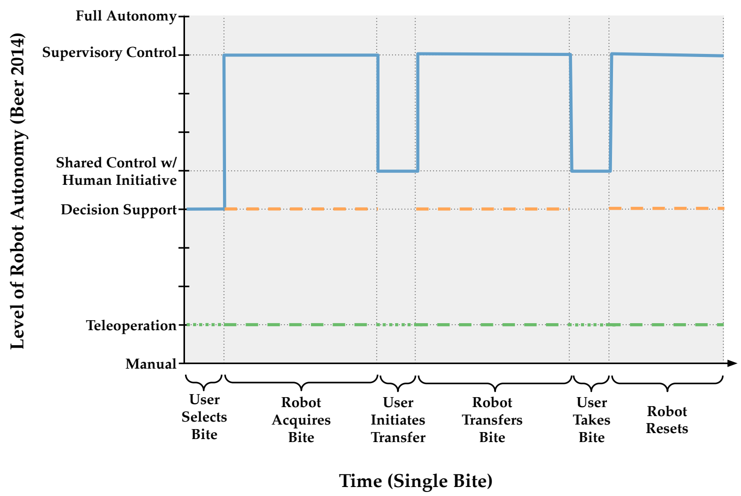


Figure B.3: The levels of control users have access to across each bite.

robot motion, and then into "teleoperation" if so desired. When waiting for the user to initiate bite transfer and to indicate that they are done with the bite, if the user toggles auto-continue on the system is in "shared control with human initiative," because the robot is automatically deciding whether to continue, but the user can override that decision. If the user has auto-continue toggled off, those stages are at the "teleoperation" level of control, since the user needs to specify when the system should move on. Finally, the user can bypass any of the perception stages altogether by teleoperating the robot from the previous stage onto the next (e.g., fully teleoperating bite acquisition removes the need for bite selection).

B.2 Health & Safety Protocols

Meals can involve health and safety risks. As a result, in both studies the research team strictly adhered to following:

Food Safety: All food was procured from: (a) a restaurant; (b) a grocery store; or (c) homemade by one of the user's caregivers. The only food preparation the research team did was re-heating, washing, cutting, and/or arranging the aforementioned food, oftentimes with direct input or supervision from the user or their caregiver. All utensils that came in contact with the food, including the robot's fork were washed with soap and water before every meal. The robot's fork was additionally washed with an alcohol wipe, in front of the user, before they began their meal. The research team washed their hands with soap and water before every meal, and used hand sanitizer before and after touching any food. At any time, the user could request the research team repeat any of the above food safety precautions.

Infection Control: All members of the research team followed the government department of health's COVID-19 prevention and safety guidelines. In addition, masks and hand sanitizers were available during meals, and at any time the user was allowed to ask team members to wear masks and/or take additional health and safety precautions.

Researcher Interventions: Any researcher was allowed to intervene in the meal on: (a) participant request; (b) unexpected or potentially dangerous robot behavior; or (c) perceptible participant distress. They were allowed to take whatever intervention necessary to rapidly resolve the issue, including but not limited to terminating the robot controllers' software or physically powering off the robot.

B.3 Comparison to Other Robot-assisted Feeding Systems

Table 6.1 presented a comparison of technical capabilities across contemporary robot feeding systems. This section provides concrete details on the criteria used for each column.

Approximate Cost. For commercial systems, if the system's website or medical tools catalogs mentioned a cost, we used that. If not, we looked through other online sources such as news articles about the technology, crowdfunded campaigns to raise money for the technology, etc. to determine the cost. For research systems, if the paper presented a cost we used that. Otherwise, we followed the aforementioned criteria for all commercially sold components of the system, and added the costs together.

Mounting. This refers to where the robot arm is placed before feeding. We determined this through textual descriptions and pictures of the system.

Autonomous Motion. This refers to whether the robot arm moves autonomously (✓) or whether users have to teleoperate its motions (✗), in nominal scenarios.

(General) Food Detection. This refers to whether the system can detect some (for "Food Detection") or any (for 'General Food Detection') bite-sized food items placed in front of the user, without requiring researcher intervention before or during the meal. Note that this column focuses on perception models that can detect masks or bounding boxes around food items, irrespective of whether the perception modules also add semantic labels to those masks. Different robot-assisted feeding works focus on different features for their food detection subsystem. For example, Jenamani et al. [145]'s food detection subsystem does not provide general food detection, because researchers must seed it with a list of food items on the plate. However, it does provide additional features not encompassed by this paper's bite selection subsystem. Specifically, their system can semantically segment foods and can segment non-bite sized food items (e.g., spaghetti or mashed potatoes).

Face Detection. This refers to whether the system can autonomously detect the user's face and mouth.

Collision Detection / Avoidance. "Collision Detection" refers to whether the system can detect a collision once it has occurred (and stop/modify its motion accordingly). "Collision Avoidance" refers to whether the system can preemptively avoid possible collisions, e.g., through its motion planning.

Portable & Self-contained. This refers to whether the system can be moved, with the user and caregiver, to the varied locations that users may eat in: e.g., at home, at a restaurant, at an outdoor picnic table, etc. Reasons a system may not be portable & self-contained include: the system requires wall power; the system is too heavy to move; the system has too large of a footprint to exist in the diverse environments people eat in; the system has wires that stretch across robot joints, which can get tangled, restrict robot motion, and be trip hazards.

User Can Stop / Restart Robot Motion. The former refers to the user's ability to stop the robot at any time, for any reason. The latter refers to the user's ability to restart the robot once they have stopped it. For example, a system that provides only an emergency stop (e-stop) button that requires a researcher to restart the system would satisfy the former but not the latter criterion. In contrast, a system that allows the user to stop robot motion via, for example, an app, and subsequently restart robot motion, satisfies both criteria.

Customizable Robot Motion. This refers to whether users can customize the autonomous motions the robot takes, for example customizing its speed, how close it gets to their mouth, the path it takes to get to their mouth, etc. Note that some systems that don't provide customizable robot motion do provide other forms of customizability, e.g., Jenamani et al. [145] allow users to customize the sequence the robot provides them bites in.

Multiple UI Modalities. This refers to whether the system allows users to use multiple interface types to interact with it, depending on their impairments and preferred assistive technologies (ATs). A system whose user interface is on a general-purpose computing device such as a smartphone or laptop implicitly satisfies this, since general-purpose computing devices typically are designed to be compatible with diverse ATs.

Note that Table 6.1 refers to the technical capabilities, as stated in the papers or websites. Table B.6, below, presents the *demonstrated* capabilities, i.e., the system's performance as demonstrated in a published user study.

B.4 Study 1: Multi-user, On-campus Study

This section contains additional details of Study 1, beyond the core details presented in Sec. 6.4.

Each user participated in a 30 minute virtual meeting before the study, followed by a 90 minute in-person session, where they ate an entire meal of their choice in an out-of-lab location. They received $25/hour compensation for their time, as well as compensation for travel to and from the study venue.

2.4.1 Virtual Session Details

We asked each participant to describe any assistive technologies that they use:

1. Do you regularly use a smartphone or tablet?
2. If so, how do you interact with the smartphone or tablet?
3. If you use an assistive technology (AT) to interact with the smartphone or tablet, where is that AT mounted?
4. If you use an AT to interact with the smartphone or tablet, how do you interact with the AT?
5. For the in-person study, would you be able to bring your smartphone or tablet, along with your preferred AT to interact with it?
6. Do you have other ATs mounted around your head or chair?

We also asked participants about their food preferences:

1. Do you have any allergies or other dietary restrictions we should be aware of?
2. As the in-person portion of the study will involve eating a full meal, what would you like to eat?

If the participant was having trouble identifying desired food items, we provided sample food items from the following list:

- **Proteins**: sandwich meats, chicken tenders, cheeses, blocks of baked or fried tofu, etc.

Figure B.4: An example of our in-person meal setup in a public cafeteria.

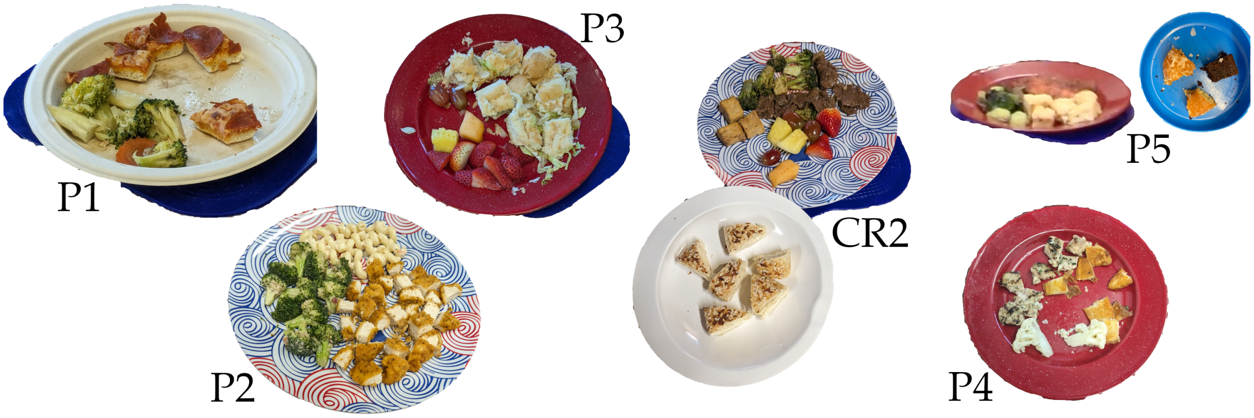


Figure B.5: Images of each user's plate of food for Study 1, taken at various points in the meal. Several users requested we serve them more food.

- **Vegetables / Fruits**: salads, roasted vegetables, crudites, fruit salad, etc.
- **Starches**: Potatoes, Rice, Bread, Noodles, etc.

Finally, we asked them about transportation logistics:

1. How do you anticipate coming to [the study venue]? What expenses are associated with your transportation?
2. Will someone be coming with you?
3. Is there anything else we can do to make the in-person study accessible to you?

2.4.2 In-Person Session Details

Figure B.4 shows a representative in-person study setup, for the meal in the cafeteria with CR2. Every in-person session had one camera zoomed in on the participant's phone or tablet, and another that captured the participant, robot, and social dining partner.

2.4.2.1 System Introduction

A researcher introduced each participant to the system as follows:

1. Set the system on a tripod next to the participant's wheelchair.
2. Mounted the emergency stop button in a location the participant could reach, and explained how to use it to stop the robot in the case of an emergency.
3. Assisted the participant in connecting their phone/tablet to the system's WiFi network and opening the system web app in a browser tab.
4. If needed, assisted the participant in setting up any assistive device to interact with the web application.
5. Walked the participant through completing one bite using the web application. Demonstrated safety features, such as the F/T sensor's ability to stop the robot when unexpected forces occur.
6. If requested by the participant, walked them through customizing how close to their mouth the robot gets.
7. Performed any necessary system adjustments, such as moving the tripod or plate, throughout the above process.

2.4.2.2 Pre-Post-Meal Questions

After the practice bite but *before* the full meal started, we asked the participant the following five-point Likert scale questions, where [AID_TYPE] is replaced by either "my caregiver," "my self," or both, depending on how the participant eats on a regular basis. At the end of the meal, we asked them these questions again, with [AID_TYPE] replaced with "the robot" and question tense shifted to past. Each answer was on the scale: Strongly Disagree, Disagree, Neutral, Agree, Strongly Agree.

1. When I eat with [AID_TYPE], I get my next bite when I want it, without waiting or feeling rushed.
2. When I eat with [AID_TYPE], I decide what food I want in my next bite.

3. When I eat with [AID_TYPE], I feel a sense of independence.
4. When I eat with [AID_TYPE], the meal requires a lot of mental energy.
5. When I eat with [AID_TYPE], the meal requires a lot of physical energy.
6. When I eat with [AID_TYPE], I am confident that I will remain safe during the entire meal.
7. When I eat with [AID_TYPE], I am confident that I will remain clean during the entire meal.

We then asked a final question (both pre- and post-meal): "How comfortable are you with the idea of being fed by a robot?" on the scale: Very Uncomfortable, Uncomfortable, Neutral, Comfortable, Very Comfortable.

Note that to accommodate their impairments, we asked all quantitative questions to participants verbally. Before asking any questions after the meal, we reminded participants that negative feedback is also very helpful for us to know how to improve the system.

2.4.2.3 Cognitive Workload (NASA-TLX)

Because we asked all quantitative questions to participants verbally, we modified the wording of the NASA-TLX [122] to be:

1. On a scale of 0-20, how mentally demanding has the task been? 0 is very low, 20 is very high.
2. On a scale of 0-20, how physically demanding has the task been? 0 is very low, 20 is very high.
3. On a scale of 0-20, how hurried or rushed has the pace of the task been? 0 is very low, 20 is very high.
4. On a scale of 0-20, how successful have you been at accomplishing the task? 0 is failure, 20 is perfect[5].
5. On a scale of 0-20, how hard have you had to work to accomplish your level of performance? 0 is very low, 20 is very high.
6. On a scale of 0-20, how insecure, discouraged, irritated, stressed, or annoyed have you been? 0 is very low, 20 is very high.

[5]When reporting this value (Table B.3), we flip it (i.e., 0 is perfect) to align with the original NASA-TLX.

2.4.2.4 System Usability Scale (SUS)

We asked the exact questions of the system usability scale [175], on a 5-point Likert Scale: Strongly Disagree, Disagree, Neutral, Agree, Strongly Agree.

1. I think that I would like to use this system frequently.
2. I found the system unnecessarily complex.
3. I thought the system was easy to use.
4. I think that I would need the support of a technical person to be able to use this system.
5. I found the various functions in this system were well integrated.
6. I thought there was too much inconsistency in this system.
7. I would imagine that most people would learn to use this system very quickly.
8. I found the system very cumbersome to use.
9. I felt very confident using the system.
10. I needed to learn a lot of things before I could get going with this system.

2.4.3 Between-Study System Patches

The order of participants in the study was: P11, P12, P13, CR2, P14, and P15. There were 6 day gaps each between P11 and P12 and between CR2 and P14[6]. This gave us the time to conduct system patches to address bugs revealed in previous meals. The applied system patches were:

1. Between P11 and P12, we:
   (a) Sped up all joint velocity limits by 66% on all kinematic motions, and sped up the cartesian motion velocity limit when moving to/from the mouth by 20%, from 0.1 to 0.12m/s.
   (b) Addressed a bug where food detection's depth readings would get skewed if the fork partially overlapped the bite.
   (c) Addressed a bug where if the user moves the arm from their mouth back to the staging configuration while "auto-continue" is checked, it will subsequently move right back to their mouth.
   (d) Modified the "MoveToMouth" action to reset the Octomap before starting; this is to account for phantom obstacles that accrue over time.

[6]There was a 0 or 1 day gap between all other participants

(e) Added the ability to zoom into the robot's camera feed during bite selection.

(f) Relaxed the deviation from goal position that is accepted when the robot is moving to the user's mouth from 0.5cm to 2.5cm.

(g) Addressed a bug where sometimes the joint state publisher's message timestamps are before the camera's, leading to failures in transforming between those frames of reference.

2. Between CR2 and P14, we:

   (a) Added an auto-restart process manager around the WebRTC signalling server, to address a known bug in a dependency that sometimes causes a segmentation fault.

   (b) Addressed a bug where food detection would return masks that have few valid depth readings (e.g., due to being too close to the image edge).

   (c) Addressed a bug where sometimes the arm would make a large, unnecessary swivel to get from one configuration to another, by rejecting plans where joints rotated being a certain threshold

   (d) Addressed a hardware issue where the bolts in one finger had loosened, resulting in the gripper holding onto the fork asymmetrically.

   (e) Replaced the e-stop button's adapter due to regular wear and tear.

   (f) Added a recovery behavior where the robot arm raises itself up 1cm if motions fail during bite acquisition (to prevent the case where the fork is left in contact with the table, causing all future actions to fail due to an unexpectedly high force sensor reading).

   (g) During bite acquisition, had the robot plan both the motion above the plate and into the food before executing. That way, any planning failures will happen while the robot is still above the plate, as opposed to after it starts moving.

Thus, P11 experienced the system with neither System Patch, P12, P13, CR2 experienced the system with System Patch 1, and P14 and P15 experienced the best version of the system, with both System Patches. Importantly, the system P11 experienced had the robot arm moving up to 66% slower than the system all other participants experienced.

2.4.4 Data Analysis

For the objective data analysis, one researcher watched the videos recorded from every sessions and tagged the timestamps of all key events in the video. A key event was defined as when the user interacted with the web app, the robot started/stopped moving, an off-nominal scenario occurred that was resolveable without researcher intervention, and an off-nominal scenario occurred that required researcher intervention. In addition, every time a bite acquisition or motion to the user's mouth ended, the researcher tagged the food type and whether or not it was successful. This resulted in a complete time profile of the meal, as well as a complete log of bite acquisition and transfer success rates. All the annotated data, with key events and timestamps per participant over the entire meal, can be found in Supplementary Materials, along with a codebook describing each key event.

For the subjective data analysis, we scaled all 5-point Likert scales to integers in the range $[-2, 2]$. For the NASA-TLX, we followed Hertzum [128]'s procedure of scaling each subscale to $[0, 100]$ and averaging them. For the SUS, we followed Lewis [175]'s procedure of setting missing values to "neutral," flipping negative subscales, transforming every subscale into the range $[0, 10]$, and summing them.

For both objective and subjective data, due to the small sample size, we do not analyze for statistical significance.

2.4.5 What Parts of the System Were(n't) Evaluated

Most of our system was evaluated in Study 1: the robot, fork holder and F/T sensor, e-stop button, all the robot code, and the web app. However, a few components were not included in Study 1, and later evaluated in Study 2. First, the only aspect of customizability included in Study 1 was: (a) customizing how close to the user's mouth the robot stops; and (b) toggling auto-continue on/off during face detection. Notably, since the other two auto-continues were not included in this evaluation, neither was the food-on-fork detection module. Second, we did not give users access to a teleoperation interface to control the robot. Thus, during robot motion, they had the "supervisory control" to drop the robot into "decision support," but could not drop it into the teleoperation level of control (Sec. 6.3.4.2). Finally, in order to improve reliability in the study, we paused online learning for deciding which action the robot should take (Sec. 6.3.3.2). As a result, the robot only executed one fixed motion primitive throughout the entire meal (except with P13, who requested

	Acquisition Success Rate Per Food
P11	Pizza: 0.78 (14/18) ; Broccoli: 1.0 (1/1)
P12	Chicken Tenders: 0.85 (11/13) ; Broccoli: 0.75 (9/12); Pasta: 0.33 (4/12)
P13	Sandwich: 0.94 (16/17) ; Strawberry: 0.24 (4/17); Brownie: 1.0 (7/7); Grape: 1.0 (2/2); Pineapple: 1.0 (1/1); Cantaloupe: 1.0 (1/1)
P14	Grilled Chicken: 1.0 (13/13) ; Potato: 1.0 (12/12); Cauliflower: 0.56 (5/9)
P15	Mac and Cheese: 0.7 (7/10); Brussel Sprouts: 0.86 (6/7) ; Salmon: 0.71 (5/7); Mushroom: 1.0 (2/2); Donut: 1.0 (2/2); Chocolate Cake: 1.0 (1/1)
CR2	Strawberry: 0.75 (3/4); Tofu: 1.0 (3/3) ; Grape: 1.0 (3/3); Cantaloupe: 1.0 (2/2); Broccoli: 0.5 (1/2); Bagel: 0.0 (0/2); Pineapple: 1.0 (1/1); Beef: 1.0 (1/1)

Table B.1: Bite acquisition success rates disaggregated by food type. Most successful foods ($\geq$ 3 bites) are highlighted.

we change the primitive so the robot could better acquire strawberries).

2.4.6 Results

2.4.6.1 Plates of Food

All food was acquired from local restaurants and grocery stores, based on the foods they had requested in the virtual session. We put these foods on a variety of plates, as shown in Figure B.5. Notably, the plates were of various colors and patterns (red, white, blue, patterned), shapes (flat, deep), and materials (paper, ceramic, melamine). The plates also did not necessarily have good color constrast with the food on it, as evidenced by the pink salmon on the red plate. This demonstrates the generalizability of the bite selection approach to diverse plate types.

2.4.6.2 Bite Acquisition

Table B.1 shows the complete bite acquisition data, disaggregated by food type. This reveals that for diverse food types, from sandwich bites to broccoli to pizza to salmon, the system is close to users' threshold acquisition success rate of 80%[7]. Further, note that because the online learning system was paused during this evaluation (Appendix 2.4.5), the robot was unable to learn from its failures and try different actions.

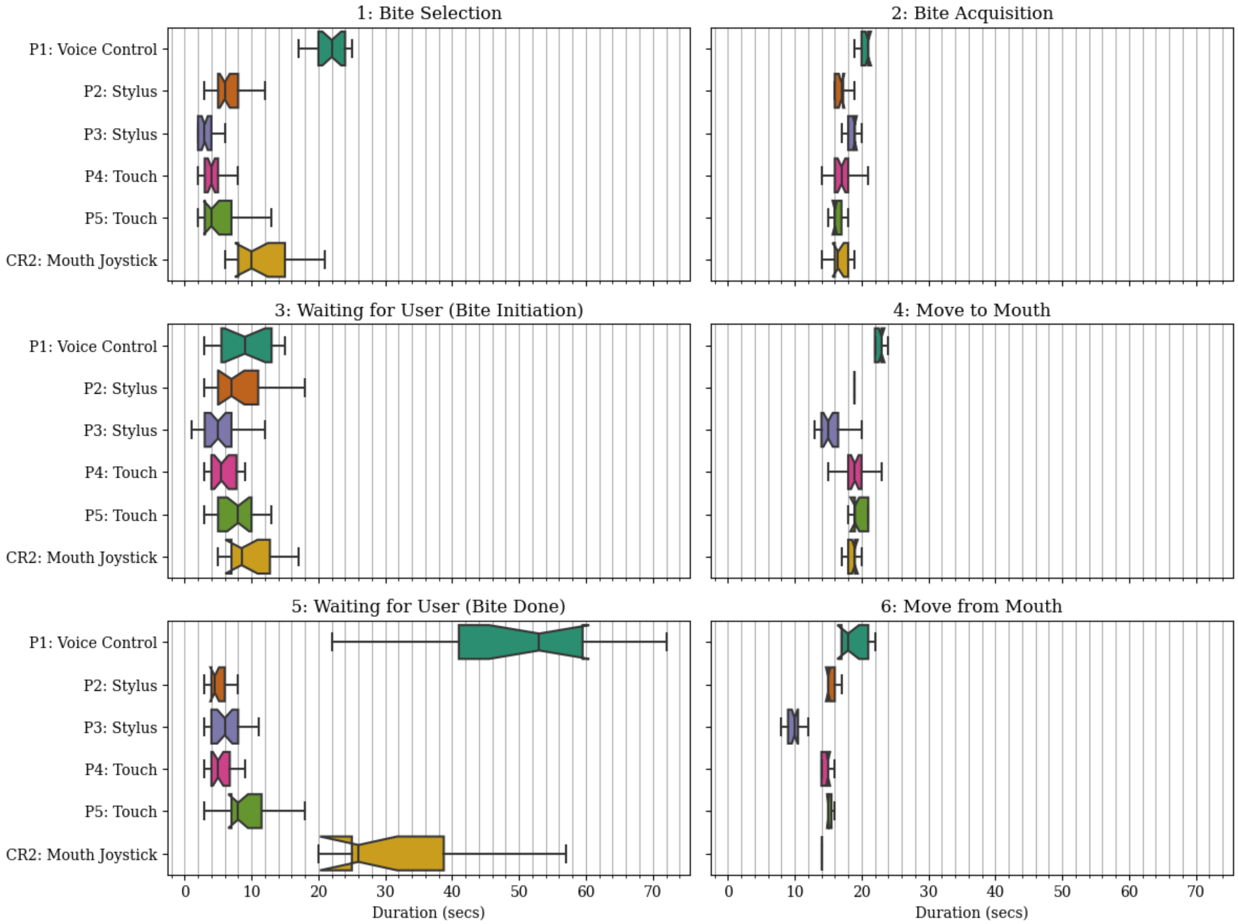


Figure B.6: A box-and-whisker plot showing the 25th, 50th, and 75th percentiles (vertical lines in box) for each of the 6 stages of robot-assisted feeding, for all users' successful bites. Notches (diagonal lines) show the 95% confidence interval around the median. Outliers excluded.

2.4.6.3 Time Profile Comparison Across Participants

Figure B.6 shows a comparison of the time each participant spent in each stage of robot-assisted feeding.

Although one might think that the reason P11 took the longest time per bite was that the robot was up to 66% slower for him than for other participants (Appendix 2.4.3), Figure B.6 reveals that the difference actually has to do with asssitive technology. Consider when the robot was "Waiting for User (Bite Done)." P11 used voice control to interact with his phone, which meant that he: (a) could not interact with his phone

[7]Although Bhattacharjee et al. [30] mention a 70% threshold bite acquisition success rate in the paper, we re-analyzed the raw data from that work, which was shared with us by the authors. The number presented in the paper is the arithmetic mean; however geometric mean tends to be more representative when the numbers are proportions or rations. The geometric meal of the user data is 80%. Thus, we use 80% as the threshold bite acquisition success rate, which also aligns with Gordon et al. [107].

while chewing; and (b) could not interact with his phone when he or others were talking. As a result, after the robot delivered a bite to his mouth, he waited until he was done chewing (which a spot check revealed took around 30 seconds per bite) and until there was a pause in the conversation before sending the robot back.

The other person who used mouth-based assistive technology, CR2, used a mouth joystick to interact with his phone. This meant that he could not interact with his phone while chewing[8], nor while he was talking, but could do so while *others* were talking. This added a layer of parallelization to the meal, which enabled CR2 to send the robot away from his mouth faster than P11 could.

The remaining users all interacted with their phone by using a stylus or touch, so they could interact with it while chewing or conversing. This added an additional layer of parallelization to the process, as those users could send the robot away from their mouth even as they chewed.

The impact of assistive technology can also be seen during "Bite Selection." Voice control is designed to click buttons. Although it is possible to tap arbitrary points, that involves a time-consuming process of zooming into a multi-layered grid to select the desired point to tap. As a result, it took P11 much longer than other participants to select his desired bite (median: 22 sec). CR2's mouth joystick is a pointer-based interface, but involves moving a cursor along the screen, which takes more time than directly tapping a point on the screen. Thus, CR2 took the second-most time on bite selection (median: 10 sec), followed by the remaining participants who used touch to interact with their devices (median: ≤ 6 sec).

This reveals the importance of not only ensuring the system *works* for diverse assistive technologies, but also considering how those assistive technologies *impact user experience*.

2.4.6.4 Time Profile Comparison to Caregiver Feeding

To compare the time profile of robot-assisted feeding to caregiver feeding, we analyzed a video of CR1 being fed a lunch of mixed berries and a protein bar by their caregiver. The motion to acquire a bite and transfer it to the user occurred in one smooth swoop, taking $1 - 3$ seconds. While the care recipient chewed, the caregiver acquired the next bite and was ready as soon as the community researcher finished chewing (around 15 seconds). This reveals a large space to improve our system's bite duration, both in terms of speed and

[8] CR2's mouth joystick requires him to suck air out of a straw to "click" a button, which can be a choking hazard if done while chewing

	P11		P12		P13			P14		P15			CR2	
	Care-giver	Robot	Care-giver	Robot	Care-giver	Robot	Self	Care-giver	Robot	Care-giver	Robot	Self	Care-giver	Robot
I get my next bite when I want it (↑)	1	0	2	2	2	1	2	-1	1	-2	1	2	1	2
I decide what food I want in my next bite (↑)	1	2	1	1	1	2	2	2	2	2	2	2	1	2
I feel a sense of independence (↑)	0	1	0	2	0	2	2	-2	2	0	1	1	-1	2
The meal requires a lot of mental energy (↓)	-1	0	0	-2	-1	2	0	0	-2	-2	0	-2	0	-1
The meal requires a lot of physical energy (↓)	-1	-1	-1	0	-2	-2	1	-2	-2	-1	-2	-2	-2	0
I am confidence that I will remain safe (↑)	2	1	2	1	2	0	2	1	2	2	1	2	1	1
I am confident that I will remain clean (↑)	0	1	-1	1	1	0	1	0	2	2	0	0	0	1

Table B.2: User responses to the pre-post questions. Highlighted values are the best per-question per-user.

parallelism. However, we note that not all people with motor impairments want their robot to feed them as fast as their caregivers: some feel that their caregiver's speed puts pressure on them [208]. The timestamped annotations from the video ofCR1's meal can be found in Supplementary Materials.

2.4.6.5 Robot-assisted vs. Caregiver Feeding

Table B.2 and Figure B.7 shows participant responses to all of the pre-post questions. Highlighted values refer to the aid type (caregiver, robot, or self) that performed highest for that participant on that question. As can be seen, the robot consistently performed as well or better than caregivers for "I feel a sense of independence" and "I decide what food I want in my next bite," and mostly outperformed caregivers on "I get my next bite when I want it." Further, note that for P14 and CR2, the robot outperforms the caregiver on nearly every question.

2.4.6.6 Cognitive Workload (NASA-TLX)

Table B.3 show participant responses to all NASA-TLX subscales, as well as their overall cognitive workload. We compare our results to Hertzum [128]'s's baseline mean and standard deviation, computed from 41 studies that used the NASA-TLX to evaluate cognitive workload during studies with "special-needs users." This reveals that for all participants but P13, our system involved less cognitive workload than the average

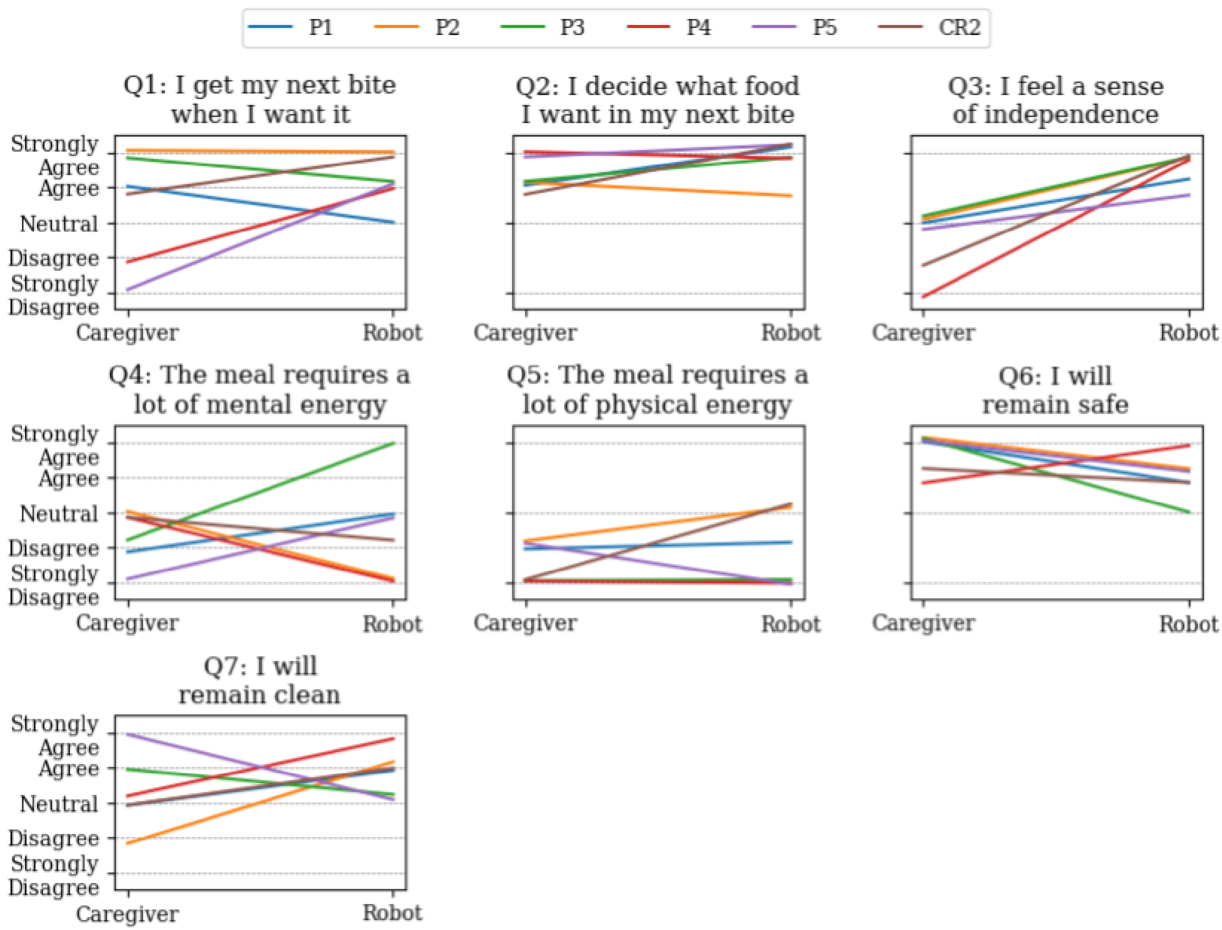


Figure B.7: All questions in users' self-reported comparison of eating with caregivers vs. the robot.

study with "special-needs users" from Hertzum's sample.

Note that the amount of physical demand required depended on the assistive device the participant used to interact with their phone or tablet. For example, in order to use her stylus to reach the top part of her phone (for bite selection), P12 needed to leverage the left armrest of her wheelchair to pull herself to the left, thereby getting the right angle to click with the stylus. Similarly, for CR2 to use his mouth joystick to tell the robot to move away from his face after he ate the bite, he had to move his head around the fork to the mouth joystick, which could be a complicated maneuver. On the other hand, P11's voice control required no additional physical demand for him to interact with his phone.

	P11	P12	P13	P14	P15	CR2	Baseline
Mental Demand (↓)	25	25	35	25	25	15	43 ± 16
Physical Demand (↓)	0	50	20	25	15	25	27 ± 12
Temporal Demand (↓)	0	25	10	50	0	25	33 ± 17
Perform-ance (↓)	30	25	40	0	25	10	44 ± 21
Effort (↓)	25	50	75	10	50	25	42 ± 17
Frustra-tion (↓)	25	0	50	10	0	15	31 ± 15
Cognitive Work-load (↓)	17.5	29.2	38.3	20	19.2	19.2	37 ± 11

Table B.3: Participants' cognitive workload (NASA-TLX) after eating with the robot. Baseline is mean $\pm$ standard deviation [128]. Highlighted values are less than baseline mean.

2.4.6.7 System Usability Scale (SUS)

Table B.4 shows participants' ratings for the subscales of the System Usability score (SUS) and the final score. We compute usability grades using the curved grading scale developed from a meta-analysis of over 400 studies that used the SUS [175]. As can be seen, 3 out of the 5 participants—P12, P14, and P15—and CR2 gave the system usability's ratings that corresponded to average or above-average usability (C or above).

2.4.6.8 Ranking Aspects of Feeding Systems

Table B.5 show user's responses when asked to rank the top three aspects of robot-assisted feeding systems to improve. 4/5 participants (excluding CR2) put speed as their top choice. Other aspects that commonly occurred across users were portability and independent use, safety, and customizability. These rankings provide pointers towards the most pressing areas of improvement necessary for the robot-assisted feeding research community.

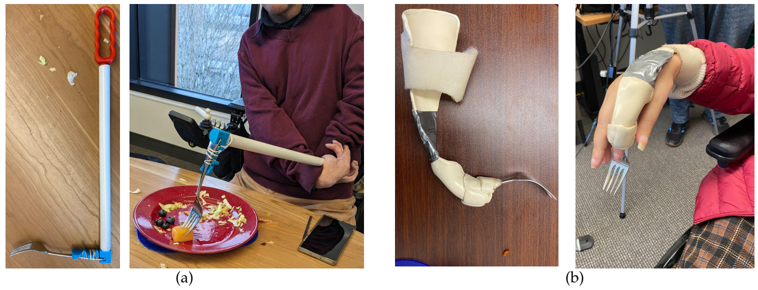


Figure B.8: (a) P13's custom 3D-printed self-feeding tool for grasping onto and maneuvering a fork. (b) P15's self-feeding tool for strapping a fork to her hand.

2.4.6.9 Self-Feeding Tools

Two participants had custom tools to enable them to feed themselves (Figure B.8). P13 had a custom 3D printed fork holder that enabled him to maneuver the fork to the plate and to his mouth, by using the table as a fulcrum, all while keeping his hands close to his lap and his face above the plate. P15 had a custom-designed strap that attached a fork tip to her hand, so she could use her arm to move the fork tip into foods and move them to her face. Importantly, all the foods P13 ate in the study were also skewerable by his self-feeding device. For P15, some pieces of brussel sprout were skewerable by her self-feeding device—if not too hard—and some pieces of fish were skewerable—if not too flaky—but overall the robot did feed her bites that she felt she may not have been able to skewer with her self-feeding device.

2.4.7 Comparison to Other Robot-Assisted Feeding Systems

Table B.6 compares the demonstrated capabilities of our system in Study 1 to the demonstrated capabilities of other research systems that had an evaluation with people with motor impairments. This shows that our system is the first to feed users entire meals of their choice in multiple out-of-lab environments. Additionally, our system fed users over 2× more food types than others (full list in Table B.1). However, the upper range of bite duration in our system is slower than most others.

[9] One user fed himself 130 bites over 6 sessions (Avg: 21.7), while the other 8 fed themselves 20 bites over 1 session.
[10] This bimanual manipulator holds the bowl in one hand and the utensil in the other, reducing the distance to traverse for acquisition and transfer.
[11] For these works, the one out-of-lab environment was an in-home environment, like that in our Study 2.

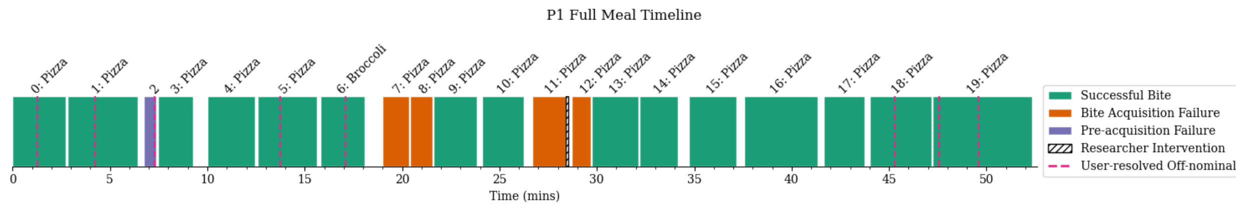


Figure B.9: The full timeline of P11's meal.

2.4.8 Off-Nominal Scenarios Per-Participant

This section contains a description of every off-nominal scenario, researcher intervention, bite acquisition failure, and successful bite whose duration was an outlier ($\geq 1.5 \cdot \text{IQR}$) relative to that participant's other successful bites.

2.4.9 P11 Study Details

Figure B.9 shows the timeline of P11's meal. Notable events include:

- **Bite 0 (User-resolved Off-nominal)**: The robot stopped too far from P11's mouth (likely due to a phantom obstacle in the Octomap; addressed with System Patch 1d). The user was able to get it to move the rest of the way to their mouth by clicking "retry" on the app.
- **Bite 1 & 6 (User-resolved Off-nominal)**: Although the robot moved close enough to the user's mouth for them to eat the bite, from the robot's perspective the action failed because it encountered an Octomap collision slightly before its goal. This is because the user leaned forward as the robot was coming in. The user resolved this by clicking "retry," letting the robot move the remaining short distance it wanted to, and then having it go back. (Addressed with System Patch 1f)
- **Bite 2 (User-resolved Off-nominal)**: The robot failed to plan to move above the food (addressed with System Patch 1g); user resolved by going back above the plate and re-selecting the bite.
- **Bite 5 (User-resolved Off-nominal)**: As with Bite 1, the robot moved close enough to the user's mouth, but thought it encountered an error (addressed with System Patch 1f). This time, instead of clicking "retry" the user mistakenly clicked "back," which took the robot back to the staging configuration. From there, "auto-continue" caused the robot arm to move back to his face (addressed with System Patch 1c), and then the user had it move back above the plate.

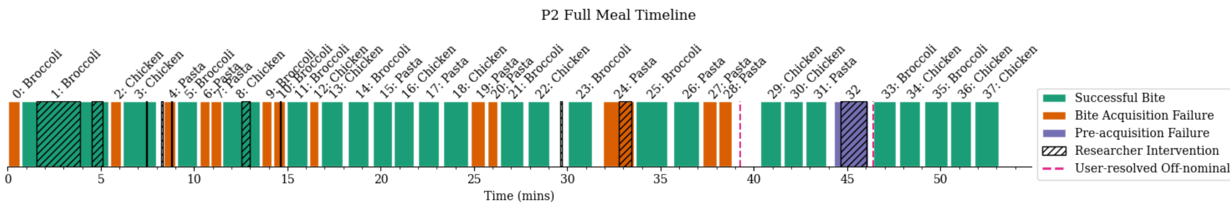


Figure B.10: The full timeline of P12's meal.

- **Bite 7, 8, & 11 (Bite Acquisition Failures)**: The robot misperceived the depth of pieces of pizza that were partially obscured by the fork, and therefore moved down too little towards the food. Addressed with System Patch 1b.
- **Bite 11-12 (Researcher Intervention)**: With the participant's consent, researchers moved the plate so that none of the pieces of food were partially obscured by the fork in the "above plate" configuration, to avoid the aforementioned issue with food depth misperception.
- **Bite 18 (User-resolved Off-nominal)**: Although bite acquisition visually succeeded, the robot thought it encountered an error. The user overcame this by having the robot move back above the plate, and then clicking "skip acquisition" on the bite selection page of the web app.
- **Bite 19 & 19 (User-resolved Off-nominal)**: First, none of the detected masks were the bite P11 wanted, so he selected another point on the robot's camera feed. Second, P11's phone mistakenly interpreted a number he was saying as part of the conversation as a button click. Thus, his phone opened the live video view of the web app. When he realized, P11 closed out of it.
- **Bite 19 (Outlier Bite Duration)**: The user continued the conversation for around 2 minutes with the robot in front of his face, before realizing and sending it back above the plate.

The additional system patches after P11 were based on his feedback that the robot should be sped up (System Patch 1a) and that the plate in the camera view is too small for bite selection (System Patch 1e).

2.4.10 P12 Study Details

Figure B.10 shows the timeline of P12's meal. Notable events include:

- **Bite 0, 12 (Bite Acquisition Failure)**: The user's selected mask was two bites of food together, leading the fork to pierce between the two.
- **Bite 1 (Researcher Intervention)**: The force-torque sensor disconnected from WiFi. To address this,

researchers lifted the backpack containing the router up off of the floor, restarted the force-torque sensor's code, and restarted the physical force-torque sensor.

- **Bite 3-4 (Researcher Intervention)**: Researchers restarted the WebRTC signalling server, which crashed due to a segmentation fault (addressed in System Patch 2a).
- **Bite 2-3, 4, 9, 10 (Bite Acquisition Failure), Researcher Intervention**: The fork wasn't centered on the bite, so researchers nudged the fork in the gripper to get it to better align with the robot's URDF model (addressed in System Patch 2d).
- **Bite 6, 7, 20, 24, 28 (Bite Acquisition Failure)**: The pasta rolled away as the robot was skewering it.
- **Bite 8, 22-23 (Researcher Intervention)**: Researchers restarted the force-torque sensor software due to it ceasing to send or receive messages.
- **Bite 19, 27 (Bite Acquisition Failure)**: The food segmentation algorithm only detected a subset of the pasta helix, not the whole piece, leading the robot to approach it not perpendicular to the piece.
- **Bite 24 (Researcher Intervention)**: Researchers reset the software because the app and robot stopped communicating.
- **Bite 25 (Outlier Bite Duration)**: Due to paying attention to the conversation, the participant waited after initiating bite selection and after the robot moved to the resting pose, thereby extending this bite.
- **Bite 28-29 (User-resolved Off-nominal)**: Force-torque sensor lost WiFi connection, but regained it shortly without requiring researcher intervention. The user re-initiated bite acquisition after force-torque connection was restored.
- **Bite 32 (Researcher Intervention)**: Researchers restarted the force-torque sensor hardware, due to an issue with receiving UDP packets.
- **Bite 32-33 (User-resolved Off-nominal)**: The robot's camera feed did not render on the app in bite selection. The user clicked the "reload video" button and then it did.

2.4.11 P13 Study Details

Figure B.11 shows the timeline of P13's meal. Unlike all other participants, P13 preferred to sit near the front of their wheelchair; this resulted in a much shorter distance to/from his mouth. Notable events in this meal include:

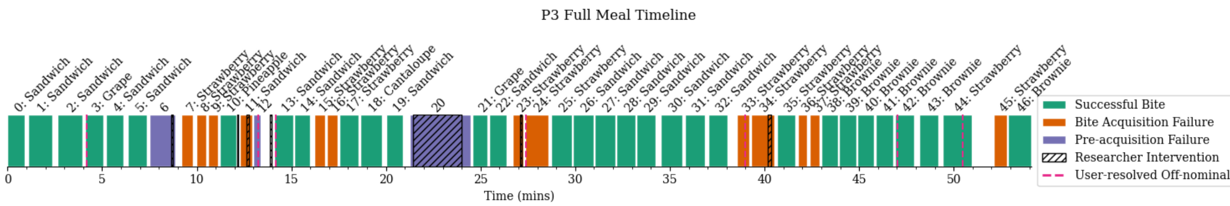


Figure B.11: The full timeline of P13's meal.

- **Bite 3, 13, 23-24, 42 (User-resolved Off-nominal)**: None of the masks returned from food segmentation aligned with the user's desired bite. To address this, they re-invoked food segmentation with a new seed point.
- **Bite 6 (Researcher Intervention)**: During acquisition, the fork missed the food and hit the plate. For some reason (perhaps being too close to a singularity) the extraction motion out of the food failed, leaving the robot in contact with the plate. This caused all future actions to fail due to the robot experiencing a higher force than the threshold. To address this, researchers briefly manually lifted the robot arm, getting it out of contact with the table, while the participant invoked the "move above plate" action on the app. Addressed in System Patch 2f.
- **Bite 7, 8, 9, 24 (Bite Acquisition Failure)**: Due to the strawberry being extremely soft, it slid off the fork as the robot was lifting the fork up.
- **Bite 10-11, 11, 23-24 (Researcher Intervention)**: Researchers nudged the fork in the gripper, so it better aligns with the robot's URDF model (addressed in System Patch 2d).
- **Bite 12 (User-resolved Off-nominal)**: Robot was unable to find a plan to move into the food. User clicked "back" and re-selected their desired food item. Partially addressed in System Patch 2g.
- **Bite 12-13 (Researcher Intervention)**: With participant consent, researchers moved the plate to be more centered on the robot arm in its "above plate" configuration, to increase the likelihood of motion success.
- **Bite 15, 34 (Bite Acquisition Failure)**: Robot arm pushed the strawberry out of the way as it was descending into it, because the curve of the strawberry aligned with the curve of the fork tines.
- **Bite 16, 33, 45 (Bite Acquisition Failure)**: Robot skewered the strawberry, but it fell off as the robot was moving to the "resting" configuration.
- **Bite 20 (Researcher Intervention)**: Robot did an extremely large, multi-part motion when moving

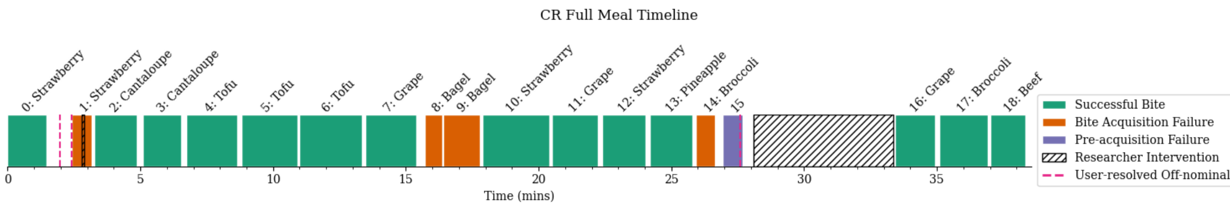


Figure B.12: The full timeline of CR2's meal.

above the plate. A researcher attempted to click the emergency stop button, which didn't register, so another resercher terminated the controllers. Swivel issue addressed in System Patch 2c, e-stop button issue addressed in System Patch 2e. Researchers restarted the software afterwards and the participant continued.

- **Bite 23 (Bite Acquisition Failure)**: Robot arm was off-center and missed the strawberry (addressed in System Patch 2d).
- **Bite 33 (User-resolved Off-nominal)**: User mistakenly initiated a bite transfer after a failed bite acquisition. He promply paused and had the robot arm go back.
- **Bite 34 (Researcher Intervention)**: On the participant's request, we re-started the code with the "vertical skewer" motion primitive hardcoded, as that acquisition action tends to have better success with strawberries.
- **Bite 44 (User-resolved Off-nominal, Outlier Bite Duration)**: The motion from the user's mouth failed soon after it started. The user his "retry" and it completed smoothly.

2.4.12 CR2 Study Details

Figure B.12 shows the timeline of CR2's meal. Although all other participants had their meals around a traditional lunchtime, CR2 had his meal between a traditional lunch and dinner time, resulting in him eating fewer bites before getting full. Notable events in this meal include:

- **Bite 0-1 (User-resolved Off-nominal)**: The robot's camera feed did not render on the app in bite selection. The user clicked the "reload video" button and then it did.
- **Bite 0-1 (User-resolved Off-nominal)**: None of the masks returned from food segmentation aligned with the user's desired bite. To address this, he re-invoked food segmentation with a new seed point.
- **Bite 1 (Bite Acquisition Failure)**: The robot was off-center and therefore missed the bite (addressed

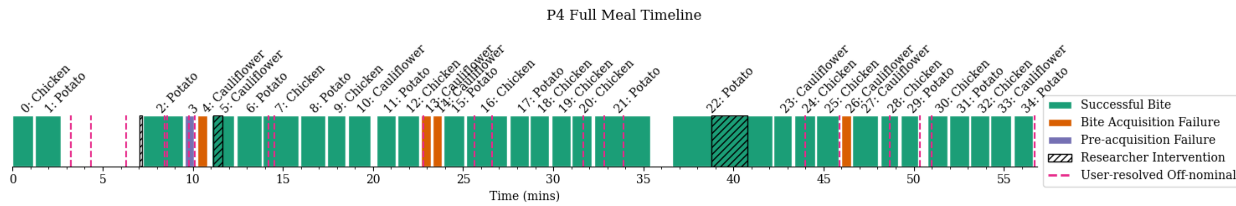


Figure B.13: The full timeline of P14's meal.

in System Patch 2d).

- **Bite 1 (Researcher Intervention)**: Researchers nudged the fork in the gripper, so it better aligns with the robot's URDF model (addressed in System Patch 2d).
- **Bite 8 (Bite Acquisition Failures)**: The fork didn't go all the way down to the food, perhaps due to inaccurate depth readings near the edge of the camera view. Addressed in System Patch 2b.
- **Bite 9 (Bite Acquisition Failures)**: The fork pushed the bagel piece to the side as it descended into the bagel, because the curve of the bagel aligned with the curve of the fork tines.
- **Bite 14 (Bite Acquisition Failure)**: The fork was off-center on the piece of broccoli (addressed in System Patch 2d).
- **Bite 15 (User-resolved Off-nominal)**: Robot was unable to find a plan to move into the food. User clicked "back" and re-selected their desired food item. Partially addressed in System Patch 2g.
- **Bite 15-16 (Researcher Intervention)**: Robot did an extremely large, multi-part motion when moving into the food. The participant clicked the emergency stop button, which immediately stopped the robot. Swivel issue addressed in System Patch 2c. Researchers restarted the software afterwards and the participant continued.

2.4.13 P14 Study Details

Figure B.13 shows the timeline of P14's meal. Notable events include:

- **Bite 1-2, 2, 7 (User-resolved Off-nominal)**: The force-torque sensor disconnected from WiFi (causing all robot motion to immediately stop), but reconnected shortly thereafter without researcher intervention.
- **Bite 1-2 (User-resolved Off-nominal)**: None of the masks returned from food segmentation aligned with the user's desired bite. To address this, they re-invoked food segmentation with a new seed point.

- **Bite 1-2 (User-resolved Off-nominal)**: When the user switched between apps on his phone, the robot web app rendered smaller than expected. He resolved this by reloading the page.
- **Bite 1-2 (Researcher Intervention)**: Since the plate location in the camera feed was too small for the user to click, researchers moved the plate as the participant zoomed into the image, ensuring the full plate was visible zoomed in.
- **Bite 1-2, 2, 3, 3-4, 7, 16, 21, 25-26 (User-resolved Off-nominal)**: A full-screen "sign in to Google" pop-up opened on the participant's browser. In two of those occasions, that caused the robot action to immediately be canceled by the web app (since it was no longer foregrounded). In all other occasions, the robot arm was already stationary, but nevertheless this off-nominal prevented the user from interacting with the system (akin to if they receive a phone call while eating). In all cases, the user closed the popup, clicked "resume" or "back" on the web app if robot action had been terminated, and continued his meal.
- **Bite 4, 13, 14, 26 (Bite Acquisition Failure)**: The fork pushed the cauliflower to the side as it descended into it, because the curve of the cauliflower top aligned with the curve of the fork tines.
- **Bite 5 (Researcher Intervention)**: Due to the force-torque sensor frequently losing WiFi connection, researchers raised the backpack containing the router up off of the ground.
- **Bite 13, 24, 28, 34 (User-resolved Off-nominal)**: The broswer's "History" tab opened full-screen, perhaps triggered by the participant's voice control due to the conversation. Details and resolution are the same as the "sign in to Google" pop-up above.
- **Bite 16 (User-resolved Off-nominal)**: The robot's code was hanging temporarily. The user used the app to pause, press back, and retry the action, which then succeeded.
- **Bite 20 (User-resolved Off-nominal)**: The broswer's "Option" tab opened full-screen, perhaps triggered by the participant's voice control due to the conversation. Details and resolution are the same as the "sign in to Google" pop-up above.
- **Bite 21 (User-resolved Off-nominal, Outlier Bite Duration)**: The robot encountered an error moving from the user's mouth (perhaps due to a phantom obstacle in the Octomap). The user left the robot there for a while as we were conversing and a researcher was serving him more food. Eventually the user hit "retry" and it resumed as expected.

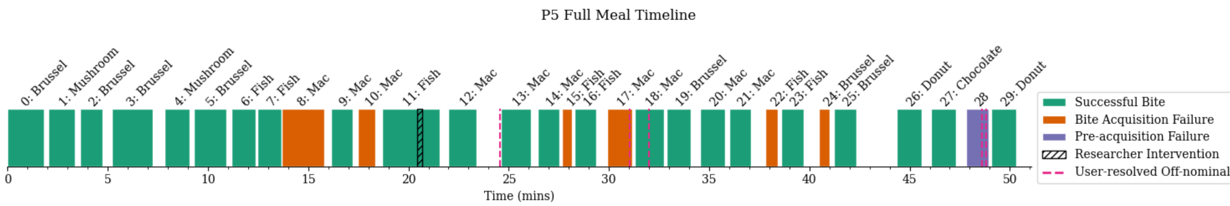


Figure B.14: The full timeline of P15's meal.

- **Bite 22 (Researcher Intervention)**: The bite acquisition ended in such a position where the only plan that could be found in the allotted time limit to move to the "resting" configuration involved a big swivel, which the threshold implemented in System Patch 2c was rejecting. Thus, a researcher terminated the code, increased the threshold, and restarted it.
- **Bite 29-30, 30 (User-resolved Off-nominal)**: The "live video" view of the app automatically popped up (perhaps the voice control running on the user's phone mistakenly heard a command and opened it). Both times, the user closed it so they could return to the main app screen and continue.

2.4.14 P15 Study Details

Figure B.14 shows the timeline of P15's meal. Notable events include:

- **Bite 3 (Outlier Bite Duration)**: The user forgot they have to tap a button to get the robot to move back from their mouth, and therefore left it at their mouth for around a minute as we were conversing, before remembering to click the button.
- **Bite 8, 10, 17 (Bite Acquisition Failures)**: The robot arm went between pieces of mac, acquiring nothing. This is also partly because those pieces of mac were oriented with the "hole side up," making it harder to skewer.
- **Bite 11 (Researcher Intervention)**: Participant mistakenly pushed the emergency stop button. To address this, researchers manually restarted the code.
- **Bite 13 (User-resolved Off-nominal)**: None of the masks returned from food segmentation aligned with the user's desired bite. To address this, the user re-invoked food segmentation with a new seed point.
- **Bite 15, 24 (Bite Acquisition Failure)**: The robot tilted the piece of food as it descended into it; thus, the fork tines did not skewer the food.

- **Bite 17 (User-resolved Off-nominal)**: User mistakenly initiated bite transfer when the robot hadn't acquired anything. On the "detecting face" screen, they clicked "move above plate" to have the robot return.
- **Bite 18, 28 (User-resolved Off-nominal)**: Robot action encounted an error. The user clicked "retry," and it proceeded smoothly.
- **Bite 22 (Bite Acquisition Failure)**: The robot was off-center and only acquired a tiny piece of fish. After this failure, the participant decided to acquire another piece of fish while that small piece was still on, which succeeded.
- **Bite 28 (User-resolved Off-nominal)**: The robot action was stalling and/or hanging. The user clicked "pause" and "back," which addressed it.

B.5 Study 2: Single-user, In-home Deployment

This section contains additional details of Study 2, beyond the core details presented in Sec. 6.5.

2.5.1 Caregiver Demographics

Table B.7 shows the demographics of the three caregivers who were present during the deployment. Although all worked with multiple care recipients, their familiarity with assistive technology came from the assistive technologies that CR2 used. Thus, they were all familiar with Alexa, voice control, mouth joystick, power wheelchairs, a hospital bed, ceiling lift, accessible van, and more. The questions "How familiar are you with assistive technology for people with motor impairments" and "How familiar are you with robots" were each on 5-point Likert scales: "Not at all familiar," "Slightly familiar," "Somewhat familiar," "Moderately familiar," and "Extremely familiar."

2.5.2 Study 2 Schedule & Overview

Figure B.15 shows the meal schedule for Study 2. Of the 5 consecutive days, 3 were wheelchair days and 2 were bed days. On wheelchair days, CR2 used the robot to feed himself breakfast and dinner. On bed days,

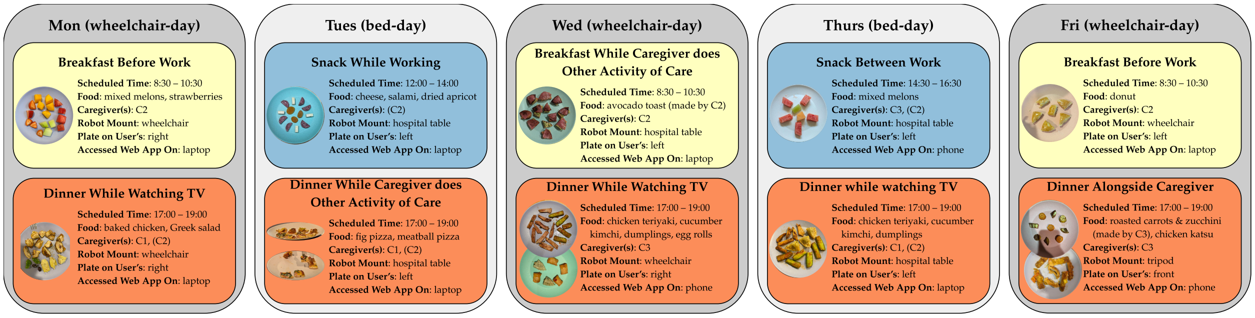


Figure B.15: An overview of the deployment schedule. Caregivers with names in parentheses were there for part, not all, of the meal.

he used the robot to feed himself snack and dinner[12]

We co-decided the meals with CR2, informed by his preferences and the robot's capabilities. We decided on the first 5 meals before the deployment began and on the latter 5 meals during the deployment week. The Wednesday breakfast (avocado toast) and part of the Friday dinner (roasted carrots and zucchini) were made by C2 and C3, respectively; all other meals were purchased from local stores and restaurants.

All meals but the Tuesday snack had a caregiver who was scheduled to be there for the entire meal. As a live-in caregiver, C2 popped into several of the meals for part of the time (typically the latter half): this occurred on the Mon dinner, Tues snack and dinner, and Thurs snack and dinner.

Over the course of the deployment, CR2 had the robot on his right, left, and front side. Across wheelchair days, he tried all three of the robot's mounts: wheelchair, hospital table, and tripod.

Over the course of the deployment, CR2 accessed the web app using his laptop and his phone. Regardless of the device, he used the mouth joystick to control the device.

2.5.3 Semi-Structured Interview Questions

The exact questions we asked CR2 during the semi-structured interviews varied based on the flow of the conversation. Below are a superset of questions we asked to start conversations (conversation-specific follow-up questions not included):

[12]Bed-days have more required activities of care in the morning. Since they are more rushed, CR2 opted to not use the robot to feed himself breakfast on those days.

2.5.3.1 Questions for the Care Recipient

After Every Meal

- What was the experience like of using the robot to eat this meal [while doing the co-occuring activity]?
- What challenges did the [co-occurring activity] introduce? How did you overcome those challenges using the robot? How could the robot be improved to better address these challenges?
- What surprised you about using the robot to eat this meal [while doing the co-occuring activity]? What would you do differently if you were to use the robot to eat this meal [while doing the co-occuring activity] in the future?
- As it stands right now, can you envision using the robot regularly to eat this meal [while doing the co-occuring activity]? If not, what would need to changed for you to envision yourself using it regularly?
- What changes in your environment or norms, if any, would you be willing to do for this robot to work?

After The Deployment

- Reflecting on this week, what went well? What went poorly? What surprised you?
- Think about your meal routine this week compared to your meal routine in past weeks. What aspects of the meal routine this week did you prefer compared to your meal routine in past weeks? What aspects of the meal routine in past weeks did you prefer compared to your meal routine this week?
- For each of the following contexts, what went well and what went poorly about eating a meal with the robot?
  - **Location**: In-bed vs. wheelchair
  - **Time**: Breakfast, Snack, Dinner
  - **Co-occurring activity**: working, watching a movie, conversing, while a caregiver does another activity of care
  - In which of these contexts would you like to continue using the robot-assisted feeding system? In which would you prefer being fed by a caregiver?
  - Do more contexts come to mind in which you'd like to try using the robot-assisted feeding system?
- Let's walk step-by-step through each part of the robot-assisted feeding system. Please share any

reflections or feedback you have on those components, both for the web app and the robot.

  - Customization
  - Bite selection
  - Bite acquisition
  - Bite transfer
  - Auto-Continue

- Is there anything else you would like to share with us?

2.5.3.2 Questions for Caregivers

- Based on what you've seen of the robot arm, what do you think are the benefits of a robot-assisted feeding system? What are the drawbacks?
- What might change in your caregiving routine if CR2 had access to a robot-assisted feeding system?
- Would you feel comfortable working in a house where the person feeds himself with a robot-assisted feeding system? Why or why not?
- What do you think of the setup procedure for the robot (explain it if need be). Would you be willing to set up the robot for CR2? What type of setup procedure would you like? How can we make the setup procedure simpler for you?
- How do you think such a system would impact CR2's health and well-being?
- Is there anything else that you would like to share with us?

2.5.4 Data Analysis

Initial transcription of all quotes was done by OpenAI's Whisper speech recognition model[13]. Subsequently, one researcher listened to all the video recordings and corrected mistranscribed participant quotes, both during the meal and after the meal. That researcher then used thematic analysis [274] to tag quotes with their key themes.

[13] https://apps.apple.com/us/app/whisper-transcription/id1668083311?mt=12

	P11	P12	P13	P14	P15	CR2
I think that I would like to use this system frequently. (↑)	1	1	-2	2	-1	2
I found the system unnecessarily complex. (↓)	1	0	0	-2	-2	-1
I thought the system was easy to use. (↑)	0	1	1	2	1	2
I think that I would need the support of a technical person to be able to use this system. (↓)	-1	-1	2	-1	-2	1
I found the various functions in this system were well integrated. (↑)	0	1	0	1	2	1
I thought there was too much inconsistency in this system. (↓)		0	1	-2	-2	-1
I would imagine that most people would learn to use this system very quickly. (↑)	1	1	-1	2	2	2
I found the system very cumbersome to use. (↓)	-1	-1	-1	-2	1	-1
I felt very confident using the system. (↑)	1	1	0	2	1	2
I needed to learn a lot of things before I could get going with this system. (↓)	-1	1	-1	-1	-1	-2
System Usability Score (SUS) (↑)	62.5	65	42.5	92.5	77.5	82.5
SUS Grade (↑)	D	C	F	A+	B+	A

Table B.4: Participants' usability ratings for the robot-assisted feeding system. Highlighted ratings are at-or-above average.

	First	Second	Third
P11	Speed	Robustness to errors	Easy user interface
P12	Speed	Robustness to errors	Safety
P13	Portability & independent operation	Customizability	Non-intrusive in daily routine
P14	Speed	Portability	Customizability
P15	Speed	Speed	Customizability
CR2	Safety	Portability & usability	Robustness to errors

Table B.5: Users' rankings for the most important aspects of robot-assisted feeding systems to work on.

Robot	Num Food Types	Num End-Users Fed (Out-of-Lab)	Num Fed Bites Per User Per Session	Avg. Bite Duration Per User (sec)	User-Decided Meal?	Entire Meal?	Arbitrary Plate?	Num Out-of-Lab Environments	Validated Metrics?	Avg. TLX Score
Song and Kim [276] / [277]	– / –	7 (–)	–	–	–	✓	✗	–	–	–
Song et al. [277]	–	14 (–)	–	–	–	✓	✗	–	–	–
Park et al. [236]	8	9 (1)	20–21.7[9]	41–78 [10]	✗	✗	✗	1[11]	TLX	18.6
Nguyen [218]	1	1 (1)	10	330	✓	✗	✓	1[11]	–	–
Bhattacharjee et al. [30]	3	10 (0)	15	90	✗	✗	✗	0	–	–
Jenamani et al. [144]	2	13 (1)	7–13	–	✗	✗	✗	1[11]	–	–
Jenamani et al. [145]	5	1 (0)	–	–	✗	✓	–	0	–	–
This paper's Study 1	22	6 (6)	14–31	62–165	✓	✓	✓	3	TLX, SUS	23.9

Table B.6: Comparison between the demonstrated capabilities from Study 1 versus other robot-assisted feeding systems' demonstrated capabilities.

ID	Age Group	Gender	Years as Caregiver	Years Worked with CR2	Live-in?	Impairments of Care Recipients	Num Deployment Meals	Familiarity with Assistive Technology	Familiarity with Robots
C1	25–34	F	0.5	0.5	✗	SCI, muscular dystrophy	3	"Somewhat familiar"	"Slightly familiar"
C2	55–64	M	25	25	✓	people with motor impairments	8	"Extremely familiar"	"Somewhat familiar"
C3	35–44	F	7	7	✗	people with motor impairments	4	"Extremely familiar"	"Not at all familiar"

Table B.7: Caregiver demographics for Study 2.