Affective interpretations of assisted driving interventions on a smart-wheelchair

An exploratory study

by

Ariadna M. Estrada Gaspar

B.Sc., Mechatronics Engineering, Universidad Autonoma de Ciudad Juarez, 2015

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF

Master of Science

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES
(Computer Science)

The University of British Columbia
(Vancouver)

December 2018

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the thesis entitled:

Affective interpretations of assisted driving interventions on a smart-wheelchair

submitted by **Ariadna M. Estrada Gaspar** in partial fulfillment of the requirements for the degree of **Master of Science** in **Computer Science**.

Examining Committee:

Ian M. Mitchell, Computer Science
Supervisor

Cristina Conati, Computer Science
Additional Examiner

Abstract

Allocating authority appropriately between humans and machines in shared control applications is crucial for the performance of the system. Particularly in the context of collaborative wheelchairs, the arbitration should be sensitive to user needs and preferences in order to avoid confusion and frustration. Current approaches to shared control for wheelchair navigation have been designed to handle objective and functional information such as goals and system states with limited analyses to subjective information such as the user's feelings when an assisted driving intervention is introduced. This thesis explores user affective responses on smart-wheelchairs as a potential communication channel through which users could interact more effectively with their smart mobility device. We present an implementation of shared control paradigms from the smart-wheelchair literature and results from a study where participants reported their affective interpretation of the emerging behaviours.

Lay Summary

Emotions play a meaningful role in social interactions and human behaviour. The same applies to interactions between human and robots when they are trying to work together to achieve a common goal. The objective of this thesis is to investigate how humans perceive non-verbal behaviours in robotic mobility assistance technologies at an emotional level. In this work, multiple assisted driving interventions were implemented on a smart-wheelchair and judged according to three basic dimensions of emotion: evaluation, potency, and activity. Based on this work, we aim to inform the design of emotion-aware collaboration strategies for smart-wheelchairs which deliver not only the right amount of assistance but also provide it in a way that best aligns with the user's preferences and emotional states.

Preface

All of the work presented henceforth was conducted in the Collaborative Robotics Laboratory at the University of British Columbia under the supervision of Dr. Ian M. Mitchell. The study reported in Chapter 4 was reviewed and approved by UBC's Behavioural Research Ethics Board (BREB) under certificate number H18-00256.

I collaborated closely with my supervisor, Dr. Ian M. Mitchell, in the development of the the work presented in this thesis. Under my supervision, Haoyu Yang, an undergraduate research assistant, and two graduate students, Jocelyn Minns and Zicong Fan, helped with the execution of the research. Haoyu contributed with early implementations of the intervention modes and the instrumentation of the powered wheelchair. Jocelyn and Zicong participated in pilot runs of the experiment described in Chapter 4 and provided valuable feedback on the experimental procedure.

Table of Contents

List of Tables

List of Figures

Glossary

AC Affective Computing

ACT Affect Control Theory

ADAS Advanced Driving Assistance Systems

AMCL Adaptive Monte Carlo Localization

DSC Dynamic Shared Control

EMFACS Emotional Facial Action Coding System

HRI human-robot interaction

ICICS Institute for Computing, Information and Cognitive Systems

MANOVA multivariate analysis of variance

PS3 Play Station 3

PWC powered wheelchair

ROS Robot Operating System

SD Semantic Differential

SDK Software development kit

SLAM Simultaneous Localization and Mapping

Acknowledgments

I want to express my sincere gratitude to my supervisor, Dr. Ian M. Mitchell for giving me the opportunity to work in this project and his continuous guidance and support during my time at UBC.

Next, I would like to thank Bryan who kept me company countless nights at the lab, proofread this thesis, listened to me practice every presentation, and continues to show me so much more love than I could ever imagine.

I would also like to thank my second reader, Dr. Cristina Conati, for taking the time to read my work and provide insightful comments.

Finally, I would like to thank my friends and family for their constant encouragement and for making this time much more pleasant. To my friends from the program, Itrat and Candice, thank you for the beautiful moments, I will always cherish our times together. A big thanks to my best friend Diana, for continually checking up on me and not letting the distance come between our friendship. A special thanks to my parents, Ivan, and Nelly for visiting me in Vancouver, and for always believing in me. I love all of you, and I would not have made it without you.

Dedication

In loving memory of my grandfather Don Simón Gaspar. Deep in my heart you will always stay, loved and remembered every day.

Chapter 1

Introduction

The world's aging population is growing, and so is the demand for technologies that support healthy aging. Mobility is fundamental for older adults to maintain their wellness and independence. Mobility limitations have been associated with negative social, mental, and physical outcomes such as poor performance of daily activities, social isolation, depression, and anxiety [1]. Older adults represent the largest group of users of wheeled mobility devices, and it is estimated that 49% of older adults living in Canadian institutional settings use a wheelchair [2, 3]. While traditional manual wheelchairs satisfy the needs of many individuals, a significant segment of the older adult community finds it difficult or even impossible to propel themselves, due to further physical, perceptual, and/or cognitive limitations [4]. The potential of powered wheelchairs to enhance the well-being of older adults and support independent mobility is well documented; some of the benefits of independent mobility include improved self-esteem [5], increased levels of activity and social participation [6], decreased dependence on caregivers and family members [7], reduced pain and discomfort, and overall improvement on quality of life [8]. Nonetheless, older adults often experience a range of additional challenges such as cognitive and sensory impairments or other motor and coordination conditions [9]. Unfortunately, difficulties related to such conditions result in the exclusion of older adults with cognitive impairments from utilizing powered wheelchairs due to safety concerns for the users and those around them [2].

To support independent mobility among this population, researchers have leveraged technologies initially developed for mobile robots to create smart-wheelchairs [4]. A smart-wheelchair typically consists of a standard powered wheelchair customized with a computer and a collection of sensors [10]. The majority of such systems aim to increase safe and independent mobility by providing assisted driving interventions for collision avoidance and navigation support [11]. However, these systems typically focus on objective and functional information such as goals, and system states with limited analyses to subjective information such as the driver's emotions when an assisted driving intervention occurs.

In this thesis, we analyze the effects of various assisted driving behaviours of a powered wheelchair at an emotional level as a potential communication channel to enhance the user experience the overall system performance. Emotional factors play a significant role in decision making [12], and have been found to be crucial for enhanced safety and comfort in driving tasks [13]. Moreover, recent research regarding older adults with dementia shows evidence that emotional processing remains considerably more intact than cognitive processing [14, 15]. In light of this finding, we propose to incorporate the user's affect as a state variable that can potentially inform the design and operation of assisted driving interventions, as it may enhance the usability of such systems.

1.1 Research goal

The long term goal of our research is to fulfill the mobility needs of older adults with mild to moderate cognitive impairments to safely and independently navigate a powered wheelchair. The aim of the work described in this thesis is to investigate users' affective response to driving intelligent wheelchairs under different navigation assistance modes. Specifically, we want to know whether different types of shared (or collaborative) control interventions have similar patterns of affective meanings, and if such responses are consistent across different emotion-measuring instruments. Through the study described herein, we aim to answer the following primary research questions:

1. What affective interpretation do users attribute to different smart wheelchair behaviours? Are those interpretations consistent across users?

2. How does the user's affective state change when interacting with various

wheelchair behaviours? Moreover, are those changes consistent across users and consistent with the affect attributed to the wheelchair?

Much of the motivation for our research lies in designing assistive technologies that adapt to people's needs by acknowledging that emotions significantly influence human behaviour. If the different wheelchair assistance modes have consistent affective ratings and we can predict the user's affective state reliably, we can design a system that switches between assistance modes according to the user's emotions in addition to traditional objectives such as safety, ease of use or timeliness. A future goal of this research is to identify the features in the driving experience that correlate with specific affective states. For example, more than recognizing very strong emotions like anger or fear, we want to answer questions such as what kind of events make users feel calm or nervous or frustrated? We intend to use the principles of affect control theory (ACT) to address these questions and potentially develop an emotion-aware smart-wheelchair.

1.2 Contributions

Within the smart-wheelchair literature, the most common way to validate a shared-control paradigm is to compare it against no intervention, full automation, or a simple blending strategy. However, little attention has been paid to evaluating the driver's feelings when the interventions are introduced. For example, the collaborative system may be able to sense that there is not enough clearance to go through a particular space, or it may know a better path (according to the planner's quality metric), but depending on how this information is delivered to and assimilated by the user, they can become confused or anxious when the wheelchair appears to act on its own. Moreover, the cognitive capabilities of older adults may change quickly over time, even as fast as in the course of the day. For example, dementia patients and elderly institutionalized patients often experience "sundowning," a phenomenon of disruptive behaviour and agitation worsening in the late afternoon or evening [16]. Therefore, we believe that a smart-wheelchair system for older adults with cognitive impairment should be more aware of the driver's current affective state in order to deliver the right amount of assistance. The contributions of this work are:

- Implementation of a set of assisted driving interventions for powered wheelchairs based on shared control policies described in the literature, including low-level collision avoidance and high-level goal pursuit.
- Design and execution of an experiment to assess the effects of these interventions on the driver's emotions by characterizing the behaviours with respect to three basic dimensions of emotional experience.

1.3 Document organization

The remainder of this thesis is organized into four chapters. Chapter 2 outlines related work in the fields of human-robot collaboration, smart-wheelchairs, and emotion-aware assistive technologies. Chapter 3 illustrates the hardware and components of our smart-wheelchair and provides details on the assisted driving algorithms implemented. Chapter 4 is dedicated to the two research questions listed before. It describes the design and findings of an exploratory study conducted to understand the affective responses to wheelchair assisted driving interventions. Chapter 5 presents our conclusions and possible future directions for this research.

Chapter 2

Background

In this chapter, we elaborate on two streams of research: shared control and affective computing. Following this is an overview of related work on addressing emotions in driving contexts as well as assistive technologies for older adults.

2.1 Shared control in human-machine collaboration

Over the last decades, the limitations on human-machine interaction have been widely reported [17]. While autonomous solutions to many well-structured situations have been accepted in our society, there are still many tasks where a high level of automation is not suitable as it can lead to undesirable effects, especially in the control of safety-critical dynamic processes in unpredictable environments [18]. For example, in the automotive field, Advanced Driving Assistance Systems (ADAS) are becoming increasingly popular while entirely autonomous vehicles still face acceptance challenges [19]. An alternative solution to complete automation is *shared control*. In a shared control framework, a single control signal is generated by combining control signals from multiple agents; for example, a human controller and some form of automation. The main objective of shared control is to keep the human in the control loop while providing continuous support [17]. Such approaches have been explored in a variety of applications; for example, minimally invasive telesurgical training [20], manipulator teleoperation [21], aircraft piloting [22], and automotive applications [23, 24].

Shared control differs from *supervisory control* where the user manages the robotic system while it behaves autonomously [25] and *switched control* in which the person and the automation exchange turns in generating the control signal according to some system states [26].

Two prominent categories of shared control can be distinguished in the literature, namely mixed-initiative shared control and haptic shared control [17]. In mixed-initiative systems, the control signal is a combination of the output of the human's control interface and the output of some autonomous system. For example, adaptive cruise control in which a drive-by-wire system adjusts the speed to maintain a time-based separation from the vehicle in front. For haptic shared control, the final control input is determined by an interface on which both the human and the automation can exert forces [18]; for example, an automobile steering wheel which pushes against the driver's hands to prevent lane departure.

2.1.1 Smart-wheelchairs and older adults

Even though powered wheelchairs are becoming an increasingly common solution to enable independent mobility, a significant portion of older adults are not allowed to use one due to cognitive, motor, and/or sensory impairment [2]. The potential of smart-wheelchairs has been recognized since the 1980s to help overcome the safety and usability concerns and enable safeguarded mobility for everyone [27].

Diverse software and hardware solutions have been developed by numerous research groups to enable safe and independent mobility using powered wheelchairs. While fully autonomous solutions have been proposed, it has been observed that powered wheelchair users want to remain active drivers, and cede only the minimum control possible to the machine [27, 28]. Furthermore, a fully autonomous wheelchair might have adverse outcomes, such as loss of residual cognitive capabilities [29], or user confusion and frustration when the automation generates motion in a situation where the user was not expecting such action. Since high levels of automation can lead to undesirable effects, many smart-wheelchairs offer a variety of semi-autonomous control modes. Most commonly, the semi-autonomous modes use shared control through input mixing, where the resulting motion is the combination of the output of some robotic planner and the user's commanded velocity coming from an input device such as a 2-axis joystick or a sip-and-puff interface. The main objective of using such a framework for wheelchair navigation, is to keep the human in the control loop while providing continuous support. For example, humans are typically good at making global plans, whereas machines are good at fine motion control. Thus, shared control complements the user's skills rather than removing them from the equation.

For collaborative systems such as the smart-wheelchair, how and when the automation intervenes are two issues that determine the overall effectiveness of the system [30]. Moreover, finding the optimal parameter which dictates how much control is allocated to the user and how much to the robotic planner is a hard problem because it depends on the kinematics of the wheelchair, the task, the environment, the control interface, and the user's capabilities. For further comprehensive reviews of smart-wheelchair development, we encourage the reader to refer to [4, 10, 31]. In Section 3.3 we will discuss the specifics of some shared control policies for wheelchair navigation which range from low level collision-avoidance to high-level way-finding assistance.

2.2 Affective computing

Although the study of emotions dates back to the nineteenth century, the field of Affective Computing (AC) is relatively young. It is an interdisciplinary field joining computer science, engineering, psychology, neuroscience, and many other disciplines. The field was introduced in 1995 by Rosalind Picard [32] who defined it as "computing that relates to, arises from, or influences emotions." Research work in the field includes building machines that have affective abilities such as recognizing, expressing, modelling, and responding to emotion to improve interactions between the sensitive human and the unemotional computer [33, 34]. The reader is referred to [34–37] for in-depth literature reviews and descriptions of AC methods and applications. The following sections elaborate on AC topics pertinent to our research.

2.2.1 Affect models

If we want to build machines that understand and react to human emotions, we need to be able to represent affect in a way that a machine can reason with. Affective computing researchers have incorporated theories of emotion proposed by psychologists in order to describe emotions and other affective states. Three affect modelling paradigms derived from emotion theories dominate the AC literature: categorical, dimensional and appraisal-based approaches [38]. Categorical approaches view emotions as discrete labels and state that there is a moderate number of discrete emotions which can be recognized universally (e.g., anger, sadness, fear). However, these approaches lack granularity, in the sense that they do not account for subtle and rather complex affective states that humans can experience [38]. In contrast, dimensional frameworks state that emotions can be described in a limited number of underlying dimensions, although the specific dimensions vary from model to model [39]. One of the most popular dimensional approaches used by AC researchers is Russell's circumplex model of affect [40]. This theory posits that affective states can be represented by a two-dimensional spatial model in which affective concepts are organized in terms of abstract dimensions called valence (pleasant vs. unpleasant) and arousal (relaxed vs. aroused). Appraisal-based approaches claim that emotions are elicited through evaluations of the individual's internal state and the state of the world [41]. In other words, specific emotions arise from the person's expectations and goals in relation to the situation [36]. The Ortony, Clore and Collins's (OCC) model of emotion [42] is one of such models widely used in affective science and one of the most influential for emotion synthesis [43].

Affect control theory

Affect Control Theory (ACT) [44, 45] is a formalized mathematical framework for modelling emotion in social interactions. ACT postulates that an individual's actions and emotional experiences are governed by the need to confirm established sentiments about their self-identities and the identities of others. ACT's models and predictions can be applied to human-computer interaction, making it suitable for the design and development of emotionally intelligent systems [46, 47]. We think of the smart-wheelchair as a social agent that should adapt to individual emotional states so we choose ACT as our emotion modelling framework. Furthermore, ACT has been used for developing emotion-aware assistive technologies for older adults with cognitive impairment (an overview of this application will be discussed in Section 2.3.2).

Evaluation, potency, and activity

The ACT framework and tools use a dimensional paradigm and have been constructed through a tradition of empirical psychological research. The three fundamental dimensions on which actors, actions, and objects are analyzed are Evaluation, Potency, and Activity (EPA). Evaluation describes whether the concept conveys a positive or negative emotion (good vs. bad). Potency describes the intensity of the emotion (powerful vs. powerless), and activity distinguishes between passive and active emotions (lively vs. calm). For example, the concept "grandparent," the general sentiment about grandparents is that they are quite good and helpful, deep and powerful, and quiet and meditative. On the other hand, most people agree that children are also good, but they are small and weak, and quite active and noisy. Each of these aspects can be felt at different levels; something can be slightly good, others moderately or even extremely good. These dimensions form a quantifiable semantic space in which the connotative meaning of any concept, object or event can be specified [48]. Moreover, the resulting EPA profiles are used in ACT for predictions on how events transform social situations.

EPA profiles can be measured using the Semantic Differential (SD), an affective-meaning measuring technique based on a combination of associational and scaling procedures. The custom is to represent goodness, powerfulness, and liveliness using positive numbers; and negative numbers for representing things that are bad, powerless, or passive. A typical range of values used in the EPA framework is $[-4.3, 4.3]$. The SD can reveal nuances in meaning which are clearly felt but hard to verbalize. Further, it yields quantitative data which are presumably verifiable, in the sense that other investigators can apply the same set of scales to similar subjects and obtain essentially the same result [48].

This theory provides the framework for the work reported in this thesis. Based on the premise that any emotion-eliciting event can be described within the semantic differential space, we used the EPA dimensions of affective meaning to characterize the behaviours of the shared control wheelchair as a first step toward building an emotionally-intelligent wheelchair.

2.2.2 Affective robotics

Given that affective processes play an essential role in human behaviour, we can intuitively think of using emotion to enhance robot performance. Arkin and Moshkina [49, 50] identify at least two relevant roles for the inclusion of affect in robotic systems: social interaction and survivability. Social robotics serves to enable robots to relate to humans in predictable and natural ways by providing a means and mechanism for increasing the bandwidth in communication, using nonverbal methods to enhance the relationship between artifact and person [50]. Its applications include education, health, quality of life, entertainment, and collaborative teamwork [51]. For example, "Pepper" is a social humanoid robot capable of recognizing faces and basic human emotions and is currently used as a service robot in retail and financial institutions[1]. Substantial research in social robotic systems has worked to facilitate and enhance human-robot interaction (HRI) through combinations of facial expressions, synthesized speech, posture and body motion (for examples, see [50, 52–55] and the citations therein).

On the other hand, survivability serves as an adaptation function potentially driving an action, which may result in the survival of an agent (human or robotic) in its environment. For example, a robot backing up or slowing down near a staircase may be perceived as *fear of falling* by an observer [50]. While this role in robotic systems has received less attention, emotion has been used to control actions in mobile robots. For example, Lee-Johnson and Carnegie [56] implemented a control architecture for autonomous mobile robots that incorporates artificial emotions as part of the robot's planning and control parameters. The emotions modeled in the robot behaviour are fear, anger, surprise, happiness, and sadness. The emotion parameters change the degree of bias toward certain driving behaviours; for example, *fear* was associated with avoiding collisions and reducing the importance of

[1] https://www.softbankrobotics.com/emea/en/pepper

the robot's current goal, whereas *anger* was linked to achieving the goal even at the expense of secondary considerations. While the robot can operate successfully with constant parameters, the inclusion of affect provides optimal performance and allows the robot to avoid dangerous behaviours by selecting appropriate settings depending on the situation.

2.3 Affect-aware assistive technologies

While there is substantial work in shared control in the context of smart-wheelchairs, addressing the user's affective state in the navigation assistance process remains understudied. The following section elaborates on previous research on emotions in two relevant contexts: automotive driving and (non-driving) assistive technologies for older adults.

2.3.1 Intelligent driver assistants

In daily driving tasks, a neutral affective state comprises the biggest portion of time; however, strong emotions such as road rage, fatigue, stress, confusion, nervousness, sadness, and boredom have the potential to endanger driving safety [13, 57]. Furthermore, affective states play a significant role in driving safety because essential cognitive processes relevant to driving are affected by emotion; for example, perception, goal generation, evaluation, decision-making, focus and attention. Ebyen and colleagues [13] motivate addressing emotional factors to enhance safety and comfort in automobiles. Exemplary use-cases and acceptance of in-car affective computing are investigated through a Wizard-of-Oz (WoZ) user study approach[2] in which subjects communicate with a simulated virtual co-driver using natural speech. Exemplary use-cases of emotion-sensitive technologies include guiding drivers from negative to neutral or even happier states to prevent road rage and to ensure a safer and more pleasant driving experience; adapting the personality of the virtual assistant to match the driver's emotion; and communicating the driver's emotional state to other road users. In their study, the intelligent driver as-

[2]The WoZ approach is widely used in human-computer interaction because it helps designers circumvent implementation challenges and explore and evaluate designs before investing considerable developmental efforts to build a functioning prototype [58].

sistance system supports users while they perform tasks including lane changing, obtaining information from the internet, handling incoming and outgoing calls and engaging in small talk. Throughout the study tasks, the wizard reacted to the subjects' responses and comments adapting his tone of voice to the user's state and the current situation. The reported findings suggest that a virtual assistant that adapts to the emotional state of the user would be accepted as long as users can maintain full control and can mute the system at any time. Furthermore, the authors highlight the need for identifying user affect by incorporating multiple modalities such as speech, facial expressions, and driving styles.

2.3.2 Cognitive assistants for older adults

In the context of assistive technologies for older adults, user affect has been explored for applications intended to assist older adults with dementia to fulfill activities of daily living (ADL) more independently. Mihailidis et al. [59] proposed the COACH system which uses a virtual assistant to monitor and prompt older adults with dementia when washing their hands. The system employs computer vision and artificial intelligence to provide verbal and visual cues at every step of the hand-washing process. Even though the system can identify whether the user needs assistance and provides the correct prompt at the right time, it is not equally accepted by all the users. A working hypothesis on the lack of effectiveness of the system is due to misalignment of the system with individual affective identities. For example, it might occur that some users respond better to more servile approaches while others prefer imperious directions; thus, a single set of pre-recorded prompts is limiting even if they are modelled after human caregivers. More recent studies involving the COACH system [60, 61] used the principles of affect control theory to build explicit models of emotion into the virtual assistant. The goal was to deliver not only the correct prompt for the given step but also the right style of prompting (e.g., imperious vs. servile) according to the affective state of the person (e.g., is the person feeling powerless, in control, angry, or depressed). In order to do so, a set of audiovisual prompts representing different personalities (e.g., big sister, boss, teenager) were implemented on a virtual character capable of displaying facial expressions and body movements. The first step toward an emotionally-aware

COACH was to obtain EPA profiles of the prompts based on real user responses. Similar to rating concepts in the EPA space such as grandparent or children, the complete behaviours of the virtual assistant were judged according to the three basic dimensions of emotion. In [61], the EPA profiles assigned to the virtual assistant and the detected affective state of the user, were used in an ACT framework to predict the prompt that aligns best with the user's affective state. Their results suggest that including the affective identities of the person and the assistant in the decision control loop has the potential to improve the effectiveness of the system.

2.4 Summary

We have reviewed related work in the smart-wheelchair context as well as addressing emotional factors in human-machine interactions. Even though shared control is generally a good approach for collaborative wheelchairs, the conditions under which it brings maximum benefit are still unclear [62]. Moreover, while shared control systems provide beneficial results, such as safer trajectories, negative effects are also reported, such as confusion and frustration when subjects are not in complete control [e.g., 2, 63]. Driven by these findings, we believe that intelligent wheelchairs should not only adapt to the user's physical and cognitive skills but also to the user's affective state. We draw inspiration from the COACH system which implements affect control theory to design assistive technology that not only gives the right information at the right time, but it also presents it in a way that has a higher likelihood of aligning with the user's affective identity. In order to implement ACT to reason over what the optimal intervention is, an EPA profile of the behaviour must be obtained.

Chapter 3

System Design

In this chapter, we elaborate on the elements of our robotic platform. The first section is an overview of the hardware used for this study, whereas the second section is an overview of the software used for mapping, localization and navigation. In the third section, we describe a set of shared control interventions drawn from the smart-wheelchair literature that are suitable for implementation in our system and our experimental protocol. We conclude with a summary on the behaviour of each control strategy.

3.1 Hardware overview

We made use of a commercially-available Permobil M300 powered wheelchair (PWC) with R-net control system and Omni controls which we customized to meet the needs for our research. The PWC is equipped with an onboard computer, LiDAR sensors, a communication interface, and a driver-facing camera (Figure 3.1). Our system runs on a ThinkPad P51 Mobile Workstation with Intel Core i7-7700HQ processor, 16GB of memory, and running Ubuntu LTS 14.04 as the operating system. We use the CoPILOT system originally developed for the study reported in [11, 64] to enable collaborative navigation of the PWC. The system incorporates a standard off-the-shelf Arduino with a custom designed shield to enable communication between the onboard computer and the PWC's motor controller [65]. Mounted on the base of the PWC are two Hokuyo UTM-30LX laser

Figure 3.1: The modified Permobil M300 powered wheelchair

rangefinders providing a 270-degree horizontal field of view at ankle-height. The lasers are used for mapping, localization, and obstacle detection. To record the participants' facial expressions during the driving sessions, we used a small web-camera mounted on an overhead boom facing toward the participant. The PWC can be controlled with two joysticks: the traditional user joystick mounted on the PWC's arm or a wireless (Bluetooth) Play Station 3 (PS3) joystick. The secondary joystick enables the experimenter to take over in case of an emergency, as well as to re-position the PWC between trials.

3.2 Software overview

A smart-wheelchair requires several software subsystems to handle perception, planning, and control. These subsystems need to be able to communicate easily with one another; to accommodate these requirements, we use the Robot Operating

System (ROS)[1]. ROS is a flexible framework to write robotic software allowing efficient integration of complex robotic algorithms into almost any velocity-controlled robot. For the work presented in this thesis, we use the ROS Indigo distribution.

A wide variety of shared control navigation algorithms for smart-wheelchairs have been proposed over the years. Rather than attempting to replicate the precise requirements and resources employed by the different research groups that have worked on this problem, we implement a few of them based on the ROS navigation stack. The robotic side of our system is handled entirely by the ROS nodes, and we implement the shared control algorithms in MATLAB 2018b using the Robotic System Toolbox version 1.3 as an interface between MATLAB and ROS.

3.2.1 Mapping and localization

One of the hardware limitations of our PWC platform is the lack of odometry information from the wheels. Driven by this consideration, we use the *hector_slam*[2] ROS package which leverages the laser rangefinder scans to generate 2D pose estimates based on scan-matching algorithms [66]. The *hector_slam* package publishes an estimate of the PWC odometry which we can use for mapping and localization.

In our experiments, the PWC operates on a known map of the indoor test environment. This map was pre-computed using a Simultaneous Localization and Mapping (SLAM) approach available in the *hector_mapping*[3] ROS package.

Once we have a map of the environment and an estimate of the PWC's odometry information, we can produce reliable state estimates with the Adaptive Monte Carlo Localization (AMCL) package. AMCL is a probabilistic approach to mobile robot localization. It uses a particle filter to track the robot pose against a known map, and is one of the most popular and robust localization algorithms in robotics [67].

3.2.2 The navigation stack

Our system aims to generate safe motion commands resulting from a collaboration between the user and an autonomous robot navigation agent. For the latter, we use

[1] http://www.ros.org/
[2] http://wiki.ros.org/hector_slam
[3] http://wiki.ros.org/hector_mapping

the ROS navigation stack[4] which takes input from the map, odometry, laser scans, and a goal location, and generates as output velocity commands which would safely drive the robot. One of the fundamental elements of the navigation stack is the move_base[5] node, which provides an interface for configuring and interacting with the navigation stack. Special consideration must be taken to include the shape and the dynamics of the mobile platform, in our case, the Permobil M300 PWC. We describe the main components of the navigation stack in the following paragraphs.

Costmaps

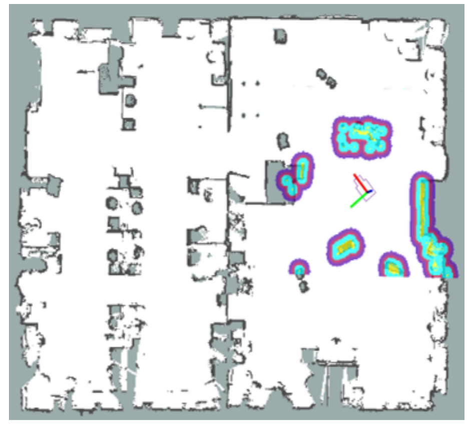
The navigation stack uses 2D occupancy grids known as costmaps to store information about obstacles in the world. The global costmap is usually constructed from a pre-computed static map and is used to create long-term plans over the entire world. On the other hand, the local costmap is a rolling window which only considers the region surrounding the robot. It has no prior knowledge of the world and is entirely constructed from recent sensor readings. The local costmap information is essential for dynamic obstacle detection and short-term planning. Each cell in the costmaps contains either free, occupied, or unknown values. Obstacle inflation can be specified, which consists of propagating the cost from each occupied cell out to a user-specified radius. The inflation radius is an inexpensive mechanism to approximate the configuration space of the robot using the costmap, and it encourages the planner to find paths that ensure a minimum clearance from obstacles. Examples of global and local costmaps are presented in Figure 3.2.

Global and local planner

Just like there are global and local costmaps, there are global and local planners. Based on the global costmap, the global planner occasionally generates a high-level plan from the current robot location to the goal location using either graph traversal or sample-based planning algorithms. The local planner generates velocity commands on every control cycle based on the high-level plan, the information

[4] http://wiki.ros.org/navigation
[5] http://wiki.ros.org/move_base

(a) Global costmap (defined on the entire map).



(b) Local costmap (defined only near the robot).

Figure 3.2: Global and local costmaps with 0.5 m inflation radius. A cell within an obstacle has unit cost, and the cost decreases toward zero as the distance from the obstacle increases.

contained in the local costmap, and the current state of the robot.

A number of specific planners are readily available to use under the ROS navigation framework. Driven by our applications in the smart-wheelchair context, we adopted the *eband_local_planner* which is an implementation of the Elastic Band method as described in [68]. The elastic band method bridges the global and local planners by generating a global plan and deforming specific regions if changes in the environment are detected. The approach avoids an expensive call to a path planner by triggering a reactive collision avoidance mechanism that works on a local level but does not limit the ability to achieve global goals.

Recovery behaviours

If the move_base node of the navigation stack fails to find a valid plan, a recovery behaviour can be executed. Traditional recovery behaviours in mobile robotics include clearing the recorded obstacles in the costmaps outside a user-specified region away from the robot, and performing in-place rotations in an attempt to gather additional sensor readings. For our experiments with the PWC, we only allow clearing the local costmap because triggering a rotation may result in counter-intuitive trajectories for the user.

3.3 Assisted driving interventions for smart-wheelchairs

In this section we describe different driver assistance interventions from the smart-wheelchair literature. A wheelchair control signal $\mathbf{u}$ is a two dimensional vector consisting of a linear velocity v and angular velocity ω:

$$\mathbf{u} = \begin{bmatrix} v \\ \omega \end{bmatrix}.$$

The high-level steps that are required to perform an intervention are:

1. Collect the user's motion command coming from the wheelchair's joystick. We call this $\mathbf{u_{user}}$; it consists of linear velocity, v_u, and angular velocity, ω_u.

2. Compute the motion command using an autonomous planner. We call this $\mathbf{u_{robot}}$; it consists of linear velocity, v_r, and angular velocity, ω_r.

3. Compute the arbitration parameter. We follow a linear blending approach, an arbitration model that has been widely adopted in the shared control literature and the assistive wheelchair community [69]. A linear blend step takes the following form:

$$\mathbf{u_{shared}} = \alpha \cdot \mathbf{u_{user}} + (1-\alpha) \cdot \mathbf{u_{robot}} \tag{3.1}$$

   Note that $\alpha \in [0,1]$ denotes the only degree of freedom, and we will refer to this parameter as the "user's control weight". A substantial effort has been carried out by several research groups to come up with optimal ways to compute this parameter. We explore a few approaches in the remainder of this section.

4. Synthesize the shared control signal. Using equation 3.1 compute the control $\mathbf{u_{shared}}$ that will be sent to the PWC's actuators.

Figure 3.3 illustrates the shared control framework. In the remainder of this section, we provide a detailed description of the set of shared control paradigms

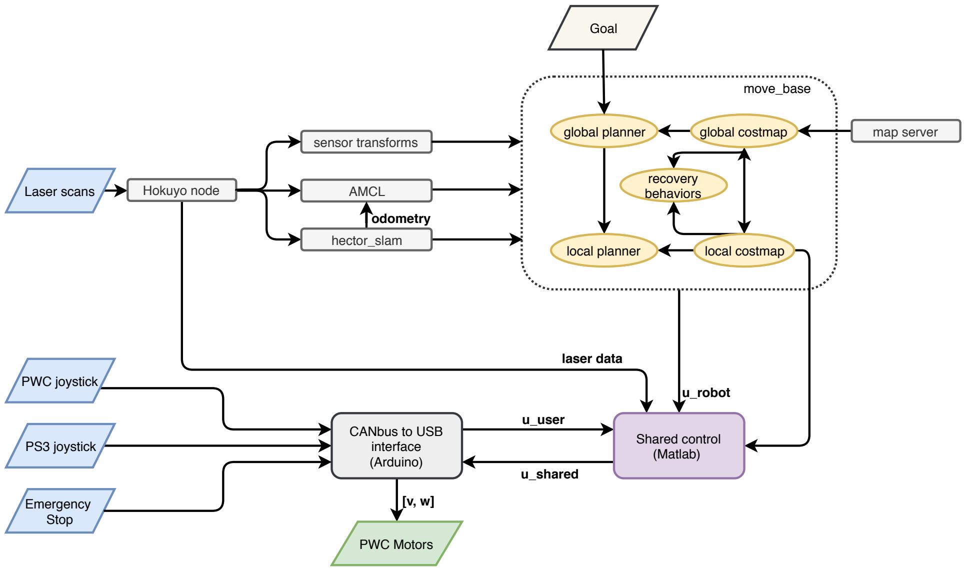


Figure 3.3: Schematic of the control architecture

considered for the experiment described in Chapter 4.

3.3.1 Dynamic Shared Control (DSC)

Li et al. [70] propose a shared control framework that optimizes the user's control weight according to three factors: *safety*, *comfort*, and *obedience*. They formulate the weight adjustment problem as a multi-objective optimization. The major elements of the proposed architecture are the reactive controller and the weight optimizer. The reactive controller is based on the minimum vector field histogram (MVFH) and the vector force field (VFF) methods.

The algorithm includes four main steps: update an obstacle map with laser data, generate an angular velocity ω_{robot} based on the reactive controller, calculate the optimal angular velocity ω^* using the proposed weight optimizer, and compute a final linear and angular velocity $[v_{\text{final}}, \omega_{\text{final}}]$. The linear velocity v_{final} will always be equal to v_{user} unless the minimum distance to an obstacle is below a threshold (0.5 m in the experiment); in that case, v_{final} will be reduced according to the nearest obstacle's distance.

The authors define the following three indices to evaluate wheelchair performance:

- *Safety* Measures the probability of collision according to the distance to the nearest obstacle.
$$\text{safety} = 1 - e^{(-\delta \cdot dist)},$$
where δ is a normalizing constant and *dist* is the minimum distance between the wheelchair and the obstacle nearest to its path. The predicted path is calculated according to the wheelchair's kinematic model. The authors used a prediction time of 4 seconds assuming the user's current control as input.

- *Comfort*: According to their experiments, frequent changes in velocity make users feel uncomfortable. Therefore, the comfort index measures the angular velocity change:
$$\text{comfort} = e^{(-\beta|\omega - \omega_0|)}$$
Where β is a normalizing constant, ω is the current angular velocity and ω_0 is the angular velocity from the previous control cycle. Given that the user is in control of the linear component of the velocity, the change in linear velocity is omitted from calculation of this index.

- *Obedience*: Measures the proximity between the user's control and the final motion command.
$$\text{obedience} = e^{(-\gamma|\xi - \xi^*|)}$$
Where γ is a normalizing constant, ξ is the orientation of the user's input, and ξ^* is the orientation determined by v^* and ω^*. This index prefers that the wheelchair remain under user control as long as *safety* and *comfort* are maintained.

The proposed principle to optimize over the three often contradictory indices is always to improve the smallest index among the three. In accordance with this principle, the multi-objective optimization problem can formulated as a simple objective problem and solved using the minimax method:

$$\max_{v,\omega}(\min(\text{safety, comfort, obedience})), \tag{3.2}$$

such that $\omega_{\text{shared}} \in [\omega_{\text{user}}, \omega_{\text{robot}}]$, and $v_{\text{shared}} = v_{\text{user}}$ unless the distance to the closest obstacle falls below 0.5 meters. This simplified optimization guarantees that the precedence among indices will naturally change when facing different situations.

Our implementation

The DSC algorithm is meant to function as a purely reactive approach to collision-free navigation; however, by using the elastic band method we can integrate a global planner and the reactive controller to enable goal-seeking assistance as well.

The emergent wheelchair behaviour depends greatly on the constants used to compute the *safety*, *obedience* and *comfort* indices. In our implementation of the DSC approach, we estimate the *safety* of a given command based on the obstacle information contained in the local costmap. Due to the inflation radius, we can rely on the costmap information since the distance to nearest obstacles is implicit in the cost of each cell. We calculate the PWC's predicted pose using the forward kinematic model, and then estimate safety by sampling the cost of the cell at the predicted location; thus, safety $= e^{-\delta \cdot \text{cost}}$. We empirically set the safety constant δ so that safety equals 0.63 when the cost of the cell at the predicted location is 0.1. We calculate the comfort constant β such that maximum comfort is achieved when there is no change in the angular velocity and minimum comfort when the angular velocity changes sharply from ω_{min} to ω_{max} or vice versa. Similarly, we compute the obedience constant γ such that maximum obedience is achieved when the user's command and the shared control command have the same orientation, and minimum obedience is achieved when the angle difference between the two commands reaches $\pi/2$.

3.3.2 Efficiency based wheelchair collaborative control

Urdiales et al. [29, 71] propose a blending approach for collaborative wheelchairs based on a continuous evaluation of the the user and planner motion commands. The aim of this approach is to allow the user to contribute as much as he/she can, and let the robot take care of the rest. The performance of the user's and robot's inputs are measured according to an efficiency metric that evaluates three main factors: *smoothness*, *directness*, and *safety*. The importance weight of each factor is controlled by the constants C_{sm}, C_{di}, C_{sf} corresponding to smoothness, directness, and safety respectively. These three factors are essentially the ones from the DSC approach; however, the methods differ in the way the arbitration parameter α is computed.

- *Smoothness*: Measured in terms of the direction of the provided motion vector, α_{dif}. Smoothness reflects that sharp changes in heading are undesirable given the wheelchair's kinematics. Smoothness is computed by:

  $$\eta_{sm} = e^{-C_{sm} \cdot |\alpha_{dif}|}$$

  Maximum smoothness is achieved when the heading changes as little as possible and minimum smoothness when the current motion command corresponds to a sharp turn (i.e., $\alpha_{dif} = \pi/2$).

- *Directness*: Measured in terms of the angular difference between the provided motion vector, α_{dif}, and the direction towards the next partial goal provided by the global planner, α_{dest}. Reflects that pursuing straight paths is desirable. Directness is computed by:

  $$\eta_{di} = e^{-C_{di} \cdot |\alpha_{dest} - \alpha_{dif}|}$$

- *Safety*: Evaluated in terms of the angle to between the output motion vector and the nearest obstacle at each instant. Safety is computed by:

  $$\eta_{sf} = 1 - e^{-C_{sf} \cdot |\alpha_{min} - \alpha_{dif}|}$$

  where α_{min} is the angle between the current heading and the direction of the closest obstacle. Thus, safety increases as the α_{min} increases, and decreases as the angle decreases.

The efficiencies are used to decide how much assistance the user needs. The better the user drives, the less effect the robot has in the intervention. A set of local efficiencies is computed for the motion command coming from the human, $\mathbf{u_{user}}$, and for the command coming from the autonomous planner, $\mathbf{u_{robot}}$. A single efficiency value for each of the human and the robot is computed by averaging the three factors, and the linear blend is defined as:

$$\mathbf{u_{shared}} = K \cdot \eta_h \cdot \mathbf{u_{user}} + (1-K) \cdot \eta_r \cdot \mathbf{u_{robot}}, \tag{3.3}$$

where η_h is the averaged human efficiencies, η_r is the averaged robot efficiencies, and K is a variable that modulates the contribution of human and machine. The value of K is chosen according to the following heuristic table:

$$K = \begin{cases} 0.75 & \text{if } (\eta_h > 0.85) \vee (\eta_h > 1.5 \cdot \eta_r; \\ 0.5 & \text{if } (\eta_h > 0.85) \wedge (0.5 \cdot \eta_r < \eta_h < 1.5 \cdot \eta_r); \\ 0.25 & \text{otherwise.} \end{cases}$$

Our implementation

The approach described in [71] extracts partial sub-goals (needed to calculate α_{dest}) by finding the point of maximum curvature of the path returned by the planner. Due to the computational load required to find such a point, we opt for finding the furthest point in the global trajectory currently lying in the local costmap.

From an initial test of the algorithm, we found that the modulation introduced by the K parameter has a tendency to significantly scale down the velocities commanded by the user. To counteract this scaling, we include a normalization step to compute the user's control weight such that the resulting user's and planner's control weight add up to 1.

Furthermore, we set all the constants C_{sm}, C_{di}, and C_{sf}, to the same value (4.5) so that all factors are equally important.

3.3.3 Blending with immediate goals (collision avoidance)

Erdogan and Argall explore the effects of four shared control paradigms for robotic wheelchairs in [72]. One of the approaches provides assistance only when an imminent collision is detected. The proposed algorithm checks for collisions by computing a forward projection of the user's current command $\mathbf{u}_{user}$ for a time Δt. If the predicted pose is found to be unsafe, the user's command is linearly blended with the planner's command in an iterative manner that takes control away from the user based on safety constraints. Thus, as the user issues unsafe commands, the user's control weight α decreases.

Our implementation

The proposed method reasons over the safety of the user command by checking whether the projected wheelchair pose lies within an obstacle. Since our experimental environment is subject to small changes in the distribution of obstacles, we sample the occupancy grid published by the local costmap (updated purely from sensor data) to determine whether a given command is safe or not. If the cost of landing in the predicted location is higher than a given threshold, we register it as unsafe and perform the blend. Otherwise, the user maintains full control of the wheelchair.

3.3.4 Blending with a high-level goal

Another navigation assistance paradigm explored in [72] consists of continuously blending the user's and the autonomous planner's commands to achieve a high-level goal. The high-level goal is determined from a set of possible locations through perception algorithms that process RGB-D data and produce a confidence measure based on the agreement with the current user input. When the confidence for a given goal is above a threshold and the user is issuing commands, the user's and planner's inputs are linearly blended in a way that steps control away from the user based on Euclidean distance to the goal. Thus, as the user gets closer to the target, the user's control weight α decreases:

$$\alpha = \frac{1}{(1+\exp(-\tau_c \cdot (d-d_c)))}, \tag{3.4}$$

where τ_c is the time constant that determines how fast the control steps away from the user, d is the Euclidean distance to the goal, and d_c is a coefficient that determines the distance at which the user and the planner have the same control weight (i.e., $\alpha = 0.5$).

If none of the possible goal locations is perceived as the high-level goal, the blending approach based on safety constraints described in Section 3.3.3 is performed.

Our implementation

Since our experimental procedure assumes a known goal location, we do not use the confidence measure and blend the user's and planner's inputs continuously. Moreover, the proposed approach does not perform further collision checks for the resulting blended command. In order to prevent a situation where the resulting command leads to a collision, we perform the same collision avoidance algorithm described in Section 3.3.3 and further decrement the user's weight if necessary.

3.3.5 Steering correction and speed limit

Mitchell et al. [26] propose three intervention policies for safe wheelchair navigation which were later tested by cognitively impaired older adults on a study conducted by Viswanathan et al. [11, 64]. The intervention mode preferred by participants in their study operates on a hysteretic principle. The user maintains complete control of the wheelchair unless the distance to the nearest obstacle falls below a threshold d_0. In that case, the planner takes over control and drives the wheelchair until the distance to the nearest obstacle is greater than d_1 (where $d_1 > d_0$); then full control is granted back to the user.

Our implementation

We empirically set $d_0 = 1$ m and $d_1 = 1.5$ m based on the affordances of our test environment. We find the distance, d, to the nearest obstacle by finding the smallest laser scan reading within a 100 degree field of view in front of the PWC.

3.3.6 Disagreement-based shared control

We have implemented an *orientation correction* control paradigm that modifies the user's angular velocity based on the disagreement between the user and the autonomous planner over a window of time: If the user remains in agreement with the planner, the planner will incrementally take control. Under normal circumstances the proposed approach only modifies the angular velocity. Control of the linear velocity is left entirely to the user unless an imminent collision is detected.

We define *disagreement* as the absolute angular difference between the two input vectors:

$$\phi_t = |\xi_r - \xi_u|, \tag{3.5}$$

where ξ_r is the deflection from forward motion of the planner's command and ξ_u is the corresponding deflection from the user's command. The operation of this assistance paradigm works by collecting a window of disagreement measures over the last k samples $\phi_{1:k}$, followed by computing a summary criterion across the window; for example, the average over the disagreement measures:

$$\phi_{avg} = \frac{1}{k}\sum_{t=1}^{k}\phi_t. \tag{3.6}$$

If the summary criterion is below a threshold λ_a, then the user is considered to be in agreement with the planner. However, if the current ϕ_t is higher than λ_b (where $\lambda_b > \lambda_a$) the user will immediately regain full control. Next, the user's control weight is computed from the current value of ϕ_{avg} such that as the agreement increases the user's control weight decreases. For example,

$$\alpha = 1 - e^{(\tau \cdot |\phi_{avg}|)}, \tag{3.7}$$

where τ and C_a are constants. For our experiment, we empirically set $k = 200$ (which corresponds to approximately 10 seconds of data), $\lambda_a = \pi/3$, $\lambda_b = \pi/2$, and $\tau = 4.5$. Similar to the algorithms described before, a safety check for collisions is evaluated for each generated command.

The proposed paradigm is expected to lead to smooth trajectories and provide guidance towards a goal. Introducing a temporal dependence allows the algorithm to be robust against small deviations from the planner's optimal path. Moreover, this approach grants full control back to the user if he/she changes direction abruptly *or* steadily disagrees with the planner, such as when he/she decide to navigate to a different location than originally planned.

3.4 Summary

We modified a commercially available PWC to leverage technologies developed for mobile robots to create our smart-wheelchair. We implemented six shared control strategies from the smart-wheelchair literature using the Robot Operating System and MATLAB. Each control strategy optimizes different sub-objectives which leads to varying levels of contribution from the autonomous system and the user. A summary of the driving assistance intervention modes is presented in Table 3.1.

Table 3.1: Summary of driving assistance interventions for smart-wheelchairs

Intervention	Description
Dynamic Shared Control (DSC)	Multi-objective constraints on safety, obedience, and comfort.
Efficiency based control (efficiency)	Optimizes smoothness, directness, and safety based on efficiency metric.
Blending with immediate goals (collision avoidance)	Iteratively steps control away from the user when an imminent collision is predicted.
Blending with high-level goals (high-level)	Smoothly steps control away from the user as they get closer to the target.
Steering correction with speed limits (steering correction)	The user maintains complete control unless the distance to the nearest obstacle falls below a threshold.
Disagreement-based shared control (disagreement)	The planner incrementally takes control if the user remains in agreement with the planner's actions.

Chapter 4

An Exploratory Study

In this study, we explored the emotional effects of six different driving assistance modes on a smart-wheelchair as experienced by the driver. Twenty able-bodied adults tested each assistance paradigm while driving the wheelchair toward goal locations in a semi-static environment.

Recognizing the affective state of a user embodies a challenge given the complexity of human emotions; therefore, we use two methods of measuring it, namely post interaction self-report (subjective component) and facial expression emotion recognition (physiological component). We adopt the EPA space of affective meaning, a quantifiable dimensional model of emotion, as our sentiment-measuring framework. The following sections describe the experimental setup of the study, the collected measures, analyses of the self-reported data, a preliminary analysis of the facial expressions of emotion, and a discussion of our findings.

4.1 Objectives and approach

The purpose of this study is to analyze the effects of various smart-wheelchair behaviours at an emotional level. If the different driving assistance modes have consistent affective ratings and we can estimate the user's affective state reliably, we can design a system that intervenes according to the user's emotions in addition to traditional objectives such as comfort, safety, ease of use or timeliness.

We emphasize that our implementation of various wheelchair assistance modes is not the novelty of this user study; instead, our aim is an exploration of a human-robot interaction model of emotion in the context of wheelchair navigation assistance.

Throughout this study we evaluate the affective responses elicited by six navigation assistance behaviours designed to enable safe and independent powered wheelchair operation. The specific objectives are:

1. Assess six different navigation modes providing low to high levels of assistance to navigate to a particular goal location.

2. Obtain and analyze self-reported affective interpretations of each intervention mode and emotions resulting from the interactions with the smart-wheelchair.

3. Collect and analyze facial expressions of emotion.

4. Obtain and analyze quantitative feedback regarding workload, perceived safety, and perceived ease of use. Obtain and analyze qualitative feedback regarding user perceived usefulness.

5. Collect and analyze joystick input data and navigation paths.

6. Determine possible consensus across users on the affective ratings of the intervention modes.

7. Identify possible significant differences in user's emotional experiences between assistance modes.

8. Determine whether the intervention modes have consistently distinguishable affective patterns.

9. Determine possible applications for the use of joystick input data and navigation path metrics to predict affective state.

In this description, we use "collect" to indicate data that will be recorded through a sensor and "obtain" to indicate data that is recorded through a survey tool.

4.2 Methods

The study was approved by the University of British Columbia (UBC)'s Behavioural Reseach Ethics Board under certificate number H18-00256. The experiment took place in the Collaborative Robotics Laboratory at the Institute for Computing, Information and Cognitive Systems (ICICS) at UBC and informed consent was obtained from all subjects. Each experimental session lasted between 70 and 97 minutes, and all participants received $15 in compensation for their participation.

4.2.1 Participants

We recruited 20 able-bodied participants through on-campus advertisements at UBC. Our participants included 9 male and 11 female, aged 22 to 68 years (mean = 30.8, SD = 12.9). Ten participants reported none to little interaction with robots in the past; seven reported having interacted with robots on a number of occasions, and three reported robot interaction on a regular basis. Sixteen participants reported absolutely no experience with powered wheelchairs, and the remaining four reported from 10 hours to 39 years of experience with powered wheelchairs as either clinicians or researchers. Regarding occupation, our sample included fourteen graduate students, two occupational therapists, one elementary school teacher, one research engineer, one consultant in assistive technology, and one participant was retired. The cultural backgrounds of our subjects include North America (6), the Middle East (4), East Asia (4), South Asia (3), Southeast Asia (1), Eastern Europe (1), and Western Europe (1).

4.2.2 Conditions

We implemented assistive navigation modes with expected affective evaluations that span the EPA space of affective meaning. We anticipate that each intervention mode will modify the user's inputs in ways that can be interpreted as a particular behaviour. For example, the word "obedient" has an average EPA representation that places it in the octant (E+, P-, A-) of the affective space. An intervention mode that resembles an obedient behaviour keeps the user in control unless an unsafe situation arises. We believe that the dynamic shared control (DSC) intervention might model this behavior because the user keeps control as long as the safety constraint is maintained. In comparison, the word "authoritarian" has average ratings (E-, P+, A+) and the corresponding intervention policy tends to generate motion commands that prioritize the control signals generated by the autonomous planner rather than the user's inputs. We believe the efficiency based blend intervention might model this behavior. We discuss our expected affective interpretations (in terms of an octant of the EPA space) for each intervention condition from Section 3.3 and the reasoning behind those expectations in Table 4.1.

Table 4.1: Affective expectations of the intervention modes

Intervention	Expected affective evaluations (EPA)	Trait
Dynamic shared control	**E**$+$ Resulting motion tends to grant high levels of control to the users as long as the safety constraints are met. The trajectories tend to be smooth. It prevents the user from getting too close to obstacles. **P**$-$ Modifications to the user commands are small if the user is navigating safely. **A**$-$ May be perceived passive if the user's inputs are not constantly and significantly modified.	Obedient

Continued on next page

Table 4.1 – *Continued from previous page*

Intervention	Expected affective evaluations (EPA)	Trait
Efficiency based blend	**E**$-$ Although this approach results in smooth trajectories, the user's input tends to be completely ignored if they want to drive to a goal different than the planner. These situations may lead to a negative evaluation. **P**$+$ Given that the planner produces approximately optimal commands, its efficiency is usually high. There are few or no occurrences when the user has more control than the planner, so the intervention feels potent. **A**$+$ The level of activity as judged by the amount of modification to the user's commands could be perceived as high.	Authoritarian
Blending with immediate goals (Collision avoidance)	**E**$+$ If the user is a good driver, the perception of the PWC's behavior will also be good. However, since this mode only provides collision avoidance and no goal-finding assistance, the wheelchair behavior might be rated as inadequate. **P**$-$ The behaviour is not potent because it allows the user to retain control most of the time. Moreover, when a possible collision is detected, the iterative approach ensures that only the minimum amount of control is taken away from the user. **A**$-$ The activity levels of the PWC may be perceived as low unless the user constantly comes close to obstacles and experiences many modifications of the commands.	Cautious

Continued on next page

Table 4.1 – *Continued from previous page*

Intervention	Expected affective evaluations (EPA)	Trait
Blending with high-level goals	**E**+ The intervention results in the user losing control as they approach the goal location. However, the control switch happens at a slow rate and could be perceived as pleasant. **P**+ Potency may be somewhat strong: as the user gets closer to the goal (this distance is determined by d_c), the wheelchair increasingly ignores the user's command, and it may be perceived as potent. **A**+ The wheelchair will guide the user continuously to the goal location resulting in high levels of activity.	Cooperative
Steering correction with speed limits	**E**− This approach provides mostly collision avoidance. However, if the user engages in a situation where the planner takes control, the wheelchair will turn in the direction to the goal. Such change can lead to jerky trajectories. **P**− We have set the threshold limits to ensure the policy is triggered in crowded spaces during the trials. However, given that this policy only intervenes when collision is imminent, we expect it to be perceived as not so potent. **A**+ The level of activity is highly affected by the values d_0 and d_1. We chose values that will trigger the intervention in several areas of the environment, so we expect it to be perceived as active.	Impatient

Continued on next page

Table 4.1 – *Continued from previous page*

Intervention	Expected affective evaluations (EPA)	Trait
Disagreement based blend	**E**+ The approach slowly steps control away from the user as the disagreement between user and planner decreases. However, it does not limit the user from changing directions which we believe will be perceived as pleasant. **P**+ Since small errors in the orientation will be corrected, the user may chose to simply move forward and the algorithm will take care of the orientation. **A**− The user may not even be aware that their inputs are constantly being modified, since any large disagreement results in the user regaining full control.	Gentle

4.2.3 Tasks

All trials occurred in an open-plan office environment with approximate dimensions of 20 m x 20 m (specifically rooms x210 and x209 in the ICICS building at UBC's Point Grey campus). Each experimental task required the participant to drive the powered wheelchair from an initial position to a marked goal destination in the other room and then return to the first room. Completing a circuit led the users to spend more time in each mode for a more accurate assessment of the PWC's behaviour. The initial and target locations were chosen to require participants to transverse through a doorway and both open and cluttered spaces, although all obstacles were stationary. To simulate a mild cognitive impairment, we asked participants to drive until they reach the goal in the other room, but we did not tell them the exact location of the goal.

Participants were asked to drive as they felt comfortable, and they were not

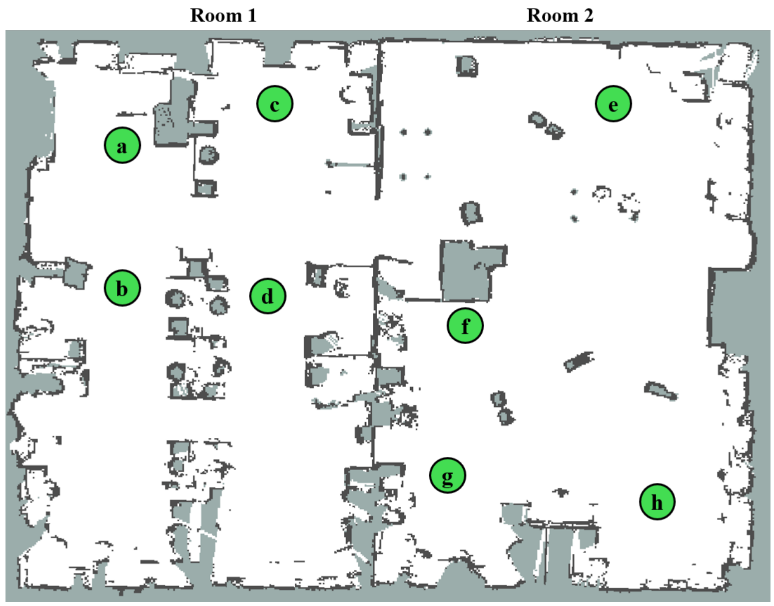


Figure 4.1: Map of the environment with goal zones labeled a to h.

informed of which intervention mode was active on a given trial. The presentation order of the modes and target locations was randomized. Each mode was used twice with different target locations and starting from different rooms to mitigate biases due to ordering, task difficulty, fatigue and/or learning effects. Refreshments were available and breaks were provided when requested by the participant to reduce participant burden.

Trajectories

The global map of our test environment is depicted in Figure 4.1. The eight green zones labeled **a** to **h** indicate the predefined locations of the possible start and goal locations for the driving tasks.

An example session of trials is described in table 4.2. A session consists of two blocks of six trials each. Half of the participants performed six navigation tasks starting from Room 1 followed by six tasks starting from Room 2. The other half of the participants began in Room 2 and then moved onto Room 1. For example, the task described by the first row of the table can be interpreted as: using the DSC

Table 4.2: Example of task start and end locations

	Begin in Room 1			Begin in Room 2		
Mode	Start	First	Second	Start	First	Second
DSC	a	f	d	e	b	g
Efficiency blend	b	g	c	f	a	e
Collision avoidance	c	f	b	h	b	e
High-level blend	d	g	a	h	c	g
Steering correction	b	e	a	g	c	h
Disagreement blend	c	h	d	g	d	f

mode, start in location **a**, navigate the wheelchair to location **f** and then to location **d**. The intervention modes were randomly assigned to each row, and the order of the interventions was randomized and counterbalanced.

4.2.4 Measures

Emotions and affect are inherently complex, so measuring an individual's affective state effectively and reliably is one of the most prevalent problems in affective science [39]. Emotions can become apparent through subjective experiences (i.e., how a person feels), internal responses (e.g., physiological signals), and behavioural manifestations (e.g., facial expressions) [38].

For this study, we used self reports in the semantic space of affect and emotion recognition from a video of the participant's facial expressions during the tasks as our emotion-measuring instruments. We also included workload, perceived safety and difficulty because we are interested to see how these measures impact the responses. Finally, we collected data from laser rangefinders, user joystick, wheelchair pose, and commanded velocities to estimate the actual levels of assistance delivered by each intervention.

Affect self-assessments

Upon completion of each experimental task, participants used the semantic differential technique to provide an affective interpretation of the wheelchair's behaviour. We used a semantic differential for each dimension of emotion (E, P, and A) with typical ranges of [-4.3, +4.3] on a continuous scale with increments of 0.1. The adjectives used for each dimension:

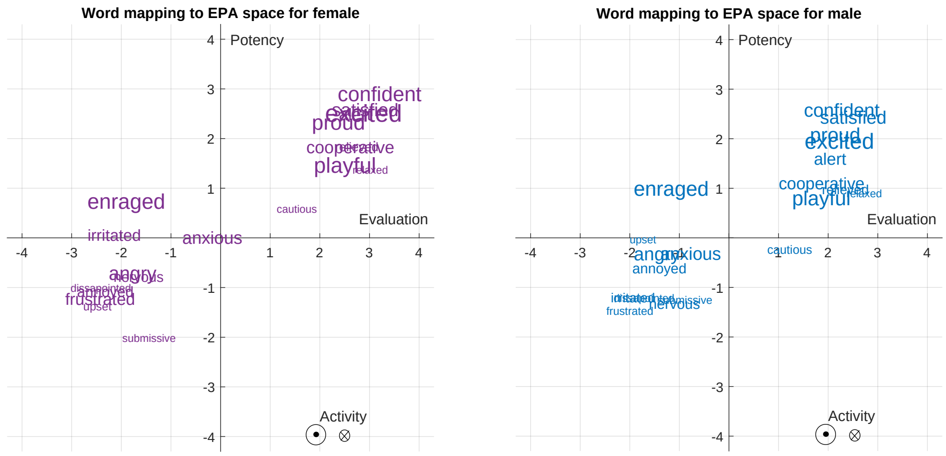
- **Evaluation**: Bad/Unpleasant vs Good/Pleasant
- **Potency**: Powerless/Weak vs Powerful/Strong
- **Activity**: Passive/Calm vs Active/Excited

After the participants evaluated the wheelchair, we asked them to describe the specific emotions elicited during the driving task. We recognize that emotion self-report is challenging because it is common to not know how to articulate or label individual feelings and people tend to differ in the extent to which they self-report emotions (i.e. when asked to report their feelings, some individuals may represent their experiences with a good deal of precision while others may prefer using global terms [73]). Driven by these considerations, we opted to ask participants to make judgments on what emotions they had been experiencing during the driving task using a set of emotion related words. We provided them with a set of 20 emotion-related adjectives with existing validated emotional mappings to the EPA space [1]. Members of the research team selected the specific words by reflecting on emotions that are most relevant for our experimental task and spread across the space associated with the EPA dimensions of emotion. Participants selected as many emotions as they needed from the following list: *alert, angry, annoyed, anxious, cautious, confident, cooperative, disappointed, enraged, excited, frustrated, irritated, nervous, playful, proud, relaxed, relieved, satisfied, submissive,* and *upset.* Projections of the corresponding EPA mappings for each word and male and female genders is presented in Figure 4.2.

Facial expressions of emotion

Exploiting visual information has been used for the detection of discrete emotions from facial expressions. Facial expressions appear to be particularly sensitive to the valence of an individual's emotional state [39]. We collected the participant's facial

[1] We use concepts from the USA-Indiana, 2003 EPA dictionary available at http://www.indiana.edu/~socpsy/ACT/data.html



(a) **(b)**

Figure 4.2: 2D visualizations of EPA mappings for 20 emotion words from American surveys (a) averaged female ratings (b) averaged male ratings. Activity is represented with the size of the word, bigger words correspond to larger values in the Activity dimension.

expressions using a driver-facing camera mounted on the PWC. After the experiment concluded, we used Affectiva Emotion Software development kit (SDK) [74] to automatically detect facial emotion indicators and overall experience valence. The software uses deep learning and computer vision to recognize emotions from facial expressions. The classifiers are based on Ekman's Emotional Facial Action Coding System (EMFACS) [75], which provides mappings between facial expressions and discrete emotions. Specifically, it measures levels of seven basic emotions: anger, contempt, disgust, fear, joy, sadness, and surprise. Furthermore, the SDK estimates valence as an indicator of the positive or negative nature of the user's experience. Each of the emotion scores has a range of [0, 100] and the estimated valence has a range of [-100, 100].

Workload

To measure perceived task workload, we used a subset of the NASA Task Load Index (NASA-TLX). The NASA-TLX is a multi-dimensional rating procedure that estimates an overall workload score based on six weighted categories: mental demand, physical demand, temporal demand, own performance, effort, and frustration [76]. This workload estimation measure has been extensively used over the past 30 years, because it is reasonably easy to deploy and reliably sensitive to experimentally important manipulations [77]. This measure was included to assess the correlation between user emotional state and subjective workload; however, we omit the physical demand subscale from the questionnaire because powered wheelchairs inherently require very little physical effort to drive. Also, we omit the effort subscale because of its overlap with other questions.

Following the driving tasks, participants evaluated the weight of each workload category or the degree to which each of the categories contributed to the subject's workload. The weight is evaluated through a set of pairwise comparison cards among the four selected factors. The procedure results in a single weighted workload measure with ranges [0, 100].

Safety and difficulty

We asked participants to rate their levels of perceived safety during their driving task using a semantic differential scale on a [0, 7] range with increments of 0.1. The lower end corresponds to "extremely unsafe" and the higher end corresponds to "extremely safe." Similarly, the difficulty to use was measured with a semantic differential scale with the lower end being "extremely difficult" and the higher end being "extremely easy." We also asked open-ended questions regarding perceived usefulness to complete the tasks and to avoid obstacles.

Motion commands and sensor data

In addition to the self-reported ratings, the following wheelchair and sensor data was collected during the driving sessions:

- velocity commands from the user joystick;
- velocity commands generated by the autonomous planner;
- velocity commands sent to the PWC's motors;
- user control weight;

- wheelchair odometry and localization data (from which its path can be reconstructed) and;
- measures from the two laser range finders mounted on the wheelchair.

4.2.5 Procedure

After providing consent, participants received a brief introduction to the semantic space of affective meaning. During this introduction, participants were asked to judge the emotional meaning of three concepts according to the evaluation, potency, and activity dimensions using a graphical user interface. After the assessment, values from the EPA dictionary for the three concepts appeared, and the respondents were able to see how their responses compared with those of the average. The purpose of the introduction was merely illustrative and no data was collected at this point. Participants then filled out a pre-experimental questionnaire, which we used to learn about the individual's affective identity, their current emotional state, and their affective expectations of smart-wheelchairs. The pre-experimental questionnaire is available in Section A.2.1.

Following the pre-experimental questionnaire, each subject spent 5 to 10 minutes familiarizing themselves with the powered wheelchair and the experiment tasks. Safety instructions were provided during this training period, and a practice trial was conducted to mitigate novelty effects; no data was collected during this initial trial and no navigation assistance was provided/applied.

Next, the assisted navigation modes were tested by the participants. Each task required them to find a randomized goal, chosen from a set of pre-selected safe goals. A general description of the goal location was given, such as "the goal is on the other side of that doorway," but the specific location of the goal was not identified. Participants were instructed to drive as they felt comfortable, but no description of the assistance was given to the participants to mitigate biases in the responses. Video recording of the participants' facial expression was active during the driving sessions using a driver-facing camera mounted on the wheelchair. The task finished when one of the following conditions occurred:

- the participant was within a 85-centimetre radius of the goal;

- the task time reached 3 minutes;
- a collision occurred or the experimenter was forced to stop the wheelchair due to an imminent collision;
- a technical failure occurred.

A bell sound informed the participants when a task had finished and the video recording stopped. Following each task, participants were asked to provide a self-report on their momentary emotional experiences and an affective interpretation of the behaviour of the PWC. Additionally, subjective metrics regarding the workload, perceived usefulness, and perceived safety were collected after each interaction. The post-interaction questionnaire is available in Section A.2.2. Participants interacted with each of the intervention modes twice for a total of 12 trials per session. Upon completion of all tasks, participants provided demographic information regarding age, gender, occupation, mother tongue, previous experience with robots, and previous experience with powered wheelchair. The demographic survey is available in Section A.2.3.

A short debriefing session was conducted after the experiment concluded, to share information related to the research goal and give participants the opportunity to ask questions about the study.

4.2.6 Hypotheses

We state our hypotheses guided by our research questions. For the first research question, regarding the affective interpretations of the smart-wheelchair behaviour, we formulate two hypotheses:

Hypothesis 1 *There will be consensus/agreement across participants of affective ratings on each dimension of emotion for each intervention mode.*

Hypothesis 2 *There will be a distinct pattern of affective ratings for each intervention mode.*

For the second research question, regarding the user's elicited affective state when interacting with various wheelchair assisted driving behaviours, we pose the following hypotheses:

Hypothesis 3 *There will be a distinct pattern of evoked emotions for each navigation mode.*

Hypothesis 4 *The wheelchair behaviour ratings will correlate with the reported elicited affective state of the participant.*

Hypothesis 5 *The resulting quality of the trajectory will correlate with the wheelchair affective ratings and/or self-reported emotional states. Specifically, there will be a significant correlation between the affect attributed to the wheelchair or the individual emotional state and*

5.a *the facial expressions of emotion;*

5.b *the degree of user input modification;*

5.c *the cognitive workload;*

5.d *the perceived safety;*

5.e *the perceived difficulty of use.*

4.3 Analyses and results

We ran a total of 240 trials (i.e., 20 participants x 6 interventions x 2 tests) out of which we discarded 19 due to technical failures. We conducted measurement reliability and participant repeatability tests to ensure consistency in the participant responses (Section A.3). Given that participants' responses were similar in roughly 75% of the ratings between test 1 and test 2, we averaged the evaluation, potency, and activity scores for each intervention into a single response leaving a total of 120 sets of EPA profiles. In the following sections we will elaborate on our analysis and findings from our exploratory study.

4.3.1 Consensus analysis

To assess the similarity in user attitudes toward the smart-wheelchair's assisted navigation behaviours, we use the consensus analysis theory and methodology developed by Borgatti et al. [78]. Consensus analysis provides a way of conceptualizing and coping with individual variability and it is based on the following three assumptions:

1. *Common truth*: there is a single right answer for every question.
2. *Conditional independence*: responses are independent from each other both across questions and across other respondents.
3. *Item homogeneity*: all questions are on the same topic for which all subjects have a uniform level of knowledge.

The consensus method consists of constructing a person-by-intervention response matrix X, in which element x_{ij} corresponds to the response from person i to intervention j. Next, a person-by-person similarity matrix M is computed in which element m_{ij} contains the correlation between respondents i and j, with the correlation being computed across the respondents' answers to all interventions. For example, 20 subjects rated the potency of 6 smart-wheelchair behaviours; the potency correlation matrix is 20 x 20, and each correlation is computed over 6 observations. The next step consists of conducting a principal component analysis on the correlation matrix resulting in a set of eigenvectors and corresponding eigenvalues. Evidence for agreement among participants can be assessed using the eigenvalues: if the largest eigenvalue is at least twice as big as the second one, the agreement is considered significant [78, 79].

We evaluate the level of consensus among participants separately for the evaluation, potency, and activity scores of the 6 intervention modes. Table 4.3 presents the eigenvalues obtained in component analyses of correlations between the participants' ratings across all modes. The first eigenvalue was at least twice as big as the second in the evaluation and activity dimensions, providing evidence that a dominant factor governed the assessments made by all respondents. Regarding the potency, the first eigenvalue was not significantly larger than the second suggesting a lack of agreement among participants. Furthermore, the ratio between the third and fourth eigenvalue was larger than the other two for this case, indicating that more than one factor was influencing the potency ratings. As detailed in Section A.3, we found evidence of a tendency for participants to duplicate potency and activity ratings.

Table 4.3: Eigenvalues obtained in principal component analyses of the correlation matrices across participants for Evaluation, Potency, and Activity.

	Eigenvalue				Ratio		
	1	2	3	4	1 to 2	2 to 3	3 to 4
Evaluation	19.24	6.68	4.63	3.82	2.88	1.44	1.21
Potency	13.40	7.14	4.36	1.78	1.86	1.64	2.45
Activity	12.53	4.92	3.91	3.15	2.55	1.26	1.24

4.3.2 Intervention effect on wheelchair E, P, and A ratings

Consensus analysis revealed that users agree on evaluation and activity factors of the intervention judgments across all modes. In the following, we assess whether those judgments are different depending on the condition. For example, we want to discard the possibility that all ratings ended up in the same octant of the affective space. A one-way repeated-measures multivariate analysis of variance (MANOVA) was conducted to evaluate whether there is a difference between the combined dependent variables (i.e., evaluation, potency, and activity) for the six smart-wheelchair behaviours. There was a statistically significant difference in the combined E, P, and A ratings depending on the intervention mode $F(15,257) = 2.802$, $p < 0.005$, Wilks' $\Lambda = 0.658$, partial $\eta^2 = 0.130$. We performed post hoc univariate analyses to determine which interventions were different and which dimension caused the distinction. A Bonferroni correction was applied, resulting in a significance level set at $p \leq 0.003$. We found significant differences in the evaluation scores of the following intervention pairs: DSC / high-level, high-level / disagreement, high-level / collision avoidance, efficiency / collision avoidance, and collision avoidance / steering correction with speed limits. We did not find significant differences in potency and activity ratings between interventions. From the pairwise comparison results between modes, and with the help of Figure 4.3, we can recognize that collision avoidance, DSC, and disagreement blend have similar patterns of evaluation scores making them practically indistinguishable from each other. On the other hand, efficiency, high-level blend, and steering correction with speed limits also have similar ratings between them.

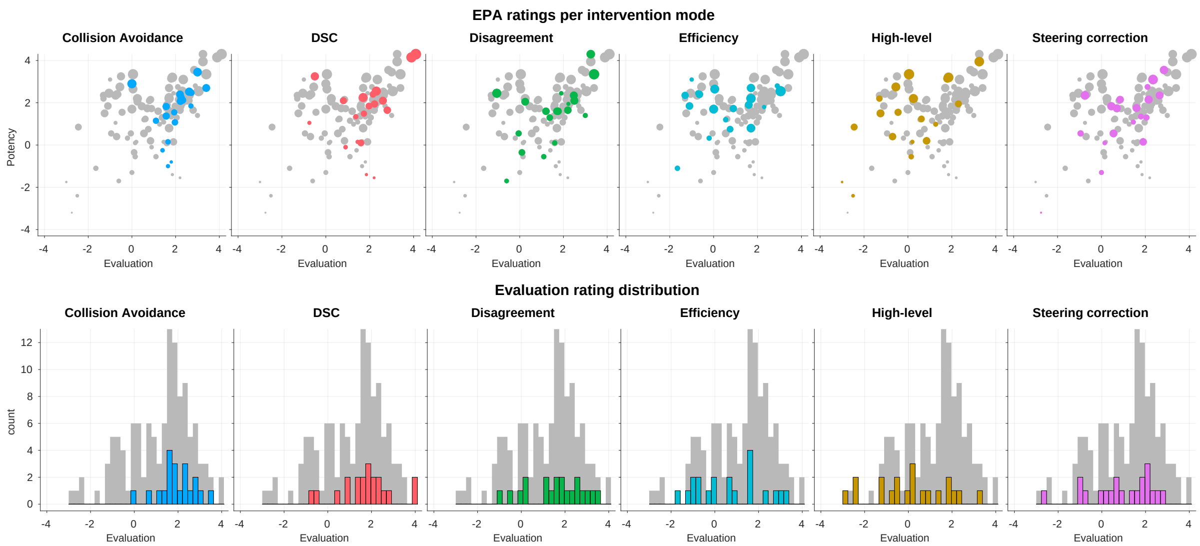


Figure 4.3: Smart-wheelchair scores in evaluation, potency, and activity dimensions of emotion. Activity scores are represented with the size of the point, larger points corresond to higher activity values. Each column corresponds to an intervention mode, data from all modes is plotted in the background in grey.

4.3.3 Intervention effect on user affective state

We analyze the elicited user affective state with two emotion-measuring instruments: a post interaction self-report (subjective component) and overall experience valence estimated using facial expression emotion recognition (physiological component). Even though we have estimates for seven discrete emotions from the facial expressions, we conduct the following preliminary analyzes using only the overall estimated experience valence.

User affect self-reports

As described earlier, participants were instructed to report their affective state by choosing as many words as needed from a set of 20 emotion-related words upon completion of the driving task. We examined the number of times each word was used to describe the elicited emotions across interventions. With 20 participants and 2 trials per intervention, the maximum number of times a word can be selected is 40. The counts per word are presented in Figure 4.4, which illustrates that overall, the six smart-wheelchair behaviours elicit more positive than negative emotions. The intervention mode that elicited the highest percentage of positively-valenced emotions (31.1%) was DSC while high-level blend elicited the least (17.3%). high-level blend also elicited the highest number of negatively-valenced emotions (12.6%) and the collision avoidance mode elicited the least (3.55%) closely followed by DSC (3.71%). Furthermore, while the count distribution across words is mostly uniform, a pattern between two groups of interventions is recognizable. The first group contains DSC, disagreement blend, and collision avoidance whereas the second group contains high-level blend, efficiency, and steering correction with speed limits. The first group tends to elicit more positively-valenced emotions (e.g., *confident* and *satisfied*) and less negatively-valenced (e.g., *irritated* and *annoyed*) than the second group.

We continued our analyses of the self-reported user affective state by constructing a resulting emotion vector for each trial. We used the EPA mappings for every emotion-related word (from the EPA dictionary), and computed a resulting emotion vector by averaging the EPA values of the selected words. Subtle differences in the EPA mappings depending on the gender of the participant were accounted for. We then conducted a one way-repeated measures MANOVA test to find differences in the participants' affective state depending on the intervention mode. Significant differences in the emotions reported by the users were found F(15, 257) = 2.746, $p = 0.001$, Wilks' $\Lambda = 0.664$, partial $\eta^2 = 0.128$. Post hoc pairwise univariate analyses with Bonferroni correction (so $p < 0.003$) revealed significant differences in the reported evaluation and potency components of the resulting emotion vectors between the following pairs of interventions: DSC / high-level, DSC / efficiency, high-level / collision avoidance, efficiency / collision avoidance, and disagreement / efficiency. DSC elicits significantly more positive and more potent emotions than high-level and efficiency. High-level and efficiency elicit more negative and less potent emotions than collision avoidance. Disagreement blend elicits more potent emotions than efficiency based blend. The activity levels of elicited emotions were similar across all intervention modes.

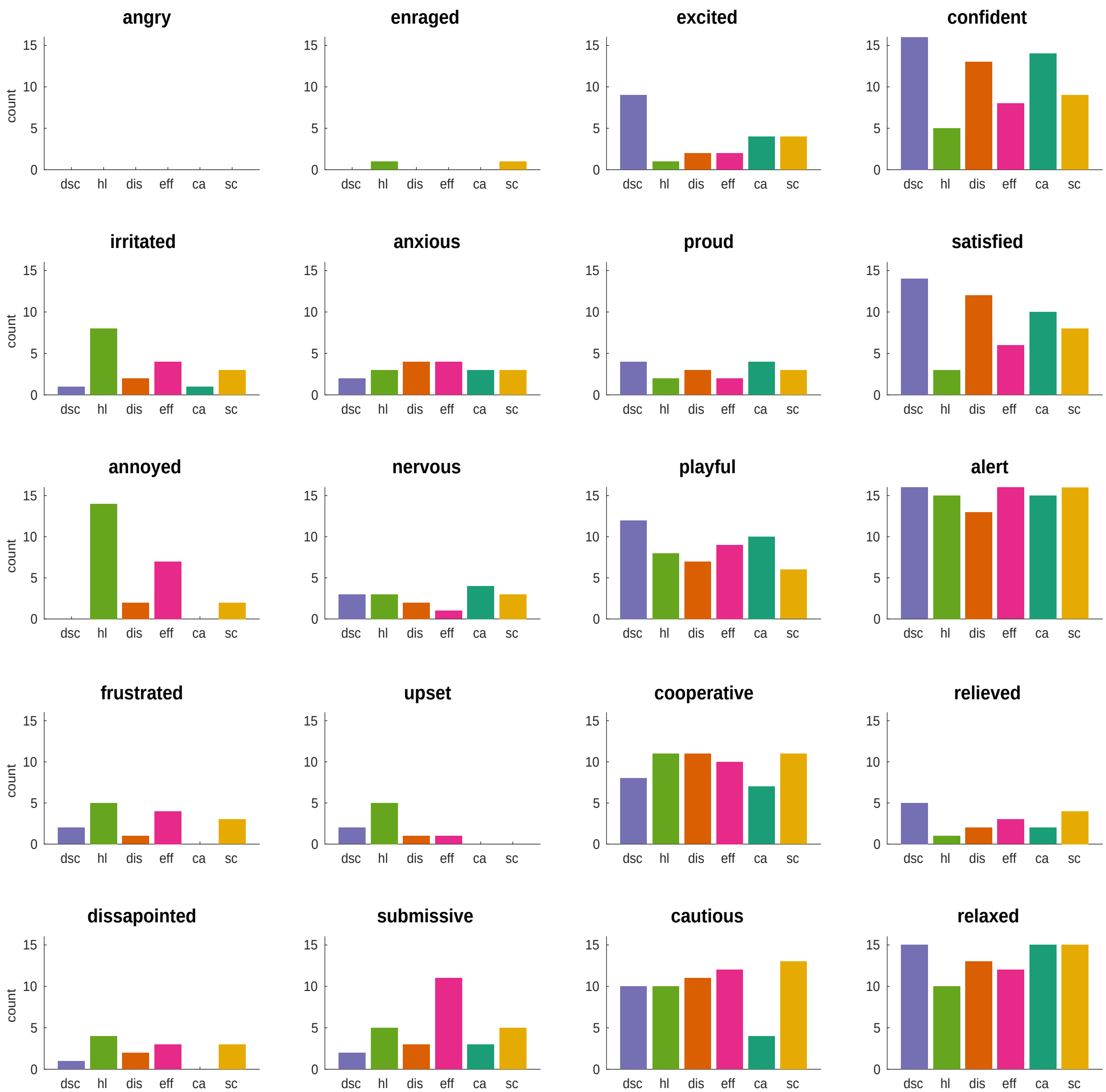


Figure 4.4: Number of times each word was selected to describe elicited affective state per intervention mode. Words ordered by valence from left to right (negative valence on the left and positive on the right) and by a combination of potency and activity from top to bottom (high potency/activity on top and low on the bottom).

Estimated experience valence

We used the Affectiva SDK to estimate the subject's valence from facial expressions recorded at 20 frames per second. The software computes a frame-by-frame valence metric based on a set of observed facial expressions; for example, if a smile is observed the likelihood of positive valence is increased, and if a brow furrow is observed, the likelihood of negative valence is increased. We averaged the resulting valence values over time for each task separately. We then looked for relationships between the averaged valence, wheelchair judgments, and self-reported emotions. A statistically significant correlation between the averaged estimated valence from facial data with the resulting emotion vector from the self-reports or the wheelchair judgments did not appear (all $p \geq 0.053$).

User affect relationship with wheelchair judgments

To assess whether the user affective state is consistent with the affect attributed to the wheelchair, we computed the correlations between the resulting emotion vector of each participant and the powered wheelchair rating separately for each dimension of emotion. We found a moderate significant linear relationship between the evaluation of the wheelchair and the evaluation of the averaged emotion vector ($r = 0.63$, $p < 0.005$), but no relationships between the activity and potency judgments ($r = 0.04$ and $r = -0.03$ respectively).

We conducted further analyses to explore how the wheelchair judgments and the self reported user affect related to the user control weight, modifications to the user commanded velocities, cognitive workload (NASA TLX), and perceived safety and difficulty. Figure 4.5 summarizes the product-moment correlation results. The first three columns represent the evaluation, potency, and activity ratings attributed to the wheelchair behaviour and the following three columns represent the computed EPA vector from the reported user emotions. The first row, labeled *Facial valence*, corresponds to the time averaged valence from the Affectiva SDK. The following row, *User weight (α)*, is the time averaged amount of control granted to the user throughout the trial. *Angular change (%)* represents the time averaged modification of the user's commanded angular speed as a percentage of the maximum possible modification. Similarly, *Linear change (%)* is the time averaged modification of the participant's linear speed as a percentage of the maximum possible modification. The final three rows—*Workload*, *Perceived safety*, and *Perceived difficulty*—are self-reported values from the post-interaction questionnaire.

As presented in Figure 4.5, the user's control weight and the degree of user command modification (both linear and angular) have an impact on the evaluation of the wheelchair. Contrary to our expectations in Table 4.1, they did not correlate with the potency and activity ratings attributed to the wheelchair. On the other hand, the same parameters have a small relationship with the reported user affective state. The reported workload and perceived difficulty have moderate negative correlations with the evaluation scores attributed to the wheelchair and the evaluation and potency components of the resulting emotion vector of each participant. Finally, the reported perceived safety has a moderate positive correlation with the evaluation score of the wheelchair.

4.4 Discussion

We conducted an experiment to explore the emotional responses to different assisted navigation interventions on a smart-wheelchair. In this section, we discuss the findings and answer our research questions. At the end of this section we elaborate on the limitations and challenges associated to this research.

4.4.1 What affective interpretation do users attribute to different smart-wheelchair behaviours?

Furthermore, are those interpretations consistent across users?

Hypothesis 1 *There will be consensus/agreement across participants of affective ratings on each dimension of emotion for each intervention mode. –Partially supported.*

We followed consensus analysis methodology to evaluate the level of consensus among participants in each dimension of emotion. By examining the eigenvalues of the correlation matrices across participants, we found evidence of agreement on two dimensions of emotion: evaluation and activity. However, we also noted a moderate tendency to duplicate potency and activity ratings. These results relate to

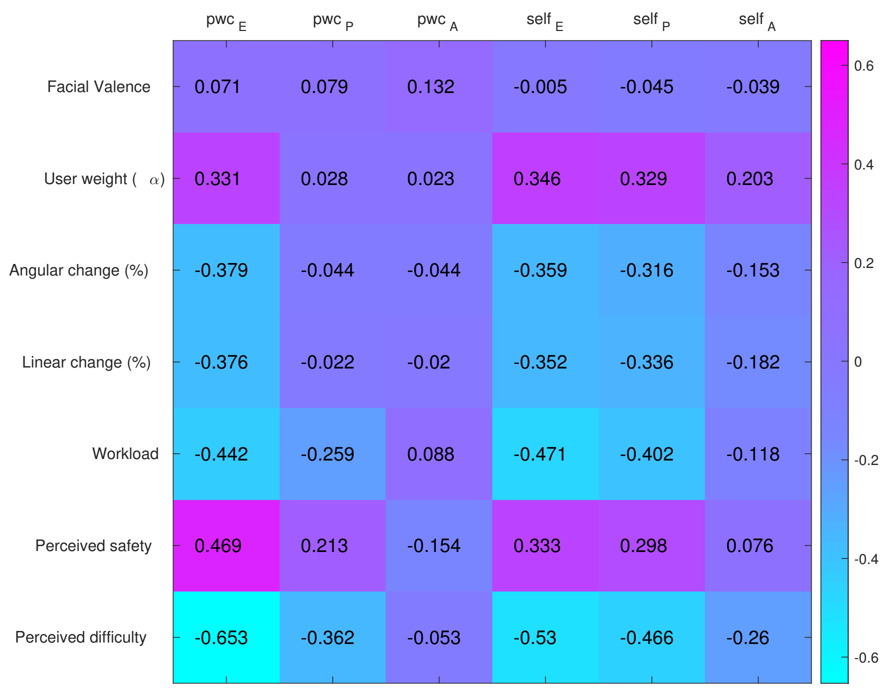


Figure 4.5: Correlation coefficients r between wheelchair judgments and reported user affect with estimated valence from facial expressions, averaged user control weight, percentage of modification to user angular and linear velocity, self-reported workload, perceived safety, and difficulty to use. All $|r| > 0.154$ found significant with $p < 0.05$.

those obtained by Malhotra et al. [60] which observed low levels of agreement in potency and activity dimensions when introducing non-verbal behaviours to a virtual character. Moreover, the lack of consensus on the potency dimension, as well as the large variance in activity scores might derive from different interpretations of the dimensions. We did not instruct the respondents to associate the dimensions of emotion with any particular characteristic of the wheelchair and left it to them to decide. Some respondents indicated that they associated activity/potency with the amount of resistance presented by the wheelchair while others associated it with an opposite concept such as the obedience of the wheelchair.

Hypothesis 2 *There will be a distinct pattern of affective ratings for each intervention mode. –Not supported*

Even though we carefully chose and tuned the shared control interventions for smart-wheelchairs to represent what we thought would be distinguishable behaviours in the EPA affective space, the participant's ratings for potency and activity were not significantly different across modes. Only for the evaluation assessments were there significant differences depending on the intervention mode. Post hoc analyses delineate two clusters of similarly rated intervention modes:

- **Group 1**. Interventions where by design the user's control weight is high for most of the trial: dynamic shared control, disagreement based blend, and collision avoidance. This group has significantly more positive scores in the evaluation dimension.
- **Group 2**. Interventions where by design the planner's control weight is high: high-level blend, efficiency-based blend, and steering correction with speed limits. This group has significantly more negative scores in the evaluation dimension.

This result corroborates that users favour control paradigms that grant higher levels of autonomy [11, 27, 28, 64]. Moreover, there is not a single control paradigm that elicits only positive emotions for every user, suggesting that it is imperative to accommodate individual preferences.

4.4.2 How does the user's affective state change when interacting with various wheelchair behaviours?

Moreover, are those changes consistent across users and consistent with the affect attributed to the wheelchair?

Hypothesis 3 *There will be a distinct pattern of evoked emotions for each navigation mode. –Partially supported*

When inspecting the reported emotions experienced by the participants, we found that interventions which grant higher levels of autonomy to the user elicit more positively-valenced and fewer negatively-valenced emotions than those that grant more control to the robotic planner. We then mapped the reported emotions into the EPA space of affective meaning and found differences in the evaluation and potency components of the DSC vs High-level, DSC vs Efficiency, High-level vs Collision Avoidance, Efficiency vs Collision Avoidance, and Disagreement vs Efficiency interventions. It should be noted that a limitation of this analysis is that the emotion-related words used for the self-report are subject to misinterpretation by the subjects.

Although the processed recordings of the participant's facial expressions were not used in depth during this analysis, they provide an insight of the user experience at a much higher sampling rate than post-trial self-reports and could prove useful in future investigation efforts.

Hypothesis 4 *The wheelchair behaviour ratings will correlate with the reported elicited affective state of the participant. –Partially supported*

Participants reported their affective state by selecting from a set of emotion-related words. We constructed a resulting emotion vector by averaging the emotional mappings of each word from the Canadian EPA dictionary. A significant correlation between the evaluation score attributed to the wheelchair behaviour and the user self-reported affect was found. However, no significant similarities between the potency and activity dimensions were apparent. It is important to note that during our analysis we mapped the reported emotions to the EPA space assuming a uniform weight (i.e., each emotion was experienced equally) which may not necessarily reflect what the user felt.

Hypothesis 5 *The quality of the trajectory will correlate with the wheelchair affective ratings and/or individual emotional states. Specifically, there will be a significant correlation between the affect attributed to the wheelchair or the participant's emotional state and:*

5.a *The facial expressions of emotion. –Not supported*
We averaged the estimated experienced valence over the trial and found no significant correlation with the affective judgments of the wheelchair or the self-reported affective state. We note that this admittedly limited form of analysis

could be improved by integrating the remaining seven discrete emotion values estimated with the Emotion SDK. Furthermore, analyzing the data throughout the whole driving task could potentially help to detect cognitive states relevant to driving such as user fatigue, confusion, and frustration.

5.b *The degree of user input modification. – Partially supported*

The time averaged user control weight over the driving task has a weak positive correlation with the evaluative dimension of wheelchair behaviour as well as the reported user affect. On the other hand, the averaged linear and angular speed modifications have a weak negative correlation with the evaluative dimension of the wheelchair scores. Moreover, as the degree of modification in the user's commands increases, the user's feelings tend toward the bad and powerless part of the EPA space.

5.c *The cognitive workload. –Supported*

The reported cognitive workload has a moderate negative correlation with the evaluative ratings of the wheelchair as well as the elicited user affect.

5.d *The perceived safety. –Supported*

The levels of perceived safety correlate positively with the evaluation of the wheelchair but not so much with the evoked user affective state.

5.e *The perceived difficulty of use. –Supported*

Difficulty of use has a strong negative correlation with evaluation and a substantial negative impact on the user's affective state. Usability should be improved to ensure a positive user experience. Open ended comments from the study participants indicate that a lack of feedback on what the wheelchair is trying to accomplish and not moving in their desired direction contribute to the difficulty of use.

4.4.3 Limitations

A significant limitation of this study is the fact that we are attempting to characterize the behaviour of a collaborative mobile robot as felt by the user sitting on it. Since the users have different driving styles, the amount of contribution from the robotic planner varies even for the same intervention mode. For example, some drivers always keep a considerable distance from obstacles; consequently, some control sharing paradigms may not intervene as strongly as with others. Moreover, we conducted our experiments in a laboratory setting which may have influenced the participants' judgments. Often we noticed that interventions that take too much control away from the user led to frustration and annoyance; however, we cannot assume that cognitively impaired older adults will have the same responses because our test subjects are quite able to perform safe navigation independently. Nonetheless, this study provides building blocks for potential future studies with the target user population.

Recognizing an individual's affective state is challenging, and there is no gold standard for the measurement of emotion. Estimates constructed from self-report or facial expressions each have their own issues, and perhaps neither should be treated as the ground truth. Self-reporting is limited in the sense that users might not be aware of their own emotional state, or they might not be willing to report it. Additionally, self-reported data can only be collected after the interaction because asking participants to provide continuous ratings while they are operating the wheelchair is infeasible and unsafe. Although self-assessments are among the most commonly used methods of evaluation in HRI studies, they lack objectivity and can be unreliable, as they reflect a large amount of individual differences [50]. Specifically for our study, we used a set of emotion-related words for users to report their experienced emotion which raises the question of whether the reports reflect differences in actual feelings or merely differences in how participants understand the words. On the other hand, facial expressions are subject to many factors such as gender, culture, expressiveness, and the inferred presence of an audience [39]. A significant challenge around facial expressions in this study was just capturing facial data consistently for every subject. When participants turned their heads to prepare for turning or backing up, the fixed camera could only catch the side of their faces. Consequently, the facial expression recognition software cannot find markers for those frames, resulting in a discontinuity in the estimated emotion values.

Finally, characterizing robotic behaviours on a three-dimensional affective space using the semantic differential technique is also limited because it only gets the connotative meaning, not the denotative one. This limitation can result in users allocating two different behaviours in approximately the same part of the affective space without implying that they are equal. The technique identifies that something is good or pleasant and strong or weak, but it does not directly identify the root cause.

Chapter 5

Conclusions

In this thesis we explored user affect as a potential input for the behaviour of a smart-wheelchair. We compared a set of shared control policies, not with the goal of finding one that beats the others, but with the purpose of characterizing the resulting behaviours in terms of the effect their behaviour had on the affect of the driver. Twenty healthy subjects tested the driving assistance interventions and judged them with respect to three basic dimensions of affective meaning: evaluation, potency, and activity. We also asked participants to report their emotions after the interaction using emotion-related words. Consensus analysis and analyses of variance were used to determine the agreement among participant judgments and the effects of the interventions on the affective state of the users. Results show that the interventions are rated consistently among participants in evaluation and activity dimensions but not so much in the potency. The different intervention modes had a significant impact on the self-reported affective state of the participants. Furthermore, judgments for interventions that take more control from the user tend to have more negative evaluations and a negative impact on the user's emotions. Adjusting the affect of the shared control intervention could potentially enable a communication channel through which users could interact more naturally with their smart mobility device and thereby enhance their own autonomy.

5.1 Future work

In our study, we collected data from the PWC sensors throughout each driving task, but we analyzed it based on a single summary value (a time average over each task). This analysis could be enriched by looking at the data as a time series. In order to get a real-time estimation of the driver's emotional state, it is crucial to take into account the interaction between the wheelchair and the user through the control interface, and also track emotions longitudinally to create an emotional profile. In the future, we could leverage the information from the user joystick, the environment (lasers and costmaps), and/or the current state of the wheelchair (pose and velocities) to give some context to the reported emotions and the estimated values from the facial recognition software. Analyzing joystick data is not uncommon in the shared control literature; for example, Carlson and Demiris [80, 81] analyze the smoothness of the joystick signals as characterized by the jerk (the change in acceleration) to measure the efficiency of the generated trajectories. We could look for relationships between joystick motion metrics, such as jerk, and emotion values from self-reported measures or facial expression emotion recognition data. There are also several additional ways to characterize the quality of a given trajectory; for example, we could look for relationships between metrics such as trajectory smoothness, fluency of the velocities or distance from obstacles and the affective ratings of the wheelchair's behaviour or the participant's reported emotions. Finally, we could use that data to make timely predictions on specific affective states that are critical for wheelchair navigation and prevalent in the older adult community; for example, fatigue, irritation, and confusion.

A key limitation of this exploratory study is that we used cognitively intact healthy adults rather than our target population of older adults with significant mobility and mild to moderate cognitive impairments. Before conducting a study with the latter population we need to improve our robotic platform and refine our experimental protocol.

5.1.1 Improvements to experimental protocol

We observed duplication between potency and activity ratings of the wheelchair behaviours. This suggests that a two dimensional model of emotion may be sufficient for our application. Furthermore, we only observed significant differences in the affective interpretations between two clusters of interventions. For future studies, we could select and fine tune only two or three representative interventions to reduce participant burden during the trials.

5.1.2 Improvements to robotic platform

The robotic platform can be enhanced in both hardware and software. We experienced multiple technical failures due to interference between the laser rangefinders. When testing with healthy adults, we simply explained the situation, fixed the issue, and moved on to the next trial. However, having constant failures during the trials may lead to lack of acceptance of robotic assistive technologies. Other research groups rely on RGB-D sensors for localization and obstacle detection, although vision-based approaches are not without their own shortcomings and challenges. Moreover, the scan-based odometry estimation is never as reliable as odometry from wheel encoders. Mounting custom wheel encoders is challenging, but it could prove useful for better state estimation. Regarding the software improvements, we need to consider that the performance of the shared control algorithm significantly depends on the quality of the incoming signals (i.e., the automation and the user). We used a robotic planner designed to control unmanned mobile robots. However, wheelchairs have humans riding the robot which means that the generated controls need to be intuitive for humans. For example, in some situations the planner would prefer to stop, turn, and then move forward, while the user would prefer to drive in a curved trajectory without stopping. We may be able to overcome such disagreements if we use a planner that better accounts for what a human would do in a similar situation.

Bibliography

[1] L. I. Iezzoni, E. P. McCarthy, R. B. Davis, and H. Siebens, "Mobility difficulties are not only a problem of old age," *Journal of general internal medicine*, vol. 16, no. 4, pp. 235–243, 2001. doi:10.1046/j.1525-1497.2001.016004235.x. → page 1

[2] T.-V. How, R. H. Wang, and A. Mihailidis, "Evaluation of an intelligent wheelchair system for older adults with cognitive impairments," *Journal of neuroengineering and rehabilitation*, vol. 10, no. 1, p. 90, 2013. doi:10.1186/1743-0003-10-90. → pages 1, 6, 13

[3] M. Shields, "Use of wheelchairs and other mobility support devices," *Health reports*, vol. 15, no. 3, p. 37, 2004. → page 1

[4] R. C. Simpson, "Smart wheelchairs: A literature review," *Journal of rehabilitation research and development*, vol. 42, no. 4, pp. 423–436, 2005. → pages 1, 2, 7

[5] I. Pettersson, G. Ahlström, and K. Törnquist, "The value of an outdoor powered wheelchair with regard to the quality of life of persons with stroke: a follow-up study," *Assistive technology*, vol. 19, no. 3, pp. 143–153, 2007. doi:10.1080/10400435.2007.10131871. → page 1

[6] A. Brandt, S. Iwarsson, and A. Stahle, "Older people's use of powered wheelchairs for activity and participation," *Journal of rehabilitation medicine*, vol. 36, no. 2, pp. 70–77, 2004. doi:10.1080/16501970310017432. → page 1

[7] R. C. Simpson, E. F. LoPresti, and R. A. Cooper, "How many people would benefit from a smart wheelchair?," *Journal of Rehabilitation Research & Development*, vol. 45, no. 1, pp. 53–72, 2008. doi:10.1682/JRRD.2007.01.0015. → page 1

[8] A. Davies, L. D. Souza, and A. Frank, “Changes in the quality of life in severely disabled people following provision of powered indoor/outdoor chairs,” *Disability and Rehabilitation*, vol. 25, no. 6, pp. 286–290, 2003. doi:10.1080/0963828021000043734. → page 1

[9] A. Mihailidis, P. Elinas, J. Boger, and J. Hoey, “An intelligent powered wheelchair to enable mobility of cognitively impaired older adults: An anticollision system,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 15, no. 1, pp. 136–143, 2007. doi:10.1109/TNSRE.2007.891385. → page 1

[10] J. Leaman and H. M. La, “A comprehensive review of smart wheelchairs: past, present, and future,” *IEEE Transactions on Human-Machine Systems*, vol. 47, no. 4, pp. 486–499, 2017. doi:10.1109/THMS.2017.2706727. → pages 2, 7

[11] P. Viswanathan, E. P. Zambalde, G. Foley, J. L. Graham, R. H. Wang, B. Adhikari, A. K. Mackworth, A. Mihailidis, W. C. Miller, and I. M. Mitchell, “Intelligent wheelchair control strategies for older adults with cognitive impairment: user attitudes, needs, and preferences,” *Autonomous Robots*, vol. 41, no. 3, pp. 539–554, 2017. doi:10.1007/s10514-016-9568-y. → pages 2, 14, 26, 52

[12] G. Loewenstein and J. S. Lerner, “The role of affect in decision making,” in *Handbook of affective science*, pp. 619–642, Oxford University Press, 2003. → page 2

[13] F. Eyben, M. Wöllmer, T. Poitschke, B. Schuller, C. Blaschke, B. Färber, and N. Nguyen-Thien, “Emotion on the road necessity, acceptance, and feasibility of affective computing in the car,” *Advances in human-computer interaction*, vol. 2010, 2010. doi:10.1155/2010/263593. → pages 2, 11

[14] J. M. Robillard and J. Hoey, “Emotion and motivation in cognitive assistive technologies for dementia,” *Computer*, vol. 51, no. 3, pp. 24–34, 2018. doi:10.1109/MC.2018.1731059. → page 2

[15] A. König, L. E. Francis, J. Joshi, J. M. Robillard, and J. Hoey, “Qualitative study of affective identities in dementia patients for the design of cognitive assistive technologies,” *Journal of Rehabilitation and Assistive Technologies Engineering*, vol. 4, pp. 1–15, 2017. doi:10.1177/2055668316685038. → page 2

[16] D. Bachman and P. Rabins, "Sundowning and other temporally associated agitation states in dementia patients," *Annual Review of Medicine*, vol. 57, pp. 499–511, 2006. doi:10.1146/annurev.med.57.071604.141451. → page 3

[17] D. A. Abbink and M. Mulder, "Neuromuscular analysis as a guideline in designing shared control," in *Advances in haptics*, InTech, 2010. → pages 5, 6

[18] D. A. Abbink, M. Mulder, and E. R. Boer, "Haptic shared control: smoothly shifting control authority?," *Cognition, Technology & Work*, vol. 14, pp. 19–28, Mar 2012. doi:10.1007/s10111-011-0192-5. → pages 5, 6

[19] J. Wei, J. M. Snider, J. Kim, J. M. Dolan, R. Rajkumar, and B. Litkouhi, "Towards a viable autonomous driving research platform," in *Intelligent Vehicles Symposium (IV)*, pp. 763–770, IEEE, 2013. doi:10.1109/IVS.2013.6629559. → page 5

[20] S. S. Nudehi, R. Mukherjee, and M. Ghodoussi, "A shared-control approach to haptic interface design for minimally invasive telesurgical training," *IEEE Transactions on Control Systems Technology*, vol. 13, no. 4, pp. 588–592, 2005. doi:10.1109/TCST.2004.843131. → page 5

[21] A. D. Dragan and S. S. Srinivasa, "A policy-blending formalism for shared control," *The International Journal of Robotics Research*, vol. 32, no. 7, pp. 790–805, 2013. doi:10.1177/0278364913490324. → page 5

[22] N. Matni and M. Oishi, "Reachability-based abstraction for an aircraft landing under shared control," in *American Control Conference*, pp. 2278–2284, IEEE, 2008. doi:10.1109/ACC.2008.4586831. → page 5

[23] M. Steele and R. B. Gillespie, "Shared control between human and machine: Using a haptic steering wheel to aid in land vehicle guidance," in *Proceedings of the human factors and ergonomics society annual meeting*, vol. 45, pp. 1671–1675, HFES, 2001. doi:10.1177/154193120104502323. → page 5

[24] D. Katzourakis, M. Alirezaei, J. C. de Winter, M. Corno, R. Happee, A. Ghaffari, and R. Kazemi, "Shared control for road departure prevention," in *IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pp. 1037–1043, IEEE, 2011. doi:10.1109/ICSMC.2011.6083811. → page 5

[25] S. P. Parikh, V. Grassi Jr, V. Kumar, and J. Okamoto Jr, "Integrating human inputs with autonomous behaviors on an intelligent wheelchair platform,"

IEEE Intelligent Systems, vol. 22, no. 2, pp. 33–41, 2007. doi:10.1109/MIS.2007.36. → page 6

[26] I. M. Mitchell, P. Viswanathan, B. Adhikari, E. Rothfels, and A. K. Mackworth, "Shared control policies for safe wheelchair navigation of elderly adults with cognitive and mobility impairments: Designing a wizard of oz study," in *American Control Conference (ACC)*, pp. 4087–4094, IEEE, 2014. doi:10.1109/ACC.2014.6859446. → pages 6, 26

[27] B. D. Argall, "Modular and adaptive wheelchair automation," in *Experimental robotics. Springer Tracts in Advanced Robotics*, vol. 109, pp. 835–848, 2016. doi:10.1007/978-3-319-23778-7_55. → pages 6, 52

[28] T. Carlson and Y. Demiris, "Human-wheelchair collaboration through prediction of intention and adaptive assistance," in *International Conference on Robotics and Automation, (ICRA)*, pp. 3926–3931, IEEE, 2008. doi:10.1109/ROBOT.2008.4543814. → pages 6, 52

[29] C. Urdiales, J. M. Peula, C. Barrué, U. Cortés, F. S. Hernández, C. Caltagirone, and R. Annicchiarico, "An adaptive scheme for wheelchair navigation collaborative control," in *AAAI Fall Symposium: AI in Eldercare: New Solutions to Old Problems*, pp. 134–141, 2008. url: http://www.aaai.org/Library/Symposia/Fall/2008/fs08-02-022.php. → pages 6, 22

[30] J. Wang and M. Lewis, "Mixed-initiative multirobot control in usar," in *Human Robot Interaction*, pp. 405–422, InTech, 2007. → page 7

[31] D. Ding and R. A. Cooper, "Electric powered wheelchairs," *IEEE Control Systems*, vol. 25, no. 2, pp. 22–34, 2005. doi:10.1109/MCS.2005.1411382. → page 7

[32] R. W. Picard, "Affective computing," in *M.I.T Media Laboratory Perceptual Computing Section Technical Report No. 321*, MIT Media Laboratory, 1995. → page 7

[33] R. W. Picard, "Affective computing: challenges," *International Journal of Human-Computer Studies*, vol. 59, no. 1-2, pp. 55–64, 2003. doi:10.1016/S1071-5819(03)00052-1. → page 7

[34] R. A. Calvo and S. Mac Kim, "Emotions in text: dimensional and categorical models," *Computational Intelligence*, vol. 29, no. 3, pp. 527–543, 2013. doi:10.1111/j.1467-8640.2012.00456.x. → page 7

[35] R. A. Calvo, S. D'Mello, J. Gratch, and A. Kappas, "Introduction to affective computing," in *The Oxford handbook of affective computing*, Oxford Library of Psychology, 2015.

[36] R. A. Calvo and S. D'Mello, "Affect detection: An interdisciplinary review of models, methods, and their applications," *IEEE Transactions on affective computing*, vol. 1, no. 1, pp. 18–37, 2010. doi:10.1109/T-AFFC.2010.1. → page 8

[37] J. Tao and T. Tan, "Affective computing: A review," in *International Conference on Affective computing and intelligent interaction*, pp. 981–995, Springer, 2005. doi:10.1007/11573548_125. → page 7

[38] H. Gunes, B. Schuller, M. Pantic, and R. Cowie, "Emotion representation, analysis and synthesis in continuous space: A survey," in *International Conference on Automatic Face & Gesture Recognition and Workshops*, pp. 827–834, IEEE, 2011. doi:10.1109/FG.2011.5771357. → pages 8, 37

[39] I. B. Mauss and M. D. Robinson, "Measures of emotion: A review," *Cognition and emotion*, vol. 23, no. 2, pp. 209–237, 2009. doi:10.1080/02699930802204677. → pages 8, 37, 38, 55

[40] J. A. Russell, "A circumplex model of affect," *Journal of personality and social psychology*, vol. 39, no. 6, p. 1161, 1980. doi:10.1037/h0077714. → page 8

[41] K. R. Scherer, A. Schorr, and T. Johnstone, *Appraisal processes in emotion: Theory, methods, research*. Oxford University Press, 2001. → page 8

[42] A. Ortony, G. L. Clore, and A. Collins, *The cognitive structure of emotions*. Cambridge university press, 1990. → page 8

[43] J. Gratch and S. C. Marsella, "Appraisal models," in *The Oxford handbook of affective computing*, Oxford Library of Psychology, 2015. → page 8

[44] D. R. Heise, *Understanding events: Affect and the construction of social action*. CUP Archive, 1979. → page 8

[45] D. R. Heise, *Expressive order: Confirming sentiments in social actions*. Springer Science & Business Media, 2007. → page 8

[46] L. Troyer, "Affect control theory as a foundation for the design of socially intelligent systems," in *Proceedings of AAAI Spring Symp. on Architechtures for Modeling Emotion*, 2004. → page 8

[47] J. Hoey, T. Schroder, and A. Alhothali, "Bayesian affect control theory," in *Humaine Association Conference on Affective Computing and Intelligent Interaction (ACII)*, pp. 166–172, IEEE, 2013. doi:10.1109/ACII.2013.34. → page 8

[48] C. E. Osgood and J. G. Snider, *Semantic differential technique; A sourcebook*. Aldine Publishing Company, 1969. → page 9

[49] R. C. Arkin, "Moving up the food chain: Motivation and emotion in behavior-based robots," in *Mobile Robot Laboratory Publications*, Georgia Institute of Technology, 2003. url: http://hdl.handle.net/1853/21306. → page 10

[50] R. C. Arkin and L. Moshkina, "Affect in human-robot interaction," in *The Oxford Handbook of Affective Computing*, Oxford Library of Psychology, 2015. → pages 10, 55

[51] C. Breazeal, D. Kerstin, and K. Takayuki, "Social robotics," in *Springer handbook of robotics*, Springer, 2016. → page 10

[52] H. Miwa, K. Itoh, M. Matsumoto, M. Zecca, H. Takanobu, S. Rocella, M. C. Carrozza, P. Dario, and A. Takanishi, "Effective emotional expressions with expression humanoid robot we-4rii: integration of humanoid robot hand rch-1," in *Proceedings of the International Conference on Intelligent Robots and Systems, (IROS)*, vol. 3, pp. 2203–2208, IEEE, 2004. doi:10.1109/IROS.2004.1389736. → page 10

[53] T. Wu, N. J. Butko, P. Ruvulo, M. S. Bartlett, and J. R. Movellan, "Learning to make facial expressions," in *International Conference on Development and Learning (DEVLRN)*, pp. 1–6, IEEE, 2009. doi:10.1109/DEVLRN.2009.5175536.

[54] S. Park, L. Moshkina, and R. C. Arkina, "Recognizing nonverbal affective behavior in humanoid robots," in *Intelligent Autonomous Systems*, vol. 11, p. 12, 2010. doi:10.3233/978-1-60750-613-3-12.

[55] C. Breazeal, "Emotive qualities in robot speech," in *Proceedings of the International Conference on Intelligent Robots and Systems (IROS)*, vol. 3, pp. 1388–1394, IEEE, 2001. doi:10.1109/IROS.2001.977175. → page 10

[56] C. P. Lee-Johnson and D. A. Carnegie, "Emotion-based parameter modulation for a hierarchical mobile robot planning and control architecture," in *Proceedings of the International Conference on Intelligent*

Robots and Systems, (IROS), pp. 2839–2844, IEEE, 2007. doi:10.1109/IROS.2007.4399080. → page 10

[57] H. Cai, Y. Lin, and R. R. Mourant, "Study on driver emotion in driver-vehicle-environment systems using multiple networked driving simulators," in *Proceedings of the Driving Simulation Conference North America (DSC-NA)*, 2007. → page 11

[58] S. Dow, B. MacIntyre, J. Lee, C. Oezbek, J. D. Bolter, and M. Gandy, "Wizard of oz support throughout an iterative design process," *IEEE Pervasive Computing*, vol. 4, no. 4, pp. 18–26, 2005. doi:10.1109/MPRV.2005.93. → page 11

[59] A. Mihailidis, J. N. Boger, T. Craig, and J. Hoey, "The coach prompting system to assist older adults with dementia through handwashing: An efficacy study," *BMC geriatrics*, vol. 8, no. 1, p. 28, 2008. doi:10.1186/1471-2318-8-28. → page 12

[60] A. Malhotra, L. Yu, T. Schröder, and J. Hoey, "An exploratory study into the use of an emotionally aware cognitive assistant," in *AAAI Workshop: Artificial Intelligence Applied to Assistive Technologies and Smart Environments*, 2015. url: https://www.aaai.org/ocs/index.php/WS/AAAIW15/paper/viewFile/10147/10147. → pages 12, 51

[61] L. Lin, S. Czarnuch, A. Malhotra, L. Yu, T. Schröder, and J. Hoey, "Affectively aligned cognitive assistance using bayesian affect control theory," in *International Workshop on Ambient Assisted Living*, pp. 279–287, Springer, 2014. doi:10.1007/978-3-319-13105-4_41. → pages 12, 13

[62] T. Carlson and Y. Demiris, "Increasing robotic wheelchair safety with collaborative control: Evidence from secondary task experiments," in *International Conference on Robotics and Automation (ICRA)*, pp. 5582–5587, IEEE, 2010. doi:10.1109/ROBOT.2010.5509257. → page 13

[63] R. H. Wang, S. M. Gorski, P. J. Holliday, and G. R. Fernie, "Evaluation of a contact sensor skirt for an anti-collision power wheelchair for older adult nursing home residents with dementia: Safety and mobility," *Assistive Technology*, vol. 23, no. 3, pp. 117–134, 2011. doi:10.1080/10400435.2010.541406. → page 13

[64] P. Viswanathan, J. Bell, R. Wang, B. Adhikari, A. Mackworth, A. Mihailidis, W. Miller, and I. M. Mitchell, “A wizard-of-oz intelligent wheelchair study with cognitively-impaired older adults: Attitudes toward user control,” in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems Workshop on Assistive Robotics for Individuals with Disabilities: HRI Issues and Beyond*, 2014. → pages 14, 26, 52

[65] M. Gerdzhev, J. Pineau, I. M. Mitchell, P. Viswanathan, and G. Foley, “On the use of modular software and hardware for designing wheelchair robots,” in *AAAI Spring Symposium Series*, 2016. url: https://www.aaai.org/ocs/index.php/SSS/SSS16/paper/viewPaper/12704. → page 14

[66] S. Kohlbrecher, J. Meyer, O. von Stryk, and U. Klingauf, “A flexible and scalable slam system with full 3d motion estimation,” in *Proceedings of the International Symposium on Safety, Security and Rescue Robotics (SSRR)*, IEEE, November 2011. doi:10.1109/SSRR.2011.6106777. → page 16

[67] S. Thrun, W. Burgard, and D. Fox, *Probabilistic robotics*. MIT press, 2005. → page 16

[68] S. Quinlan and O. Khatib, “Elastic bands: Connecting path planning and control,” in *International Conference on Robotics and Automation (ICRA)*, pp. 802–807, IEEE, 1993. doi:10.1109/ROBOT.1993.291936. → page 18

[69] P. Trautman, “Assistive planning in complex, dynamic environments: a probabilistic approach,” in *International Conference on Systems, Man, and Cybernetics (SMC)*, pp. 3072–3078, IEEE, 2015. doi:10.1109/SMC.2015.534. → page 19

[70] Q. Li, W. Chen, and J. Wang, “Dynamic shared control for human-wheelchair cooperation,” in *International Conference on Robotics and Automation (ICRA)*, pp. 4278–4283, IEEE, 2011. doi:10.1109/ICRA.2011.5980055. → page 20

[71] C. Urdiales, M. Fernández-Carmona, J. M. Peula, R. Annicchiaricco, F. Sandoval, and C. Caltagirone, “Efficiency based modulation for wheelchair driving collaborative control,” in *International Conference on Robotics and Automation (ICRA)*, pp. 199–204, IEEE, 2010. doi:10.1109/ROBOT.2010.5509606. → pages 22, 24

[72] A. Erdogan and B. D. Argall, “The effect of robotic wheelchair control paradigm and interface on user performance, effort and preference: an

experimental assessment," *Robotics and Autonomous Systems*, vol. 94, pp. 282–297, 2017. doi:10.1016/j.robot.2017.04.013. → pages 24, 25

[73] L. F. Barrett, "Feelings or words? understanding the content in self-report ratings of experienced emotion," *Journal of personality and social psychology*, vol. 87, no. 2, pp. 266–281, 2004. doi:10.1037/0022-3514.87.2.266. → page 38

[74] D. McDuff, A. Mahmoud, M. Mavadati, M. Amr, J. Turcot, and R. e. Kaliouby, "Affdex sdk: a cross-platform real-time multi-face expression recognition toolkit," in *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, pp. 3723–3726, ACM, 2016. doi:10.1145/2851581.2890247. → page 39

[75] P. Ekman and W. V. Friesen, *Manual for the facial action coding system*. Consulting Psychologists Press, 1978. → page 39

[76] S. G. Hart and L. E. Staveland, "Development of nasa-tlx (task load index): Results of empirical and theoretical research," in *Advances in psychology*, vol. 52, pp. 139–183, Elsevier, 1988. doi:10.1016/S0166-4115(08)62386-9. → page 40

[77] S. G. Hart, "Nasa-task load index (nasa-tlx); 20 years later," in *Proceedings of the human factors and ergonomics society annual meeting*, vol. 50, pp. 904–908, HFES, 2006. doi:10.1177/154193120605000909. → page 40

[78] S. P. Borgatti and D. S. Halgin, "Consensus analysis," in *A companion to cognitive anthropology*, John Wiley & Sons, 2015. → pages 43, 44

[79] D. R. Heise, *Surveying cultures: Discovering shared conceptions and sentiments*. John Wiley & Sons, 2010. → pages 44, 89

[80] T. Carlson and Y. Demiris, "Collaborative control in human wheelchair interaction reduces the need for dexterity in precise manoeuvres," in *Workshop on Robotic Helpers: User Interaction, Interfaces and Companions in Assistive and Therapy Robotics*, pp. 59–66, University of Hertfordshire, 2008. → page 58

[81] T. Carlson and Y. Demiris, "Collaborative control for a robotic wheelchair: evaluation of performance, attention, and workload," *IEEE Transactions on Systems, Man, and Cybernetics, Part B*, vol. 42, no. 3, pp. 876–888, 2012. doi:10.1109/TSMCB.2011.2181833. → page 58

Appendix A

Supporting Materials

A.1 Study forms

A.1.1 BREB approval certificate

The University of British Columbia
Office of Research Services
Behavioural Research Ethics Board
Suite 102, 6190 Agronomy Road, Vancouver, B.C. V6T 1Z3

CERTIFICATE OF APPROVAL - MINIMAL RISK

PRINCIPAL INVESTIGATOR:	INSTITUTION / DEPARTMENT:	UBC BREB NUMBER:
Ian Mitchell	UBC/Science/Computer Science	H18-00256

INSTITUTION(S) WHERE RESEARCH WILL BE CARRIED OUT:

Institution	Site
UBC	Vancouver (excludes UBC Hospital)

Other locations where the research will be conducted:
N/A

CO-INVESTIGATOR(S):
Ariadna Estrada Gaspar

SPONSORING AGENCIES:
Aging Gracefully across Environments to Ensure Well-being, Engagement and Long Life (AGEWELL) - Networks of Centres of Excellence (NCE) - "CoPILOT - collaborative power mobility for an aging population"
Natural Sciences and Engineering Research Council of Canada (NSERC) - "Numerical algorithms for verification, design, analysis and operation of shared control cyber-physical systems"

PROJECT TITLE:
Affective user response to shared control interventions on a smart wheelchair: An exploratory study

CERTIFICATE EXPIRY DATE: October 11, 2019

DOCUMENTS INCLUDED IN THIS APPROVAL:

DATE APPROVED:
October 11, 2018

Document Name	Version	Date
Protocol:		
Study Protocol	7	October 4, 2018
Consent Forms:		
Consent Form	6	October 4, 2018
Advertisements:		
Call for Participation	5	October 4, 2018
Questionnaire, Questionnaire Cover Letter, Tests:		
EPA introduction	2	September 20, 2018
Post-trial questionnaire	1	September 19, 2018
Baseline questionnaire	1	September 18, 2018
Demographics questionnaire	1	September 19, 2018
Letter of Initial Contact:		
Recruitment email	5	October 4, 2018
Other Documents:		
E-mail list usage authorization	1	October 10, 2018

Other:
baseline-questionnaire webhost: https://ubc.ca1.qualtrics.com/jfe/form/SV_9WWSRjK3NHoSfCl post-trial and demographics questionnaire webhost: https://ubc.ca1.qualtrics.com/jfe/form/SV_6nDXp3O5cGG5IpL

The application for ethical review and the document(s) listed above have been reviewed and the procedures were found to be acceptable on ethical grounds for research involving human subjects.

This study has been approved either by the full Behavioural REB or by an authorized delegated reviewer

A.1.2 Consent form

a place of mind
THE UNIVERSITY OF BRITISH COLUMBIA

Faculty of Science
Department of Computer Science
2366 Main Mall
Vancouver, B.C., V6T 1Z4

Consent Form

Affective user response to shared control interventions on a smart wheelchair: An exploratory study

Principal Investigator: Dr. Ian M. Mitchell, Professor, Department of Computer Science, University of British Columbia, mitchell@cs.ubc.ca, 778-223-7538

Co-Investigator: Ariadna Estrada, M.Sc. Student, Department of Computer Science, University of British Columbia, aestra42@cs.ubc.ca, 604-704-1845

This study is being funded by the Aging Gracefully across Environments to Ensure Well-being, Engagement and Long Life (AGEWELL) Network Centre of Excellence (NCE), and the Natural Sciences and Engineering Research Council of Canada (NSERC).

Introduction: Thank you for participating in this study. This work is affiliated with the Collaborative Robotics Laboratory at UBC. Please note that we are seeking people aged 19 and over who are able to sit and drive a powered wheelchair using a joystick for up to 90 minutes. You are being invited to take part in this research study to help us understand how intelligent powered wheelchairs and their autonomous behaviors are perceived at an emotional level. The results may help us to design smart wheelchairs which support older adults with cognitive and mobility impairments to more safely and effectively use powered wheelchairs.

Purpose: The overall purpose of this study is to explore whether different types of assistance behaviors of intelligent wheelchairs are interpreted consistently across users. For each assistance behaviour, we are seeking to measure users' affective (emotional) response, level of physical and mental effort, and perception of performance, usefulness and safety while driving the wheelchair to a goal location.

What you will be asked to do: After you have read this document, the experimenter will respond to any questions or concerns that you may have. Once you have signed this consent form, you will be asked to:

- **Familiarize yourself with driving a powered wheelchair.** You will be given a 10 minute training session to explain the features of the wheelchair and so you can get used to driving the wheelchair in the test environment without navigation assistance. We will not collect any data during the training session.
- **Drive an intelligent powered wheelchair with different assistance modes.** After you are trained, you will complete up to 12 driving tasks where you will drive the wheelchair to a goal location. In all cases we will provide only a general location of the goal. The smart wheelchair <u>may or may not</u> modify your input commands to assist you in the navigation task. The task will terminate once you

are near the goal, when the driving time reaches 3 minutes, if you get too close to an obstacle, or if you give up, whichever happens first. During your driving sessions we will collect data about the location of nearby obstacles, your joystick commands, the wheelchair's motion and any modification of your commands, and a video of your face (for facial emotion recognition analysis).

- **Complete a questionnaire.** After each driving task, you will be asked to provide an assessment of the wheelchair behavior on three affective (emotional) dimensions: **E**valuation (good/bad), **P**otency (powerful/powerless), and **A**ctivity (Active/Passive). Similarly, you will be asked to provide an assessment of your own emotional state based on your ability to perform the task. Finally, you will be asked to rate the task workload, as well as the safety and difficulty of the interaction. Once you have completed every task, you will be asked to complete a brief demographic survey.

Time commitment: This study should take 60 to 90 minutes and be completed in 1 session.

Potential risks: There is a possibility to bump into obstacles when you are operating the wheelchair. Bumping into obstacles is not likely and the potential for injury to yourself or others has been minimized by setting the maximum speed of the wheelchair to 0.6 m/s (a slow walking pace) to give you enough time to react to potential collisions and by having an automated collision avoidance algorithm and a researcher with a remote control ready to stop the wheelchair if a collision appears imminent. In addition, you will be asked to keep your arms inside the wheelchair and your feet on the pedals, so that any obstacles in the environment make contact with the wheelchair rather than your body. You may experience discomfort when sitting on the wheelchair used in this study. You may take breaks between the driving sessions to mitigate fatigue and discomfort.

Potential benefits: There are no direct benefits to you from your involvement in this study; however, your responses may contribute to the design of new technologies that may benefit others in the future.

How the data collected will be used: Data collected will be used for analysis and may also be used by the student investigator to form the basis of thesis research which might be submitted as a research publication and/or presentation. The data could be used in the future to develop emotionally-aware intelligent wheelchairs.

Video recordings: During each driving task, we will record a video of your face. After the experiment session, the video will be processed to estimate the type and level of emotion you were feeling based on your facial expression at each time during the task. This estimated emotion data will be de-identified and used with the other data from the experiment. The original video excerpts will not be used in any reports, presentations or publications.

Confidentiality: Your confidentiality will be respected. Any information that could identify you as a participant in this study will be kept confidential. All information that you provide will be stored in Canada. Your identity will not be revealed in reporting the study results.

Data Retention: Identifiable data and video recordings will be stored securely in a locked metal cabinet or on an encrypted and password protected computer storage device. All data from individual participants will be coded so that their anonymity will be protected in any reports, research papers, thesis documents, and presentations that result from this work. All data will be destroyed or deleted after 5 years.

Compensation: You will receive monetary compensation of $15 for this session.

Contact for information about the study: If you have any questions, concerns, or desire for further information about the study before or during participation, you may contact Ariadna Estrada at 604-704-1845 or aestra42@cs.ubc.ca.

Contact for information about the rights of research subjects: If you have any concerns or complaints about your rights as a research participant and/or your experiences while participating in this study, contact the Research Participant Complaint Line in the UBC Office of Research Ethics at 604-822-8598 or if long distance e-mail RSIL@ors.ubc.ca or call toll free 1-877-822-8598.

Indicate your agreement to collect a video of your facial expression by providing your **initials**:

- I consent to being video recorded for this study. _______

I, _______________________________, have read the explanation about this study. I have been given the opportunity to discuss it, and my questions have been answered to my satisfaction. I hereby consent to take part in this study. However, I realize that my participation is entirely voluntary and that I am free to withdraw at any time.

Participant's Signature____________________________________Date_____________________

A.1.3 Call for participation form

a place of mind
THE UNIVERSITY OF BRITISH COLUMBIA

Faculty of Science
Department of Computer Science
2366 Main Mall
Vancouver, B.C., V6T 1Z4

Call for Participation:

Affective user response to shared control interventions on a smart wheelchair: An exploratory study

Principal Investigator: Dr. Ian M. Mitchell, Professor, Department of Computer Science, University of British Columbia, mitchell@cs.ubc.ca, 778-223-7538

Co-Investigator: Ariadna Estrada, M.Sc. Student, Department of Computer Science, University of British Columbia, aestra42@cs.ubc.ca, 604-704-1845

You are invited to participate in a research study involving intelligent wheelchairs and their emotional effects conducted by the Collaborative Robotics Lab in the Department of Computer Science at the University of British Columbia.

Who can participate?

You must be 19 years of age or older and be able to:

- sit and drive a powered wheelchair for up to 90 minutes;
- understand and respond to tablet-based questionnaires written in English;
- operate a joystick.

What is involved?

You will be asked to:

- Drive an intelligent powered wheelchair to different locations in a laboratory with different navigation assistance modes helping you to accomplish the task. We will be recording your facial expression during the driving sessions.
- Rate the wheelchair behaviors with respect to three basic dimensions of emotional experience: evaluation, potency, and activity.
- Report the emotions that you felt while driving the powered wheelchair.
- Provide feedback on the workload, helpfulness, safety and difficulty of using each intervention mode.
- Complete a demographic survey.

What is the time commitment? This study should take 60 to 90 minutes and be completed in 1 session.

Is there compensation? You will be compensated with $15.

Interested in Participating?

Please contact Ariadna Estrada at aestra42@cs.ubc.ca

A.2 Participant response forms

A.2.1 Baseline questionnaire

Introduction

Introduction

Thank you for participating in this study. Our overall goal is to understand the emotional effects of driving intelligent wheelchairs under different navigation assistance modes.

The purpose of the following questions is not to test you in any kind of way; we just need to get an idea of your personality.

Preliminary Self-Evaluation

Please identify yourself using the following scales:

Evaluation

Bad / Unpleasant **Good / Pleasant**

Infinitely Extremely Quite Slightly Neutral Slightly Quite Extremely Infinitely

Potency

Powerless / Weak **Powerful / Strong**

Infinitely Extremely Quite Slightly Neutral Slightly Quite Extremely Infinitely

Activity

Passive / Calm **Active / Excited**

Infinitely Extremely Quite Slightly Neutral Slightly Quite Extremely Infinitely

Wheelchair affective expectations

How would you describe a "smart-wheelchair"?

Evaluation

Bad / Unpleasant **Good / Pleasant**

Infinitely Extremely Quite Slightly Neutral Slightly Quite Extremely Infinitely

Potency

Powerless / Weak **Powerful / Strong**

Infinitely	Extremely	Quite	Slightly	Neutral	Slightly	Quite	Extremely	Infinitely

Activity

Passive / Calm **Active / Excited**

Infinitely	Extremely	Quite	Slightly	Neutral	Slightly	Quite	Extremely	Infinitely

Current mood

How are you feeling today? Select all that apply.

Alert	Angry	Annoyed	Anxious	Cautious
Confident	Cooperative	Disappointed	Enraged	Excited
Frustrated	Irritated	Nervous	Playful	Proud
Relaxed	Relieved	Satisfied	Submissive	Upset

A.2.2 Post-interaction questionnaire

Wheelchair behavior evaluation block 1

How would you describe the wheelchair's behaviour based on your most recent interaction?

Evaluation

Bad / Unpleasant **Good / Pleasant**

Infinitely Extremely Quite Slightly Neutral Slightly Quite Extremely Infinitely

Potency

Powerless / Weak **Powerful / Strong**

Infinitely Extremely Quite Slightly Neutral Slightly Quite Extremely Infinitely

Activity

Passive / Calm **Active / Excited**

Infinitely Extremely Quite Slightly Neutral Slightly Quite Extremely Infinitely

Post-interaction self-evaluation 1

How would you describe your mood based on your experience in completing the driving task? Select all that apply.

Alert	Angry	Annoyed	Anxious	Cautious
Confident	Cooperative	Disappointed	Enraged	Excited
Frustrated	Irritated	Nervous	Playful	Proud
Relaxed	Relieved	Satisfied	Submissive	Upset

Workload 1

How much mental activity was required (e.g. thinking, deciding, calculating, remembering, looking, searching, etc.)? Was the task easy or demanding?

Easy | Mental Demand | Demanding

How much time pressure did you feel due to the pace of the task? Was the pace slow and leisurely or rapid and frantic?

Slow | Temporal Demand | Rapid

How successful were you in accomplishing what you were asked to do? How satisfied were you with your performance in accomplishing this goal?

Good | Performance | Poor

How irritated, stressed and annoyed versus content, relaxed and complacent did you feel during the task?

Content / Relaxed / Complacent | Frustration Level | Irritated / Stressed / Annoyed

Safety and difficulty 1

Did you feel safe or unsafe in the wheelchair? How safe/unsafe?

Unsafe **Safe**

Extremely Quite Slightly Neutral Slightly Quite Extremely

Did you find the wheelchair easy or difficult to use? How easy/difficult?

Difficult **Easy**

Extremely Quite Slightly Neutral Slightly Quite Extremely

Did you feel like the wheelchair helped (or made it difficult for) you to avoid obstacles? How?

Did you feel like the wheelchair helped (or made it difficult for) you to complete the task? How?

Workload Comparison Cards

Sources of workload

You will be presented with a series of pairs of rating scale titles (for example, Performance vs. Mental Demand) and asked to choose with of the items was more important to **your** experience of workload in the tasks that you just performed.

Please consider your choices carefully and make them consistent with how you used the rating scales during the particular tasks that you were asked to evaluate. Don't think that there is any *correct* pattern: we are only interested in your opinions.

Each of the six boxes below contains two different sources of workload. In each of the six boxes click on the source which you feel was the more important contributor to your workload.

Mental Demand or Temporal Demand	Performance or Frustration	Temporal Demand or Performance
Frustration or Temporal Demand	Mental Demand or Frustration	Performance or Mental Demand

A.2.3 Demographic questionnaire

Demographics

Demographics

How old are you?

What is your gender?

- Male
- Female
- Non-binary/third gender
- Prefer to self-describe
- Prefer not to say

What is your occupation?

What is your mother tongue?

How often do you interact with robotic systems?

- I never interact with robots
- I have interacted with robots once or twice
- I have interacted with robots on a number of occasions
- I interact with robots regularly
- I do research on robots

What is your experience driving powered wheelchairs?

- This is the first time I have ever used a powered wheelchair
- I have used a powered wheelchair in the past

How long have you spent driving a powered wheelchair?

A.3 Data verification

A.3.1 Measurement reliability

We followed the approach presented in Chapter 8 of Heise's book *Surveying Cultures* to assess the reliabilities of the Evaluation, Potency, and Activity measures across stimuli. Table A.1 summarizes the test-retest data for the ratings attributed to the wheelchair behaviour. The column *Intervention* shows which intervention mode was rated, the column *Dim* shows whether the statistics are for evaluation, potency, or activity ratings. The columns under *Test 1* give the means and variances of ratings and the number of subjects who rated the intervention during the first block of trials. For example, 17 participants rated the Dynamic shared control intervention during the first block and their mean Evaluation was 1.48 with a variance of 2.68. The columns under *Test 2* gives the corresponding statistics of the repeated measure. Small discrepancies occur since some respondents did not rate an intervention due to technical failures with the platform. The column titled N under *T1, T2* gives the number of participants who rated each behavior at both test 1 and test 2. The column titled *r* shows Pearson's correlation coefficient between individuals' test 1 and test 2 ratings, and the column titled *p-value* shows the level of significance for that coefficient.

Overall, no significant differences were found on the variances between test 1 and test 2 of Evaluation, Potency, and Activity ratings considering the 6 intervention modes. Furthermore, 10 out of the 18 correlations between test 1 and test 2 ratings were significantly greater than 0. Correlations of 0.48 or more are significant with $p \leq 0.05$ in a two-tailed test. The mean overall correlation was 0.49, the mean correlation for Evaluation was 0.35, the mean correlation for Potency was 0.46, and the mean for Activity was 0.67.

The individual variance in Activity ratings was significantly higher ($p \leq 0.05$) than those of Evaluation and Potency. The large variance in Activity might derive from different interpretations of the dimension. We did not instruct the respondents to associate the Activity dimension with a particular characteristic of the wheelchair. We found that some respondents associated it with the amount of resistance presented by the wheelchair while others associated it with an opposite

Table A.1: Statistics for E, P, and A ratings of wheelchair behaviours

		Test 1			Test 2			T1, T2		
Intervention	Dim	N	Mean	Var	N	Mean	Var	N	r	p-value
Dynamic shared control	E	17	1.48	2.68	18	1.98	1.87	16	0.50	0.05
	P	17	1.35	3.62	18	1.61	3.09	16	0.73	0.00
	A	17	0.56	6.68	18	0.93	5.07	16	0.75	0.00
High Level	E	20	0.12	3.84	18	0.14	5.18	18	0.40	0.10
	P	20	1.10	3.32	18	1.47	4.96	18	0.45	0.06
	A	20	0.69	4.86	18	1.29	5.06	18	0.83	0.00
Disagreement	E	20	1.86	2.57	17	1.19	2.91	17	0.28	0.28
	P	20	1.59	3.10	17	1.39	3.73	17	0.20	0.44
	A	20	0.98	4.60	17	0.78	5.20	17	0.48	0.05
Efficiency Blend	E	20	0.25	4.35	18	1.19	2.73	18	0.39	0.11
	P	20	1.65	2.71	18	2.08	1.22	18	0.12	0.63
	A	20	1.06	4.17	18	0.98	5.32	18	0.54	0.02
Collision Avoidance	E	17	1.86	2.93	17	2.05	0.40	15	0.08	0.78
	P	17	1.91	2.51	17	1.26	2.71	15	0.57	0.03
	A	17	0.79	4.64	17	0.89	4.00	15	0.81	0.00
Steering Correction	E	19	0.95	2.55	18	0.86	4.13	17	0.43	0.09
	P	19	1.23	2.00	18	1.32	4.57	17	0.67	0.00
	A	19	0.52	3.69	18	1.22	5.95	17	0.63	0.01

concept such as the obedience of the wheelchair.

A.3.2 Participant repeatability

To assess the repeatability of each subject, we computed two measures: the number of times a respondent judged an intervention within the same octant and the Euclidean distance between the first and second trial ratings. For both measures, we only considered the ratings where the participant tested each mode twice. Out of the 240 trials (20 subjects, 6 modes, 2 blocks), 19 trials had no rating due to some technical failure leaving 202 trials (i.e., $240 - 19 * 2$) or 101 complete pairs of ratings.

Out of the 101 pairs of ratings, 52 (51.49%) were rated within the same octant for both trials. The average euclidean distance between the two ratings was 2.83. Considering the distance between the center of an octant to the origin (3.724) as a

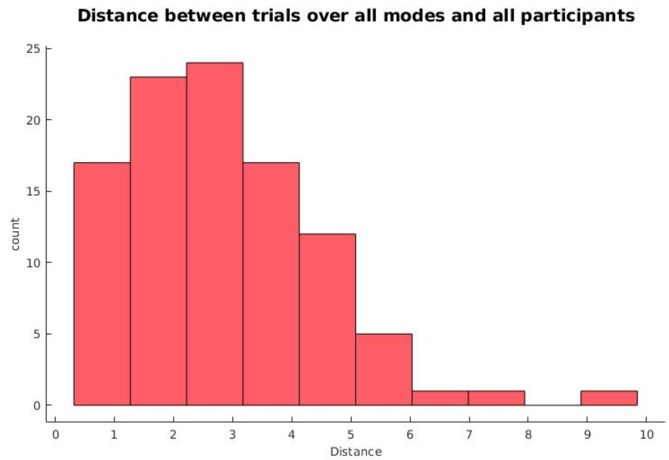


Figure A.1: Distribution of rating distance between trials across participants

reasonable margin of error roughly 75% of the trials were rated within a reasonable distance between trial 1 and trial 2. Figure A.1 shows a summary of the distances across modes and across participants.

We followed an approach similar to the one described by Heise in [79] to examine whether the Potency and Activity ratings were affected by more than one factor. We computed the Pearsons correlation between Evaluation-Potency, Evaluation-Activity, and Potency-Activity ratings across all subjects. The correlation between Evaluation-Potency ($r = 0.41$, $p < 0.05$) and Potency-Activity ($r = 0.57$, $p < 0.05$) show a moderate linear relationship. Figure A.2 shows the scatter plots of the relationships between E, P, and A ratings.

Summarizing, the somewhat large second eigenvalues in the Potency and Activity component analyses reflect the fact that Potency or Activity ratings among some respondents were moderately predictable from the Evaluation and Potency ratings respectively.

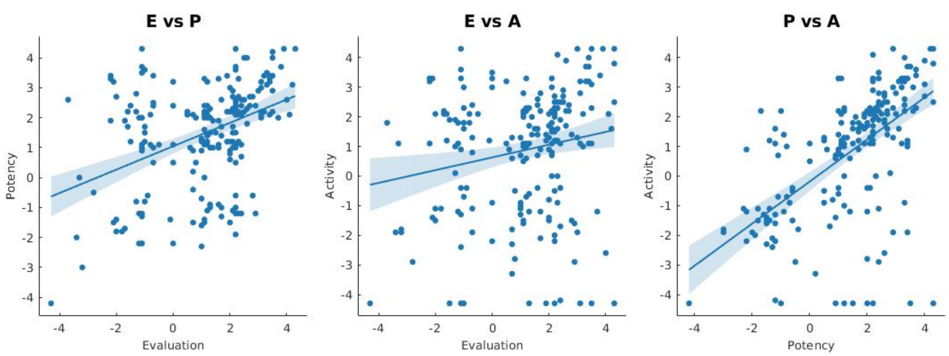


Figure A.2: Relationships between E, P, and A with a linear model fit to the data. The shaded area represents the 95% confidence bounds.