robotics MDPI

Article

Got It? Comparative Ergonomic Evaluation of Robotic Object Handover for Visually Impaired and Sighted Users

Dorothea Langer [1,*], Franziska Legler [1], Pia Diekmann [2], André Dettmann [1], Sebastian Glende [2] and Angelika C. Bullinger [1]

1 Chair of Ergonomics and Innovation, Department of Mechanical Engineering, Chemnitz University of Technology, 09111 Chemnitz, Germany; franziska.legler@mb.tu-chemnitz.de (F.L.); andre.dettmann@mb.tu-chemnitz.de (A.D.); angelika.bullinger-hoffmann@mb.tu-chemnitz.de (A.C.B.)
2 YOUSE GmbH, 13187 Berlin, Germany; pia.diekmann@youse.de (P.D.); sebastian.glende@youse.de (S.G.)
* Correspondence: dorothea.langer@mb.tu-chemnitz.de; Tel.: +49-371-531-32217

Abstract: The rapidly growing research on the accessibility of digital technologies has focused on blind or visually impaired (BVI) users. However, the field of human–robot interaction has largely neglected the needs of BVI users despite the increasing integration of assistive robots into daily life and their potential benefits for our aging societies. One basic robotic capability is object handover. Robots assisting BVI users should be able to coordinate handovers without eye contact. This study gathered insights on the usability of human–robot handovers, including 20 BVI and 20 sighted participants. In a standardized experiment with a mixed design, a handover robot prototype equipped with a voice user interface and haptic feedback was evaluated. The robot handed over everyday objects (i) by placing them on a table and (ii) by allowing for midair grasping. The usability target was met, and all user groups reported a positive user experience. In total, 97.3% of all handovers were successful. The qualitative feedback showed an appreciation for the clear communication of the robot's actions and the handover reliability. However, the duration of the handover was seen as a critical issue. According to all subjective criteria, the BVI participants showed higher variances compared to the sighted participants. Design recommendations for improving robotic handovers equally supporting both user groups are given.

Keywords: human–robot interaction; robot handover; accessibility; usability; user experience; blind or visually impaired users

check for updates

Citation: Langer, D.; Legler, F.; Diekmann, P.; Dettmann, A.; Glende, S.; Bullinger, A.C. Got It? Comparative Ergonomic Evaluation of Robotic Object Handover for Visually Impaired and Sighted Users. *Robotics* **2024**, *13*, 43. https://doi.org/10.3390/robotics13030043

Academic Editors: Weitian Wang, Michael Bixter and Quanjun Song

Received: 31 December 2023
Revised: 19 February 2024
Accepted: 25 February 2024
Published: 5 March 2024

1. Introduction

Research on the accessibility of digital technologies has rapidly grown in recent years. In 2022, about 330 million people worldwide were estimated to be blind or have moderate to severe vision impairment, and the projections show that this number will rapidly grow within the next 30 years due to aging societies [1]. Assistive technologies are expected to greatly impact the lives of BVI people [2]. Almost half of the published studies on accessibility have BVI users as the main group of interest; nevertheless, the research focus has remained on fields like web browsing, education, mobility, and technical categories like interaction with mobile screens or access to traditional graphical user interfaces [3].

In contrast, the BVI user group is widely neglected in the research on human–robot interaction (HRI). Soon, service robots will be part of our daily lives, being employed in shared/public spaces and carrying out tasks like assistance, cleaning and inspection, transporting objects, and guiding in unfamiliar environments. Hence, interactions with very diverse users are to be expected [4], also including humans with impairments. One instance is the shortage of staff and skills in the healthcare industry. This issue could be addressed by introducing assistive robots, which would benefit individuals with various impairments. To ensure that, besides non-disabled users, BVI users can also benefit from assistive robots, accessibility research has to take into account the special needs of this user group. This paper focuses on one specific interaction: robotic object handover.

1.1. Robotic Object Handover

Object handover is a basic capability of assistive robots that seems simple and commonplace to humans but remains a complex task with major challenges for robots [5]. This complexity also includes the requirements of diverse user groups. For an overview of the current research and the core concepts of robotic object handover, refer to [6]. They also describe the basic characteristics of human-to-human handover and define handover as "a joint action between a human giver and a human receiver" (p. 1856). There are two possible handover directions: robot-to-human and human-to-robot. As the robot-to-human direction is especially relevant to robot applications assisting diverse users, e.g., a robot delivering beverages to patients in a hospital or a robot vending food at a train station, we will go into more detail below.

The handover process between humans can be divided into two distinct exchange phases, which can also be used for robotic object handover: (1) the Prehandover phase and (2) the Physical Handover phase [6]. They can be further divided into subphases. (1) The Prehandover phase begins with the Inaction subphase. This is followed by the Initiation subphase, which starts when one agent initiates the handover by either making a handover request or assigning a task that naturally involves an object handover. The Prehandover phase progresses to the Preparation subphase, in which the giver grasps the object and transports it to the receiver. Subsequently, (2) the Physical Handover phase starts with the Physical Transfer subphase, beginning with the receiver's initial contact with the object and ending when the giver withdraws their hand, and the object is securely in the receiver's grip. The Physical Handover phase ends with the Performance subphase, where the receiver uses the object. The handover phases are shown in Table 1 (adapted from [6]), along with the relevant user requirements for robotic object handover, which are described subsequently.

When planning a robotic handover, there are several important factors to consider in order to ensure a successful grasp. These are, for example, the shape of the object and its function, the safety of the user [7], and the object orientation [8], while dynamics [5] and the subsequent task the receiver performs after handover [6] represent crucial situational factors. Previous studies have indicated that in robotic handovers, people generally prefer the object to be oriented in a manner consistent with conventional human practices [8]. An example of this preference for easy handling is termed the "etiquette factor", where a convenient grasp, such as presenting a cup horizontally with the handle first, is emphasized in robot-to-human handovers [9].

Some object surfaces pose a risk of injury during handover, possibly due to pointed edges or sharp contours [10]. These objects require even greater adherence to social conventions to prevent injury to the receiver, such as covering hazardous surfaces and orienting hazardous parts of objects away from the receiver [6,9,11].

The physical handover of an object can take two forms: direct or indirect [6]. A direct handover involves passing the object directly from the giver's hand into the receiver's hand. Such midair passing of an object is the most immediate and practical method in certain scenarios, e.g., when a mechanic beneath a car requests a tool. Direct handovers reduce the effort for the receiver in terms of the motions needed to obtain the object [12]. They also reduce the risk of touching hazardous object surfaces when the object is handed over with a predefined orientation and covered hazardous surfaces [11]. Conversely, during an indirect handover, the giver places the object on a surface, such as a table. Indirect handovers offer more flexibility to the receiver regarding timing and the grasp type for acquiring the object [6]. With indirect object handover, in addition, particularly in HRI, the robotic giver is provided greater flexibility in selecting its grasping strategy because it does not have to consider the areas on the object that must be left free for the human to grasp, and therefore, it has more surfaces available for a stable grip [11].

The success of a joint action like an object handover relies heavily on coordination [6]. Planned coordination originates from shared visions of desired outcomes, individual tasks, and goals. This includes, for example, a shared representation of both partners, which object will be handed over, where it will happen, and what task the receiver will subsequently perform with the object. In contrast, emergent coordination is not tied to predefined joint plans, but arises from the interplay between perception and action. Active, coordinating gaze behavior is an example of an emergent coordination mechanism and is regarded as a prerequisite for effective HRI [13]. For a seamless robotic object handover, it is therefore recommended that the robot check for human eye contact as it approaches [14]. Accordingly, technical sensor concepts detect human eye contact as a prerequisite for starting the interaction to ensure safe HRI, e.g., [15]. Integrating gaze into robotic handovers yields positive effects, including faster object reaching and a more naturally perceived interaction by human receivers [6]. Moreover, gaze can influence cooperation by reducing human response times [16].

In contrast, individuals with visual impairments face inherent challenges as they cannot establish eye contact with the robot, and thus have been excluded from potential user groups. This exclusion is further exacerbated in scenarios like constrained working postures during a production task [17], highlighting a critical need for more inclusive research involving BVI and sighted users. We will subsequently describe current research regarding accessibility in robotic object handover as well as special needs of BVI users in more detail.

1.2. Accessibility of Human–Robot Handover

Requirements of non-disabled users might differ from or even be opposed to those of disabled users [18]. Moreover, disabled users are a heterogeneous group differing individually, e.g., in abilities, use cases, and preferences, whose combination causes individually unique challenges [18]. Apart from that, disabled users prefer interfaces designed as close as possible to the norm and feel uncomfortable being dependent on special solutions [19]. Hence, despite differing requirements, design should always consider the inclusion aspect [19].

Research on accessible HRI is scarce despite extensive studies on human–robot handovers. Literature often focuses on robots designed for specific user groups, such as educational robots for children with learning disabilities [20] or healthcare robots [21,22]. For instance, ref. [12] explored robotic handovers with patients diagnosed with Amyotrophic Lateral Sclerosis (a neurodegenerative disease that leads to motor impairment), reporting high success rates and satisfaction. So far, no comparable study has included non-disabled or visually impaired users for comprehensive comparisons. Such studies, tailored to specific user groups, may not address the broader context of diverse user interactions. Furthermore, accessibility research often lacks comparisons with non-disabled users, e.g., [23,24], or does not even include the disabled target user group, e.g., [25].

BVI users wish to employ robots to locate and transfer objects to aid them in their daily activities [24]. Against this, specific handover requirements of BVI users can be derived from a limited number of studies focusing on robot-to-BVI user interactions. Additionally, it is necessary to supplement this information with insights from the accessibility literature addressing other technologies or human-to-human handovers.

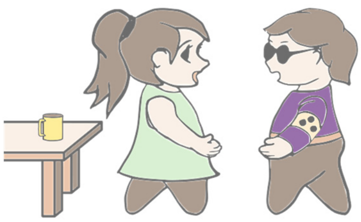
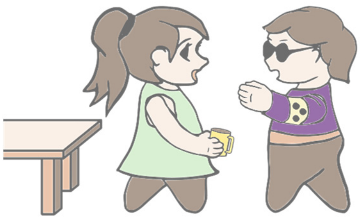
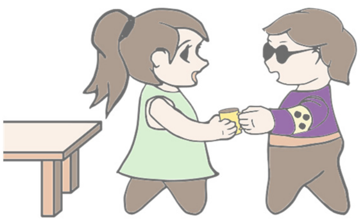
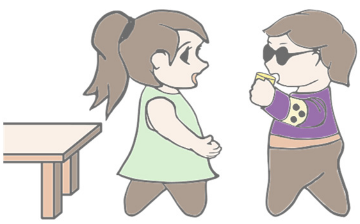
Studies investigating robot-to-human handovers with sighted users primarily concentrated on a restricted set of handover subphases rather than the complete process. In the context of research involving robot-to-human handovers with BVI users, early stages within the Prehandover phase were identified as particularly crucial for BVI users and are supported by a Voice User Interface (VUI) [26]. These stages, namely 'welcoming and start of interaction' and 'obtaining order information', can be assigned to the Inaction and Initiation subphases, respectively. Due to the verbal nature of these stages, vital preconditions that facilitate the handover process later on can be captured in advance. For instance, the verbal welcoming and introduction of the robot by name can assist BVI users in locating the robot and establishing a social connection with it. The latter is especially important, as handovers are joint actions [6], and BVI users can only perceive anthropomorphic cues through verbal support provided by VUI. Hence, a robotic system should actively provide them with ongoing status updates or enable retrieval of this information through voice commands at any given time [27]. Furthermore, access to active and precise auditory status information becomes especially critical at crucial points in the handover process where BVI users wish to actively control the process themselves, e.g., to confirm when they have safely grasped the object before the robot opens its gripper [26]. During the performance subphase, the robot should return to a known base position to avoid becoming a dangerous obstacle after handover [26].

These preferences align with specific strategies employed by BVI individuals in everyday handovers between humans [28]: (1) the expectation of verbal information about the type, orientation, and position of the object, along with potential sources of harm; (2) a preference for only one partner to move at a time during the handover process; and (3) a desire to remain passive during delivery but actively engage in grasping the object during the physical handover phase. Hence, the robot's movement should alternate with BVI users, allowing them to remain passive during delivery but to actively control physical handover. Additionally, if possible, BVI users tend to avoid midair handovers and instead favor indirect handovers with contact on a stable surface, such as a table, providing haptic support for better orientation [28]. Consequently, during midair handovers, providing haptic support for better orientation should be considered to facilitate the process. For example, an airflow could help to determine the object's position.

Some additional requirements can be derived from studies that included sighted participants simulating blindness, e.g., by wearing an eye mask. For example, one study found that blinded participants performed handovers more carefully and much more slowly than sighted participants [29]. Hence, the robot should be capable of accommodating waiting times, considering that BVI users may execute certain processes slower than sighted users. Additionally, ref. [11] developed and evaluated a robotic grasping strategy for object handover with potentially hazardous and non-hazardous objects and found that object orientation and grasp planning along human conventions resulted in a high success rate and high satisfaction for blinded and sighted participants. Hence, the orientation of delivered objects should adhere to the social conventions observed in everyday human handovers, particularly when dealing with potentially hazardous items. However, those requirements should be taken with some caution because engaging nondisabled participants as a replacement for disabled users should be performed only in the early stages of development rather than in late evaluation cycles, as disabled users have different education, knowledge, and mental representations [30].

Despite the specific needs of BVI users, the design of the robot should aim for equal support for both BVI and sighted users, mitigating any potential for stigmatization [19]. To determine whether the handover robot fulfills those user needs and is equally proficient in assisting both BVI and sighted users, we require a theoretical foundation for selecting evaluation methods to assess robot-to-human handovers.

Table 1. Handover phases (adapted from [6]) and relevant user requirements.

Handover Phase			General User Requirements	BVI User Requirements
Throughout all phases			apply emergent coordination arising from the interplay between perception and action [6], e.g., active, coordinating gaze behavior [13]	design the interface as close as possible to the norm [19] coordinate handovers without eye contact [26] allow only one partner to move at a time [28]
Prehandover Phase	Inaction			use verbal welcoming and introduction of the robot by name to help locate the robot and establish a social connection with it [26]
	Initiation		initiate the handover through a handover request or a task that naturally involves an object handover [6] use planned coordination by forming a shared representation of, e.g., which object will be handed over, handover location, and what task the receiver will subsequently perform with the object [6]	support obtaining order information by a Voice User Interface [26]
	Preparation		when planning a grasp, pay attention to object shape, object function and user safety [7], object orientation [8], situation, such as dynamics [5], and the subsequent task the receiver performs after handover [6] let the robot check for human's eye contact as it approaches [14]	plan object orientation and grasping along human conventions [11] provide verbal information about the type, orientation, and position of the object, along with potential sources of harm when approaching [28] actively provide ongoing status updates or enable retrieval of this information through voice commands at any given time [26,27]
Physical Handover Phase	Physical Exchange		let the orientation of delivered objects adhere to the social conventions observed in everyday human handovers [8,9], particularly covering hazardous surfaces and orienting hazardous parts of the object away from the receiver when dealing with potentially hazardous items [11]	provide direct and indirect handover modalities [12,28] provide haptic support for better orientation during midair handovers [28] (e.g., airflow) allow handover processes to be slowed down for more careful execution [29] let users actively control the process themselves, e.g., to confirm when they have safely grasped the object before the robot opens its gripper [26]
	Performance			let the robot return to a known base position to avoid becoming a dangerous obstacle after handover [26]

1.3. Methods for Evaluating Robot-to-Human Handover

Due to the diverse taxonomies in HRI and a broad array of evaluation metrics, there is no universal method for the assessment of HRI [31]. ISO 9241-11:2018 [32] seems to be an appropriate foundation to assess whether a robot can equally support interaction with BVI and sighted users. It is the common standard for the evaluation of human–system interaction in general and has also been applied to the ergonomic evaluation of HRI [31]. Its core concept is usability, but it further explicitly addresses accessibility. In the ISO norm [32], usability is defined as "the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use" (p. 11). Accessibility is seen as a goal of system design to make the system available to a wide range of people in diverse contexts of use. According to the standard, usability and accessibility are an outcome of interaction rather than an inherent property of a product.

To assess usability according to ISO 9241-11:2018 [32], it is essential to establish a context of use and have users engage in tasks within this context, whether in a real or simulated environment. Table 2 presents the definition of usability subcomponents, namely effectiveness, efficiency, and satisfaction, along with potential metrics for their evaluation according to ISO 9241-11:2018 [32]. The standard does not prescribe a general rule for the selection or combination of metrics, but the selection should be derived from the task objectives, the specific context, and/or the perceived outcomes. Nonetheless, usually, at least one metric per usability component is used, and combining objective and subjective measures is seen as valuable to measure usability entirely. One example of subjective evaluation of usability in its entirety is the System Usability Scale (SUS) [33], for which interpretative normative data have been established [34,35].

Specifically for the evaluation of robotic handovers, ref. [6] recommends a minimal set of metrics and suggests incorporating a specific number of objects into each handover study. This standardization should facilitate fair comparisons across various studies. The review of [6] was just about to be published when this study was conducted. Still, it is obvious that both frameworks critically overlap in metrics. Table 2 compares the suggested metrics of [6] to ISO 9241-11:2018 [32].

ISO 9241-11:2018 [32] also considers other desired results of an interaction, the so-called human-centered quality components: user experience and possible user-related damage. User experience is more clearly defined in ISO 9241-210:2019 [36] as "A person's perceptions and responses that result from the use and/or anticipated use of a product, system, or service" (p. 11). Consequently, user experience encompasses all the impacts that interacting with an interactive system has on the user, covering the period before, during, and after use. Ref. [37] refers to this as hedonic quality aspects, describing it as the contribution of perceived fun and enjoyment to user satisfaction. Hence, user experience and satisfaction are closely related. It can be assessed by standardized surveys, like the User Experience Questionnaire [38]. In ISO 9241-11:2018 [32], possible user-related damage is characterized as negative results that could arise from inappropriate usability. These damages can relate to the user, the using organization, or the developing organization of an interactive system. This study addresses possible damages to the user, which is associated with interaction effectiveness.

Table 2. Comparison of possible metrics for the evaluation of robot-to-human handovers.

Subcomponent of Usability According to ISO 9241-11:2018 [32]	Metrics in ISO Norm	Metrics in [6]
Effectiveness "accuracy and completeness with which users achieve specified goals" (p. 18)	**objective** success rate number of errors	**objective task performance** success rate [1] errors occurred
	subjective perceived level of target achievement	**subjective** robot relative contribution
Efficiency "resources used in relation to the results achieved" (p. 19) e.g., time, effort, costs, materials	**objective** task completion time physical effort	**objective task performance** (receiver's) task completion time [1] interaction force total handover time [1] reaction time fluency (concurrent activity vs. idle time)
	subjective perceived time perceived effort	**subjective** workload (NASA-TLX) HR-fluency [1]
Satisfaction "extent to which the user's physical, cognitive and emotional responses that result from the use of a system [. . .] meet the user's needs and expectations" (p. 20)	**objective** repeated use of a system (long-term)	**objective task performance** psycho-physiological measures
	subjective satisfaction with task completion overall impression of the system	**subjective** improvement preference
		Additional concepts trust in the robot [1] working alliance [1] positive teammate traits safety

[1] metrics belong to the 'minimal set' of metrics suggested by [6].

1.4. Research Objectives and Paper Structure

The research objective of this paper is to evaluate whether a handover robot designed to meet specific user needs is equally proficient in assisting BVI and sighted users. To investigate this, the following assumptions are derived based on ISO 9241-11:2018 [32] and the previously outlined user requirements:

Overall usability:

Hypothesis 1. *As the handover robot is an early prototype, we adhere to [34] and aim for a System Usability Scale (SUS) score above 70, which is considered acceptable. There is no discernible difference in overall usability ratings between BVI and sighted users.*

Effectiveness:

Hypothesis 2. *The handover success rate is consistently high in both groups, with low failure rates independent of the user group.*

Hypothesis 3. *Objects handed over are equally easy to grasp for BVI and sighted users, regardless of whether the objects are potentially hazardous or nonhazardous.*

Hypothesis 4. *In everyday situations, BVI users prefer indirect handovers on a table. Therefore, BVI users perceive objects as easier to grasp in indirect handovers, while sighted users find objects easier to grasp in the more immediate direct handover in midair.*

Efficiency:

Hypothesis 5. *BVI users may exhibit slower actions during active physical handover stages, resulting in longer time measures compared to sighted users.*

Hypothesis 6. *Subjective appropriateness of handover time is independent of the user group.*

Hypothesis 7. *Since direct handover in midair requires more coordination steps, time measures are higher compared to indirect handover on a table.*

Hypothesis 8. *Higher duration of direct handover in midair might lead to lower ratings of the appropriateness of handover time for direct handover in midair compared to indirect handover on a table.*

Hypothesis 9. *There is no difference in handover time between potentially hazardous and non-hazardous objects.*

User Experience and Satisfaction:

Hypothesis 10. *There is no difference in user experience ratings between BVI and sighted users. Open feedback is positive and reflects instances where the interaction met the requirements of both user groups. Haptic feedback aids BVI users in locating the object.*

Hypothesis 11. *In line with Hypothesis 4, BVI users choose indirect handover on a table more often, whereas sighted users choose direct midair handover more often.*

2. Materials and Methods

2.1. Sample

Overall, 40 participants were involved in the evaluation study (26 female, 14 male), of which 20 belonged to the sighted and 20 belonged to the BVI user group. Mean age was M = 41.2 (SD = 17.5), ranging between 18 and 71 years. Both groups did not differ in age ($t = -0.34$, $p = 0.736$) or prior experience with robots ($X^2 = 2.06$, $p = 0.151$) and showed the same gender distribution.

BVI individuals are diverse; individual capabilities cannot be solely determined by visual test results. In this study, BVI participants are included based on legal definitions. This includes "blind" individuals with less than 2% remaining sight (14 participants), "strongly visually impaired" individuals with up to 8% (2 participants), and "visually impaired" individuals with up to 30% remaining sight in the less impaired eye (1 participant). The remaining 3 participants stated another kind of visual impairment, e.g., homonymous restriction of their visual field to 3%. Regarding the duration of impairment, 9 participants in the sample were congenitally blind, 9 participants were visually impaired for at least 14 years, and 2 participants did not specify. Regarding remaining abilities, 12 participants stated being able to visually differentiate colors, 8 participants said shapes, and 5 participants were able to identify people (multiple answers possible). BVI participants were acquired by distributing invitation letters through local Associations for the Blind.

2.2. Materials and Context of Use

A Universal Robots UR10 collaborative robotic arm [39] was used for object handover. It was installed on an Automated Guided Vehicle platform, which stayed stationary during the study. The robot's end effector was equipped with a stereo camera system, although real-time vision was not employed during this experiment. Moreover, a flexible three-finger gripper type BH8 280 of the brand BarrettHand^{TM} (Barrett Technology, Newton, MA 02460, USA) was attached to the robot's end effector. The three fingers of the gripper can be flexibly oriented and combined to achieve a wide variety of gripping poses. In addition, tactile sensors are built into the finger surfaces and the palm of the gripper, which enable sensory monitoring of the grip. Robotic grasping strategies for the different objects were implemented according to [11]. A decision tree was utilized to determine the robot's grasping strategy, evaluating the appropriateness of a standard grasp. If deemed appropriate for the given object category, this grasp was executed; otherwise, the system assessed the next prioritized grasp until a feasible one was identified. Grasp suitability was gauged based on estimations of user safety in line with human conventions, grasp stability, and applicability across various objects. Object category and orientation were preset by an experimenter before the robot grasping phase. The robot arm moved autonomously along a free path, with trajectory limitations imposed only by furniture in the testing environment and a safety space around the participants. However, the algorithm designed to estimate human posture and hand position, based on processing and segmenting face features compared with face and palm color matching, proved unreliable when users wore an air-filtering facepiece respirator, mandatory during the study due to pandemic restrictions. Therefore, a fixed handover position was established as the target for the robot's trajectory. A VUI was implemented as described in [26]. The handover phases were delineated using specific predefined sentences or questions for robot–user interaction. These prearranged speech outputs were recorded digitally using the synthetic voice 'Claudia' from the acapela group repertoire (https://www.acapela-group.com/de/voices/repertoire/, accessed on 19 July 2022). The Voice User Interface (VUI) was independently developed, utilizing Voice over IP (VoIP) technology through PhonerLite freeware (https://lite.phoner.de/index_de.htm, version 3.23, accessed on 19 July 2022). It encompassed a greeting and introduction by the robot, followed by a brief manual, voice dialogue for handover tasks, and additional short commands like "status" and "continue." Short commands were prioritized and, when used, interrupted ongoing dialogue, confirmed by a tone upon user input processing. Integrating the VUI into the robot necessitates functional alignment of hardware, software, and user interfaces, achieved through an event-based distributed state machine linking all robot subcomponents. To help users locate the object for physical handover, the robot was equipped with a simple haptic feedback system. It was realized through a seamless airflow mechanism attached to the gripper, generating an airflow of 550 L per hour and an additional buzzing sound. Due to pandemic restrictions, this feedback system was a preliminary prototype that could not be evaluated and improved in an iterative User-Centered Design process. It was implemented to obtain initial exploratory insights into robot handovers for BVI users, incorporating an additional sensory channel for haptic information. All of the robot's subcomponents were interlinked through an event-based distributed state machine. The entire process of the robot can be subdivided into the processes 'ordering' (supported by VUI), 'searching for the object' inside the room, 'grasping the object', 'delivery' towards the user, 'handing over', and 'parking' by relocating to a predefined home position. An overview of the experimental setup with the robot for handover tasks is given in Figure 1.

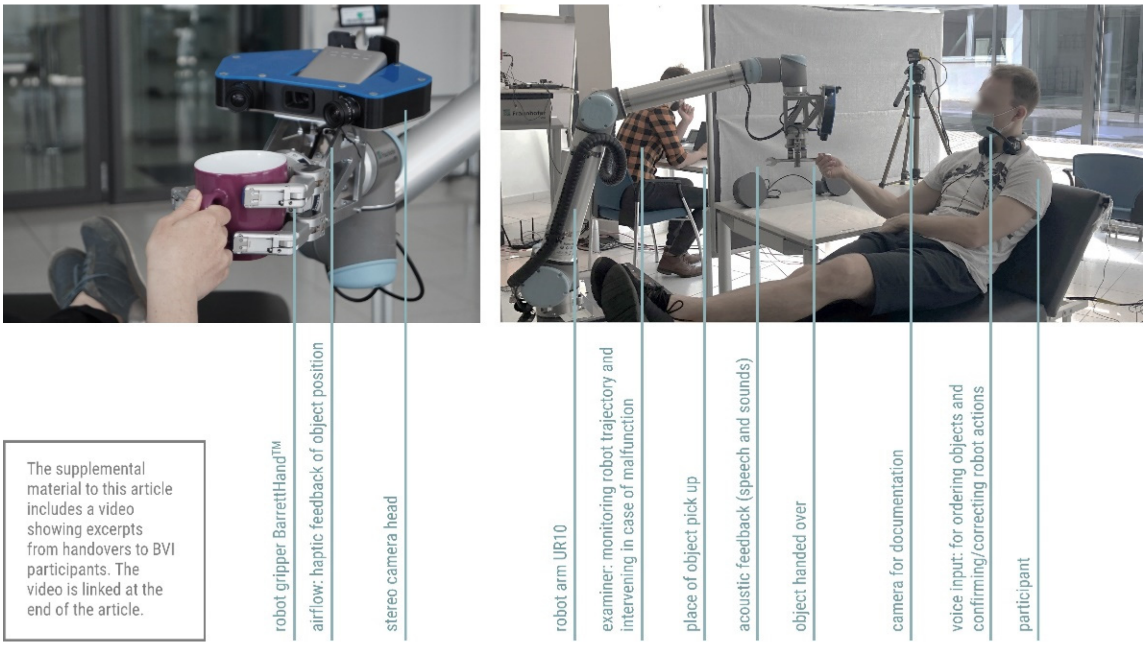


Figure 1. Experimental setup with robot for handover tasks.

Building on ISO 9241-11:2018 [32], a context of use was defined. It corresponds to the one described in [26]. A use case was defined for each user group, establishing the necessary criteria for interaction strategies and user interfaces. Also, the handover objects were determined based on these use cases. One use case involved the hospitalization of a blind person, where the robot delivered various objects, including potentially hazardous ones such as a knife, to the bedside. The benefit of such a robot for BVI users lies in providing support in an unfamiliar environment and offering additional autonomy and independence. In that use case, the robot's task is to hand over an object, but additionally, the robot has to relocate to a predefined home position to ensure that the robot is not an obstacle for the blind person. Hence, the robot task is only completed when the robot arrives at the home position. Another use case focused on sighted users, specifically a mechatronics technician working in an unfavorable position. In this scenario, the robot delivered tools, such as a spanner, to the technician. The benefit of such a robot for the sighted person in this context lies in locating awkwardly positioned objects and avoiding inefficient work steps, such as groping or standing up. The hazardousness of the objects handed over to the users by the robot in this study were classified equivalently to that in [11]; the knife was categorized as hazardous, the spanner as unhazardous, and the cup as neutral.

2.3. Design, Procedures and Measures

Design. The evaluation study was conducted as a standardized quasi-experiment with a 2 × 3 × 2 mixed design. The first independent variable was the user group (BVI vs. sighted), which was assessed between groups. The second independent variable was object (knife vs. cup vs. spanner), which was assessed within groups. The third independent variable was handover modality (direct midair handover vs. indirect handover by placing the item on a table), which was assessed within groups. The experimental design resulted in 6 experimental conditions. With the research aim of an integrated evaluation, subjective measures (perceived easiness and perceived efficiency) and objective measures of effectiveness (success rate, joint action failure rate, robotic failure rate), efficiency (task completion time and handover time), and satisfaction (objective preference for modality) were examined for all six experimental conditions. Additionally, subjective measures of overall usability and user experience were collected once at the end of the evaluation study. Audio and video recordings were used.

Procedure. Before commencing the experiment, participants received detailed information about the study procedure and privacy policy via written participant information sheets and provided consent by signing a declaration. All documents were compatible with screen readers. Participants were categorized into either the BVI or the sighted group based on their visual abilities. Experimenters greeted participants in an anteroom and ensured their compliance with hygiene protocols necessitated by the pandemic. Participants were compensated for their time and completed a pre-survey on demographics. To maintain consistency, all surveys were read aloud, with participants verbally responding while experimenters digitally recorded answers. Once inside the experimental room, participants were encouraged to familiarize themselves with the robot by tactile exploration, accompanied by verbal explanations of its functions. Initially, the robot did not engage in communication. Subsequently, participants were seated and provided with a headset microphone for speech input. The robot then introduced itself, explained its capabilities, and instructed participants on speaking clearly and the short commands available (see [26] for further details). Experimenters directed participants to request objects from the robot (cup, knife, spanner), choosing between 'midair handover' or 'placing objects on the table' (block 1). In a subsequent block (block 2), participants repeated the process with the opposite handover modality. The order of handover objects was predetermined for each participant and randomized. Each trial encompassed the entire robotic process, from object 'ordering' to the robot's return to a predefined home position ('parking'). In the case of handover on the table, participants were instructed to pick up the object from the table to fulfill the handover task. After each trial, participants answered questionnaires on perceived easiness and perceived efficiency. In the end, participants verbally completed a post-experimental survey covering subjective overall usability and user experience measures. After all handover trials, a post-experimental survey and interview were performed with the participants. Further measures were assessed during the course of the study, such as evaluation of the VUI used (see [11] for further details on VUI evaluation) and trust in robots (to be published), which are not within the scope of this paper. For the robotic subtasks 'object identification' and 'object grasping', the methodology of Wizard of Oz was applied. The second experimenter placed the predefined object within the robotic gripper and manually sent the command to close the gripper. The sighted participants were not able to see the manual actions of the experimenter behind a partition wall. Owing to pandemic-related restrictions, the permitted test duration was restricted to 90 min per participant. Consequently, in certain instances, specific handovers could not be executed. Additionally, in consideration of these constraints, the study opted for the most efficient survey instruments available.

Measures. In the pre-survey, demographic information regarding gender, age, prior experience with robots, and sight impairments was captured. If participants did receive the specific ordered object, this was defined as a success. Several trials could not be fully or partially performed due to technical issues with the robot (e.g., problem of opening the robotic gripper or server performance) or constrained test duration. Overall, for 40 participants performing 6 experimental trials each, 240 trials were planned during the experiment, but only 220 trials (further denoted as 'total number of trials') could be performed. Objective time data were read out based on a script from the robot logfiles. *Success rate* for the human–robot interaction was measured by summing up all successful handover trials and dividing it by the total number of all trials. Failures led to unsuccessful trials, classified as either 'joint action failures' or 'robotic failures.' A *joint action failure* was defined as a failure in coordinating the handover or a misunderstanding between the human and the robot during the physical exchange phase, for instance, if the user commanded the robot to open its gripper before they had actually grasped the object. In addition, *robotic failures* were defined as failures purely resulting from the robot, e.g., the gripper losing an object while approaching the user during the preparation phase. The

Supplemental Material to this article includes a video showing excerpts from handovers to BVI participants, including examples of joint action failures and robotic failures. The video is linked at the end of the article. Subjective measures (perceived easiness and perceived efficiency) and objective measures of effectivity (success rate, joint action failure rate, robotic failure rate), efficiency (task completion time, handover time), and satisfaction (objective preference for modality) were examined following all six experimental conditions. As unsuccessful trials sometimes did not have a predefined endpoint (e.g., because trials had to be stopped during handover), these were excluded in some measures (see Table 3). Section 2.4 gives further information about the analysis of objective measures. For post-trial measurement of subjective effectivity and efficiency, two single-item measures were used. Participants verbally rated their answers on a ten-point Likert scale. The effectiveness item was adopted from [17]. At the end of the evaluation study, the System Usability Scale (SUS) [33] was used as a subjective measure of overall usability. Standardized Cronbach's alpha showed α = 0.76, which is lower than the reliabilities found in other studies [35]. The User Experience Questionnaire [38] in its short version (UEQ-S) [40] was used as a subjective measure of user experience. The UEQ-S contains 8 items scored on two subscales: pragmatic quality (4 items) and hedonic quality (4 items). Standardized Cronbach's alpha showed $\alpha_{\text{pragmatic}} = 0.47$ and $\alpha_{\text{hedonic}} = 0.72$. Due to the poor reliability of the pragmatic quality subscale, user experience was analyzed overall across 8 items. Cronbach's alpha showed $\alpha_{\text{UEQ-Soverall}} = 0.72$. Satisfaction with the interaction was additionally subjectively evaluated by several open questions regarding the HRI: 'What did you particularly like and what did you dislike?', 'Could you find an adjective describing the entire interaction process with the robot?', 'During midair handover, an airflow indicated the position of the object. How did you experience the airflow? How did it support the handover?'. Following the results of [28], BVI participants were additionally asked, 'Which handover modality did you prefer: at the table or midair? Please explain why'. An overview of dependent measures can be found in Table 3.

Table 3. Measures for dependent variables.

Variable	s/o [1]	Measure	Description	Unit	Time Assessed
Overall usability	s	SUS	System Usability Scale (SUS) [33], 10 items	from 1 = strong disagreement to 5 = strong agreement	End of study
effectiveness	o	success rate	ratio of successful trials (participants successfully received the object independent of their grasping behavior) and the total number of trials (see Section 2.3)	Percent	Across all trials
		joint action failure rate	number of unsuccessful trials caused by failures during joint actions of the human and the robot divided by the total number of trials (see Section 2.3)	Percent	Across all trials
		robotic failure rate	number of unsuccessful trials caused by robotic failures divided by the total number of trials (see Section 2.3)	Percent	Across all trials
	s	perceived easiness [2]	single item 'getting the object was easy' [17]	from 1 = strongly disagree to 10 = strongly agree	After each trial [2]

Table 3. *Cont.*

Variable	s/o [1]	Measure	Description	Unit	Time Assessed
efficiency	o	task completion time	the time period between the robot asking for participants' order and the robot arriving at home position again after handing over the object	Minutes	Each trial [2]
		handover time (in minutes)	the time period between the robot requesting handover readiness of participants and the final opening of the robotic gripper	Minutes	Each trial [2]
	s	perceived efficiency	single item 'the duration of the object handover was acceptable'	from 1 = strongly disagree to 10 = strongly agree	After each trial [2]
User Experience	s	UEQ-S	User Experience Questionnaire [38] in its short version [40], 8 items	semantic differential rated on a 7-point scale (scale range −3 to 3)	End of study
Satisfaction	o	Objective preference for modality	proportion of chosen handovers with modality direct handover 'midair' and indirect handover 'on table' during block 1		First 3 trials
	s	qualitative feedback	Open questions	-	End of study

[1] objective (o) or subjective (s) assessment. [2] unsuccessful handovers excluded from measure.

2.4. Data Analysis

For analysis of quantitative evaluation data, statistics software R [41] version 4.2.2 was used.

Bayes analyses were used for inference statistics. Therefore, the R package 'BayesFactor' [42] was used for group comparison and to create Bayes Models containing all independent variables of interest. According to [42], a parameter value of 0.707 for 'rscale' is defined as medium. Yet, with the aim of equivalence testing, in this paper, 'rscale' was set to 0.3 for all analyses to test the null hypothesis against a rather small effect. BF_{01} represents the relative likelihood of the null hypothesis over the alternative hypothesis. If $BF_{01} > 1$, the null hypothesis is more likely than the alternative. In contrast, BF_{10} represents the relative likelihood of the alternative hypothesis over the null hypothesis. Both measures are convertible by calculating the inverse value of $1/BF_{01}$ and $1/BF_{10}$, respectively. For example, a $BF_{01} = 2$ shows that the null hypothesis is two times likelier than the alternative hypothesis.

For the calculation of Bayes Models covering the design of the experiment, missing data were imputed. The imputation method 'MissForest' [43] was applied. The capabilities of 'MissForest' under various conditions were demonstrated and recommended by [44], based on a simulation study. The R package 'missForest' [45] was used. The relative number of imputed data per measure is given in the results sections.

To analyze the qualitative data, inductive and deductive category formation were applied. Therefore, qualitative data were audio and/or video recorded to enable a retrospective in-depth analysis. The relevant sense of the conversations was summarized directly while listening and later from audio data. For interviews, participants' answers were directly assigned to the pre-defined category of interest from the interview guideline (deductive).

3. Results

Results showed acceptable overall usability (Section 3.1) and high effectiveness (see Section 3.2), shortcomings regarding efficiency (Section 3.3), but interestingly also high satisfaction (Section 3.4). For an overview of the results, the table in Appendix A gives descriptive statistics for all dependent variables for BVI and sighted participants.

3.1. Overall Usability

Figure 2 visualizes the results on usability as well as Bayes Factors for comparison of groups (for a description of the Bayes Factor, see Section 2.4). Data showed relatively small variances, but in group 'BVI', one outlier is visible. The Bayes Factor was calculated for the groups 'BVI' (M_{BVI} = 70.25) and 'sighted' ($M_{sighted}$ = 72.50). For the SUS Score, a value of BF_{01} = 1.74 ± 0% resulted, showing the null hypothesis to be 1.7 times likelier than the alternative. Therefore, data support Hypothesis 1 of equal usability for both user groups. Additionally, means of both groups exceeded the target value of 70, which is acceptable usability according to [34].

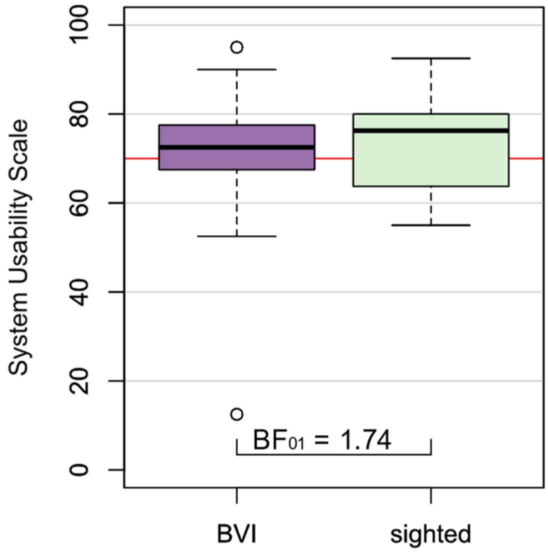


Figure 2. Results of subjective usability, with target value marked red. The circles denote outlier values. A value is considered an outlier if it exceeds 1.5 times the interquartile range above the third quartile or falls below 1.5 times the interquartile range from the first quartile.

3.2. Effectiveness

Table 4 summarizes descriptive statistics of effectiveness measures dependent on object and modality for both groups as well as Bayes Factor comparison results between groups.

Table 4. Descriptive statistics for dependent variables of effectiveness.

Variable	Handover Modality	Object	BVI M (SD)	Sighted M (SD)	BF_{01}
success rate	midair	cup	100%	100%	-
		knife	90%	100%	1.06 ± 0.01%
		spanner	89%	100%	1.05 ± 0.01%
	on table	cup	100%	100%	-
		knife	94%	100%	1.37 ± 0.00%
		spanner	94%	100%	1.37 ± 0.00%
joint action failure rate [1]	midair	cup	-	-	-
		knife	5%	-	-
		spanner	5%	-	-
	on table	cup	-	-	-
		knife	-	-	-
		spanner	-	-	-

Table 4. *Cont.*

Variable	Handover Modality	Object	BVI M (SD)	Sighted M (SD)	BF_{01}
robotic failure rate [1]	midair	cup	-	-	-
		knife	5%	-	-
		spanner	5%	-	-
	on table	cup	-	-	-
		knife	6%	-	-
		spanner	6%	-	-
perceived easiness [2] (scale range 1 to 10)	midair	cup	7.75 (2.40)	8.16 (2.17)	1.71 ± 0.0%
		knife	6.35 (3.35)	9.35 (0.75)	0.06 ± 0.0%
		spanner	6.58 (3.34)	8.84 (2.06)	0.67 ± 0.0%
	on table	cup	9.65 (0.70)	9.68 (0.67)	1.80 ± 0.0%
		knife	9.06 (2.19)	9.78 (0.55)	1.40 ± 0.0%
		spanner	8.82 (2.40)	9.72 (0.58)	1.26 ± 0.0%

[1] no group comparison performed due to rare events and as data are included in measure 'success rate'. [2] unsuccessful handovers excluded from measure.

3.2.1. Success Rate

Across conditions, a high success rate for all handover trials resulted. Only six trials out of two hundred and twenty were classified as unsuccessful, resulting in an overall success rate of 97.3%. In the group 'sighted', no unsuccessful trials occurred, resulting in a success rate of 100%. Against this, the success rate in group 'BVI' was 94.55%. The resulting Bayes Factor was BF_{01} = 0.74 ± 0%, showing that the alternative hypothesis was 1.3 times likelier than the null hypothesis and showing a higher success rate in the group 'sighted'.

In group 'BVI', the success rate varied with object and modality. The lowest success rate occurred during midair handovers of the spanner, still showing a success rate of 89%. For both modalities, handing over the cup was more successful than the spanner and knife.

A Bayes Model was created based on 7.1% imputed missing data. The Bayes Model showed a mean effect for the group (BF_{10} = 3.08), whereas object (BF_{01} = 2.57) and modality (BF_{01} = 3.49) favored the null hypothesis. Therefore, data support Hypothesis 2 but do not support Hypothesis 3. Additionally, an interaction effect of group × object (BF_{10} = 1.22) occurred. Therefore, first, 'BVI' participants experienced a lower success rate than 'sighted' participants, and second, the success rate was especially low for some objects (knife and spanner) in group 'BVI'.

3.2.2. Joint Action Failure Rate

In group 'BVI', a total of two failures in joint actions occurred, resulting in an overall joint action failure rate of 1.81%. These failures occurred during midair handovers of the knife and the spanner (see Table 4). In contrast, in the group 'sighted', no failures occurred in joint actions. Due to the low incidence rate of failures in joint actions, as these are also included within the measure 'success rate', no Bayes Factors were calculated for group comparison.

3.2.3. Robotic Failure Rate

In group 'BVI', a total of four robotic failures occurred, resulting in an overall robotic failure rate of 3.63%. These failures occurred during handovers of knives and spanners in both handover modalities (see Table 4). In contrast, in the group 'sighted', no robotic failures occurred. As robot failures did not include HRI, errors occurred randomly within the group of 'BVI' participants, and therefore, Bayes Factors are not calculated.

3.2.4. Perceived Easiness

Figure 3 and means in Table 4 show that modality 'on table' was rated as being easier in both groups, showing values only slightly below the scale maximum. Still, variances in group 'BVI' were higher. Table 4 also shows Bayes Factors for between-group comparison. Overall, mean values were high for all objects, and both modalities remained far above the theoretical scale mean. The cup handed over 'midair' and all objects handed over 'on table' showed BF_{01} values in favor of the null hypothesis, showing no difference between the 'BVI' and 'sighted' groups. In contrast, Bayes Factors for objects knife (M_{BVI} = 6.35, $M_{sighted}$ = 9.35, BF_{10} = 17.71) and spanner (M_{BVI} = 6.58, $M_{sighted}$ = 8.84, BF_{10} = 1.49) handed over 'midair' showed the alternative hypothesis to be likelier. Therefore, BVI participants perceived less easiness during these handovers compared to sighted participants, and data showed higher variances in this group.

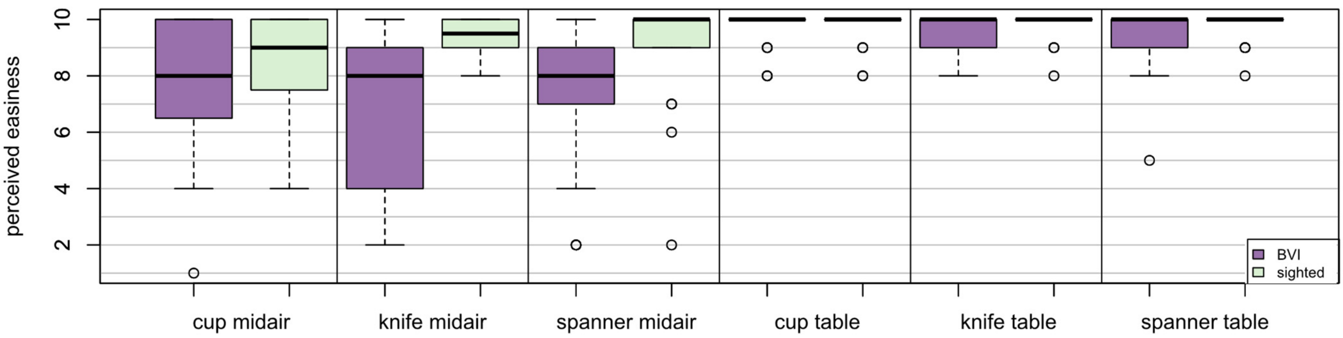


Figure 3. Subjective ratings of perceived easiness (subjective effectiveness) (unsuccessful trials excluded). The circles denote outlier values. A value is considered an outlier if it exceeds 1.5 times the interquartile range above the third quartile or falls below 1.5 times the interquartile range from the first quartile.

A Bayes Model was created based on 7.1% imputed missing data. The Bayes Model showed a mean effect for group (BF_{10} = 363.31) and modality (BF_{10} = 951,280.60), whereas object (BF_{01} = 7.20) favored the null hypothesis. The main effect of modality does not support Hypothesis 4. Additionally, interaction effects of group $\times$ object (BF_{10} = 51.73), group $\times$ modality (BF_{10} = 1,104,519,191.00), and object $\times$ modality (BF_{10} = 135,560.40) occurred. Therefore, first, 'BVI' participants perceived less easiness during handovers than 'sighted' participants. Second, 'midair' handovers were perceived as less easy than 'on table' handovers. Third, group 'BVI' experienced less easiness during the handover of the objects knife and spanner. Fourth, group 'BVI' experienced less easiness during 'midair' handovers, and fifth, easiness was lowest during 'midair' handovers of the objects knife and spanner.

Figure 3 does not include unsuccessful handover trials. For group 'BVI', six unsuccessful handover trials occurred. All six participants answered the scale with 1.

3.3. Efficiency

Table 5 summarizes descriptive statistics of efficiency measures dependent on object and modality for both groups, as well as results of between-group comparisons.

Table 5. Descriptive statistics for dependent variables of efficiency.

Variable	Handover Modality	Object	BVI M (SD)	Sighted M (SD)	BF_{01}
task completion time [1] (in minutes)	midair	cup	5.09 (0.73)	5.17 (0.69)	1.76 ± 0.0%
		knife	5.98 (0.76)	5.97 (1.36)	1.82 ± 0.0%
		spanner	6.29 (0.97)	6.19 (0.87)	1.74 ± 0.0%
	on table	cup	4.31 (0.51)	4.89 (1.23)	0.82 ± 0.0%
		knife	4.67 (0.66)	4.92 (0.87)	1.43 ± 0.0%
		spanner	4.78 (0.73)	4.96 (0.98)	1.62 ± 0.0%
handover time [1,2] (in minutes)	midair	cup	1.39 (0.40)	1.31 (0.47)	1.65 ± 0.0%
		knife	1.93 (0.62)	1.66 (0.55)	0.71 ± 0.0%
		spanner	2.07 (0.90)	1.72 (0.43)	0.79 ± 0.0%
perceived efficiency [1] (scale 1 to 10)	midair	cup	5.85 (2.66)	6.21 (2.74)	1.77 ± 0.0%
		knife	5.65 (2.96)	6.50 (2.40)	1.65 ± 0.0%
		spanner	5.47 (3.04)	5.74 (2.16)	1.80 ± 0.0%
	on table	cup	6.41 (2.85)	6.63 (2.52)	1.78 ± 0.0%
		knife	6.41 (2.85)	6.28 (2.47)	1.77 ± 0.0%
		spanner	5.88 (3.12)	6.83 (2.20)	1.56 ± 0.0%

[1] unsuccessful handovers excluded from measure. [2] no comparison for modality 'on table' as time differences only resulted from technical variance.

3.3.1. Task Completion Time

Figure 4 shows the overall task completion time dependent on object, modality, and group. On average across groups and objects, task completion time was 5.76 min for 'midair' handovers and 4.76 min for handovers 'on table'. Task completion time was composed of ordering (M = 67 s), searching for the object (M = 49 s), grasping the object (M = 43 s), delivery (M = 32 s), handing over (M = 87 s), and parking (M = 42 s). In the Prehandover phase, this equates to a total average of 202 s, consisting of 67 s for the Initiation subphase and 135 s for the Preparation subphase. In the Physical Handover phase, the total average is 102 s, with 87 s for the Physical Exchange subphase and 42 s for the Performance subphase.

Bayes Factors for group comparison favored the null hypothesis for all objects and modalities, except for the object cup handed over 'on table' (M_{BVI} = 4.31, $M_{sighted}$ = 4.89, BF_{10} = 1.22). Trials with 'midair' handovers showed that task completion time was lowest for the cup.

A Bayes Model was created based on 12.1% imputed missing data. The Bayes Model showed a mean effect for object (BF_{10} = 734.81) and modality ($BF_{10} = 7.71 \times 10^{13}$), whereas group ($BF_{01}$ = 3.66) favored the null hypothesis. Additionally, interaction effects of group × object (BF_{10} = 224.08), group × modality ($BF_{10} = 2.61 \times 10^{13}$), and object × modality ($BF_{10} = 1.54 \times 10^{18}$) occurred. Therefore, first, different objects resulted in different task completion times. Second, handing over objects 'on the table' was faster than 'midair'. Third, the highest task completion times resulted for group 'BVI' during handovers of the knife and the spanner. Fourth, the highest task completion times resulted for group 'BVI' during 'midair' handovers, and fifth, the object spanner resulted in the highest task completion time during 'midair' handovers.

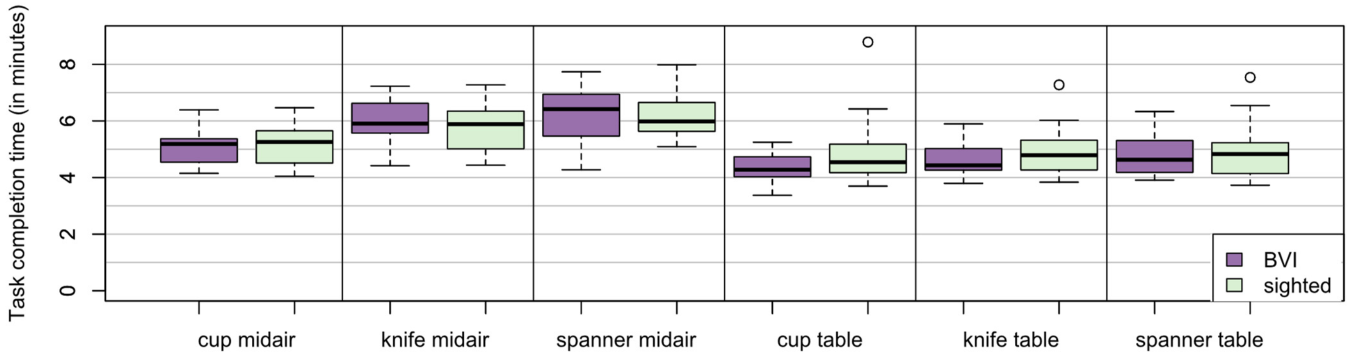


Figure 4. Task completion time dependent on object, modality, and group (unsuccessful trials excluded). The circles denote outlier values. A value is considered an outlier if it exceeds 1.5 times the interquartile range above the third quartile or falls below 1.5 times the interquartile range from the first quartile.

3.3.2. Handover Time

Figure 5 shows handover times dependent on object, modality, and group. Handover times varied with modality, where midair handovers took longer than handovers on the table except for the object cup. On average, 'midair' handover time varied between 1.31 min ('sighted'—'cup') and 2.07 min ('BVI'—'spanner'). Handover times 'on the table' resulted in about 1 and a half minutes. As handovers on the table did not include HRI, Bayes Factors were not calculated for modality 'on the table'.

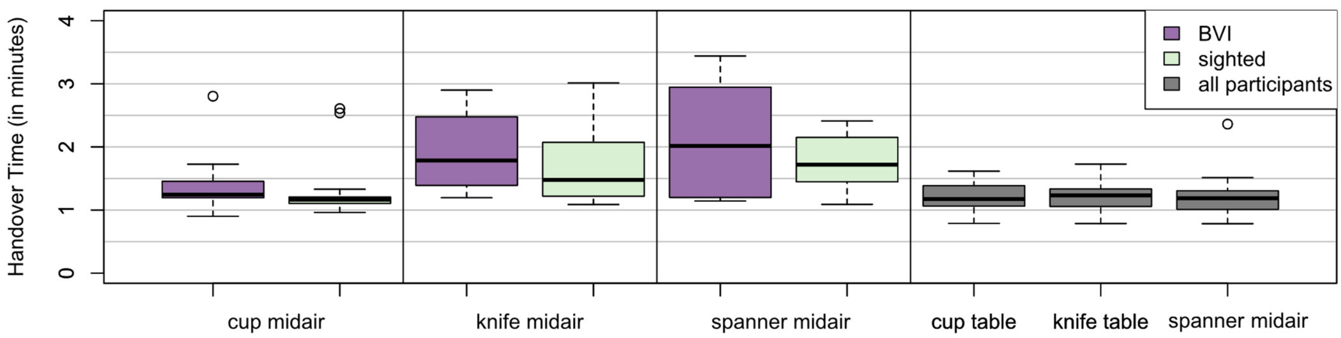


Figure 5. Handover time dependent on object, modality, and group (unsuccessful trials excluded). The circles denote outlier values. A value is considered an outlier if it exceeds 1.5 times the interquartile range above the third quartile or falls below 1.5 times the interquartile range from the first quartile.

Bayes Factors for group comparison favored the null hypothesis for the cup, but favored the alternative hypothesis for objects knife (M_{BVI} = 1.93, $M_{sighted}$ = 1.66, BF_{10} = 1.41) and spanner (M_{BVI} = 2.07, $M_{sighted}$ = 1.72, BF_{10} = 1.26) handed over midair.

A Bayes Model was created based on 14.2% imputed missing data. The Bayes Model showed a mean effect for object (BF_{10} = 6543.86) and modality (BF_{10} = 15,104.78), whereas group (BF_{01} = 2.20) favored the null hypothesis. The missing main effect of group supports Hypothesis 5, while the main effect of modality supports Hypothesis 7. On the contrary, the main effect of object does not support Hypothesis 9.

Additionally, interaction effects of group × object (BF_{10} = 3220.36), group × modality (BF_{10} = 7445.15), and object × modality (BF_{10} = 489,762,237) occurred. Therefore, first, different objects resulted in different handover times. Second, handing over objects 'on the table' was faster than 'midair'. Third, for group 'BVI', handing over the knife and the spanner resulted in especially higher handover times. Fourth, the highest handover times resulted from ingroup 'BVI' during 'midair' handovers. Fifth, the highest handover time resulted for objects knife and spanner during 'midair' handovers.

3.3.3. Perceived Efficiency

Figure 6 shows the subjective measures of efficiency and perceived efficiency, dependent on object, modality, and group. In comparison to perceived easiness, mean and especially median values are lower. Mean values remained above the theoretic scale center for both groups, but median values (see Figure 6) in group 'BVI' even fell below the theoretical scale center. Additionally, high variances are visible in both groups. Bayes Factors for group comparison (see Table 5) favored the null hypothesis for all conditions.

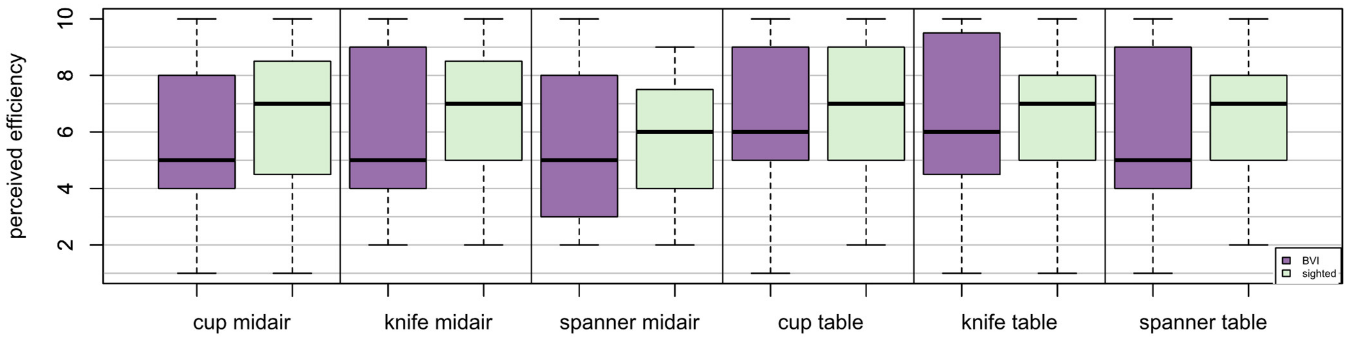


Figure 6. Subjective ratings of perceived efficiency (subjective efficiency) (unsuccessful trials excluded).

A Bayes Model was created based on 7.1% imputed missing data. The Bayes Model did not show main effects ($BF_{01} \geq 2.29$) nor interaction effects ($BF_{01} \geq 6.04$). The missing main effect of the group supports Hypothesis 6. On the contrary, the missing main effect of modality does not support Hypothesis 8.

Figure 6 does not include unsuccessful handover trials. For group 'BVI', six unsuccessful handover trials occurred. Three participants answered the scale with 1, one participant with 5, one participant with 6, and one participant with 8.

3.4. User Experience and Satisfaction

Figure 7 shows group comparisons for user experience for UEQ-$S_{overall}$. Visually, user experience was higher for group 'BVI' compared to 'sighted' participants.

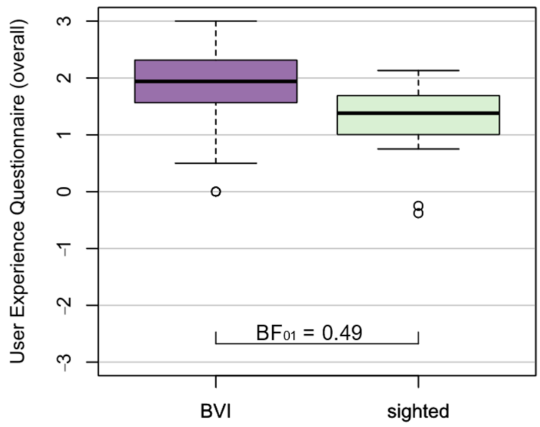


Figure 7. Results of user experience ratings for UEQ-$S_{overall}$. The circles denote outlier values. A value is considered an outlier if it exceeds 1.5 times the interquartile range above the third quartile or falls below 1.5 times the interquartile range from the first quartile.

For UEQ-$S_{overall}$ ($M_{BVI} = 1.79$, $M_{sighted} = 1.27$), the Bayes Factor showed $BF_{01} = 0.49 \pm 0\%$, showing the alternative hypothesis to be two times likelier than the null hypothesis. Therefore, the data do not support Hypothesis 10, with higher ratings of 'BVI' participants compared to sighted participants.

Data of objective satisfaction (objective preference for modality) were analyzed over time, as participants were free to choose a modality for the first three handovers. Hence, we report the relative proportion of direct handover 'midair' and indirect handover 'on table' trials during block 1 (trial numbers 1 to 3; see Procedure in Section 2.3). Overall, 0.5% of trials in block 1 were performed using the modality 'midair', while 19.5% were performed 'on the table'. Preferences for modality dependent on group are given in Appendix A. A Bayes Model was created with the objective preference of modality as the dependent variable, based on 1.6% imputed missing data. The Bayes Model did not show main effects for group ($BF_{01} \geq 3.20$), trial number ($BF_{01} \geq 4.52$), object ($BF_{01} \geq 1.75$), nor interaction effects for group x order ($BF_{01} \geq 14.85$) or order x object ($BF_{01} \geq 8.02$). The missing main effect of group contradicts Hypothesis 11.

In the following, we present the qualitative satisfaction feedback among the different user groups. When queried about aspects of the interaction they particularly liked or disliked, some sighted participants commended the positive communication with the robot, while others praised the reliability of the handovers. However, some sighted participants negatively pointed out that trajectories during object handovers are ambiguous, and a few noted that the overall process duration was too lengthy. Nevertheless, most sighted participants perceived their interaction with the robot positively. They characterized the robot by using adjectives such as 'interesting', 'exciting', 'pleasant', and 'helpful', with only a minority of interactions feeling 'boring' and 'inflexible'. BVI participants expressed preferences and concerns regarding the handover and voice interaction features. Some BVI participants were satisfied with the handover process, describing it as 'effective', while others perceived it as 'generally safe'. Positive feedback was also received for the robot's voice recognition, with a few participants appreciating the clear communication of the robot's actions, providing them with a sense of awareness. Some participants specifically praised the robot's quiet operation and friendly voice. On the flip side, certain BVI participants highlighted challenges, noting that not all handovers were successful. For instance, one participant mentioned an extended period of time searching for the item during a midair handover. One participant reported difficulty in distinguishing the metal material of the robot gripper from the spanner by touch. Dissatisfaction with the robot's speed was raised by a few participants who found it 'too slow'. Additionally, some BVI participants expressed problems with the voice interaction, citing occasional misunderstandings of speech inputs and perceiving the interaction as 'excessively time-consuming'. A few participants were dissatisfied with the inability to process voice input in complete sentences. Some were also unsure about the appropriateness of speaking loudly and vigorously to the VUI in certain environments, such as hospitals. A few participants also felt that the robot's volume was too low. Despite these challenges, the overall interaction with the robot received positive feedback from most BVI participants. They described the experience as 'interesting', 'exciting', 'pleasant', and 'supportive'. However, a minority found the interaction 'tedious', 'tiring', and somewhat 'unreliable'.

The airflow from the haptic feedback system went largely unnoticed by most sighted participants, providing little to no assistance during the handover process. However, some found the use of an airflow for orientation during robotic object handover to be a good idea. The BVI participants were also not or hardly supported by the airflow during the handover. The position of the airflow in relation to the object was not communicated clearly enough for them. Some BVI participants noticed the airflow, which was perceived as partially supportive. Others stated that they had perceived little or no airflow, and therefore, it provided them with little or no support. A few participants stated that the airflow was irritating rather than supportive. For example, they described the position as unfavorable, as the airflow was not close enough to the object. One participant found the idea of the airflow 'interesting'.

BVI participants were also asked about their preferred handover modality. Almost all BVI participants (18 out of 20) preferred indirect handovers on the table. Some stated that during indirect handovers on the table, they knew where to search for the object, which gave them a sense of security. Two participants preferred direct midair handover. One of these participants emphasized the immediacy of receiving the object, while the other stated that searching for the object located on the robot was safe and not a problem. However, for some other BVI participants, midair handover generated a sense of insecurity. They did not know where they could grasp the object or whether they had grasped it correctly. In general, they were concerned that the handover would not work smoothly in all circumstances.

3.5. Relationship between Dependent Variables

Table 6 shows the results of descriptive correlational analyses. First, relationships between overall usability ratings and measures of usability subcomponents were assessed. The System Usability Scale values did not correlate with success rate (effectiveness), indicated by the rank correlation coefficient Kendall's tau.

Table 6. Relationships between dependent variables.

Relationship	Repeated Measures Correlation [1]	Kendall's Tau	z	p-Value
overall usability and specific measures				
SUS × success rate [2]	-	−0.006	−0.045	0.964
SUS × UEQ-S overall	-	0.302	2.61	<0.001
objective and subjective measures				
effectiveness:				
success rate × perceived easiness [2]	0.544	-	8.33	<0.001
efficiency:				
task completion time × perceived efficiency [2]	−0.242	-	3.41	<0.001
handover time × perceived efficiency [2]	−0.191	-	2.69	0.007
human idle time × perceived efficiency [2]	0.076	-	1.06	0.287

[1] repeated measures correlation coefficient due to dependent data. [2] imputed data used for correlational analysis.

Since the perceived easiness (a subjective measure of effectiveness) and time-related objective measures of efficiency varied across experimental conditions, calculating a correlation between overall usability and means of those measures was omitted. Between the System Usability Scale values and the overall user experience, the rank correlation coefficient Kendall's tau indicated a positive relationship.

The two subcomponents of usability, effectiveness, and efficiency were assessed using objective and subjective measures. These correlated for effectiveness, but for efficiency, only the objective measures of active stages of the handover process showed significant negative correlations with subjective measures.

4. Discussion

This study provides important insights into the usability of robot-to-human handovers for differing user groups, namely BVI and sighted users. We evaluated whether a handover robot designed to meet specific user needs is equally proficient in assisting BVI and sighted users. This section discusses the study's results, examining how various measures relate to each other and findings from other studies. However, as highlighted by the authors in [6], comparing results across different studies on robotic object handover is challenging due to the absence of established standards regarding the objects used, metrics employed, and handover phases considered.

4.1. Overall Usability

First, we evaluated overall usability, aiming to determine if the robot could be equally effectively used by both BVI and sighted users for object handover. We anticipated no significant difference in overall usability ratings between both groups (Hypothesis 1). The data from the System Usability Scale ratings supported this expectation, confirming the achievement of the primary goal of our study.

We also aimed to achieve a minimum usability quality for both groups. Since the handover robot is an early prototype and not a market-ready product, setting expectations based on consumer software products, typically used in System Usability Scale evaluations, was not sensible. Therefore, we aimed for an SUS score above 70, which is considered acceptable usability according to [34]. This goal was met, as means of both groups exceeded the predefined value of 70. However, the robot requires several enhancements to fully meet user requirements and be considered a product with good usability for supporting everyday life, necessitating a minimum SUS score of 80 or higher [35]. The extent to which these requirements are already met and where there is room for improvement is explored in more detail in the following sections.

4.2. *Effectiveness*

Effectiveness as a subcomponent of usability describes the extent to which goals are achieved. The goals were equally successful handovers and easy grasping for BVI and sighted users. Overall, the effectiveness of the robot handover can be described as high.

The handover **success rate** was expected to be consistently high in both groups, with low failure rates independent of the user group (Hypothesis 2). The success rate for all handover trials across conditions was high, with an overall success rate of 97.3%. This aligns with findings from other handover studies, such as [17], which reported an overall success rate of 94%, or [12], which observed a success rate of 97% for indirect handover and 78% for direct handover. If participants did not receive the specifically ordered object, the trial was defined as unsuccessful. The few cases without a successful handover only occurred in the group of BVI users. They occurred due to errors by the robot (a total of four out of one hundred and ten (3.6%) handovers in this group) or failures in joint action (a total of two out of one hundred and ten (1,8%) handovers). Other studies have documented similar failures, such as those outlined in [12], where the robot misinterpreted user inputs or failed to open its gripper. From failures, negative results can arise for the user, characterized as possible user-related damage in ISO 9241-11:2018 [32]. For example, if BVI users have to pick up objects lost by the robotic gripper or due to failures in joint actions, this can lead to injuries on unexpected obstacles. An accessible handover robot should, therefore, have the functionality to provide precise information about the error and the ability to pick up lost objects.

Despite the high success rate, Bayes analyses indicated a higher likelihood of a difference between the two groups than no difference, as errors were exclusively observed in the BVI group. However, it is important to note that two-thirds of the observed errors were caused by the robot and were independent of user actions. Therefore, this result should be viewed with caution. Additionally, the Bayes Model revealed an interaction effect with an increased probability of errors in the BVI group for specific objects, particularly narrow ones like the knife or spanner, suggesting a need for improvement in the robot's grasping strategy for such objects.

Also, **perceived easiness** was expected to be consistently high in both groups independent of the user group (Hypothesis 3). Corresponding with results regarding success rate, ratings of perceived easiness were high, remaining far above the theoretical scale mean. Similarly, ref. [9] reported high ratings for a comparable item assessing comfortable handover, particularly when the handle of the object handed over was oriented towards the user, mirroring the implementation of object handover in our study. However, in detail, our results show a more differentiated picture. Contrary to expectations and despite the high average ratings, BVI participants rated the handover as less easy than sighted participants. In addition, the assessment of BVI participants was significantly more heterogeneous than that of sighted participants. It should be noted that unsuccessful handovers are not included in this measure. This means that the expectations of BVI users regarding the easiness of receiving the object from the robot were only partially met.

An explanation for this outcome can be derived by examining the results of the other test conditions. Initially, for the handover modality, we anticipated an interaction effect wherein BVI users would rate indirect handover on the table as easier, while sighted users would find direct handover in midair easier (Hypothesis 4). Surprisingly, both groups perceived indirect handover on the table as easier. However, this difference was more pronounced among BVI users, partially aligning with expectations. This was evident in a considerable main effect identified by the Bayes analysis for the group factor and a substantial interaction effect for the group x modality factors. Consequently, the findings of [28], indicating that visually impaired individuals tend to avoid midair object handovers in daily life and prefer indirect handovers on a table for better orientation, can be extended to robotic handovers.

Our expectation was that the objects, regardless of their potential risk of injury, would be equally easy to grasp (Hypothesis 3). This assumption was based on the robot's grasping strategy, which adhered to social conventions in human object handover, ensuring that potentially hazardous surfaces were covered by the gripper and handed over oriented away from the receiver. Surprisingly, the hazardousness of the objects did not influence perceived easiness ratings. However, two interaction effects revealed by Bayes analysis indicated that both the knife and spanner were rated as more challenging to receive, both for BVI users and in the case of midair handover in general. This could be attributed to difficulties some BVI users faced in distinguishing the gripper's material from the object. Additionally, the success rate indicated that robotic failures occurred most frequently with these two objects, suggesting that although the transfer might have been successful overall, the gripper did not grasp the object very stably. Unfortunately, this stability was not separately measured. However, similar issues with narrow objects were already highlighted in the results of [11] during the development of the applied robotic grasping strategy.

Discussed results regarding perceived easiness so far did not include unsuccessful handover trials. In all unsuccessful trials, participants consistently rated the perceived easiness with the lowest possible value. This serves as an effective control, indicating that participants based their ratings on objective success.

4.3. Efficiency

Efficiency, as a component of usability, refers to the effort needed to accomplish a goal. In summary, the efficiency of the robot handover must be deemed to require improvement.

Time-related measures were used to objectively assess efficiency. We expected higher times for BVI users in active handover phases than sighted users, as they potentially perform their movements more slowly and cautiously and have to invest more effort in localizing the object (Hypothesis 5). This was not confirmed. The data showed no group difference in the task completion or handover time. These are encouraging results for the goal of equally good assistance of the handover robot for BVI and sighted users. However, this initial impression is diminished if we take a closer look at the results of the other test conditions.

For handover modalities, we expected higher times for midair handover compared to indirect handover on the table, as direct handovers require more coordination effort (Hypothesis 7). This expectation was very clearly confirmed. On average, across all objects, the task completion time for midair handovers was 5 min and 46 s. Indirect handovers were around 1 min faster than direct handovers. This difference was also evident in the Bayes analysis of all three time-related measures, which revealed a main effect for modality, which was particularly high for handover time. The findings contrast with those of [12], which found direct handover to be quicker, with a mean total time of 2 min and 26 s, compared to indirect handover, which averaged 4 min and 39 s. However, comparing the studies is challenging. In [12], participants had Amyotrophic Lateral Sclerosis, leading to motor impairment, not BVI users. Also, differences exist in the handover phases. For example, in [12], initiation involved laser pointer ordering, with durations of 14 s for direct handover and 12 s for indirect, unlike our verbal ordering, averaging 63 s. Moreover, ref. [12]'s Preparation did not involve searching and grasping but included calculation time, resulting in transport times of 98 s for direct handover and 184 s for indirect, compared to our mean total time of 135 s. Further, transport times are influenced by setup distances. In the Physical Exchange subphase, ref. [12] included different robot steps. For instance, placing or handing time was 31 s for direct and 66 s for indirect handover in [12], including calculating handover position and grasping time, whereas we did not differentiate those times, and our timing included VUI processing for verbal coordination, averaging 102 s. This timing might increase if a system had to perform both calculations, making the robot slower. Additionally, our time measure included the Performance subphase with parking time, unlike [12].

In addition, the Bayes analyses of the time-related measures revealed interaction effects between group and modality. This was again particularly pronounced for handover time. Here, it was shown that in midair handovers, blind or visually impaired participants took significantly longer to receive the object than sighted participants. The values of the BVI group also varied much more widely for midair handovers than in the sighted group. This can be seen as confirmation of the previously discussed Hypothesis 7, which is limited to midair handovers. Once again, this shows that midair handovers are challenging for BVI users.

As with the effectiveness measures, we did not expect any differences between the different objects for the time-related efficiency measures (Hypothesis 9). Similar to the effectiveness measures, this expectation was not confirmed by the data, and interaction effects of objects were found, both for BVI users and for the midair handover modality. BVI users took longer to complete the task and receive the object when handed over directly midair. In addition, handing over the knife and the spanner in midair took longer than handing over the cup.

Subjective efficiency was also measured with reference to time by asking whether the time required for the handover was perceived as acceptable. This is where the results showed the greatest weakness of the interaction with the handover robot. Across all conditions, the mean rating of time acceptability was significantly lower than the subjective easiness rating and close to the theoretical center of the scale. In some conditions, the median was even below the theoretical scale center. The overarching conclusion is that the handover by the robot was perceived as unacceptably slow. However, the participants were very divided in their ratings; the values varied very widely in all conditions, often across the entire scale range. This cannot be explained by the variance of the objectively measured times but can only be explained by different user expectations. As with the other dependent variables, we did not expect any difference between the groups (Hypothesis 6). However, we expected lower values of subjective efficiency for direct handover midair compared to indirect handover on the table, as the latter should be faster due to less coordination effort (Hypothesis 8). As reported in the discussion of objective efficiency measures, direct handover to the table was indeed faster. However, this did not affect whether this time was perceived as more acceptable. Instead, the Bayes analysis showed no difference for any experimental condition, which was certainly due to the generally poor mean rating of perceived efficiency combined with the broad variance in all conditions.

4.4. User Experience and Satisfaction

In simple terms, user experience represents the contribution of perceived fun and enjoyment to user satisfaction. Overall, the user experience of the robot handover can be described as good and largely satisfying.

User Experience: We expected no difference in user experience ratings between BVI and sighted users. Therefore, the data do not support Hypothesis 10 for overall user experience, with BVI participants rating user experience higher than sighted participants. This difference could be attributed to the infrequent need for assistance in daily tasks among sighted users, as opposed to BVI individuals, who often require such support.

Consequently, BVI users may hold different expectations of the robot, potentially resulting in higher ratings for the hypothetical benefits compared to sighted users.

Satisfaction: Repeated preference for one modality was meant to serve as a short-term objective measure of satisfaction. Hence, we expected BVI users to choose more often indirect handover on a table, whereas sighted users were expected to choose more often direct midair handover. However, in around 80% of cases, users chose direct handover midair when given the choice, regardless of user group and trial number. This result contradicts our hypothesis but should be taken cautiously. Participants were only allowed to choose the modality three times, and the object handed over changed each time. Therefore, the few repetitions may have been unsuitable as an objective measure of satisfaction, and the choice may have been overridden by curiosity about the novelty of the next delivery. This assumption is supported by the qualitative results on the users' preferred modality. As expected, most BVI users stated they prefer indirect handover on a table. This aligns with handover strategies employed by BVI individuals in everyday handovers between humans, as documented by [28].

We expected positive overall subjective satisfaction in qualitative feedback, reflecting instances where the interaction met the requirements and preferences of the user groups but also where it did not. As expected, users reported a positive experience with the robot's handover interaction and expressed satisfaction with its performance. However, the feedback from BVI participants exhibited more variability than that of sighted participants. This also aligns with our findings on overall usability, showing positive SUS values in both groups, but at the same time, some significant outliers in the BVI group, corresponding to a substantial portion of negative feedback from this group. BVI users also provided suggestions for improving object location and distinguishability of objects from the gripper. Nevertheless, across all participants, common issues arose regarding the performance of the VUI. Users in both groups encountered challenges where their speech inputs were not accurately understood, necessitating repetition or correction of processed order information. This issue is particularly significant for the BVI user group, as they rely heavily on acoustic coordination and status information [27]. Additionally, a consensus emerged across all participants that the robot's speed was perceived as too slow.

Contrary to our expectations, haptic feedback did not aid BVI users in locating the object. This was mainly due to the haptic feedback system's air flow being too weak to be noticeable. However, when noticed, it was mostly rated as helpful. This can be explained by the fact that the haptic feedback system was an initial prototype, lacking iterative development and improvement in a user-centered design process.

4.5. Relationship between Dependent Variables

We had no specific assumptions in advance about the relationships between the dependent variables in this study. The results should, therefore, be considered exploratory.

On the one hand, our design followed the recommendation of ISO 9241-11:2018 [32] to combine objective and subjective measures with regard to effectiveness and efficiency. ISO 9241-11:2018 [32] also points out that the results of these measures can correspond with each other, but can also differ from each other. This had to be checked for the measures of effectiveness and efficiency used in this study. For effectiveness, findings of objective and subjective measures corresponded to each other in many respects: high overall effectiveness, but more challenges in the BVI group as well as with the narrow objects knife and spanner. This aligns with relationship analysis, indicating a high relationship between both measures. On the contrary, data are not so clear for the efficiency measures. Here, the objective time-related measures showed a common pattern, namely that the indirect handover on the table was faster, whereas for direct midair handover, the BVI group and the objects knife and spanner showed longer times. This pattern was not reflected in the subjective efficiency ratings, which can be explained by their wide variance. However, the relationship analysis revealed a correlation between task completion time and handover time, respectively, with the subjective ratings of the time required. The longer the objective time measure, the less acceptable it was rated, which could indicate that interaction times have an influence on the perception of efficiency.

The various measures of usability and its subcomponents also revealed recurring patterns that cannot be represented by correlation measures, but which contribute to the overall evaluation of the interaction. There was a broad consensus among the participants that the robot took too long to hand over objects. This was reflected in the subjective assessment of efficiency and was also reflected in the qualitative data on satisfaction. Nevertheless, this had little impact on the overall assessment of the participants, which was reflected in an acceptable rating for usability overall and a predominantly good rating for user experience and satisfaction. The users in the BVI group experienced more failures than the sighted group, some of which could negatively affect them. This may also partly explain the greater variation in the subjective assessments of this group. However, this is certainly also due to the greater heterogeneity of the BVI group. Nevertheless, their subjective assessment of the interaction in terms of effectiveness and satisfaction and the user experience was good. Furthermore, across various measures of effectiveness and efficiency, it was shown that the direct midair handover posed more challenges for BVI users than for sighted users and that the direct midair handover posed more problems for the objects knife and spanner.

4.6. Limitations and Future Research

This study is limited to accessibility for BVI users compared to sighted users. As most research on accessibility exists for this user group in general [3], this study could build on extensive previous findings in this research field. However, future studies on the accessibility of robotic handovers should be extended to other user groups with disabilities. Another limitation is its short-term perspective. The few assessed interactions in this study do not allow inferences about the long-term development of user perception of the robotic handover. The level of novelty could have boosted user ratings in a positive way. Therefore, future research should assess long-term use of accessible handover robots in everyday settings. In addition, there remains the exciting question of the extent to which interaction strategies for HRI that meet the needs of users with disabilities can also be helpful for users under temporarily restrictive conditions, such as poor lighting or temporary limitations due to illness.

This study was carried out before the release of the review on robotic object handover by [6], and as such, the subjective and objective performance measures employed here do not align with the minimal set proposed by their review. However, our chosen method systematically aligns with ISO 9241-11:2018 [32] as a widely established and coherent theoretical basis. Accordingly, the objects for handover in this study were selected based on the use cases for the two user groups. It remains to be seen whether the methodology proposed by [6] can prove itself in a similar way, for which its embedding in a coherent theoretical foundation would be helpful. Despite differing in metrics and object selection, we endorse the goal of achieving greater comparability across studies on robotic object handover, whose current challenges also became clear in the discussion of our results. Hence, we recommend future research to reconsider the accessibility of robotic object handover, also incorporating the measures and object classes proposed by [6].

Although conventions for Bayes Factors exist [42], the absolute height of results depends on the priors set. Our assumptions were mainly equivalence hypotheses according to the study objective for equal support for BVI and sighted users. Therefore, the parameter 'rscale' was set to a low value when calculating the Bayes Factors in all analyses of this study (for details, see Section 2.4). In the case of testing null hypotheses, high Bayes Factors are difficult to achieve in general and a higher priors set would have resulted in Bayes Factors showing higher evidence for the null hypothesis. Therefore, conventions for Bayes Factors should be used with caution when interpreting the results of this study.

Moreover, SUS-Scale received lower reliability values than reported in previous research [35]. For the UEQ-S, no reliability values for comparison are available in [40].

However, its subscale for pragmatic quality exhibited very poor reliability. This diminished reliability of both scales may result from an oral presentation of items and the method of answering chosen for this study, which was done for equal accessibility reasons. Future research should investigate how this method of presentation and answering influences the reliability metrics of these generally approved scales. Regarding the measurement of user experience, future research can avoid this problem by using the long version of the UEQ [38], which was not used in this study due to the time constraints imposed by pandemic regulations.

From the results, the potential for improving the handover robot and for future research can be derived. Overall, the time required for the processes was repeatedly rated negatively by all groups. This is especially problematic, as Wizard of Oz was applied in this study for object recognition and object grasping, reducing the task completion time for all participants. There are technical challenges here to speed up various autonomous robot actions. However, the high safety standards applied were also partly responsible for the time required. It is, therefore, an interesting question for future research. To what extent are users willing to accept uncertainty in favor of a faster process? Ethical, legal, and technical criteria will also need to be examined in the future. Integrating object and hand recognition into the robot could prevent incorrect or perceived threatening path planning and avoid cases in which blind people have difficulties locating objects. In this respect, the results on tactile feedback also hold potential for future research. The airflow must be technically amplified to be more perceptible and fulfill the intended orientation function. It also needs to be positioned more precisely in relation to the object. With regard to the Voice User Interface, the accuracy of speech recognition should be technically improved, e.g., to avoid repeated inputs. Further technical improvements are needed in the robotic grasping strategies to prevent narrow objects from falling out of the robot's hand. In addition, in case the object and the gripper consist of the same material, it should be made easier for BVI users to distinguish between the robot hand and the object, e.g., by embossing the material of the gripper in a distinctive way.

5. Conclusions

The presented research contributes to the goal of more accessibility of handover robots. Robots are expected to play a crucial role in aging societies, e.g., through automated object handovers. While extensive research has focused on robotic object handover, the literature addressing the specific needs of BVI users is scarce. This study aimed to address this gap by exploring the usability of robot-to-human handovers involving BVI and sighted users. The evaluation was conducted with a handover robot prototype featuring a Voice User Interface and haptic feedback. The interaction concept developed for robotic object handovers with BVI and sighted users can be rated as successful. The usability ratings reflected this success by meeting the set target across all groups. Almost all handovers were completed successfully. This confirms the implemented grasping strategy with averted or concealed hazardous surfaces. Additionally, the applied VUI seems sufficient to support the handover process. A high user experience was also attested, and the interaction was predominantly perceived as 'pleasant', 'interesting', and 'helpful'. This confirms the process and information content applied in the interaction concept, with specific position details and instructions for action. However, concerns were raised about the time required for handovers, and feedback from BVI participants varied more compared to sighted participants.

The key takeaway is to design the interface of collaborative robots for BVI users as close as possible to the norm. This includes planning object orientation and grasping based on human conventions, coordinating handovers without eye contact using a Voice User Interface, and offering direct and indirect handover options. Providing haptic support enhances orientation during midair handovers. Ensuring that the robot returns to a known base position post-handover prevents safety issues. Additionally, considering the needs of BVI users in interface design can improve interactions for sighted users as well. These findings highlight opportunities and challenges in enhancing assistive robots, making a valuable contribution to more accessible and broader applicable robotic assistance.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/robotics13030043/s1, Video S1: excerpts from handovers to BVI participants.

Author Contributions: Conceptualization, D.L. and F.L.; Data curation, D.L. and F.L.; Formal analysis, D.L. and F.L.; Funding acquisition, A.D. and A.C.B.; Investigation, D.L., F.L. and P.D.; Methodology, D.L., F.L. and P.D.; Project administration, D.L. and F.L.; Resources, D.L., F.L., P.D. and A.C.B.; Supervision, D.L. and F.L.; Validation, D.L. and F.L.; Visualization, F.L.; Writing—original draft, D.L. and F.L.; Writing—review and editing, D.L., F.L., P.D., A.D., S.G. and A.C.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by German Federal Ministry of Education and Research, grant number 16SV7969K. The APC was funded by Chemnitz University of Technology.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to restricted consent provided by participants on the use of personal data pursuant to the General Data Protection Regulation in the European Union and the European Economic Area (EU General Data Protection Regulation).

Acknowledgments: We thank Daniel Wimpff from Sikom Software GmbH for the technical implementation of the speech interface. Additionally, we thank the researchers of the department of Cognitive Human-Machine-Systems at Fraunhofer IWU Chemnitz for the realization of the robotic demonstrator. Furthermore, we thank HFC Human-Factors-Consult GmbH for the cooperation in conceptualization and carrying out the UCD steps during robot development. We appreciate the efforts and valuable input of our participants in this study. We would like to extend a special thanks to those who gave their consent for the publication of images and video material to illustrate the study results.

Conflicts of Interest: Authors Pia Diekmann and Sebastian Glende were employed by the company YOUSE GmbH. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Appendix A. Overview on Descriptive Statistics for all Dependent Variables

Table A1. Descriptive statistics for overall dependent variables of the evaluation study.

Variable	BVI M (SD)	Sighted M (SD)
Overall		
SUS (scale range 0 to 100)	70.25 (16.68)	72.50 (11.47)
Effectiveness		
success rate	94.55%	100%
joint action failure rate	1.81%	0%
robotic failure rate	3.63%	0%
perceived easiness (scale range 1 to 10) [1,2]	7.95 (2.83)	9.25 (1.43)
Efficiency		
Task completion time [1,2] (in minutes)		
—midair	5.75 (0.96)	5.78 (1.10)
—on table	4.58 (0.66)	4.92 (1.03)
Handover time [1,2,3] (in minutes)		
—midair	1.78 (0.71)	1.56 (0.51)
perceived efficiency (scale range 1 to 10) [1,2]	5.93 (2.87)	6.36 (2.39)
Satisfaction/User Experience		
Objective preference for modality (midair vs. table)	81.7% vs. 18.3%	79.3% vs. 20.7%
UEQ-$S_{overall}$ (scale range −3 to 3)	1.79 (0.81)	1.27 (0.67)

[1] measures averaged across either objects and modality or across objects. [2] unsuccessful handovers excluded from measure. [3] no comparison for modality 'on table' as time differences only resulted from technical variance.

References

1. IAPB. International Agency for the Prevention of Blindness's Vision Atlas. Available online: https://www.iapb.org/learn/vision-atlas/magnitude-and-projections/ (accessed on 11 December 2023).
2. Bhowmick, A.; Hazarika, S.M. An insight into assistive technology for the visually impaired and blind people: State-of-the-art and future trends. *J. Multimodal User Interfaces* **2017**, *11*, 149–172. [CrossRef]
3. Mack, K.; McDonnell, E.; Jain, D.; Lu Wang, L.; Froehlich, J.E.; Findlater, L. What Do We Mean by "Accessibility Research"? In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, Yokohama, Japan, 8–13 May 2021; Kitamura, Y., Quigley, A., Isbister, K., Igarashi, T., Bjørn, P., Drucker, S., Eds.; ACM: New York, NY, USA, 2021; pp. 1–18. [CrossRef]
4. Babel, F.; Kraus, J.; Baumann, M. Findings from a Qualitative Field Study with an Autonomous Robot in Public: Exploration of User Reactions and Conflicts. *Int. J. Soc. Robot.* **2022**, *14*, 1625–1655. [CrossRef]
5. Kupcsik, A.; Hsu, D.; Lee, W.S. Learning Dynamic Robot-to-Human Object Handover from Human Feedback. In *Robotics Research*; Bicchi, A., Burgard, W., Eds.; Springer Proceedings in Advanced Robotics; Springer International Publishing: Cham, Switzerland, 2018; pp. 161–176. [CrossRef]
6. Ortenzi, V.; Cosgun, A.; Pardi, T.; Chan, W.P.; Croft, E.; Kulic, D. Object Handovers: A Review for Robotics. *IEEE Trans. Robot.* **2021**, *37*, 1855–1873. [CrossRef]
7. Kim, J.; Park, J.; Hwang, Y.K.; Lee, M. Three handover methods in esteem etiquettes using dual arms and hands of home-service robot. In Proceedings of the ICARA 2004: Proceedings of the Second International Conference on Autonomous Robots and Agents, Palmerston North, New Zealand, 13–15 December 2004; Mukhopadhyay, S.C., Sen Gupta, G., Eds.; Institute of Information Sciences and Technology, Massey University: Palmerston North, New Zealand, 2004; pp. 34–39.
8. Cakmak, M.; Srinivasa, S.S.; Lee, M.K.; Forlizzi, J.; Kiesler, S. Human preferences for robot-human hand-over configurations. In Proceedings of the 2011 IEEE/RSJ International Conference on Intelligent Robots and Systems, San Francisco, CA, USA, 25–30 September 2011; pp. 1986–1993. [CrossRef]
9. Aleotti, J.; Micelli, V.; Caselli, S. An Affordance Sensitive System for Robot to Human Object Handover. *Int. J. Soc. Robot.* **2014**, *6*, 653–666. [CrossRef]
10. Haddadin, S.; Albu-Schaffer, A.; Haddadin, F.; Rosmann, J.; Hirzinger, G. Study on Soft-Tissue Injury in Robotics. *IEEE Robot. Automat. Mag.* **2011**, *18*, 20–34. [CrossRef]
11. Langer, D.; Legler, F.; Krusche, S.; Bdiwi, M.; Palige, S.; Bullinger, A.C. Greif zu—Entwicklung einer Greifstrategie für robotergestützte Objektübergaben mit und ohne Sichtkontakt. *Z. Arb. Wiss.* **2023**, *30*, 297–316. [CrossRef]
12. Choi, Y.S.; Chen, T.; Jain, A.; Anderson, C.; Glass, J.D.; Kemp, C.C. Hand it over or set it down: A user study of object delivery with an assistive mobile manipulator. In Proceedings of the RO-MAN 2009—The 18th IEEE International Symposium on Robot and Human Interactive Communication, Toyama, Japan, 27 September–2 October 2009; pp. 736–743. [CrossRef]
13. Cochet, H.; Guidetti, M. Contribution of Developmental Psychology to the Study of Social Interactions: Some Factors in Play, Joint Attention and Joint Action and Implications for Robotics. *Front. Psychol.* **2018**, *9*, 1992. [CrossRef]
14. Strabala, K.W.; Lee, M.K.; Dragan, A.D.; Forlizzi, J.L.; Srinivasa, S.; Cakmak, M.; Micelli, V. Towards Seamless Human-Robot Handovers. *J. Hum.-Robot. Interact.* **2013**, *2*, 112–132. [CrossRef]
15. Bdiwi, M. Integrated Sensors System for Human Safety during Cooperating with Industrial Robots for Handing-over and Assembling Tasks. *Procedia CIRP* **2014**, *23*, 65–70. [CrossRef]
16. Boucher, J.-D.; Pattacini, U.; Lelong, A.; Bailly, G.; Elisei, F.; Fagel, S.; Dominey, P.F.; Ventre-Dominey, J. I Reach Faster When I See You Look: Gaze Effects in Human-Human and Human-Robot Face-to-Face Cooperation. *Front. Neurorobot.* **2012**, *6*, 3. [CrossRef]
17. Prada, M.; Remazeilles, A.; Koene, A.; Endo, S. Implementation and experimental validation of Dynamic Movement Primitives for object handover. In Proceedings of the 2014 IEEE/RSJ International Conference on Intelligent Robots and Systems, Palmer House Hilton, Chicago, IL, USA, 14–18 September 2014; pp. 2146–2153. [CrossRef]
18. Newell, A.F.; Gregor, P.; Morgan, M.; Pullin, G.; Macaulay, C. User-Sensitive Inclusive Design. *Univ. Access Inf. Soc.* **2011**, *10*, 235–243. [CrossRef]
19. Costa, D.; Duarte, C. Alternative modalities for visually impaired users to control smart TVs. *Multimed. Tools Appl.* **2020**, *79*, 31931–31955. [CrossRef]
20. Da Guia Torres Silva, M.; Albuquerque, E.A.Y.; Goncalves, L.M.G. A Systematic Review on the Application of Educational Robotics to Children with Learning Disability. In Proceedings of the 2022 Latin American Robotics Symposium (LARS), 2022 Brazilian Symposium on Robotics (SBR), and 2022 Workshop on Robotics in Education (WRE), São Bernardo do Campo, Brazil, 18–21 October 2022; pp. 448–453. [CrossRef]
21. Kyrarini, M.; Lygerakis, F.; Rajavenkatanarayanan, A.; Sevastopoulos, C.; Nambiappan, H.R.; Chaitanya, K.K.; Babu, A.R.; Mathew, J.; Makedon, F. A Survey of Robots in Healthcare. *Technologies* **2021**, *9*, 8. [CrossRef]
22. Morgan, A.A.; Abdi, J.; Syed, M.A.Q.; Kohen, G.E.; Barlow, P.; Vizcaychipi, M.P. Robots in Healthcare: A Scoping Review. *Curr. Robot. Rep.* **2022**, *3*, 271–280. [CrossRef]
23. Tsui, K.M.; Dalphond, J.M.; Brooks, D.J.; Medvedev, M.S.; McCann, E.; Allspaw, J.; Kontak, D.; Yanco, H.A. Accessible Human-Robot Interaction for Telepresence Robots: A Case Study. *Paladyn. J. Behav. Robot.* **2015**, *6*, 000010151520150001. [CrossRef]
24. Bonani, M.; Oliveira, R.; Correia, F.; Rodrigues, A.; Guerreiro, T.; Paiva, A. What My Eyes Can't See, A Robot Can Show Me. In Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility, Galway, Ireland, 22–24 October 2018; Hwang, F., Ed.; ACM: New York, NY, USA, 2018; pp. 15–27. [CrossRef]

25. Qbilat, M.; Iglesias, A.; Belpaeme, T. A Proposal of Accessibility Guidelines for Human-Robot Interaction. *Electronics* **2021**, *10*, 561. [CrossRef]
26. Langer, D.; Legler, F.; Kotsch, P.; Dettmann, A.; Bullinger, A.C. I Let Go Now! Towards a Voice-User Interface for Handovers between Robots and Users with Full and Impaired Sight. *Robotics* **2022**, *11*, 112. [CrossRef]
27. Leporini, B.; Buzzi, M. Home Automation for an Independent Living. In Proceedings of the 15th International Web for All Conference, Lyon, France, 23–25 April 2018; ACM: New York, NY, USA, 2018; pp. 1–9. [CrossRef]
28. Walde, P.; Langer, D.; Legler, F.; Goy, A.; Dittrich, F.; Bullinger, A.C. Interaction Strategies for Handing Over Objects to Blind People. In Proceedings of the Annual Meeting of the Human Factors and Ergonomics Society Europe Chapter, Nantes, France, 2–4 October 2019; Available online: https://www.hfes-europe.org/wp-content/uploads/2019/10/Walde2019poster.pdf (accessed on 11 December 2023).
29. Käppler, M.; Deml, B.; Stein, T.; Nagl, J.; Steingrebe, H. The Importance of Feedback for Object Hand-Overs Between Human and Robot. In *Human Interaction, Emerging Technologies and Future Applications III*; Ahram, T., Taiar, R., Langlois, K., Choplin, A., Eds.; Advances in Intelligent Systems and Computing; Springer International Publishing: Cham, Switzerland, 2021; pp. 29–35. [CrossRef]
30. Brulé, E.; Tomlinson, B.J.; Metatla, O.; Jouffrais, C.; Serrano, M. Review of Quantitative Empirical Evaluations of Technology for People with Visual Impairments. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, Honolulu, HI, USA, 25–30 April 2020; Bernhaupt, R., Mueller, F., Verweij, D., Andres, J., McGrenere, J., Cockburn, A., Avellino, I., Goguey, A., Bjørn, P., Zhao, S., et al., Eds.; ACM: New York, NY, USA, 2020; pp. 1–14. [CrossRef]
31. Nelles, J.; Kwee-Meier, S.T.; Mertens, A. Evaluation Metrics Regarding Human Well-Being and System Performance in Human-Robot Interaction—A Literature Review. In Proceedings of the 20th Congress of the International Ergonomics Association (IEA 2018), Florence, Italy, 26–30 August 2018; Bagnara, S., Tartaglia, R., Albolino, S., Alexander, T., Fujita, Y., Eds.; Advances in Intelligent Systems and Computing. Springer International Publishing: Cham, Switzerland, 2019; pp. 124–135. [CrossRef]
32. *ISO_9241-11:2018*; ISO International Organization for Standardization. DIN EN ISO 9241-11:2018-11, Ergonomie der Mensch-System-Interaktion_-Teil_11: Gebrauchstauglichkeit: Begriffe und Konzepte (ISO_9241-11:2018). Deutsche Fassung EN_ISO_9241-11:2018; Beuth Verlag GmbH: Berlin, Germany, 2018.
33. Brooke, J. SUS: A 'Quick and Dirty' Usability Scale. In *Usability Evaluation in Industry*; Jordan, P.W., Thomas, B., Weerdmeester, B.A., McClelland, I.L., Eds.; Taylor and Francis: Bristol, PA, USA; London, UK, 1996; pp. 189–194.
34. Bangor, A.; Kortum, P.T.; Miller, J.T. An Empirical Evaluation of the System Usability Scale. *Int. J. Hum.-Comput. Interact.* **2008**, *24*, 574–594. [CrossRef]
35. Lewis, J.R. The System Usability Scale: Past, Present, and Future. *Int. J. Hum.-Comput. Interact.* **2018**, *34*, 577–590. [CrossRef]
36. *ISO_9241-210:2019*; ISO International Organization for Standardization. DIN EN ISO 9241-210:2020-03, Ergonomie der Mensch-System-Interaktion_-Teil_210: Menschzentrierte Gestaltung Interaktiver Systeme (ISO_9241-210:2019). Deutsche Fassung EN_ISO_9241-210:2019; Beuth Verlag GmbH: Berlin, Germany, 2019.
37. Hassenzahl, M. The Effect of Perceived Hedonic Quality on Product Appealingness. *Int. J. Hum.-Comput. Interact.* **2001**, *13*, 481–499. [CrossRef]
38. Laugwitz, B.; Held, T.; Schrepp, M. Construction and Evaluation of a User Experience Questionnaire. In *HCI and Usability for Education and Work*; Holzinger, A., Ed.; Lecture Notes in Computer Science; Springer Berlin Heidelberg: Berlin/Heidelberg, Germany, 2008; pp. 63–76. [CrossRef]
39. Universal Robots. Die Zukunft ist Kollaborierend. 2016. Available online: https://www.universal-robots.com/de/download-center/#/cb-series/ur10 (accessed on 8 December 2021).
40. Schrepp, M.; Hinderks, A.; Thomaschewski, J. Design and Evaluation of a Short Version of the User Experience Questionnaire (UEQ-S). *Int. J. Interact. Multimed. Artif. Intell.* **2017**, *4*, 103. [CrossRef]
41. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2022; Available online: https://www.R-project.org/ (accessed on 19 September 2022).
42. Morey, R.D.; Rouder, J.N. BayesFactor: Computation of Bayes Factors for Common Designs. R Package Version 0.9.12-4.2. 2018. Available online: https://CRAN.R-project.org/package=BayesFactor (accessed on 19 September 2022).
43. Stekhoven, D.J.; Bühlmann, P. MissForest–non-parametric missing value imputation for mixed-type data. *Bioinformatics* **2012**, *28*, 112–118. [CrossRef] [PubMed]
44. Platias, C.; Petasis, G. A Comparison of Machine Learning Methods for Data Imputation. In Proceedings of the 11th Hellenic Conference on Artificial Intelligence, Athens, Greece, 2–4 September 2020; Spyropoulos, C., Varlamis, I., Androutsopoulos, I., Malakasiotis, P., Eds.; ACM: New York, NY, USA, 2020; pp. 150–159. [CrossRef]
45. Stekhoven, D.J. missForest: Nonparametric Missing Value Imputation using Random Forest. R Package Version 1.5 (2022-04-14). Available online: https://cran.r-project.org/web/packages/missForest/index.html (accessed on 19 September 2022).