
**Deliverable D5.4: Report on performance metrics and final evaluation study**

Release date: 26.08.2016
Status: Public

Document name: MOBOT_WP5_D5.4_final version
Executive Summary

The deliverable D5.4 documents the activities and work performed within Work Package 5 of the MOBOT project. It comprises the report of the final evaluation studies on the MOBOT prototypes which were conducted for three weeks at the BETHANIEN-hospital/Geriatric Centre at the University of Heidelberg in Germany and for two weeks at the rehabilitation centre DIAPLASIS in Greek.

The D5.4 is separated into two parts: The first part describes the study performed at BETHANIEN from 22nd April to 12th May 2016, and the second part the study performed at DIAPLASIS from 15th to 27th July.

In the first part, we provide details on the subjective recruitment procedure, different test scenarios to evaluate specific functionalities of the MOBOT rollator-type mobility assistant, adequate quantitative and qualitative performance measures, and we present preliminary results of the evaluation study performed at BETHANIEN. The evaluation study at BETHANIEN was carried out with 42 participants who met predefined inclusion criteria and who were recruited from geriatric rehab/acute care clinic of BETHANIEN, hospital-associated nursing homes or rehab sports club. The aim of this evaluation study was to validate the cognitive (navigation) assistance system and the audio-gestural human-robot communication. According to these functionalities, two specific test scenarios with “tailored” assessment methods and qualitative as well as quantitative performance metrics were developed. Major parts of performance metrics are based on technical data derived from the data flow of integrated systems on the MOBOT rollator-type mobility assistant. As these data are currently still being post-processed by technical partners, only a limited number of results can be presented in this report. Preliminary results from the clinical partner BETHANIEN on the users’ navigation performance and on the subjective user satisfaction with the MOBOT cognitive assistance system are quite promising, especially in cognitively impaired people. It is planned to present results on the performance metrics obtained from post-processed technical data in later publications, which will be produced also after the end of the project in collaboration between the technical and clinical partners.

The second part describes the validation study that took place in DIAPLASIS at the end of the project. The two different scenarios examined using the MOBOT rollator type device are described in detail. An analysis of the qualitative and quantitative measures used during the evaluation is also reported.

The final evaluation of the MOBOT rollator type device conducted in DIAPLASIS with 30 patients. All patients met the inclusion criteria defined at the beginning of the project. An extended recruitment phase was carried out in DIAPLASIS using both inpatients, outpatients of the DIAPLASIS rehabilitation hospital as well as patients from other collaborating facilities (e.g. geriatric centres). A total number of 425 patients was initially screened and interviewed in order to recruit the 30 patients participated in the evaluation study. Two scenarios were tested: (i) the cognitive assistance functionality, and (ii) the audio-gestural interaction. The diversification of the pathologies of the subjects participated in the study can be considered as a unique advantage. Moreover, the development of a new psychometric scale called “PYTHEI A”, able to measure the subjective satisfaction of the end users of the used assistive robotic device is also another major advantage and outcome of the project. According to a review conducted there is no such scale available till now. The results from the final evaluation in DIAPLASIS are still in post processing and thus only some descriptive results are reported in the current deliverable. As soon as the technical partners deliver the processed collected measurements we will publish an in depth analysis of the evaluation phase conducted in DIAPLASIS with the MOBOT rollator-type device.
**Deliverable Identification Sheet**

<table>
<thead>
<tr>
<th>IST Project No.</th>
<th>FP7 – ICT for Cognitive Systems and Robotics – 600796</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acronym</td>
<td>MOBOT</td>
</tr>
<tr>
<td>Full title</td>
<td>Intelligent Active MObility Assistance RoBOT integrating Multimodal Sensory Processing, Proactive Autonomy and Adaptive Interaction</td>
</tr>
<tr>
<td>Project URL</td>
<td><a href="http://www.mobot-project.eu">http://www.mobot-project.eu</a></td>
</tr>
<tr>
<td>EU Project Officer</td>
<td>Michel Brochard</td>
</tr>
</tbody>
</table>

**Deliverable** | D5.4 Report on performance metrics and final evaluation study  
**Work package** | WP5 Specification of Use Cases, Performance Metrics, and User Evaluation Studies  

<table>
<thead>
<tr>
<th>Date of delivery</th>
<th>Contractual</th>
<th>M42</th>
<th>Actual</th>
<th>26-08-2016</th>
</tr>
</thead>
<tbody>
<tr>
<td>Status</td>
<td>Final version</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nature</td>
<td>Report</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dissemination Level</td>
<td>Public</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Authors (Partner)**

Klaus Hauer (BETHANIEN), Christian Werner (BETHANIEN), Yiannis Koumpouros (DIAPLASIS), Costas Tzafestas, George Mountris, Petros Maragos, Vassilis Pitsikalis, Nikos Kardaris and Isidoros Rdomagoulakis (ICCS)

**Responsible Author**

Klaus Hauer  
Email: khauer@bethanien-heidelberg.de  
Partner: BETHANIEN  
Phone: +49-6221-319 1532

**Keywords**

final user evaluation study, subject recruitment, scenarios, results

**Version Log**

<table>
<thead>
<tr>
<th>Issue Date</th>
<th>Rev No.</th>
<th>Authors</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>19-07-2016</td>
<td>1.0</td>
<td>Klaus Hauer, Christian Werner</td>
<td>First draft</td>
</tr>
<tr>
<td>20-07-2016</td>
<td>1.1</td>
<td>Costas Tzafestas</td>
<td>Added ICCS input on section 3</td>
</tr>
<tr>
<td>28-07-2016</td>
<td>1.2</td>
<td>George Mountris, Petros Maragos, Vassilis Pitsikalis, Nikos Kardaris and Isidoros Rdomagoulakis</td>
<td>Added ICCS input on section 4</td>
</tr>
<tr>
<td>05/09-08-2016</td>
<td>1.3</td>
<td>Yiannis Koumpouros</td>
<td>Added DIAPLASIS input on part II</td>
</tr>
<tr>
<td>22/26-08-2016</td>
<td>1.4</td>
<td>Klaus Hauer, Christian Werner</td>
<td>Integration of corrections after internal review</td>
</tr>
</tbody>
</table>
TABLE OF CONTENTS

INTRODUCTION........................................................................................................................................................................... 7

I  FINAL EVALUATION STUDY AT BETHANIEN ............................................................................................................................ 8

1  Subject recruitment (BETHANIEN) .............................................................................................................................................. 8
   1.1  Recruitment strategy ................................................................................................................................................................. 8
   1.2  Inclusion criteria ........................................................................................................................................................................... 9
   1.3  Proceeding .................................................................................................................................................................................. 9
   1.4  Results .......................................................................................................................................................................................... 9
       1.4.1  Recruitment ........................................................................................................................................................................ 9
       1.4.2  Sample description .......................................................................................................................................................... 12

2  Scenario 1: Cognitive assistance – navigation trail (BETHANIEN, ICCS) .................................................. 13
   2.1  Background ................................................................................................................................................................................ 13
   2.2  Design and experimental setup .................................................................................................................................................. 13
   2.3  Performance metrics ................................................................................................................................................................. 15
   2.4  Statistical analysis ....................................................................................................................................................................... 16
   2.5  Results ........................................................................................................................................................................................ 17
       2.5.1  Success rate ...................................................................................................................................................................... 17
       2.5.2  Task completion time ......................................................................................................................................................... 17
       2.5.3  Subjective user satisfaction ............................................................................................................................................... 20

3  Scenario 2: Audio-gestural human-robot communication (ICCS) .............................................................................. 21
   3.1  Background ................................................................................................................................................................................ 21
   3.2  Design and experimental setup .................................................................................................................................................. 22
       3.2.1  Audio mode ........................................................................................................................................................................ 22
       3.2.2  Gestural mode ....................................................................................................................................................................... 22
       3.2.3  Audio-gestural mode .......................................................................................................................................................... 23
       3.2.4  “Self-chosen/-free” mode .................................................................................................................................................. 23
   3.3  Performance metrics ................................................................................................................................................................. 24
   3.4  Results ........................................................................................................................................................................................ 24

4  Summary (BETHANIEN, ICCS) ................................................................................................................................................... 25

II  FINAL EVALUATION STUDY AT DIAPLASIS (DIAPLASIS) ........................................................................................... 27

1  Subject recruitment ...................................................................................................................................................................... 27
   1.1  Recruitment strategy ................................................................................................................................................................. 27
   1.2  Results ........................................................................................................................................................................................ 30
       1.2.1  Recruitment ...................................................................................................................................................................... 30
       1.2.2  Sample description .......................................................................................................................................................... 30
2 Scenario 1: Cognitive assistance ................................................................. 31
  2.1 Design and experimental setup ............................................................ 31
  2.2 Performance metrics ........................................................................... 33
3 Scenario 2: Audio-gestural human-robot communication ....................... 34
  3.1 Design and experimental setup ............................................................ 34
    3.1.1 Audio mode .................................................................................. 34
    3.1.3 Audio-gestural mode ................................................................. 35
    3.1.4 Free mode .................................................................................. 36
  3.2 Performance metrics ........................................................................... 36
  3.3 Subjective evaluation ........................................................................... 36
4 Summary .................................................................................................. 37

MAIN SUMMARY AND CONCLUSION .......................................................... 38

References ................................................................................................. 39
### LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>WP</td>
<td>Work Package</td>
</tr>
<tr>
<td>REGE (e.V.)</td>
<td>REhabilitation in GERiatric Medicine (registered association)</td>
</tr>
<tr>
<td>MMSE</td>
<td>Mini-Mental Status Examination</td>
</tr>
<tr>
<td>POMA</td>
<td>Performance Oriented Mobility Assessment</td>
</tr>
<tr>
<td>SD</td>
<td>Standard Deviation</td>
</tr>
<tr>
<td>GDS</td>
<td>Geriatric Depression Scale</td>
</tr>
<tr>
<td>SF-12</td>
<td>12-item Short Form Health Survey</td>
</tr>
<tr>
<td>FES-I</td>
<td>Standard Deviation</td>
</tr>
<tr>
<td>TSQ-WT</td>
<td>Telehealthcare Satisfaction Questionnaire – Wearable Technology</td>
</tr>
<tr>
<td>ANOVA</td>
<td>Analysis of Variance</td>
</tr>
<tr>
<td>$\eta^2$</td>
<td>Effect size eta-squared</td>
</tr>
<tr>
<td>$r_s$</td>
<td>Spearman’s rank correlation coefficients</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Partner Abb.</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BETHANIEN</td>
<td>BETHANIEN KRANKENHAUS - GERIATRISCHES ZENTRUM - GEMEINNUTZIGE GMBH</td>
</tr>
<tr>
<td>DIAPLASIS</td>
<td>DIAPLASIS REHABILITATION CENTER SA</td>
</tr>
<tr>
<td>ICCS</td>
<td>INSTITUTE OF COMMUNICATION AND COMPUTER SYSTEMS</td>
</tr>
</tbody>
</table>
INTRODUCTION

A major task of the clinical partners BETHANIEN and DIAPLASIS in WP5 is to validate the different prototypes as developed during the MOBOT project from a clinical as well as technical perspective including adequate clinical as well as technical expertise. In this report, we document planning, organisation, and results of the final user evaluation studies of the final prototype of the rollator-type mobility assistant.

The evaluation of this prototype was based on two systematic reviews conducted during the MOBOT project with focus for both the technical perspective (Geravand et al. “Mobility Assistance Robots: A Survey”, submitted to Journal of Intelligent and Robotic Systems) as well as the clinical perspective (Werner et al. (2016) “Evaluation Studies of Robotic Rollators by the User Perspective: A Systematic Review”, Gerontology, Epub ahead of print [1]). In addition, we were able to draw experiences gained from the successful first evaluation study performed at BETHANIEN with the mid-term prototype of the rollator-type mobility assistant. This allowed us to develop an advanced, multifaceted assessment strategy with performance metrics specifically tailored to the functionalities to be tested.

Previous evaluation studies of robotic rollators often lack adequate objective and task-specific assessment strategies. A solution for this substantial limitation could be based on the use of the performance metrics derived from the technical data flow of the robot-integrated technical systems. Such a superior “tailored” assessment was so far performed only in a very limited number of studies [1] . To achieve such a tailored, high-tech assessment within the MOBOT project, a task force was installed within the project with Dr. Costas Tzafestas (ICCS) as the leading technical partner and Prof. Dr. Klaus Hauer (BETHANIEN) as the leading clinical partner.

The major part of the performance metrics in a tailored assessment strategy was derived from the MOBOT-integrated technical systems. For the data thus obtained, comprehensive, time consuming, post-processing analysis is mandatory. As these data are currently still being processed by technical partners, this report is mainly based on simple quantifiable outcomes (e.g. success rate, task completion time). The completion of the analysis of more technique-based performance metrics will be postponed to later data exploitation for publication.
I FINAL EVALUATION STUDY AT BETHANIEN

For the series of evaluation studies as projected in the BETHANIEN-hospital/Geriatric Centre of University of Heidelberg a positive vote from the ethical committee of the Medical Faculty of Heidelberg (Ethikkommission I der Medizinischen Fakultät Heidelberg) was achieved covering previous assessments and both projected evaluation studies of the prototypes of the rollator-type mobility assistant (proposal number S-358/2013).

1 Subject recruitment (BETHANIEN)

1.1 Recruitment strategy

In contrary to other clinical studies, the participation in evaluation measures, whose focus lies primarily on the technical verification of the MOBOT functionalities, was not associated with direct incentives for study participants such as improved physical function when taking part in training programs of intervention studies or receiving detailed diagnostic results in studies using advanced diagnostic tools.

Participation was strenuous for the frail multi-morbid target sample, with the risk to overtax participants by study tasks and associated risk of drop out. No other incentives such as a financial honorarium were available to motivate participants.

In the project’s DOW, a sample size of approximately 30 participants was planned for the final user evaluation study of the rollator-type mobility assistant. The main scenario of this study (cognitive assistance on a navigation trail) covered a between-subject study design with two different conditions (2x2: cognitively intact vs. impaired, assisted vs. unassisted navigation). Multiple trials for one subject (within-subject study design) with two different conditions was, however, not possible as the scenario required the lack of knowledge about the navigation trail for both conditions. To have a sufficient sample size for the statistical analysis of this scenario, the recruitment goal was raised to 40 participants. Given the limited duration of three weeks for this last validation, the successful recruitment of 40 participants and the organization of this study was a challenge.

BETHANIEN developed the following strategies to increase the number of participants and allow a sample size of 40 participants:

1. Increasing duration of recruitment period: Recruitment settings were contacted not only during the assessment period but also 4 weeks before. By this strategy the pool of potential participants could be increased and a longer period of time was available for screening and recruitment purpose.

2. Recruitment at different settings. As the target sample of potential users of the MOBOT project was described as patients in geriatric rehab and institutionalized, recruitment mainly takes place in geriatric rehab wards and hospital-associated nursing homes. However, we also screened members of an associated sports club for geriatric rehabilitation (REGE e.V.) which includes former patients of geriatric ward rehabilitation, now taking part in post-ward rehab. In addition, the recruitment was extended from rehab wards also to the acute care clinic in order to recruit also “subsequent” rehab patients.

3. Inclusion criteria: To increase the success rate of recruitment we used the general clinical criteria: “current use of rollator” rather than the impairment defined inclusion criteria regarding motor performance (“habitual gait speed ≤0.6m/s”). The clinical criteria “use of rollator” allowed inclusion of persons representing the predefined, impaired target sample of MOBOT who actually used a supportive device as tested in the evaluation study.
1.2 Inclusion criteria

For recruitment purpose we used inclusion criteria based on the predefined target group of the MOBOT rollator-type mobility assistant as previously documented in WP5 reports. The specific inclusion criteria as well as their detailed design and analysis can be found in deliverable D5.1.

For clinical inclusion criteria we used:

- Current use of rollator and
- No or mild to moderate cognitive impairment (MMSE score 17-26) in addition to following general criteria:
  - Age ≥ 65 yrs. and
  - Written informed consent and
  - Ability to perform assessments as requested in the study and
  - Patient of geriatric rehab within hospital or
  - Resident of hospital associated nursing homes
  - Member of rehab sport club or previous patients of hospital with residence in vicinity of study centre (≤ 15km)

1.3 Proceeding

Potential participants were screened according to predefined inclusion criteria. The screening process included a personal interview, several screening tests, reviewing the patient’s medical charts, documenting of descriptive data, and contacting care- or therapeutic personnel and relatives/proxies. Persons who met inclusion criteria were informed on the assessment and legal consequences and were asked to provide their written-informed-consent. In most cases repeated personal contacts with potential participants and their relatives were mandatory. After giving consent, patients were scheduled for assessments in coordination with assessors. Participants which were not acute or rehab patients of the hospital were transported with a transportation service to assessments at the hospital and back home.

1.4 Results

1.4.1 Recruitment

A total number of 616 persons were screened: 270 nursing home residents, 201 patients of the rehab/acute care wards, and 145 members of the rehabilitation sports club. 488 of these persons did not meet the predefined inclusion criteria with 135 patients of the rehab/acute care wards, 125 rehab sport club members and 228 nursing home residents being excluded in the different settings. Main exclusion criteria for rehab/acute care wards were medical contraindications (n = 47) and patients’ inability to walk (with rollator) (n = 46). For rehab patients, which were already dismissed at the start of the user evaluation study, also the place of residence represented a relevant exclusion criterion (n = 13). Other rehab/acute patients had to be excluded due to severe cognitive impairment (n = 12). For members of the rehabilitation sports club, who generally have a higher motor-functional performance level than rehab /acute patients, the lack of the habitual use of a rollator (documenting the high motor-functional level) was the only relevant exclusion criteria. The main exclusion criterion for nursing home residents were related to their walking ability: 91 residents did not use a rollator because they were still quite confident in unassisted walking, whereas 64 were unable to walk with a rollator (e.g. wheelchair use). Another 60 nursing home residents were excluded due to severe cognitive impairment (see Fig. 1).
Out of the 128 persons who met the predefined inclusion criteria, 72 persons refused the study participation and 12 patients of rehab/acute care wards were excluded from the ongoing recruitment process due to medical reasons (e.g. unexpected deterioration in health, acute illnesses). Written informed consent was obtained from 44 persons. Two potential study participants assigned to the group without navigation assistance dropped out before testing due to acute illness and to a falling accident, respectively. The navigation scenario was finally performed with 42 participants; 22 persons in the group with navigation assistance and 20 persons in the group without navigation assistance.

For one participant of the assisted group, the MOBOT’s navigation assistance system broke down during the second part of the navigation trail (chapel to admission centre) due to technical problems. Thus, the assisted group for the second part of the navigation scenario comprised 21 persons.

In agreement with the MOBOT consortium, the scenario for evaluating the audio-gestural human-robot interaction (HRI) was performed with a reduced number of 20 persons.
Fig. 1. Flow chart of recruitment, screening, enrolment, allocation, and data analysis.
1.4.2 Sample description

The study sample comprised the user group of frail older persons with functional impairments for the MOBOT rollator-type mobility assistant as described in previous WP5 deliverables. Study participants were all acute or previous geriatric patients of the rehab/acute care wards, nursing home residents, or current members of the rehabilitation sports club at BETHANIEN. The sample is defined by advanced age (82.5 ± 8.7 yrs.), frailty as expressed by impaired motor status (habitual gait speed = 0.56 ± 0.22 m/s, POMA score = 19.2 ± 5.9), and high risk of falling (59.5 % reported one or more falls in the last year). Five participants (11.9 %) met criteria for depression (GDS score > 5) according to the results of the GDS screening test for depression, around half of the participants (n = 20, 47.6 %) met criteria for cognitive impairment (MMSE score ≤ 26), and two third (n = 28, 66.7 %) reported fear of falling. One half of the participants (n = 22, 52.4 %) was still living independently at home, while the other half (n = 20, 47.6 %) was institutionalized.

Self-reported data concerning participants’ functional status (Barthel-Index) seem less valid in the sample including also patients with cognitive impairment, when the relatively high Barthel-Index (86.4 ± 15.0; documented by self-report in 76.2 % of participants, based on expert rating only available for acute geriatric in-patients) is compared to more impaired motor status as assessed by more objective performance-based/oriented measures (gait speed, POMA).

For the cognitive assistance scenario, study participants were randomly allocated into two groups: with vs. without navigation assistance matched for cognitive status (MMSE score). The subgroups did not differ with respect to age, functional status, depressive symptoms, health-related quality of life, history of falls, and fear of falling, living situation, and motor performance. The group with navigation assistance comprised more females (81.2 %) than the unassisted group (50.0%) (see Tab. 1).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Total sample (n = 42)</th>
<th>Assisted navigation (n = 22)</th>
<th>Unassisted navigation (n = 20)</th>
<th>p-valuea</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, years, mean (SD)</td>
<td>82.5 (8.7)</td>
<td>84.1 (7.7)</td>
<td>80.7 (9.5)</td>
<td>0.204</td>
</tr>
<tr>
<td>Female, n, (%)</td>
<td>28 (66.7)</td>
<td>18 (81.2)</td>
<td>10 (50.0)</td>
<td>0.029</td>
</tr>
<tr>
<td>MMSE, score, mean (SD)</td>
<td>25.9 (3.6)</td>
<td>25.9 (4.0)</td>
<td>25.9 (3.3)</td>
<td>0.958</td>
</tr>
<tr>
<td>Barthel-Index, score, mean (SD)</td>
<td>86.4 (15.0)</td>
<td>85.2 (12.6)</td>
<td>87.8 (11.5)</td>
<td>0.512</td>
</tr>
<tr>
<td>GDS, score, median (range)</td>
<td>2.0 (0-11)</td>
<td>2.0 (0-11)</td>
<td>1.5 (0-9)</td>
<td>0.912</td>
</tr>
<tr>
<td>SF-12, score, mean (SD)</td>
<td>40.4 (9.4)</td>
<td>39.8 (9.4)</td>
<td>41.2 (9.6)</td>
<td>0.652</td>
</tr>
<tr>
<td>Mental component</td>
<td>52.8 (6.8)</td>
<td>53.3 (6.2)</td>
<td>52.2 (7.5)</td>
<td>0.596</td>
</tr>
<tr>
<td>Recent history of falls, n (%)</td>
<td>25 (59.5)</td>
<td>14 (63.6)</td>
<td>11 (55.0)</td>
<td>0.707</td>
</tr>
<tr>
<td>Fear of Falling, n (%)</td>
<td>10.0 (7-20)</td>
<td>9.5 (7-20)</td>
<td>10.0 (7-17)</td>
<td>0.731</td>
</tr>
<tr>
<td>Living Situation, n (%)</td>
<td>28 (66.7)</td>
<td>16 (72.7)</td>
<td>12 (60.0)</td>
<td>0.677</td>
</tr>
<tr>
<td>Community-dwelling</td>
<td>22 (52.4)</td>
<td>12</td>
<td>10</td>
<td>0.768</td>
</tr>
<tr>
<td>Institutionalized</td>
<td>20 (47.6)</td>
<td>10</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>POMA, score, mean (SD)</td>
<td>19.2 (5.9)</td>
<td>19.5 (6.4)</td>
<td>18.9 (5.6)</td>
<td>0.766</td>
</tr>
<tr>
<td>Habitual gait speed, m/s (SD)</td>
<td>0.56 (0.22)</td>
<td>0.54 (0.22)</td>
<td>0.60 (0.21)</td>
<td>0.381</td>
</tr>
</tbody>
</table>

Note. n = number of cases; MMSE = Mini-Mental State Examination; GDS = Geriatric Depression Scale, 15 items; SF-12 = 12-item Short Form Health Survey; FES-I = Falls Efficacy Scale – International, 7-item version; POMA = Performance Oriented Mobility Assessment.

† Based on interview-led self-reports; †† no vs. some to very much fear of falling; ††† Assessed by 4-Meter-Walk-Test.

*p-values are given for t-tests (age, MMSE, Barthel-Index, SF-12, POMA, habitual gait speed), Chi-square tests (gender, recent history of falls, fear of falling, living situation) and Mann-Whitney-U tests (GDS, FES-I) applied to test for differences between assisted and unassisted group.
2 Scenario 1: Cognitive assistance – navigation trail (BETHANIEN, ICCS)

2.1 Background

The first scenario implemented in this validation study aimed at assessing the performance of an audial cognitive assistance mode of operation. This mode is supported by a subset of the autonomous robot navigation modules that have been previously developed and successfully integrated on the MOBOT rollator-type mobility platform, including mapping, localisation and path planning (technical details are provided in other MOBOT deliverables: see D3.2 and D3.3). This audial cognitive assistance module operates by providing pre-specified audio cues to the user, essentially associated to navigation instructions, while walking with the rollator. The module assumes a known (previously captured) map of the indoor environment along with a set of pre-set guard points (2D positions) that have audio tokens associated with them. The cognitive assistance module was also augmented to accommodate loops and overlapping guard points within higher-level paths integrating multiple intermediate target points. This module is designed to help the user move around and orient him/herself inside an indoor environment, by providing localisation information as well as by giving assistive directions and issuing navigation instructions in the form of audial messages. The research hypothesis to be tested with this scenario is related to the effect that such a navigation assistance modality can have with respect to the cognitive status of the user (that is, for cognitively impaired vs. cognitively intact subjects).

2.2 Design and experimental setup

A 2x2 study design was used comparing two groups (cognitively intact [MMSE >26] vs. impaired [MMSE ≤26]) and two conditions of different navigation assistance levels (assisted vs. unassisted navigation).

In cooperation with the technical partner ICCS, BETHANIEN designed a navigation trail at the ground floor of the BETHANIEN-hospital. This navigation trail was divided in two sections. The first section led from the starting position in front of the elevator of the acute care clinic along a corridor and through the main entrance hall to the first target position in the hospital’s chapel (shortest distance: about 45 meter). The second section led from the chapel through the main entrance hall and along a corridor to the final target position at the hospital’s admission centre (shortest distance: about 55 meter) (see Fig 2).

![Fig. 2: Layout of the navigation trail at the ground floor of the BETHANIEN-hospital](image-url)
Both groups performed the navigation trial with the MOBOT rollator-type mobility assistant. Participants allocated to the group with navigation assistance were, however, assisted by the MOBOT’s navigation assistance system which provided audio cues at critical waypoints during both sections of the navigation trail. They were instructed to reach the target positions (chapel/admission centre) by following the navigational audio cues as provided by the MOBOT. The navigation assistance system provided the appropriate audio cue as soon as they entered a critical area. Each audio cue was repeated every 3 seconds until the participant has left the critical area. The audio cues and their corresponding positions are presented in Table 2 and Figure 3, respectively.

Tab. 2: Audio cues provided to the group with navigation assistance (in consecutive order) on both sections

<table>
<thead>
<tr>
<th>Section 1: starting position → chapel</th>
<th>Section 2: chapel → admission centre</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0: “Walk straight ahead along the corridor.”</td>
<td>2.0: “Leave the chapel.”</td>
</tr>
<tr>
<td>1.1: “Keep straight on.”</td>
<td>2.1: “Turn right and walk into the entrance hall.”</td>
</tr>
<tr>
<td>1.2: “Keep straight on.”</td>
<td>2.2: “Turn diagonally to the right and walk to the end of the entrance hall.”</td>
</tr>
<tr>
<td>1.3: “Turn diagonally to the right and walk through the entrance hall.”</td>
<td>2.3: “Turn left and walk straight ahead along the corridor.”</td>
</tr>
<tr>
<td>1.4: “Turn left and walk towards the elevator.”</td>
<td>2.4: “Turn right and walk straight ahead.”</td>
</tr>
<tr>
<td>1.5: “Turn left in front of the elevator and walk into the chapel.”</td>
<td>2.5: “Keep straight on.”</td>
</tr>
<tr>
<td>1.6: “You have achieved your goal.”</td>
<td>2.6: “You have achieved your goal.”</td>
</tr>
</tbody>
</table>

Fig. 3: Position of the audio cues provided to the group with navigation assistance on the two sections of the navigation trail. Blue circles represent audio cues to the chapel (section 1), and purple circles audio cues to the admission centre (section 2). The size of circles stands for the area in which the audio cues were provided.

In contrast, the unassisted group performed the navigation trail without assistance of the MOBOT’s navigation assistance system. They were asked to walk from the starting position to the target position (section 1: in front of the elevator to chapel; section 2: chapel to admission centre) solely by using the classical signs at the walls or ceilings of the hospital for orientation. Asking other people for help was not allowed. The position of each sign was presented in Figure 4.
Irrespective of group allocation, no instructions were given on participants’ walking. Each participant could perform the navigation trail by his/her self-selected walking speed. After reaching the first target position (chapel), a break was provided before starting the second section of the navigation trail. For section 2, both groups received exactly the same instructions, except for the next target to be reached (admission centre).

All participants were strictly accompanied and supervised by the test administer during the navigation trail in order to ensure participants’ safety and to have the opportunity to intervene in critical or unexpected situations as quickly as possible.

2.3 Performance metrics

Initial performance metrics of this scenario comprised:

- success rate
- task completion time recorded by stopwatch
- subjective user satisfaction assessed by questionnaire

More technique-based objective performance metrics (e.g. walking distance, walking speed, gait parameters, interruption of walk etc.) captured by the robot-integrated technical systems (e.g. laser range finder, odometry sensors) are currently still being post-processed by technical partners (ICCS). Results on these performance metrics could therefore not be presented in this report but will be part of later publications.

Telehealthcare Satisfaction Questionnaire – Wearable Technology

After completion of the navigation trail, the subjective user satisfaction with the navigation assistance system was evaluated by using the Telehealthcare Satisfaction Questionnaire – Wearable Technology (TSQ-WT). The TSQ-WT was developed by the Robert-Bosch Hospital Stuttgart, Department of Clinical Gerontology and has already been used in other EU-projects (FARSEEING – FAll Repository for the design of Smart and sElf-adaptive Environments prolonging INdependent livinG). The paper including validation results is being currently prepared by the RBK Stuttgart for submission. The questionnaire was customized for application with robotic systems allowing the assessment of user satisfaction with such systems at five different dimensions (benefit, usability, self-concept, privacy & loss
of control, quality of life) [2]. In addition, the TSQ-WT can be adapted to specific functionalities of robotic systems and thus enables an individually tailored subjective assessment of different functionalities (see Tab. 3). The dimension “privacy & loss of control” focus on long-term use of robotic systems. As the MOBOT rollator-type mobility assistant was evaluated during a specific test scenario, we only used four dimensions in this evaluation study (benefit, usability, self-concept, quality of life). Each of the dimensions contained five questions. Rating was scaled from 0 to 4, which entailed a maximum score of 20 for each dimension. Higher scores indicate a more positive rating.

**Tab. 3:** TSQ-WT adapted to the navigation assistance system of the MOBOT rollator

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Content</th>
</tr>
</thead>
</table>
| Benefit       | 1. I can benefit from this navigation assistance system*.  
2. The effort of using this navigation assistance system.  
3. I am confident I’m getting the most out of this navigation assistance system.  
4. This navigation assistance system is helping me to reach my goal.  
5. I would recommend this rollator with navigation assistance system to other people in my situation. |
| Usability     | 1. The use of this rollator with navigation assistance system requires effort.  
2. The navigation assistance system is reliable according to my estimation and experience so far.  
3. This rollator with navigation assistance system is easy to use.  
4. I feel safe when using this rollator with navigation assistance system.  
5. I feel good while using this rollator with navigation assistance system. |
| Self-concept  | 1. The use of this rollator with navigation assistance system is an interesting challenge for me.  
2. This navigation assistance system reminds me of losing my independence.  
3. The use of this navigation assistance system is making me feel older than I am.  
4. I (would) feel embarrassed using this navigation assistance system visible around others.  
5. I like to use technological products or systems like this navigation assistance system system. |
| Quality of life | 1. Using this rollator with navigation assistance system navigation assistance system could improve my physical well-being.  
2. This navigation assistance system evokes unpleasant feelings.  
3. This navigation assistance system could enhance my social contacts.  
4. This navigation assistance system could help me to maintain or increase my independence.  
5. The use of this rollator with navigation assistance system has a positive effect on me. |

* highlighted in *italics* is an indication for the adaption of the questionnaire to the specific functionality of the MOBOT rollator-type mobility device

2.4 Statistical analysis

The performance metric “success rate” was analysed by Fisher’s exact chi-square test allowing the analysis of associations between the dichotomous variables “navigation assistance” (assisted vs. unassisted) and the success rate (yes vs. no). Unpaired t-tests were calculated to compare completion times between the assisted and unassisted group.

To test for differences in completion times between group (cognitively impaired vs. non-impaired) and navigation assistance (assisted vs. unassisted navigation), we used a two-way analysis of variance (2x2 ANOVA) with Bonferroni’s post-test. Effect sizes were calculated as eta-squared ($\eta^2$) and were considered as small ($\eta^2 = 0.01$-$0.06$), moderate ($\eta^2 = 0.06$-$0.14$), and large ($\eta > 0.14$) [3]. The level of significance was set at $p = 0.05$ for all analyses.

To evaluate potential relations between the participants’ cognitive status and their subjective satisfaction with the MOBOT’s navigation assistance system, Spearman’s rank correlation coefficients ($r_s$) between the TSQ-WT item score and the MMSE score were calculated. The correlation coefficients was considered as low ($r_s < 0.2$), moderate ($r_s = 0.2$-$0.5$), and high ($r_s >0.5$) [3].
2.5 Results

2.5.1 Success rate

To evaluate the dichotomous performance metric “success rate” (yes vs. no), Fisher’s exact chi-square test was used allowing the analysis of associations between dichotomous variables such as “navigation assistance” (assisted vs. unassisted) and the success rate. For none of the two parts of the navigation trail, results of statistical data analysis showed significant associations between the navigation assistance and the target achievement ($p > .05$). Thus, success rates did not depend on whether or not participants were assisted by the MOBOT’s navigation assistance system.

**Tab. 4:** Success rates of both sections depending on the two different navigation assistance levels

<table>
<thead>
<tr>
<th>Navigation assistance</th>
<th>Section 1: starting position → chapel</th>
<th>Section 2: chapel → admission centre</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>yes (90.1) no (9.9)</td>
<td>$p$-value</td>
</tr>
<tr>
<td>Assisted, n (%)</td>
<td>20</td>
<td>2</td>
</tr>
<tr>
<td>Unassisted, n (%)</td>
<td>19</td>
<td>1</td>
</tr>
</tbody>
</table>

*p-values are given for Fisher’s exact chi-square tests applied to test for differences between assisted and unassisted group.

2.5.2 Task completion time

Navigation assistance (assisted vs. unassisted)

Results of the descriptive statistic showed a trend of shorter mean completion times of the assisted group for both sections (section 1: $91.2 \pm 31.2$ vs. $103.0 \pm 56.1$ s; section 2: $106.2 \pm 37.5$ vs. $157.2 \pm 103.6$ s) and in total ($193.5 \pm 67.2$ vs. $250.0 \pm 129.1$ s) (see Fig. 5). However, t-tests did not reveal any of the differences to be significant ($p = .062-.427$).

![Fig. 5: Task completion times for both sections and the total navigation trail depending on the two different navigation assistance levels (assisted vs. unassisted)](image)

Navigation assistance (assisted vs. unassisted) x cognitive impairment (intact vs. impaired)

*Section 1:* The 2x2 ANOVA showed that main effect of the navigation assistance (NAVI) on the time to reach the chapel was not significant ($F(1,35) = 1.268$, $p = .268$, $\eta^2 = .035$) but the main effect of cognitive status (COGN) was significant with a large effect size ($F(1,35) = 7.935$, $p = .008$, $\eta^2 = .185$) such that the cognitively intact group was significantly faster than the cognitively impaired group (see Tab. 5). There was a significant interaction between COGN and NAVI) with a large effect size ($F(1,35) = 9.137$, $p = .005$, $\eta^2 = .207$). Bonferroni-adjusted post hoc analyses of simple main effects showed that for cognitively impaired participants, the average time to reach the chapel was significantly shorter with navigation assistance than without navigation assistance ($F(1,35) = 8.000$, $p = .008$, $\eta^2 = .186$). For
cognitively intact participants, the use of the MOBOT’s navigation assistance system had no significant effect on the time to reach the chapel \( (F(1,35) = 1.946, p = .172, \eta^2 = .053). \)

**Fig. 6:** Descriptive statistic: Task completion time from starting position to chapel depending on the navigation assistance level (assisted vs. unassisted) and cognitive impairment level (intact vs. impaired)

**Tab. 5:** Results of the 2x2 ANOVA: Task completion time from starting position to chapel depending on the navigation assistance level (assisted vs. unassisted) and the cognitive impairment level (intact vs. impaired)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Navigation assistance</th>
<th>2x2 ANOVA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>unassisted</td>
<td>assisted</td>
</tr>
<tr>
<td>MMSE &gt;26: n = 10</td>
<td>MMSE &gt;26: n = 11</td>
<td>2x2 ANOVA</td>
</tr>
<tr>
<td>MMSE ≤26: n = 9</td>
<td>MMSE ≤26: n = 9</td>
<td></td>
</tr>
<tr>
<td>Time to chapel [s], mean ± SD</td>
<td></td>
<td>Factor</td>
</tr>
<tr>
<td>Intact (MMSE &gt;26)</td>
<td>69.1 ± 8.3</td>
<td>COGN</td>
</tr>
<tr>
<td>Impaired (MMSE ≤26)</td>
<td>140.6 ± 63.0</td>
<td>NAVI</td>
</tr>
<tr>
<td></td>
<td>92.3 ± 35.1</td>
<td>.008</td>
</tr>
<tr>
<td></td>
<td>89.8 ± 27.6</td>
<td>.268</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.185</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.035</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.207</td>
</tr>
</tbody>
</table>

**Section 2:** For the time to reach the admission centre, the 2x2 ANOVA showed a significant large effect of the factor COGN \( (F(1,34) = 9.438, p = .004, \eta^2 = .217) \) such that cognitively intact participants reached the admission centre significantly faster than cognitively impaired participants (see Tab. 6). There was also a significant large effect for the factor NAVI \( (F(1,34) = 7.621, p = .009, \eta^2 = .183) \). The participants assigned into the navigation-assisted group were significantly faster than those of the unassisted group. Additionally, a significant interaction with a large effect size between the two factors COGN and NAVI was found \( (F(1,34) = 8.076, p = .008, \eta^2 = .192) \). Bonferroni-adjusted post hoc analyses of simple main effects showed that for cognitively impaired participants, the navigation assistance led to a significantly shorter average time to reach the admission centre \( (F(1,34) = 14.822, p = .001, \eta^2 = .304) \). For cognitively intact participants, the use of the MOBOT's navigation assistance system had no significant effect on the time to reach the admission centre \( (F(1,34) = 0.003, p = .953, \eta^2 = .001) \).
Fig. 7: Descriptive statistic: Task completion time from chapel to admission centre depending on the navigation assistance level (assisted vs. unassisted) and cognitive impairment level (intact vs. impaired)

Tab. 6: Results of the 2x2 ANOVA: Task completion time from chapel to admission centre depending on the navigation assistance level (assisted vs. unassisted) and cognitive impairment level (intact vs. impaired)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Navigation assistance</th>
<th>2x2 ANOVA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>unassisted</td>
<td>assisted</td>
</tr>
<tr>
<td></td>
<td>MMSE &gt;26: n = 10</td>
<td>MMSE &gt;26: n = 10</td>
</tr>
<tr>
<td></td>
<td>MMSE ≤26: n = 8</td>
<td>MMSE ≤26: n = 10</td>
</tr>
<tr>
<td>Time to admission centre [s], mean ± SD</td>
<td>COGN</td>
<td>NAVI</td>
</tr>
<tr>
<td>Intact (MMSE &gt;26)</td>
<td>102.0 ± 26.4</td>
<td>103.7 ± 40.7</td>
</tr>
<tr>
<td>Impaired (MMSE ≤26)</td>
<td>226.1 ± 124.1</td>
<td>108.6 ± 36.2</td>
</tr>
</tbody>
</table>

Total completion time: For the total completion time, the 2x2 ANOVA revealed a significant large effect for the factor COGN ($F(1,33) = 10.547, p = .003, \eta^2 = .242$) such that cognitively intact participants completed the navigation trail faster (mean ± SD = 266.4 ± 129.5s) than participants with cognitive impairment (mean ± SD = 182.3 ± 56.4s) (see Tab. 7). There was also a significant moderate effect for the factor NAVI ($F(1,33) = 5.869, p = .021, \eta^2 = .151$). The participants assigned into the assisted group (mean ± SD = 193.53 ± 60.4s) had significant shorter completion times (mean ± SD = 193.53 ± 60.4s) compared to those in the unassisted group (mean ± SD = 249.8 ± 129.1s). Additionally, a significant interaction between these two factors was found, with a large effect size ($F(1,33) = 10.483, p = .003, \eta^2 = .241$). Bonferroni-adjusted post hoc analyses of simple main effects showed that for cognitively impaired participants, the navigation assistance provided by MOBOT’s navigation system led to significantly shorter total completion times ($F(1,33) = 14.794, p = .001, \eta^2 = .310$). For cognitively intact participants, the use of the MOBOT’s navigation assistance system had no significant effect on the total completion time ($F(1,33) = 0.362, p = .551, \eta^2 = .011$).
Fig. 8: Descriptive statistic: Total task completion time depending on the navigation assistance level (assisted vs. unassisted) and cognitive impairment level (intact vs. impaired)

Tab. 7: Results of the 2x2 ANOVA: Total task completion time depending on the navigation assistance level (assisted vs. unassisted) and cognitive impairment level (intact vs. impaired)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Navigation Assistance</th>
<th>2x2 ANOVA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unassisted</td>
<td>Assisted</td>
</tr>
<tr>
<td>MMSE &gt;26: n = 10</td>
<td>MMSE &gt;26: n = 10</td>
<td></td>
</tr>
<tr>
<td>MMSE ≤26: n = 8</td>
<td>MMSE ≤26: n = 9</td>
<td></td>
</tr>
<tr>
<td>Total completion time [s],</td>
<td>MMSE &gt;26: 171.2 ± 25.4</td>
<td>193.4 ± 76.1</td>
</tr>
<tr>
<td>mean ± SD</td>
<td>MMSE ≤26: 348.1 ± 140.6</td>
<td>193.7 ± 60.4</td>
</tr>
</tbody>
</table>

2.5.3 Subjective user satisfaction

The evaluation of the subjective user satisfaction was performed with all participants that used the MOBOT’s navigation assistance system to complete the navigation trail (n = 22). Median values of all questionnaire items were in the upper third of the scale (≥13.5), indicating a positive overall satisfaction with the MOBOT’s navigation assistance system (see Fig. 5).

Benefit: The median value of 13.5 for the item “benefit” (range 8-20) indicated that the MOBOT’s navigation assistance system provided a benefit for participants by helping them to reach the target positions. The effort of using this functionality seemed to be worthwhile for participants.
the participants and they would also recommend the navigation assistance systems to other people in similar situations.

**Usability:** The evaluation of “usability” aspects reached a median of 17 out of 20 points (range 7-20), suggesting that the navigation assistance system was perceived as an aid easy to use which did not require much effort and did not cause feelings of insecurity or indisposition. According to the participants’ estimation and experience so far, the system worked well and reliable.

**Self-concept:** High agreement could also be reported for the item “self-concept” (median = 16, range 9-19), indicating that the use of the MOBOT rollator with its navigation assistance system was an interesting challenge for the participants. They did not associate the use of the system with the loose of their independence or the feeling of being older than they actually were. They would also not feel embarrassed when using the MOBOT rollator with its navigation assistance system in public.

**Quality of life:** The median score of 13.5 (range 7-18) emphasized the potential of the MOBOT’s navigation assistance system for improving users’ physical well-being, enhancing their social contacts, and maintaining or even increasing their independence. The use of the system seemed not to be associated with unpleasant feelings, and may have a positive effect on the user’s quality of life.

Although high median values correspond to a positive user perception of the navigation assistance system, it also should be noted that the wide ranges, which could be found across all items, also indicated great heterogeneity of results.

All correlations between the MMSE score on the different questionnaire items were low ($r_s = -.138-.033$) and consistently insignificant ($p = .541-.981$), indicating no association of the cognitive status on the user perception of the navigation assistance system.

### 3 Scenario 2: Audio-gestural human-robot communication (ICCS)

#### 3.1 Background

Recent technological and scientific progress in assistive living and computer vision have led to significant interdisciplinary advances in assistive human-robot interfaces (for a review see: [4-9]). Human-computer/robot interfaces can offer natural communication channels to elderly subjects by putting emphasis on making their interaction easier and increasing the multimodal communication throughput.

In this scenario we aim to validate our developed action recognition system and models with actual elderly subjects, in collaboration with the clinical partners. This task still entails several challenges, as shown by the lack of similar works in the recent interdisciplinary literature. Visual and multimodal audio-gestural action detection and recognition comprise some of the most challenging areas of computer vision. For instance, action/gesture recognition techniques, despite recent progress [10], still do not incorporate rich temporal information efficiently. Moreover, the lack of adequate training data in currently available datasets seriously limits the potential of core machine learning techniques. In addition, the task of multimodal gesture/action recognition becomes much more challenging when applied on elderly subjects because their voice and gestures/actions may be significantly different compared to non-elderly subjects.

Towards this direction, the data collection effort, analysis, and post-processing of results for the two validation studies are still on-going. This is due to the large volumes of data that have been recently acquired. The complete results of these experiments and studies are to be
published subsequently. Until then, we present in the following subsections the experimental setup as well as some preliminary results and discussion.

3.2 Design and experimental setup

Three different user modes concerning the communication between the participants and the MOBOT rollator-type mobility assistance were tested during this scenario:

1. Audio mode
2. Gestural mode
3. Audio-gestural mode
4. “Self-chosen/-free” mode

For each mode, the following five commands were tested:

1. Come near.
2. Stand up.
3. Help
4. Where am I?
5. Park.

The participants were seated on a standard chair 1-2 m behind the MOBOT rollator-type mobility assistance during the whole testing scenario.

3.2.1 Audio mode

Before starting the experiment, the supervisor introduced the five audio commands to be tested by showing cards with the written commands in a predefined order (1.-5.). The commands were simultaneously read aloud to the participant. After this first step, the cards were presented again in the same order and the participant was now asked to read aloud the commands by themselves. If necessary, reading was corrected and repeated until the participant was able to read the commands as intended.

For the experiment, the supervisor showed the commands once more in the same order and the participant was again told to read aloud the commands as they were written there. After reading each command, a short brake was made in order to give the system the opportunity to respond on the command. In case of successful recognition, the system responded with an audio message. The maximum system response time was approx. 5-7 seconds for audio commands. If the system did not recognize the audio command in this time interval, the supervisors continued with presenting the next audio command. All five commands were presented subsequently in the same order for three times. There was no significant audio noise or other persons talking to or around the participant.

3.2.2 Gestural mode

Prior to the testing scenario, the gestures to be tested were also introduced by the supervisor. In a first step, the supervisor demonstrated the gestures for each command. After the demonstration, the supervisor presented the gestures once more and asked the participant to imitate them simultaneously. Each gesture was trained until it was performed in the way as it was intended.

For the experiment, the supervisor demonstrated the gestures once more and the participant was again told to imitate them in the way it was performed by the supervisor. After performing the gesture, the supervisor and participant had to wait for system response also for approx. 5-7 seconds (= system response time for gestural commands). In case of successful recognition, the system also responded with an audio message. If the system did not recognize the gesture in this time interval, the supervisors continued with demonstrating the next gesture.
All five gestures were presented subsequently in the same order for three times. There was no extra person moving close to, in the field of view or touching the subject.

Fig. 10: Participant performing gestural commands (left: come near; right: stand up)

3.2.3 Audio-gestural mode

For testing the audio-gestural mode, the same test procedure was used as for the other modes; expect that there were now two test supervisors (A, B). Initially, for introduction of the combination of audio and gestural commands, supervisor A showed and read out aloud the audio commands and supervisor B simultaneously demonstrated the gestures. After that, the combination of the audio and gestural commands was shortly trained with the participants until they were able to perform the combination correctly.

For the experiment, the participants were asked to read aloud the written commands shown by supervisor A and simultaneously perform the corresponding gestures as demonstrated by supervisor B. System response time for the audio-gestural mode was approx. 10 seconds, and in case of successful recognition, the system also responded with an audio command. All five audio-gestural commands were presented subsequently in the same order for three times. There was no significant audio noise or other persons talking to or around the participant, and no extra person moving close to, in the field of view or touching the subject.

Fig. 11: Illustration of the audio-gestural mode. The patient performs the gesture (top row) and utters the audio command (waveform on the second row) at the same time. Video frames are overlaid with part of the visual processing pipeline described in D1.3.

3.2.4 “Self-chosen/-free” mode

In this last experiment, the participant was asked to perform as many as possible of the commands he/she could remember within a time interval of 3 minutes, no matter in which order or which kind of modality of the command (i.e. audio, gestural, or audio-gestural) was performed.
In contrast to the other experiments, no demonstrations of the gestural commands were given and no cards with the written audio commands were shown prior to the experiment. The data collected in this mode shall be exploited for future research.

3.3 Performance metrics

In this preliminary reporting, since large quantities of experimental validation data are still being acquired, and/or post-processed we report a few average technical performance results such as the automatic recognition rate performed by system recognizing the subjects’ commands.

We should note here that the above up-to-now results are based on the assumption that the subjects have uttered and performed the exact audio-gestural commands as the ones in the provided scripts. However, this is not the case at all due to practical reasons. Thus, due to such issues the reported results may be revised after very detailed manual annotation of the ground truths and the involved multimodal data.

The quantitative measure employed at this stage is computed by comparing the sequence of ground truth labels to the recognized audio or gestural commands. In this way we measure the false alarms as “insertions” yielding the accuracy as \((N-D-S-I)/N * 100\%\); \(N\) is the number of ground-truths, \(S\) the substitution errors, \(D\) the deletion errors and \(I\) the insertion errors.

“Self-chosen/-free” mode

For this experiment, we report also the number of commands (i.e. audio, gestural, audio-gestural commands) the participants were able to recall in a time interval of 3 minutes.

3.4 Results

By the quantity of data that has up to now been processed we should first note that there is a short percentage of missing scenario experimental cases of data, that is approx. 11-15\% due to several reasons, e.g. the subject did not manage to conduct the experiment (“self-chosen/-free” mode).

The median automatic recognition rate in terms of accuracy is currently reported for the top performing subjects processed up to now and is 53.3\% for the audio and 40\% for the visual modality. Nevertheless, there are a few subjects performing as low as 6 or 26\% in either the audio or visual modalities. However these, and probably all will be carefully reviewed not only automatically but also by manual annotation, so as to refine the ground-truth labels, and find out how accurate these results are.

In any case we should note that this task is rather challenging and the elderly subjects, even the top performing ones, would perform variations of the audio-gestural commands or altered in ways that the system cannot sometimes recognize. This signifies that there is still more interdisciplinary research to follow, especially for the (multimodal) audio-gestural mode.

**Tab. 8**: Top performing subjects (ID in terms of system’s automatic recognition with respect to the audio (left part of the table), and gesture (right part of the table) modalities.

<table>
<thead>
<tr>
<th>ID</th>
<th>audio [%]</th>
<th>gesture [%]</th>
<th>ID</th>
<th>audio [%]</th>
<th>gesture [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td>53.33</td>
<td>46.67</td>
<td>7</td>
<td>53.33</td>
<td>40</td>
</tr>
<tr>
<td>25</td>
<td>53.33</td>
<td>33.33</td>
<td>14</td>
<td>26.67</td>
<td>46.67</td>
</tr>
<tr>
<td>12</td>
<td>60</td>
<td>26.67</td>
<td>13</td>
<td>53.33</td>
<td>46.67</td>
</tr>
</tbody>
</table>
Number of recalled commands
Although each command was intensively introduced, trained before testing, and repeatedly performed for 3 times during experiments, on average, participants were able to recall only 3.5 (23.3%) out of 15 possible commands.
When differentiating between the individual modalities, only 1.3 ± 1.3 (26.0%) audio, 1.3 ± 1.4 (26.0%) gestural, and 1.6 ± 1.6 (32.0%) audio-gestural commands could be recalled out of the five possible commands for each modality.
When not considering the modality, participants were, however, able to recall 4.2 ± 2.2 (84.0%) out of the five commands.
To evaluate potential relations between the participants’ cognitive status and the ability to learn the commands, Spearman’s rank correlation coefficients (low: $r_s < 0.2$; moderate $r_s = 0.2-0.5$; high: $r_s >0.5$) between the number of recalled commands and the MMSE score were calculated. The correlations coefficients were low ($r_s = .061-.260$) and consistently insignificant ($p = .268-.798$), indicating no association of the cognitive status and the ability to learn the commands.

4 Summary (BETHANIEN, ICCS)
With the last evaluation study at BETHANIEN, we evaluated the cognitive assistance system and the audio-gestural human-robot interaction of the final prototype of the rollator-type mobility assistant. In collaboration of BETHANIEN and ICCS, a task force with clinical as well as technical expertise could be established developing a feasible and adequate study design and assessment strategy that covers the specific MOBOT functionalities to be tested in the study.
By using a broad and timely recruitment strategy, the clinical partner BETHANIEN was able to even oversample ($n = 42$) the recruitment goal of 40 participants within a narrow test period of about 3 weeks.
Preliminary results from the clinical partner BETHANIEN on the users’ navigation performance with the MOBOT’s cognitive assistance system are quite promising, showing the added value of such robotic functionality for cognitively impaired participants. By using the navigation assistance system of the MOBOT rollator-type mobility assistant, cognitively impaired participants were able to complete navigation tasks significantly faster than without the MOBOT's navigation assistance. In addition, an overall positive user satisfaction with the assistant system could be observed.
Concerning the audio-gestural scenario, we are still in the processing stage of the large data quantities acquired up to now. The preliminary results so far indicate an automatic recognition accuracy for the top performing elderly subjects of approximately 53.3% for the audio and 40% for the gestural modality. Note that these preliminary results are based on several assumptions concerning the ground-truth annotations, which need to be revised and examined in detail manually by annotation experts. Thus, a more detailed analysis and post-processing of the results as well as of the assumed ground-truths shall shed more light on our analysis and lead us to more mature conclusions. Based on these preliminary results that depend on the timewise approximate annotations, there is still room for improvement of the performance of the audio-gestural recognition system so that it could provide a robust multimodal interface that enables elderly people to interact with the robot in an unconstrained way. This, of course, is also related to the fact that the field of computer vision and multimodal interaction for assistive robotics is still in its infancy, as shown by the lack of published results on similar challenging tasks that involve elderly people under realistic conditions. Therefore, we regard our results as very promising, paving the way for further research in a field which has high impact on the development of assistive robots.
Study results of the audio-gestural scenario addressing the performance of users in interacting with the MOBOT rollator showed that for an adequate human-robot communication, a more extensively training in learning the audio and gestural commands seemed to be necessary in the user group.
II FINAL EVALUATION STUDY AT DIAPLASIS (DIAPLASIS)

The final evaluation study took place in DIAPLASIS rehabilitation hospital in Kalamata, Greece. During this evaluation we examined the cognitive assistance and the audio-gestural modules/functionalitys of the MOBOT rollator type device. The evaluation was based both on the experience gained during the evaluation studies in BETHANIEN, as well as on previous research conducted in the frames of the project [11].

The evaluation scenarios were tailored to meet the needs of the project while using both objective and subjective criteria. To this end, we established a multi-disciplinary team coordinated by Prof. Costas Tzafestas, Prof. Yiannis Koumpouros and Dr. Alexandra Karavasili. For the objective evaluation of the MOBOT we collected various data from the multiple sensors integrated in the MOBOT rollator-type device. The subjective assessment of the device was also a major issue. This was achieved by using three different already validated questionnaires [12-15].

Since the evaluation in DIAPLASIS was conducted during the last days of the project, the data collected are still in processing. The process of data cleaning and post-processing is a very time consuming and demanding one, since the volume of the data is huge. The data collected from the several sensors are being processed by the different technical teams of the project, while the ones collected from the questionnaires are being processed by DIAPLASIS.

Due to the fact that the final evaluation study in DIAPLASIS was conducted in the period from the 15th of July till the 27th of July, the current deliverable includes only some basic early results. We are still waiting for the post-processing of the collected data from the technical partners in order to be able to report more comprehensive and statistically analysed results. This task will be performed as soon as the data are available in order to produce the research insights.

The evaluation study in DIAPLASIS was approved by the Bioethic and Deontology Committee of the Technological Educational Institute of Athens (application number 2832/09-05-2014).

1 Subject recruitment

1.1 Recruitment strategy

The recruitment process was very time consuming since we had to identify the right subjects that would fit the MOBOT evaluation strategy, while not giving them any incentives. The participation was on a volunteering basis with no other incentives provided. Due to the health conditions of the target group there was also a great risk of drop out that could be caused due to fatigue by the requested tasks.

Even though in the DoW it was planned the rollator-type mobility assistant to undergo a user evaluation with fifteen (15) people engaged in rehabilitation programs in DIAPLASIS, we extended this scenario and included thirty (30) subjects to participate in the final evaluation of the MOBOT rollator-type device. Two scenarios were examined:

- cognitive assistance
- audio-gestural communication

To be able to achieve the desired number of participants we had an extended period of screening and evaluation of the potential subjects, their medical files, etc. We examined both inpatients and outpatients of our rehabilitation hospital and in case of a limited number identified, we contacted also other settings (i.e. collaborating geriatric centers).
As far as the inclusion criteria are concerned, the subjects should be users of an assistive device (rollator type) for at least a month. Moreover, they should meet also the inclusion criteria reported in deliverable D5.2:

- Persons with moderate motor impairment (habitual use of rollator, gait speed <0.6 m/sec unassisted)
- Persons with a moderate impairment (Person is unable to stand up and sit down unassisted on a normal chair (standardized 100% leg length) without problem, 5-chair stand >16.7 sec or able to stand up from a chair from elevated chair (120% lower leg length)
- Persons aged ≥ 65 years old
- No cognitive impairment (MMSE ≥26-30) or mild to moderate impairment (MMSE 17-25)

Apart from the above clinical criteria, all participants should also sign an informed consent, and they should be able to perform the required tests/assessments.

In order to identify the potential participants, DIAPLASIS clinical team screened the patient files (inpatients and outpatients) and interviewed the selected patients. In many cases, they needed to contact also the relatives of the selected patients in order to inform them and explain them the requirements of the evaluation study. The subjects that finally agreed to participate in the evaluation sessions were scheduled accordingly. Many outpatients of DIAPLASIS as well as the ones coming from other settings (e.g. collaborating geriatric centres) were transported to and from DIAPLASIS hospital premises with DIAPLASIS transportation service. Figure 12 depicts the selection process.
Fig. 12. Flow chart of selection process for the DIAPLASIS validation study

Initial screening
n = 425
- Inpatients n = 185
- Outpatients n = 200
- Collaborating settings n = 40

Excluded after screening according to predefined inclusion criteria
n = 328
- Inpatients
  - n = 127
  - Medical reasons: 63
  - Unable to walk with rollator: 42
  - MMSE < 17: 15
  - Other reasons: 7
- Outpatients
  - n = 170
  - Medical reasons: 41
  - No use of rollator: 104
  - MMSE < 17: 3
  - Other reasons: 22
- Collaborating settings
  - n = 31
  - No use of rollator: 8
  - Unable to walk with rollator: 17
  - MMSE < 17: 2
  - Other reasons: 4

Eligible based on predefined inclusion criteria
n = 97
- Inpatients n = 58
- Outpatients n = 30
- Collaborating settings n = 9

Consented to study participation
n = 30
- Inpatients n = 13
- Outpatients n = 15
- Collaborating settings n = 2

Excluded
n = 67
- Inpatients
  - n = 45
  - Refused to participate: 19
  - Medical reasons: 4
  - Other reasons: 22
- Outpatients
  - n = 15
  - Refused to participate: 15
- Collaborating settings
  - n = 7
  - Refused to participate: 7

Navigation & Audio-gestural scenarios
- Participants with MMSE < 26
  - n = 15
- Participants with MMSE ≥ 26
  - n = 15
1.2 Results

1.2.1 Recruitment

As presented in Figure 12, we initially screened 425 subjects. 185 were inpatients, 200 were outpatients in DIAPLASIS rehabilitation hospital and 40 of them were found in collaborating settings (i.e. geriatric centers). 328 out of the 425 were excluded after the first screening due to the fact that they could not meet the inclusion criteria used for our study. After studying the patient files 127 inpatients were found ineligible (the medical condition of 63 of them didn’t fit the characteristics of the study, 42 could not walk using a rollator, 15 had a MMSE less than 17, and 7 could not participate for other reasons: e.g. gait speed > 0.6m/sec). 170 outpatients were also found ineligible (41 due to medical contradictions, 104 had stopped using a rollator, 3 scored in MMSE with less than 17, and 22 for other reasons: e.g. gait speed > 0.6m/sec or/and 5 sit-to-stand<16.7sec). From the collaborating centres, 31 found ineligible (2 had stopped using a rollator, 17 could not walk with a rollator, 2 scored less than 17 in MMSE, and 4 for other reasons: e.g. gait speed >0.6m/sec, or age <65 years).

The remaining 97 subjects were potential participants in our study. However, 19 inpatients refused to participate due to several reasons (e.g. would leave for vacations, or their relatives did not want, etc.), 4 even though they met the inclusion criteria could not participate due to unexpected fatigue, illness, and 22 more for other reasons. In total, 45 inpatients were excluded in a second step. From the 30 outpatients, half of them were also excluded in the second phase because they refused to participate (mostly because their program didn’t meet the MOBOT evaluation schedule). Finally, 7 subjects from the collaborating centres refused to participate in the study.

All remaining subjects (13 inpatients + 15 outpatients + 2 from collaborating centres) were again informed about the study. They were given a written informative package about the study and the scope and their participation. After that, they all agreed and signed the appropriate informed consent. They also signed the appropriate forms in order to allow us to use photos and videos from their participation during the evaluation of MOBOT.

In total, 30 subjects participated in the evaluation of both the scenarios. The subjects were appropriately distributed in order to have two equal groups of participants: 15 with cognitive impairment (MMSE<26) and 15 being cognitive intact (MMSE>25).

1.2.2 Sample description

The participants in the evaluation in DIAPLASIS all met the inclusion criteria as they were defined in previous deliverables (current users of a rollator, aged above 65 years old, MMSE>17, gait speed<0.6 m/sec and 5 sit-to-stand > 16.7 sec). The study did not focus on pathologies rather on the functional status of the participants as described in the inclusion criteria. Most of them were inpatients or outpatients in DIAPLASIS rehabilitation hospital and only 2 of them were from collaborating geriatrics centres. The sample basic descriptive characteristics are presented in Table 9.
Tab. 9: Sample descriptive characteristics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Total sample (n = 30)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, years, mean (SD)</td>
<td>78.6 (6.9)</td>
</tr>
<tr>
<td>Female, n, (%)</td>
<td>17 (56.7)</td>
</tr>
<tr>
<td>MMSE score, mean (SD)</td>
<td>24.8 (4.2)</td>
</tr>
<tr>
<td>Barthel-Index score, mean (SD)†</td>
<td>76.8 (14.5)</td>
</tr>
<tr>
<td>GDS score, median (range)</td>
<td>4.0 (0-14)</td>
</tr>
<tr>
<td>BBS score, mean (SD)</td>
<td>29.2 (10.7)</td>
</tr>
<tr>
<td>Recent history of falls, n (% ††)</td>
<td>22 (73.3)</td>
</tr>
<tr>
<td>FES-I score, median (range)</td>
<td>13.0 (7-28)</td>
</tr>
<tr>
<td>Fear of Falling, n (% †††)</td>
<td>17 (56.7)</td>
</tr>
<tr>
<td>FIM score, median (range)</td>
<td>104.0 (75-121)</td>
</tr>
<tr>
<td>Living Situation, n (%)</td>
<td></td>
</tr>
<tr>
<td>Community-dwelling</td>
<td>15 (50.0)</td>
</tr>
<tr>
<td>Institutionalized</td>
<td>15 (50.0)</td>
</tr>
<tr>
<td>POMA score, mean (SD)</td>
<td>15.0 (4.2)</td>
</tr>
<tr>
<td>Habitual gait speed, m/s (SD) †††</td>
<td>0.24 (0.13)</td>
</tr>
</tbody>
</table>

Note. n = number of cases; MMSE = Mini-Mental State Examination; GDS = Geriatric Depression Scale, 15 items; SF-12 = 12-item Short Form Health Survey; BBS = Berg Balance Scale; FES-I = Falls Efficacy Scale – International, 7-item version; POMA = Performance Oriented Mobility Assessment; FIM = Functional Independence Measure.
† Based on interview-led self-reports; †† no vs. some to very much fear of falling; ††† Assessed by 4-Meter-Walk-Test.

2 Scenario 1: Cognitive assistance

The objective of the cognitive assistance scenario was to assess the integrated functionality of MOBOT related to the assistance of the end user when navigating in a trail. MOBOT provides an audio assistance to the end user at pre-specified “critical” waypoints. The technical details of this functionality are reported in details in other deliverables (D3.2 and D3.3). A pre-mapping of the space and the navigation trail and a localization of the MOBOT rollator-type device are required prior initiating the experiments.

According to the research hypothesis, the cognitive assistance functionality would assist subjects from Group A (cognitive impairment group) to navigate in unfamiliar spaces and achieve comparable results as compared to Group B (cognitive intact group).

As already described in Part I (section 3) of this deliverable, the navigation assistance system of the MOBOT rollator-type device provides audio cues to the end user in order to help him/her navigate while walking with MOBOT in unfamiliar spaces. To this end, a pool of pre-set audio commands is associated with certain points in the space where the patient walks with MOBOT. For more information on this module please consult the technical descriptions provided in other deliverables as well the description already provided in previous sections of the current deliverable.

2.1 Design and experimental setup

For the purposes of the study we formed two groups:

- Group A: cognitively intact (MMSE >26)
- Group B: cognitively impaired (MMSE ≤26)

The cognitive assistance was turned on in all patients (both groups).
The appropriate navigation trail was designed in DIAPLASIS in cooperation with ICCS. The final navigation trail was complex enough in order to test the capabilities and the assistance provided by the system. It was divided in three different sections:

- 1st section: From ICU (point 1) to the occupational therapy room (point 4).
- 2nd section: From the occupational therapy room (point 4) to the entrance (point 6).
- 3rd section: From entrance (point 6) to nursing stop in ward C (point 12).

The navigation trail is depicted in Figure 13.

![Fig. 13: Layout of the navigation trail at the ground floor of the DIAPLASIS rehabilitation hospital](image)

After having mapped the navigation space using the MOBOT localisation and mapping modules, the technical team programmed the appropriate audio cues for each point on the navigation trail. A list of the audio cues is presented in Table 10.

<table>
<thead>
<tr>
<th>Tab. 10: Audio cues list</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Section 1: starting position → occupational therapy room</strong></td>
</tr>
<tr>
<td>Point 1: “Walk straight ahead”</td>
</tr>
<tr>
<td>Point 2: “Turn right and walk straight”</td>
</tr>
<tr>
<td>Point 3: “Turn right and walk straight”</td>
</tr>
<tr>
<td>Point 4: “Turn back and walk straight ahead”</td>
</tr>
<tr>
<td>Point 5: “Walk straight ahead towards the front door”</td>
</tr>
<tr>
<td>Point 6: “Turn back and walk straight towards the reception desk”</td>
</tr>
<tr>
<td>Point 7: “Turn right and walk straight”</td>
</tr>
<tr>
<td>Point 8: “Turn left and walk straight”</td>
</tr>
</tbody>
</table>
Both Group A and Group B performed the navigation trial with the MOBOT rollator-type device. All participants were instructed about the final targets prior to initiating the experiment. They were instructed to try to reach the targets either by following the signs at the walls and ceiling in DIAPLASIS hospital or by hearing and following the audio commands provided by MOBOT. The staff of DIAPLASIS was not allowed to give any hint or help during trials. MOBOT’s navigation assistance was provided when MOBOT reached a certain area defined around each one of the points. Each audio cue was repeated every 3 seconds until the participant has left the critical area of the point.

During the experiments, all participants were accompanied and supervised by our health professionals in DIAPLASIS in order to help or stop the experiments in case a subject was tired and prevent falling or other unexpected situations. There was no stop for the patients between the targets and the system was recording the time elapsed for each target. These data are still not available to us in order to report on them and analyse them statistically. This will be done as soon as we have them in order to export the insights of the experiment, which will be part of future publications.

2.2 Performance metrics

The MOBOT rollator-type device was assessed during the experiments using both objective and subjective performance measures.

Objective Measures:
- success rate (first/intermediate/final target position)
- task completion time (first/intermediate/final target position)
- stopping time
- walking trajectories
- gait parameters captured by laser sensor

Subjective Measures:
- ATD PA Greek version [14]
- QUEST 2.0 Greek version [12]
- PYTHEIA questionnaire [13,15]

The data regarding the objective measurement were collected by the system and are still being processed by the technical partners. After providing these data we will be able to analyse and publish them accordingly.

In order to be able to measure the subjective satisfaction of the users we conducted a lot of research regarding already valid and reliable questionnaires that could be used in the frames of the MOBOT experiments. As reported in [11] most of the studies are utilizing either custom-made questionnaires or interviews that are neither valid nor reliable instruments to represent the subjective opinion and perception of the end users. There is therefore a great gap in the subjective assessment of rehabilitation or assistive robot devices. The absence of standard scales/questionnaires for the subjective assessment of robot-based devices makes it difficult to design products that meet exactly the needs of the intended end users, to further improve prototypes, or to compare the results from different researchers. Based on the findings of the review we proceeded in developing a new questionnaire that could meet the needs of our study but also could be used in other studies working in the development of rehabilitation or assist-
tive robot devices. The new questionnaire is called PYTHEIA and prior to the evaluation study in DIAPLASIS was already tested for its reliability and validity [13,15].

PYTHEIA’s unique characteristic is that it can be used for the measurement of satisfaction of the end users with assistive and robotic technologies, while being able to evaluate any individual features-functionalities implemented. This is not present in any other scale, at least to the knowledge of the authors. More details on the subjective measurement are provided in next sections.

3 Scenario 2: Audio-gestural human-robot communication

The audio-gestural human-robot communication scenario covered the interaction of the end user with MOBOT using only audio and/or gestural commands. The human-robot communication model was built as a structured tree of possible multimodal action-reaction interactions engaging both audio and gestural signals. The end user can interact therefore with the MOBOT device in a more natural way using oral commands and gestures that he/she is commonly use in his/her everyday life.

3.1 Design and experimental setup

During this scenario we evaluated four different setups:

- Audio mode commands
- Gestural mode commands
- Audio-gestural mode commands
- Free mode commands (the user could choose any combination of the commands already used in the previous modes)

In order to test the integrated functionality, five different commands were used:

- “MOBOT, come near”
- “MOBOT, I need help”
- “MOBOT, I want to stand up”
- “MOBOT, where am I?”
- “MOBOT, go park”

During the experiments, the participants were seated on a standard chair 1 m behind the MOBOT device.

3.1.1 Audio mode

Prior to starting the experiments, the supervisor instructed the participant on the whole process and the way he/she should communicate with MOBOT. The supervisor presented to the participant the five phrases, then the five signs and then the combination of them. All commands in the first three modes (audio mode, gestural mode and audio-gestural mode) were tested in the same order.

Upon recognition of the oral command the system provided an appropriate audio feedback, letting the user know that MOBOT understood the desire of the end user and will react accordingly. A minimum of five seconds was needed before the user was allowed to give another command to MOBOT. This time was required by the system in order to try to identify the commands given by the user. In order to guarantee the best possible setting for the experiment, there was no external noise by the environment.

3.1.2 Gestural mode
The gestural scenario was tested again using the same setting as in the previous section (audio mode). Again, the user was seated 1 m in front of MOBOT. Then the user was moving his/her hands as prompted by the supervisor, in order to calibrate MOBOT. After calibration was achieved, the user was nodding to MOBOT according to the instructions given by the supervisor. The system was supposed to respond upon recognition of the gestural command. In that case, MOBOT provided an appropriate audio feedback to the end user, making him/her understand that MOBOT recognised the gesture. A time of approximately five seconds was required before the user performed the next gesture. This time was needed by the MOBOT in order to try to recognise the gesture. The environment close to the end user was clear in order for the sensors of MOBOT to catch his/her gestures without any interference.

![Fig. 14: Participant performing gestural commands](image)

3.1.3 Audio-gestural mode

The third mode tested was a combination of the previous two modes. The user was instructed again by the supervisor how to perform the audio and gestural commands simultaneously. Then, the end user was performing the combination of audio gestural commands and was waiting for approximately ten seconds prior to performing the next one. This time was needed by the system in order to try to recognise the performed commands. Upon successful recognition the system provided an appropriate audio response.
3.1.4 Free mode

During this experiment, the end user could either provide an oral command or gesture or even a combination of both. The participant had to remember the commands by himself/herself. No further help was given to him/her by the supervisor.

3.2 Performance metrics

The evaluation of the second scenario (audio-gestural human robot communication module) was again evaluated against both objective and subjective criteria.

Objective Measures:
- success rate (% of successful detection)
- \( \frac{(N-D-S-I)}{N} \times 100\% \) where \( N \) is the number of ground-truths, \( S \) the substitution errors, \( D \) the deletion errors and \( I \) the insertion errors

Subjective Measures:
- ATD PA Greek version [14]
- QUEST 2.0 Greek version [12]
- PYTHEIA questionnaire [13,15]

Since the collected objective data are still in post-processing by the technical partners, it is not possible to report any outcomes on that. After providing these data we will be able to analyse and publish them accordingly.

3.3 Subjective evaluation

In the frames of the evaluation in DIAPLASIS we used three different questionnaires, as already reported in the previous sections.

Each participant, after completing each scenario, was asked to complete the QUEST 2.0 Greek version questionnaire, the ATD PA Greek version questionnaire and the PYTHEIA questionnaire. Using the Part A of PYTHEIA he/she evaluated the MOBOT system as a whole, and then by using Part B of the questionnaire he/she assessed independently the (i) cognitive module (navigation assistance) and (ii) the audio-gestural module of the MOBOT rollator-type device.

Since the data collected are still being processed it is not possible to present in this deliverable the outcomes of the evaluation. The analysis of the qualitative data will be available soon, and will be reported in future publications. The PYTHEIA questionnaire is presented in the next table.

PYTHEIA questionnaire

The PYTHEIA questionnaire is presented in the next table.

Tab. 11: PYTHEIA questionnaire

<table>
<thead>
<tr>
<th>Item #</th>
<th>Question</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PART A</strong></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Rate your satisfaction with the supporting device and the services provided in relation to the adaptability in the spaces you spend your everyday life (home, work).</td>
</tr>
<tr>
<td>2</td>
<td>Rate your satisfaction with the supporting device and the services provided in relation to its contribution to the improvement of your everyday life.</td>
</tr>
<tr>
<td>Item #</td>
<td>Question</td>
</tr>
<tr>
<td>-------</td>
<td>----------</td>
</tr>
<tr>
<td>3</td>
<td>Rate your satisfaction with the supporting device and the services provided in relation to the ease of learning all individual functions.</td>
</tr>
<tr>
<td>4</td>
<td>Rate your satisfaction with the supporting device and the services provided in relation to the ease of learning the basic functions (the functions that concern me more).</td>
</tr>
<tr>
<td>5</td>
<td>Rate your satisfaction with the supporting device and the services provided in relation to the ease of use (complexity, required effort).</td>
</tr>
<tr>
<td>6</td>
<td>Rate your satisfaction with the supporting device and the services provided in relation to how secure it is.</td>
</tr>
<tr>
<td>7</td>
<td>Rate your satisfaction with the supporting device and the services provided in relation to the dimensions (height, width, length).</td>
</tr>
<tr>
<td>8</td>
<td>Rate your satisfaction with the supporting device and the services provided in relation to the weight.</td>
</tr>
<tr>
<td>9</td>
<td>Rate your satisfaction with the supporting device and the services provided in relation to if the functionalities existing are sufficient.</td>
</tr>
<tr>
<td>10</td>
<td>I will feel more secure (protected, confident) when using this assistive device.</td>
</tr>
<tr>
<td>11</td>
<td>I will feel more autonomous when using this assistive device.</td>
</tr>
<tr>
<td>12</td>
<td>I will need help from another person to use the assistive device.</td>
</tr>
<tr>
<td>13</td>
<td>I will feel comfortable to use the assistive device around the community.</td>
</tr>
<tr>
<td>14</td>
<td>I will feel comfortable to use the assistive device among my colleagues (working environment).</td>
</tr>
<tr>
<td>15</td>
<td>I will feel comfortable to use the device around friends and family.</td>
</tr>
</tbody>
</table>

**PART B**

IF1 Rate your satisfaction with the specific feature of your assistive device in relation the ease of use.

IF2 Rate your satisfaction with the specific feature of your assistive device in relation to the help it provides in your everyday life.

IF3 Rate your satisfaction with the specific feature of your assistive device in relation to how safe/secure it is.

IF4 Rate your satisfaction with the specific feature of your assistive device in relation to its reliability (i.e. whether it applies always correctly).

IF5 Rate your satisfaction with the specific feature of your assistive device in relation to the feeling of safety (I will feel more secure, protected, confident when using it).

**4 Summary**

During the evaluation that took place in DIAPLASIS we assessed both the cognitive and the audio-gestural functionalities of MOBOT rollator-type device. A total of 30 subjects participated in the evaluation study in DIAPLASIS and completed the experiments. A longer recruitment process took place in order to satisfy the needs of the evaluation.

Since the evaluation was conducted during the last days of the project and the collected data are still not available, it is impossible to report any statistically processed results. This will be done in the short future in order to provide insights of the MOBOT evaluation and publish the results in peer reviewed scientific journals and conferences. It is important though to mention that the diversity of the participants in terms of pathology, age and health condition can be considered as an asset for the evaluation of MOBOT under real conditions.
MAIN SUMMARY AND CONCLUSION

The final evaluation studies conducted at BETHANIEN and DIAPLASIS aimed to validate the navigation assistance system of the MOBOT rollator-type mobility assistant as well as the audio-gestural human-robot interaction with potential end-users. In collaboration between technical and clinical partners of the MOBOT consortium, adequate test scenarios were developed which addressed the specific robotic functionalities to be tested. With an appropriate recruitment strategy, both clinical partners BETHANIEN and DIAPLASIS were able to recruit a sufficient number of study participants representing the predefined, impaired target user group of the MOBOT rollator-type mobility assistant and to oversample the sample sizes projected in the DoW for the final evaluation studies. Both studies could be carried out successfully without any critical events during test scenarios and no participant who objected to the test procedure. Serious technical issues that could have affected the execution of the evaluation studies were successfully prevented by providing permanent technical support from the technical partners of the consortium over the entire study period.

Preliminary results on objective performance metrics of the cognitive assistance scenario are quite promising, indicating that cognitively impaired participants achieved a better user performance (i.e. task completion time) with the MOBOT’s cognitive assistance system when completing a navigation task in a complex real-world application scenario. Results on the subjective user satisfaction are also very positive and suggest, for instance, that the use of this MOBOT functionality provides a benefit for the users, do not cause feelings of insecurity, is an interesting challenge for the users, and may have a positive effect on the user’s quality of life. Further analyses of the more technique-based objective performance metrics (e.g. walking distance, gait parameters) captured by the robot-integrated technical systems which are currently being post-processed by technical partners may have the potential to substantiate these positive results and will be part of future publications produced after the end of the MOBOT project. Although preliminary, these results impressively show the potential added value of such robotic functionality for cognitively impaired, frail older people.

Audio-gestural human-robot interaction enables new modes of natural communication between users and assistive robotic devices and is therefore highly relevant for the development of such devices with focus on the user perspective. Audio-gestural action detection and recognition comprise, however, some of the most challenging areas of computer vision, especially in frail old people which might perform the commands with some variations. Current available preliminary results on the automatic recognition accuracy of the MOBOT’s multimodal action recognition system also reflected this issue, signifying that there is still more technical research to follow. One the other hand, results of the audio-gestural scenario focusing more on the performance of users highlight that for an adequate human-robot communication, also a more extensively training in learning the commands may be necessary in the user group.

For the data thus obtained, comprehensive, time consuming, post-processing analysis is mandatory. As these data are currently still being processed by technical partners, this report is mainly based on simple quantifiable outcomes (e.g. success rate, task completion time). The completion of the analysis of more technique-based performance metrics will be postponed to later data exploitation for publication.

As both evaluation studies were conducted at the end of MOBOT project and the majority of the predefined performance metrics of the test scenarios based on a time-consuming analysis and post-processing of the collected data, we were not able to provide a comprehensive presentation of study results. However, the MOBOT consortium planned to present further results to the research community in later publications.
References