Cognitive Decline Detection for Alzheimer's Disease Patients Through an Activity of Daily Living (ADL)

G. Palacios-Navarro, J. Buele, S. Gimeno Jarque, and A. Bronchal García

***Abstract*— There are conventional screening instruments for the detection of cognitive impairment, but they have a reduced ecological validity and the information they present could be biased. This study aimed at evaluating the effectiveness and usefulness of a task based on an activity of daily living (ADL) for the detection of cognitive impairment for an Alzheimer's disease (AD) population. Twenty-four participants were included in the study. The AD group (ADG) included twelve older adults (12 female) with AD (81.75±7.8 years). The Healthy group (HG) included twelve older adults (5 males, 77.7 ± 6.4 years). Both groups received a ADL-based intervention at two time frames separated 3 weeks. Cognitive functions were assessed before the interventions by using the MEC-35. The test-retest method was used to evaluate the reliability of the task, as well as the Intraclass Correlation Coefficient (ICC). The analysis of the test-retest reliability of the scores in the task indicated an excellent clinical relevance for both groups. The hypothesis of equality of the means of the scores in the two applications of the task was accepted for both the ADG and HG, respectively. The task also showed a significant high degree of association with the MEC-35 test (rho = 0.710, p = 0.010) for the ADG. Our results showed that it is possible to use an ADL-based task to assess everyday memory intended for cognitive impairments detection. In the same way, the task could be used to promote cognitive function and prevent dementia.**

***Index Terms*— ADL, Alzheimer's disease, cognitive decline, dementia, memory assessment.**

I. Introduction

ALZHEIMER'S disease (AD) is the most common type of dementia, which produces a gradual decline in control over memory, thought and language, abruptly limiting the independence of the patient [1]. Thus, the majority of people who have mild cognitive impairment (MCI) will develop AD in the future [2], [3]. In order to identify the appearance of problems in human cognition, several tools have been developed for their evaluation [4], [5]. However, some procedures are invasive and others are part of pharmacological trials, which causes discomfort and stigma [6]. When the patient prefers not to perform them, early detection of the disease and the respective treatment are difficult. Therefore, the specialist must choose the best option according to the needs of the patient, nature of the injury and cognitive impairment.

In a conventional way, there are screening instruments that make it possible to establish whether a person suffers from a cognitive problem, even if it is not visible. In the review conducted by Roeck *et al.* [7] it is mentioned that these tests might not be as sensitive and therefore effective for the detection of AD. Examples like the Mini-Mental State Examination (MMSE), Scenery Picture Memory Test (SPMT), Memory Impairment Screen (MIS), Alzheimer Quick Test (AQT), Montreal Cognitive Assessment (MoCA), Short Test of Mental Status (STMS) and Diagnostic utility of the Addenbrooke's Cognitive Examination - III (ACE-III), stand out among the main ones. [8]. Another disadvantage of these tools is their reduced ecological validity, since they are laboratory tests that do not evaluate their interaction with situations and conditions of the daily life (ADL) [9], [10]. This could skew the obtained data, since those people with AD have difficulty performing these types of activities, as a result of cognitive impairment.

It is important that these instruments have the support of technology, making better use of the results obtained [11]. For example, *et al.* [12] pointed out in their review the potential that designs based on Innovative Assistive Technologies have to compensate for functional decline. O'Leary *et al.* [13] developed a platform that has a touch screen for cognitive evaluation in rodents, which can be compared with human neuropsychological tests, obtaining greater advantages than using standard mazes. Other studies have used classical tests as the basis for the design of new proposals to be compared. Pitteri *et al.* [14] presented a video game that sought to identify whether there was a cognitive impairment in the information processing speed (IPS). The results revealed that those patients with multiple sclerosis (MP) showed a lower performance compared to healthy ones. Significant correlations with the Symbol Digit Modalities Test (SDMT) were found,

Manuscript received 22 December 2021; revised 16 June 2022; accepted 1 August 2022. Date of publication 4 August 2022; date of current version 12 August 2022. This work was supported in part by the University of Zaragoza under Grant OTRI 2015/0409. *(Corresponding author: G. Palacios-Navarro.)*

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Community of Aragón Institutional Review Board (CEICA) under Protocol No. PI21/326 in July 2021.

G. Palacios-Navarro is with the Department of Electronic Engineering and Communications, University of Zaragoza, 44003 Teruel, Spain (e-mail: guillermo.palacios@unizar.es).

J. Buele is with the Department of Electronic Engineering and Communications, University of Zaragoza, 44003 Teruel, Spain, and also with the Centro de Investigaciones de Ciencias Humanas y de la Educación, Universidad Tecnológica Indoamérica, Ambato 180103, Ecuador (e-mail: jorgebuele@uti.edu.ec).

S. Gimeno Jarque and A. Bronchal García are with the Teruel Center for Alzheimer's Disease, 44002 Teruel, Spain, and also with SOCIAL-CORE Ltd., 44002 Teruel, Spain (e-mail: sandragj16@gmail.com; andrea_bronchal@hotmail.com).

Digital Object Identifier 10.1109/TNSRE.2022.3196435

verifying its validity as a diagnostic tool. On the other hand, Khaligh-Razavi *et al.* [15] developed a new integrated cognitive assessment called CGN_ICA whose application lasts 5 minutes and is based on a visual categorization activity. Various natural images were presented to 448 participants to assess their cognitive performance and compared with the standard SDMT and MoCA tests. Similarly, Kalafatis *et al.* [16] used an artificial intelligence model to perform a cognitive test for cognitive impairment detection validating their efficacy in 230 participants. However, as pointed out by Marson *et al.* [17] in their review, more studies are required to evaluate the potentialities and limitations of these alternative tools during the treatment process of neurodegenerative diseases.

The aim of the study is to investigate new technological-based methods to detect cognitive decline in AD patients. The literature shows us that in spite of the acceptable results found when using technology to replace conventional methods, these alternative methods have a reduced ecological validity in predicting real-world performance [18], [19]. To accomplish that, this research proposes a more ecological task than traditional tests for cognitive impairment detection. The task represents an activity of daily living (ADL) in the sense that it is based on an activity that users can perform in their daily life. We can distinguish between basic ADL and instrumental ADL (iADL). The former include the fundamental skills needed to basic physical needs (hygiene, dressing, toileting, etc.). The latter include higher level activities such as the ability to use the telephone, shopping, food preparation, housekeeping, laundry, etc. [20]. According to Lawton [20] food preparation is an iADL and the proposed task is very close to this task. As secondary objectives, we intend to look for relationships among the obtained results and variables such as age, level of education, etc., in order to add more knowledge to the existing literature in cognitive decline.

II. Materials and Methods

This non-experimental study investigated the creation of new paradigms to detect changes in cognitive functions, such as cognitive impairment, in older adults with MCI, dementia or Alzheimer's disease through a game that emulates an activity of daily life. The study has been carried out in accordance with the Declaration of Helsinki. The protocol used in this study was approved by the Community of Aragón Research Ethics Committee (CEICA) on July 7, 2021 (protocol PI21/326) with the understanding and written consent of all the participants before the study.

A. Materials

For the development of the experiment we used a cupboard (figure 1) with the following dimensions: 120 cm width, 50 cm height, 24.5 cm depth. The cupboard was placed 102 cm above the ground so that all the participants were able to watch the inside properly. It consisted of four compartments of the same size (30 cm width, 50 cm height, 24.5 cm depth). Elements related to the kitchen, personal hygiene, sewing elements and elements found in the classroom where the participants attend

Fig. 1. Cupboards compartments with kitchen elements.

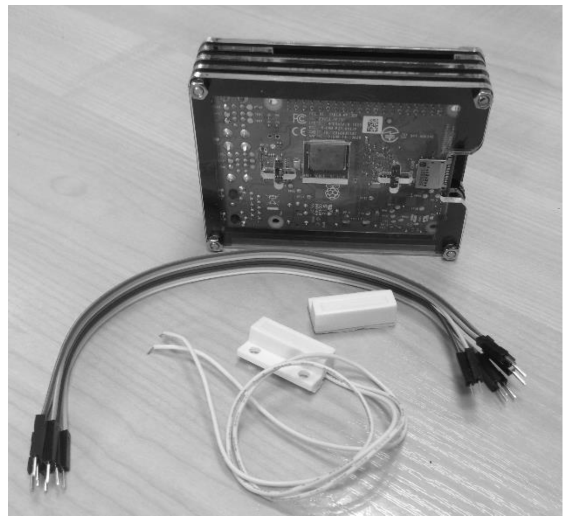

Fig. 2. Components used in the experiment: a Raspberry Pi, door sensors and cables.

their daily therapy session were also used. These elements were placed inside the compartments in the way indicated in the procedure section. Figure 1 shows the cupboard populated with elements one can find in a regular kitchen at home. Likewise, sensors were installed on the doors and a Raspberry Pi was placed under the cupboard in order to receive the inputs from each of the sensors and thereby store the data. Figure 2 depicts the main components used in the experiment (a Raspberry Pi 3, door sensors and cables).

B. Participants

Our design included two groups in our experiment. On one hand, the ADG, including participants diagnosed with Alzheimer's disease. On the other hand, we included a second group to take advantage from a second measurement using healthy people performing the same task (HG). By doing so, we wanted to check the ability of the test to differentiate between healthy people, and people with MCI, respectively.

A total of 13 older adults (all women) were recruited from the Teruel Center for Dementia and Alzheimer's disease. Inclusion criteria were: (1) AD diagnosis; and (2) ability to understand and follow verbal instructions. Exclusion criteria were: (1) secondary chronic disease that may affect cognitive functioning; (2) auditory and visual problems that may affect communication; and (3) physical limitations to access a one

TABLE I
PARTICIPANT DEMOGRAPHICS AND CLINICAL DATA (ADG)

Part. #	Age	Educational level	Months since diagnosis	MEC-35	Co-morbidity
1	87	11	24	22	No
2	75	8	48	19	Yes
3	89	7	48	28	Yes
4	77	8	136	21	Yes
5	87	8	27	22	Yes
6	87	8	114	25	Yes
7	62	11	42	27	No
8	81	8	152	26	Yes
9	85	10	30	21	No
10	87	7	84	17	Yes
11	87	8	36	28	Yes
12	77	9	48	29	Yes

TABLE II
PARTICIPANT DEMOGRAPHICS AND CLINICAL DATA (HG)

Part. #	Age	Gender	Educational level	MEC-35	Co-morbidity
1	78	M	6	35	Yes
2	70	F	6	35	No
3	88	F	6	35	Yes
4	73	F	6	35	No
5	80	M	6	35	No
6	82	F	6	35	No
7	65	M	13	35	No
8	77	F	10	35	No
9	78	F	6	35	No
10	83	M	6	35	No
11	84	F	6	35	Yes
12	74	M	6	35	No

meter-height cupboard and open the compartments. Based on the inclusion and exclusion criteria, only one subject was excluded due to her inability to reach the cupboard in a suitable way. All the participants had been diagnosed with AD.

The HG consisted of 12 participants with a maximum MEC-35 score as an inclusion criterium and we tried to get an age matched as close as possible.

C. Measurement

Demographic data of the participants were collected, including gender, age, educational level (years), time since diagnosis (months), cognitive function assessment by using the Cognitive mini-test (MEC-35) [21] which is the Spanish adaptation and validation of the MMSE [22] in its 35-item version. Comorbidity was also collected, including in this item the presence of visual difficulty, lack of inhibition, anxiety, processing speed, pulmonary fibrosis and use of cane, respectively. Participants in the ADG were aged between 62 and 89 years, (81.75 ± 7.8), with and educational level ranged between 7 and 11 (8.58 ± 1.38) years. The months since the disease was diagnosed ranged from 24 to 152. Table I shows the demographic and clinical data of the participants belonging to the ADG. Regarding the HG, participants were aged between 65 and 88 years, (77.7 ± 6.4), with and educational level ranged between 6 and 13 (6.92 ± 2.23) years. 41.66 % were male. Table II shows the data related to the HG.

D. Procedure

The task consisted of administering 4 subtasks of 8 items each (32 items in the global task). It was accomplished on four consecutive days, each day dedicated to a subtask. In the first subtask, 8 objects that can be found in a kitchen were placed inside the cupboard, while in the second subtask the items were related to personal hygiene items. The third was dedicated to sewing items while in the last subtask the participants had to memorize items found in the classroom where they work every day. Each of the subtasks consists of two phases. In the first phase (coding phase), the user was

Fig. 3. AD patient performing the task.

placed in front of the cupboard and was asked to memorize the elements of each of the compartments during 10 seconds [23], (40 seconds for the total of the cupboard). In the second phase (recovery phase), the participant had to locate the object required by the therapist by selecting a door. For example, they were asked questions like: Where is the milk? For any of the subtasks, the user has a time limit to choose the corresponding door which was fixed to 10 seconds. If this time was exceeded, then a failure was written down and the evaluation continued. Items were randomly placed in the compartments and shelves and doors could be opened individually. Accuracy responses and reaction times were recorded in plain text files to be saved in the local database and sent to a remote server (SFTP) for a subsequent analysis. The test was conducted in the room where the AD patients attend their daily therapy. The test was repeated after an interval of three weeks in order to perform a test-retest reliability analysis. Figure 3 shows an AD patient performing the task (the person depicted agreed to the use of her image).

E. Statistical Analysis

Statistical analysis was carried out using SPSS 16. Shapiro-Wilk's test was performed to check for data normality (the sample size less than 30 participants). The variables age, months from diagnosis and level of education do not follow a normal distribution (Shapiro-Wilk's test no significant, $p > 0.05$), so Spearman's rho correlation was performed when checking for any association between accuracy response and age, years of education, time since diagnosis, as well as in the case of any association between reaction time and age. A Spearman's correlation analysis was also used to check any association between the MCE-35 test results and our task results. A repeated measures ANOVA was performed to check for differences in accuracy response among the different subtasks, whereas a paired t-test was performed to check for differences between the test-retest measures. The t student was used to check for differences between groups (HG vs ADG) as well as one-way ANOVA with repeated measures. We also performed a reliability analysis by reckoning the intraclass coefficient (ICC) with a two-way mixed effects model and absolute agreement, aimed at assessing the concordance of the test-retest measures.

III. Results

In this section we present the results obtained in terms of test performance (accuracy response), task completion times (reaction times), reliability analysis using the test-retest method, and the degree of association between the MEC-35 test scores (only for the ADG) and our task results, respectively. As secondary results, we try to analyze the relationships or degree of association between the different independent variables with the accuracy response (AR) and the reaction times (RT), respectively. Table III shows the obtained results.

TABLE III
Statistical Results for Both the ADG and HG

Measure	ADG	HG
accuracy vs age	rho=0.182, p=0.572	rho=-0.736, p=0.006.
accuracy vs years o. education	rho=-0.035, p=0.913	rho=0.410, p=0.185
accuracy vs time s. diagnosis	rho=0.094, p=0.772	-
Accuracy vs MEC-35	rho=0.710, p=0.01	-
accuracy vs gender	-	t=0.394, p=0.702
Co-morb vs no-co-morb. (accuracy)	t(10) = -6.41, p=0.535	t(10) = 1.051, p=0.318
Repeated measures ANOVA different subtasks (accuracy)	F=0.819, p=0.385	F=35.203, p<0.001
RT vs age	rho=0.472, p=0.121	rho=0.340, p=0.280
RT vs. time s. diagnosis	rho=-0.212, p=0.508,	-
Co-morb vs no-co-morb (total task time)	t(10) = -2.419, p=0.036	t(10) = -0.169, p=869
ICC	0.757	0.821
Paired t-test	t(11) = 1.52, p=0.156	t(11)=1.125, p=0.272

A. Accuracy Response (AR)

With respect to the ADG, we found that age is not related to the accuracy response either in any of the subtasks or in the total task (adding all the subtasks scores). Performance in the task does not depend either on the level of training of the participants or the time since diagnosis of the disease. The participants did not show significant differences in the accuracy response regarding whether or not they had any type of morbidity ($t\ (10) = -6.41$, $p = 0.536$). The analysis of repeated measures to see if there were differences in the accuracy response of each of the subtasks revealed that there were no significant differences between them ($F = 0.819$, $p = 0.385$). Figure 4 depicts the accuracy response results in percentage terms versus the different subtasks for the ADG at baseline and after 3 weeks. Finally, the Spearman's correlation coefficient showed a significant (large and positive) association between the results of the MEC-35 scale and the results obtained with our task ($p = 0.01$, $rho = 0.710$), respectively. This coefficient is large, according to Cohen [24].

With respect to the HG, we found that age was related to the accuracy response in the whole task ($rho=-0.736$, $p=0.006$. Therefore, the older the participant, the worse performance. Performance in the task did not depend either on the level of training of the participants or gender. The participants did not show significant differences in the accuracy response regarding whether or not they had any type of morbidity ($t\ (10) = 1.051$, $p = 0.318$). Unlike the ADG, the analysis of repeated measures revealed significant differences in the accuracy response among the subtasks ($F = 35.203$, $p < 0.001$). Finally, we present the accuracy response results versus subtasks and group in figure 5. Figure 6 depicts the accuracy results for the whole sample (24 participants).

B. Reaction Time (RT)

Regarding the possible associations with the independent variables and the completion times of the respective tasks, we can highlight the following. As far as the ADG is concerned, we did not find any association between the age of the participants and the completion times of the different tasks and the total task. There was also no association between the months since diagnosis and such RT. Regarding co-morbidity and RT, we found significant differences between participants with some type of morbidity in the classroom subtask ($t\ (10) = -2.62$, $p = 0.026$) as well as in the time spent in the whole task ($t\ (10) = -2.419$, $p = 0.036$). In cases of co-morbidity, execution times have been longer in comparison with the no co-morbidity condition.

Regarding the HG, we did not find any association between the age of the participants and the completion times of the different tasks and the total task. Co-morbidity was not either associated to the total completion time of the task (RT). In cases of co-morbidity, execution times have been longer in comparison with the no co-morbidity condition, but non-significant.

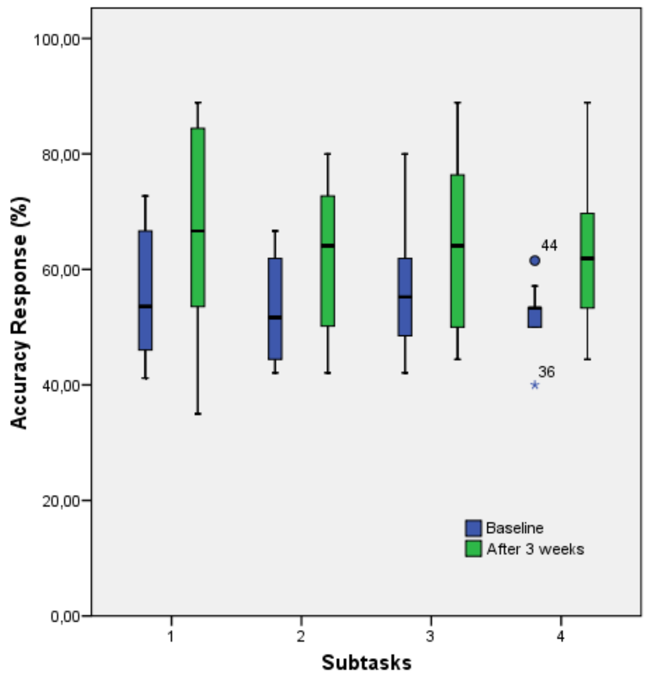

Fig. 4. Accuracy response results (%) vs. subtasks for the ADG at baseline and after 3 weeks. Subtask 1: kitchen, subtask 2: personal hygiene, subtask 3: sewing elements, subtask 4: classroom elements.

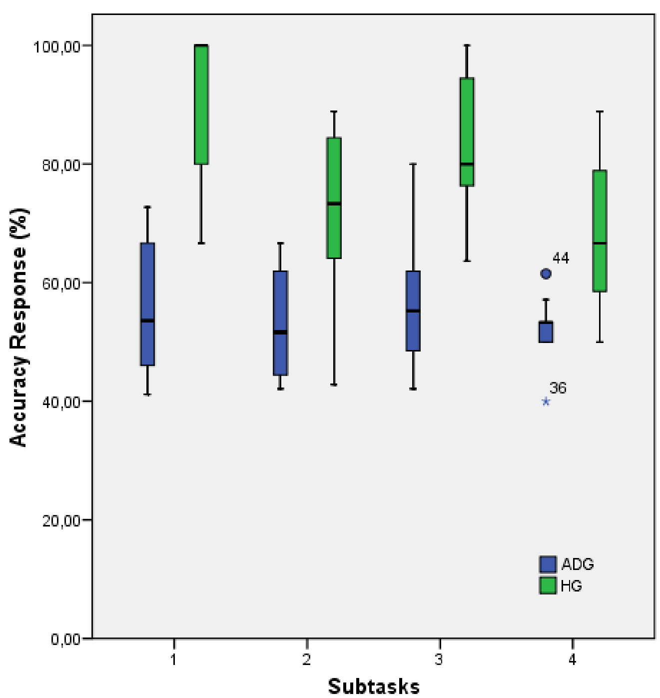

Fig. 5. Accuracy response results (%) versus subtasks for both the ADG and the HG, respectively. Subtask 1: kitchen, subtask 2: personal hygiene, subtask 3: sewing elements, subtask 4: classroom elements.

C. Test Retest Analysis

In order to evaluate the degree of reliability of our task, we performed a test-retest method, applying the same test three weeks after doing it for the first time. This is an adequate analysis since both memory and cognitive impairment

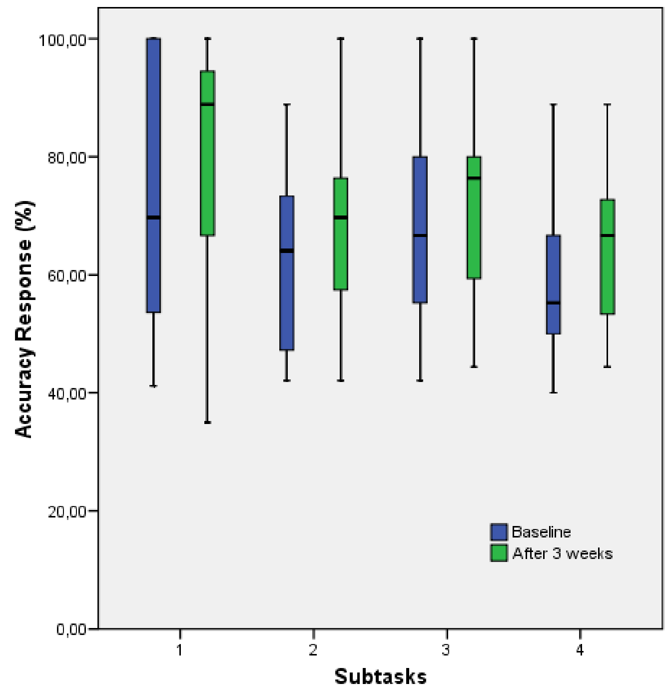

Fig. 6. Accuracy response results (%) for the 24 participants at baseline and after 3 weeks. Subtask 1: kitchen, subtask 2: personal hygiene, subtask 3: sewing elements, subtask 4: classroom elements.

are stable attributes. The degree of agreement between the measures reflected an Intraclass Correlation Coefficient (ICC) of 0.757 for the ADG, which corresponds to an excellent level of clinical significance according to Cicchetti [25]. No significant differences were found between the means of the scores of the task in the two moments in which it was applied ($t\ (11) = 1.52$; $p = 0.156$). With respect to the HG, we obtained an ICC equals to 0.821, also considered as excellent. In the same way, no significant differences were found between the means of the scores of the task in the two moments in which it was applied ($t(11)=1.125, p=0.272$). The ICC for the whole group was 0.904, an excellent level of clinical significance, too.

D. Pairwise Comparison

We found significant differences between groups in accuracy response in the whole task ($t=5.969$, $p<0.001$). The HG accuracy response mean in the whole task was 78.35 % whereas the ADG accuracy response mean was 54.10 %, respectively. Consequently, the resulting effect size ($d=2.42$) was very large according to Sawilowsky [26]. In the same way the repeated measures one-way ANOVA revealed between subject differences ($F=45.02$, $p<0.001$). We also found statistically significant differences within subjects ($F=27.6$, $p<0.001$) in the different subtasks. The interaction subtask*group was also significant ($F=27.608$, $p<0.001$). Figure 7 shows the estimated marginal means in accuracy response versus subtasks and group. Regarding completion times, we found statistically significant differences between groups ($t=-4.064$, $p=0.001$), with a large effect size ($d=1.66$) according to Cohen [24].

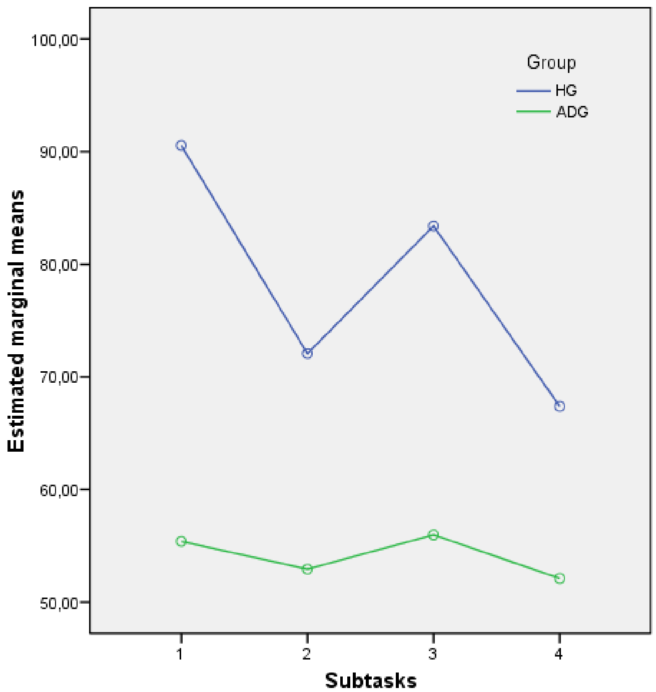

Fig. 7. Accuracy response estimated marginal means vs subtasks and group: ADG and HG. Subtask 1: kitchen, subtask 2: personal hygiene, subtask 3: sewing elements, subtask 4: classroom elements.

IV. Discussion

The main objective of our study was to evaluate the usefulness of a task based on an ADL for the detection of cognitive impairment. The results highlighted a strong positive and significant association with respect to the MEC-35 scale (Spearman rho=0.71) according to Cohen [24]. This results aligns with other works, like the study of Pitteri *et al.* [14], who also found significant correlations between the SDMT and the accuracy response of their task, thus verifying its validity as a diagnostic tool. Similarly, Kalafatis *et al.* [16] demonstrated a convergent validity with the MOCA test (Pearson r = 0.58) and the ACE test (Pearson r = 0.62), respectively. In both cases they found a slightly lower association than that achieved in our study. In the same way, the reliability values (calculated through the ICC) obtained in the test-retest analysis for both separate groups and considered as a whole sample, reflect an excellent level of clinical significance and therefore give robustness to our proposal.

Although a high correlation is not necessarily synonymous with agreement between methods, this task could hint signs of a cognitive impairment. As mentioned by Nakhla *et al.* [10], people with dementia tend to show lower performance in applications that involve ADLs, particularly in activities that involve memory and verbal fluency. Furthermore, some studies have already shown their ability to detect executive dysfunctions in daily live conditions that were underestimated by these classical tests [27].

The second measurement, using healthy participants performing the same task, has highlighted the usefulness of the test to differentiate between healthy people and people with MCI. The results obtained with the HG have been congruent with those obtained with the ADG, which indicates the consistency of the developed task. Nevertheless, we found some ceiling effects in accuracy performance for the HG only in the kitchen subtask. That is, more than 20% of participants scored 100% [28]. In particular, 7 out of 12 participants got a 100 %, but no one exceeded a 90% accuracy response in the total task (sum of the 4 subtasks), so there was no ceiling effect in the whole task. We find a plausible reason in the fact that participants may be more familiar with the objects in the kitchen and of course they do not have any kind of cognitive impairment. Similarly, Holmes and Shea [28] also found an important ceiling effect when using the instrument for cognitive abilities screening (CASI) to assess global cognitive function in people with dementia. Specifically, 4 cognitive domains were affected by this effect (4 out of 7). However, as in our case, the total score did not show any ceiling effect.

The ceiling effect is quite common when the tests are applied to nonclinical populations, such as the notable ceiling effects found in the RAVLT test [29], which limits its use in populations without evident memory impairments. Therefore, it is often necessary to introduce modifications to such tests or memory tasks. Hale *et al.* had to introduce modifications in their study in order to avoid these effects [30]. Something similar was done by Ivanoiu *et al.* [31] with the free delayed recall test, one of the tests with the highest sensitivity for early diagnosis of dementia. The authors designed a new memory test based on cued recall to avoid ceiling effects in early diagnosis of AD. Wester *et al.* found a reduction in the problem of ceiling and floor effects in the revised Rivermead Behavioral Memory Test (RBMT-3) versus the traditional RBMT [32].

A conclusion that we can obtain from our task is that it is suitable that the objects are known but they must be varied. To reduce the ceiling effect, we could add a little more difficulty, especially in cases of mild cognitive impairment. However, in general we can say that the tool has shown an acceptable discriminative validity when we administer it jointly (as a whole task). Nevertheless, we also agree with many authors that claim the necessity of more research to improve the ecological validity of this kind of memory tasks [32].

Regarding reaction times, the large effect size found between the groups precisely reflects the good cognitive status of the HG participants with respect to the ADG, in addition to the already reflected effect of comorbidity, which affects much more (and in a significant way) to the second group. In addition, the mean age of the ADG was slightly higher than the HG. We found studies in which similar memory tasks were validated (although in virtual environments), such as the one performed by Ouellet *et al.* [33], in which older adults spent more time than young adults in the completion of an everyday memory assessment task. Similarly, Moffat *et al.* [34] showed that the youngest age group performed a spatial memory task much faster than two other older groups of participants.

With respect to the already known age-related differences effect observed on episodic memory tasks, in our task, age was not independently associated with the accuracy response, either in any of the subtasks or in the total task, but only in the ADG. Nevertheless, the HG exhibited a strong (and negative) association between both variables. In the ADG, the low variability of age and the cognitive decline may explain this lack of association. By contrast, similar studies involving healthy participants have corroborated an age-related decrease in performance when dealing with spatial memory [35], [34] or episodic memory [36], [37]. In these cases, there was also a lot of variability in the age of the participants. This fact may lead us to think that certain tasks may be sensitive between age groups with very little variation within the same age group. Therefore, this detail should be taken into account in the design of these tasks, since both age and gender (including cultural background) could have a relevant influence on the results obtained, as described by Coutrot *et al.* [35].

Regarding the level of education of the subjects and the accuracy response, in both cases (individual subtasks and global task) and both groups there was no significant association between both variables. This result is in line with the study of García-Magariño *et al.* [38] despite the fact that the average level of training of the participants in their study (15.09 $\pm$ 4.83) was much higher than ours (8.58 $\pm$ 1.38), which makes our task very robust in this regard. This contrasts with other studies such as that of Chua *et al.* [11], who found that the range in the educational level of participants was so extraordinarily wide that it became a limitation of their study. To try to tackle the problem, they proposed in the future to stratify their participants by their level of education.

A. Clinical Implications

The introduction of this task as an exercise to be carried out within the daily activity of the AD patient can be very beneficial. On one hand, it would allow the therapist to better monitor the progress of the disease in an automatic way. On the other hand, it could also warn of a possible cognitive deterioration without resorting to the performance of traditional cognitive tests, since the deficiencies in the development of these activities can represent an important and consistent predictor of cognitive deterioration. Early detection of possible cognitive impairment is especially critical in the early stages of the disease and when certain symptoms are not so noticeable, because the pre-dementia stage is quite extensive [39]. For these reasons, such a task could be used when the patient is still relatively young, where the disease evolves more quickly and immediate treatment would have better results [15].

B. Limitations of the Study

The present study has obvious limitations. On one hand, the small size of the sample means that the results, although they are very promising, have to be confirmed in RTC studies with larger samples, in order to offer better evidence. In fact, there are few studies with sufficient evidence that present validated proposals for the diagnosis of cognitive problems. In the same way, cognitive training requires greater support so that it can be considered an effective strategy to promote the recovery of these functions [40], [41].

One of the problems we have found when developing the experience, is that people with some type of co-morbidity (especially cane carriers) have presented higher response times. This directly affects the performance of people when developing ADLs as mentioned by De Vriendt *et al.* [39], since the literature shows that reaction time can be an important variable to be considered [11], [34], [36].

C. Future Work

To overcome the abovementioned problems due to the comorbidity of some patients, we propose in the future the use of virtual reality (VR), either on its non-immersive mode (Tablet or desktop computer) or through an immersive virtual environment (via Head-mounted displays), to overcome these disadvantages. In this case, the subject would be sitting in a chair with total comfort and could be completely concentrated on the task to be carried out. In both cases, the aim is to automate this process in such a way that the therapist has the data through a mobile application and can carry out a more detailed follow-up of each of the patients, with the establishment of alert mechanisms. Thus, the implementation of advanced ADLs is also proposed as future work, since as mentioned by Fieo *et al.* [42], they produce better results than basic ones. We also want to bring this first experience to other groups such as stroke patients, due to their evident needs to promote cognitive functions and prevent dementia. Finally, we see a clear potential of our task to train people with some cognitive impairment in ADLs, as carried out by Foloppe *et al.* in his study [9].

V. Conclusion

With this work, we have demonstrated the usefulness of a memory task based on an ADL for the detection of cognitive impairments. Despite having a small sample of participants, the inclusion of a group of healthy older adults has shown the ability of the task to differentiate between healthy people and people with cognitive impairments. Consequently, the study could be used for early detection of the disease, when certain symptoms that are normally overlooked can be detected, and with this, treatment can be administered in time. The excellent level of clinical significance obtained through the test-retest reliability test give robustness to our proposal.

Acknowledgment

The authors want to thank all the participants for their effort and commitment, the area of Social Services, Family, Equality of the Teruel Town Council as well as SOCIALCORE Ltd. for their support.

References

[1] C. A. Lane, J. Hardy, and J. M. Schott, "Alzheimer's disease," *Eur. J. Neurol.*, vol. 25, pp. 59–70, Jan. 2018, doi: 10.1111/ene.13439.
[2] R. C. Petersen, "Mild cognitive impairment as a diagnostic entity," *J. Internal Med.*, vol. 256, no. 3, pp. 183–194, 2004, doi: 10.1111/j.1365-2796.2004.01388.x.
[3] Y. Liu, W. Tan, C. Chen, C. Liu, J. Yang, and Y. Zhang, "A review of the application of virtual reality technology in the diagnosis and treatment of cognitive impairment," *Frontiers Aging Neurosci.*, vol. 11, p. 280, Oct. 2019, doi: 10.3389/fnagi.2019.00280.

[4] B. Dubois, A. Padovani, P. Scheltens, A. Rossi, and G. Dell'Agnello, "Timely diagnosis for Alzheimer's disease: A literature review on benefits and challenges," *J. Alzheimer's Disease*, vol. 49, no. 3, pp. 617–631, Oct. 2015, doi: 10.3233/JAD-150692.

[5] J. M. F. Montenegro and V. Argyriou, "Cognitive evaluation for the diagnosis of Alzheimer's disease based on Turing test and virtual environments," *Physiol. Behav.*, vol. 173, pp. 42–51, May 2017, doi: 10.1016/j.physbeh.2017.01.034.

[6] G. D. Stanciu *et al.*, "Alzheimer's disease pharmacotherapy in relation to cholinergic system involvement," *Biomolecules*, vol. 10, no. 1, p. 40, Dec. 2019, doi: 10.3390/biom10010040.

[7] E. E. De Roeck, P. P. De Deyn, E. Dierckx, and S. Engelborghs, "Brief cognitive screening instruments for early detection of Alzheimer's disease: A systematic review," *Alzheimer's Res. Therapy*, vol. 11, no. 1, pp. 1–14, Feb. 2019, doi: 10.1186/s13195-019-0474-3.

[8] M. Senda *et al.*, "Diagnostic utility of the Addenbrooke's cognitive examination—III (ACE-III), mini-ACE, mini-mental state examination, Montreal cognitive assessment, and Hasegawa dementia scale-revised for detecting mild cognitive impairment and dementia," *Psychogeriatrics*, vol. 20, no. 2, pp. 156–162, Mar. 2020, doi: 10.1111/psyg.12480.

[9] D. A. Foloppe, P. Richard, T. Yamaguchi, F. Etcharry-Bouyx, and P. Allain, "The potential of virtual reality-based training to enhance the functional autonomy of Alzheimer's disease patients in cooking activities: A single case study," *Neuropsychol. Rehabil.*, vol. 28, no. 5, pp. 709–733, Jul. 2018, doi: 10.1080/09602011.2015.1094394.

[10] M. Z. Nakhla, D. Banuelos, C. Pagán, A. G. Olvera, and J. Razani, "Differences between episodic and semantic memory in predicting observation-based activities of daily living in mild cognitive impairment and Alzheimer's disease," *Appl. Neuropsychol., Adult*, pp. 1–12, Mar. 2021. [Online]. Available: https://www.tandfonline.com/doi/full/10.1080/23279095.2021.1893172, doi: 10.1080/23279095.2021.1893172.

[11] S. I. L. Chua *et al.*, "Virtual reality for screening of cognitive function in older persons: Comparative study," *J. Med. Internet Res.*, vol. 21, no. 8, Aug. 2019, Art. no. e14821, doi: 10.2196/14821.

[12] B. Thordardottir, A. M. Fänge, C. Lethin, D. R. Gatta, and C. Chiatti, "Acceptance and use of innovative assistive technologies among people with cognitive impairment and their caregivers: A systematic review," *BioMed Res. Int.*, vol. 2019, pp. 1–18, Mar. 2019, doi: 10.1155/2019/9196729.

[13] J. D. O'Leary, O. F. O'Leary, J. F. Cryan, and Y. M. Nolan, "A low-cost touchscreen operant chamber using a raspberry Pi," *Behav. Res. Methods*, vol. 50, no. 6, pp. 2523–2530, Dec. 2018, doi: 10.3758/s13428-018-1030-y.

[14] M. Pitteri, C. Dapor, S. Ziccardi, M. Guandalini, R. Meggiato, and M. Calabrese, "Visual-attentional load unveils slowed processing speed in multiple sclerosis patients: A pilot study with a tablet-based videogame," *Brain Sci.*, vol. 10, no. 11, p. 871, Nov. 2020, doi: 10.3390/brainsci10110871.

[15] S.-M. Khaligh-Razavi *et al.*, "Integrated cognitive assessment: Speed and accuracy of visual processing as a reliable proxy to cognitive performance," *Sci. Rep.*, vol. 9, no. 1, pp. 1–11, Jan. 2019, doi: 10.1038/s41598-018-37709-x.

[16] C. Kalafatis *et al.*, "Validity and cultural generalisability of a 5-minute AI-based, computerised cognitive assessment in mild cognitive impairment and Alzheimer's dementia," *Frontiers Psychiatry*, vol. 12, p. 1155, Jul. 2021, doi: 10.3389/fpsyt.2021.706695.

[17] F. Marson, S. Lasaponara, and M. Cavallo, "A scoping review of neuromodulation techniques in neurodegenerative diseases: A useful tool for clinical practice?" *Medicina*, vol. 57, no. 3, p. 215, Mar. 2021, doi: 10.3390/medicina57030215.

[18] A. Ruet and B. Brochet, "Cognitive assessment in patients with multiple sclerosis: From neuropsychological batteries to ecological tools," *Ann. Phys. Rehabil. Med.*, vol. 63, no. 2, pp. 154–158, Mar. 2020, doi: 10.1016/j.rehab.2018.01.006.

[19] T. D. Parsons and T. Mcmahan, "An initial validation of the virtual environment grocery store," *J. Neurosci. Methods*, vol. 291, pp. 13–19, Nov. 2017.

[20] M. P. Lawton and E. M. Brody, "Assessment of older people: Self-maintaining and instrumental activities of daily living," *Gerontologist*, vol. 9, pp. 179–186, Oct. 1969.

[21] A. Lobo, J. Ezquerra, F. G. Burgada, J. M. Sala, and A. S. Díaz, "El miniexamen, cognoscitivo (un test sencillo, práctico, para detectar alteraciones intelectuales en pacientes médicos)," *Actas Luso-Espanolas de Neurologia Psiquiatria y Ciencias Afines*, vol. 7, no. 3, pp. 189–202, 1979.

[22] M. F. Folstein, S. E. Folstein, and P. R. McHugh, "'Mini-mental state'. A practical method for grading the cognitive state of patients for the clinician," *J. Psychiatric Res.*, vol. 12, no. 3, pp. 189–198, Nov. 1975, doi: 10.1016/0022-3956(75)90026-6.

[23] F. González-Landero, I. García-Magariño, R. Amariglio, and R. Lacuesta, "Smart cupboard for assessing memory in home environment," *Sensors*, vol. 19, no. 11, p. 2552, Jun. 2019, doi: 10.3390/s19112552.

[24] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed. Hillsdale, NJ, USA: Erlbaum Associates, 1988.

[25] D. V. Cicchetti, "Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology," *Psychol. Assessment*, vol. 6, no. 4, pp. 284–290, 1994.

[26] S. S. Sawilowsky, "New effect size rules of thumb," *J. Modern Appl. Stat. Methods*, vol. 8, no. 2, pp. 597–599, Nov. 2009, doi: 10.22237/jmasm/1257035100.

[27] O. Rouaud *et al.*, "Contribution of ecological evaluation of executive disorders in multiple sclerosis," *Revue Neurologique*, vol. 162, no. 10, pp. 964–969, Oct. 2006, doi: 10.1016/s0035-3787(06)75106-2.

[28] W. C. Holmes and J. A. Shea, "Performance of a new, HIV/AIDS-targeted quality of life (HAT-QoL) instrument in asymptomatic seropositive individuals," *Qual. Life Res.*, vol. 6, no. 6, pp. 561–571, 1997.

[29] V. Drolet *et al.*, "A comparison of the performances between healthy older adults and persons with Alzheimer's disease on the rey auditory verbal learning test and the test de rappel libre/rappel indicé 16 items," *Gériatrie et Psychologie Neuropsychiatrie du Vieillissement*, vol. 12, no. 2, pp. 218–226, Jun. 2014, doi: 10.1684/pnv.2014.0469.

[30] C. Hale *et al.*, "The ModRey: An episodic memory test for nonclinical and preclinical populations," *Assessment*, vol. 26, no. 6, pp. 1154–1161, Sep. 2019, doi: 10.1177/1073191117723113.

[31] A. Ivanoiu *et al.*, "Memory evaluation with a new cued recall test in patients with mild cognitive impairment and Alzheimer's disease," *J. Neurol.*, vol. 252, no. 1, pp. 47–55, Jan. 2005, doi: 10.1007/s00415-005-0597-2.

[32] A. J. Wester, P. Leenders, J. I. M. Egger, and R. P. C. Kessels, "Ceiling and floor effects on the rivermead behavioural memory test in patients with alcohol-related memory disorders and healthy participants," *Int. J. Psychiatry Clin. Pract.*, vol. 17, no. 4, pp. 286–291, Oct. 2013, doi: 10.3109/13651501.2013.813553.

[33] É. Ouellet, B. Boller, N. Corriveau-Lecavalier, S. Cloutier, and S. Belleville, "The virtual shop: A new immersive virtual reality environment and scenario for the assessment of everyday memory," *J. Neurosci. Methods*, vol. 303, pp. 126–135, Jun. 2018, doi: 10.1016/j.jneumeth.2018.03.010.

[34] S. D. Moffat, A. B. Zonderman, and S. M. Resnick, "Age differences in spatial memory in a virtual environment navigation task," *Neurobiol. Aging*, vol. 22, no. 5, pp. 787–796, Sep. 2001, doi: 10.1016/S0197-4580(01)00251-2.

[35] A. Coutrot *et al.*, "Global determinants of navigation ability," *Current Biol.*, vol. 28, no. 17, pp. 2861–2866, Sep. 2018, doi: 10.1016/j.cub.2018.06.009.

[36] N. C. Lecavalier, É. Ouellet, B. Boller, and S. Belleville, "Use of immersive virtual reality to assess episodic memory: A validation study in older adults," *Neuropsychol. Rehabil.*, vol. 30, no. 3, pp. 462–480, Mar. 2020, doi: 10.1080/09602011.2018.1477684.

[37] M. O. Pflueger, R.-D. Stieglitz, P. Lemoine, and T. Leyhe, "Ecologically relevant episodic memory assessment indicates an attenuated age-related memory loss—A virtual reality study," *Neuropsychology*, vol. 32, no. 6, pp. 680–689, Sep. 2018, doi: 10.1037/neu0000454.

[38] I. García-Magariño, J. T. Fox-Fuller, G. Palacios-Navarro, A. Baena, and Y. T. Quiroz, "Visual working memory for semantically related objects in healthy adults," *Rev Neurol*, vol. 71, no. 8, pp. 277–284, 2020.

[39] P. De Vriendt, E. Cornelis, W. Cools, and E. Gorus, "The usefulness of evaluating performance of activities in daily living in the diagnosis of mild cognitive disorders," *Int. J. Environ. Res. Public Health*, vol. 18, no. 21, p. 11623, Nov. 2021, doi: 10.3390/ijerph182111623.

[40] T. B. Cumming, K. Tyedin, L. Churilov, M. E. Morris, and J. Bernhardt, "The effect of physical activity on cognitive function after stroke: A systematic review," *Int. Psychogeriatrics*, vol. 24, no. 4, pp. 557–567, 2012, doi: 10.1017/S1041610211001980.

[41] M. N. McDonnell, A. E. Smith, and S. F. Mackintosh, "Aerobic exercise to improve cognitive function in adults with neurological disorders: A systematic review," *Arch. Phys. Med. Rehabil.*, vol. 92, no. 7, pp. 1044–1052, Jul. 2011, doi: 10.1016/j.apmr.2011.01.021.

[42] R. Fieo, L. Zahodne, M. X. Tang, J. J. Manly, R. Cohen, and Y. Stern, "The historical progression from ADL scrutiny to IADL to advanced ADL: Assessing functional status in the earliest stages of dementia," *J. Gerontol., Ser. A*, vol. 73, no. 12, pp. 1695–1700, Nov. 2018, doi: 10.1093/gerona/glx235.