Journal of the International Neuropsychological Society (2014), **20**, 1–10

doi:10.1017/S1355617714000344

Detecting Everyday Action Deficits in Alzheimer's Disease Using a Nonimmersive Virtual Reality Kitchen

Philippe Allain,[1,2] Déborah Alexandra Foloppe,[1,3], Jérémy Besnard,[1] Takehiko Yamaguchi,[4] Frédérique Etcharry-Bouyx,[1,2] Didier Le Gall,[1,2] Pierre Nolin,[5] AND Paul Richard[3]

[1]LUNAM Université, Laboratoire de Psychologie des Pays de Loire (UPRES EA 4638), Université d'Angers, France
[2]Unité de Neuropsychologie, Département de Neurologie, CHU Angers, France
[3]LUNAM Université, Laboratoire Angevin de Recherche en Ingénierie des Systèmes (UPRES EA 7315), Université d'Angers, France
[4]Department of applied electronics, Faculty of Industrial Science and Technology, Tokyo University of Science, Tokyo, Japan
[5]Laboratoire de Recherche Interdisciplinaire en Réalité Virtuelle, Université du Québec à Trois-Rivières, Trois-Rivières, Quebec, Canada

(RECEIVED May 20, 2013; FINAL REVISION March 8, 2014; ACCEPTED March 17, 2014)

Abstract

Alzheimer's disease (AD) causes impairments affecting instrumental activities of daily living (IADL). Transdisciplinary research in neuropsychology and virtual reality has fostered the development of ecologically valid virtual tools for the assessment of IADL, using simulations of real life activities. Few studies have examined the benefits of this approach in AD patients. Our aim was to examine the utility of a non-immersive virtual coffee task (NI-VCT) for assessment of IADL in these patients. We focus on the assessment results obtained from a group of 24 AD patients on a task designed to assess their ability to prepare a virtual cup of coffee, using a virtual coffee machine. We compared performance on the virtual task to an identical daily living task involving the actual preparation of a cup of coffee, as well as to global cognitive, executive, and caregiver-reported IADL functioning. Relative to 32 comparable, healthy elderly (HE) controls, AD patients performed worse than HE controls on all tasks. Correlation analyses revealed that NI-VCT measures were related to all other neuropsychological measures. Moreover, regression analyses demonstrated that performance on the NI-VCT predicted actual task performance and caregiver-reported IADL functioning. Our results provide initial support for the utility of our virtual kitchen for assessment of IADL in AD patients. (*JINS*, 2014, *20*, 1–10)

Keywords: Alzheimer's disease, Virtual reality, Action, Activities of daily living (ADL), Instrumental activities of daily living (IADL), Neuropsychological assessment

INTRODUCTION

Alzheimer's disease (AD) is defined by memory deficits and a decline in another cognitive domain (e.g., executive functions), severe enough to impede functioning in everyday life. Initially, AD patients have difficulty with complex Instrumental Activities of Daily Living [IADL: making a cup of tea (Rusted & Sheppard, 2002)]. These deficits are followed by a progressive decline in basic Activities of Daily Living [ADL: toileting (Millán-Calenti et al., 2012)]. Giovannetti, Schmidt, Gallo, Sestito and Libon (2006) noted that IADL/ADL impairments are associated with serious consequences, including caregiver burden (De Bettignies, Mahurin, & Pirozzolo, 1993), institutionalization (Mast, Azar, MacNeill, & Lichtenberg, 2004), depression (Adam, Van Der Linden, Juillerat, & Salmon, 2000), and death (Noale et al., 2003).

IADL/ADL assessment in AD patients is mainly done with functional scales based on self- or informant-report, such as the Lawton-Brody IADL scale (Lawton & Brody, 1969). In AD, caregiver ratings on these scales have been found to correlate with cognitive and behavioural impairments (Barberger-Gateau, Fabrigoule, Rouch, Letenneur, & Dartigues, 1999; Boyle et al., 2003; Cahn-Weiner, Ready, & Malloy, 2003; Lechowski et al., 2003; Senanarong et al., 2005; Tekin, Fairbanks, O'Connor, Rosenberg, & Cummings, 2001). However, this method poses several challenges (Mitchell et al., 2011). AD patients may underestimate their functional impairment in everyday situations because of anosognosia (Starkstein, Brockman, Bruce, & Petracca, 2010). It also offers a very gross

Correspondence and reprint requests to: Philippe Allain, Unité de Neuropsychologie, Département de Neurologie, Centre Hospitalier Universitaire, 4 rue Larrey, 49033 Angers, Cedex 09, France. E-mail: phallain@chu-angers.fr

assessment of performance, instead of a more nuanced characterization of ADL-IADL deficits (Giovannetti, Bettcher, Brennan, Libon, Burke, et al., 2008; Giovannetti, Bettcher, Brennan, Libon, Kessler, 2008).

In neuropsychology, few objective methods for assessing ADL-IADL are available. Most traditional tests were developed for assessing performance in a specific domain (e.g., executive functions) or the ability to use a particular cognitive function (e.g., planning). These tests measure cognitive abilities in isolated and artificial situations, which bear little similarity to the situation that patients encounter in their daily life. These tests have low ecological validity and limited ability to predict actual functioning in ADL-IADL (Chaytor, Schmitter-Edgecombe, & Burr, 2006).

The Naturalistic Action Test (NAT; Schwartz, Segal, Veramonti, Ferraro, & Buxbaum, 2003) is one of the tests that overcome these problems. It is a standardized measure that requires completion of three multiple-step everyday tasks (e.g., preparing coffee with cream and sugar). Performance is usually videotaped, and the NAT yields detailed performance variables reflecting the percent of steps accomplished, number of errors committed (omission and commission), and time to completion. Giovannetti, Bettcher, Brennan, Libon, Kessler, et al. (2008) reported that AD patients demonstrated difficulty accomplishing steps involved in each task, performing steps accurately, and inhibiting irrelevant, off-task actions. Inconsistent with the conclusion of Chaytor et al. (2006), Giovannetti, Bettcher, Brennan, Libon, Kessler, et al. (2008) found that action errors in the NAT were related to classical neuropsychological measures in AD patients, with omission errors predicted by global cognitive efficiency and episodic memory measures and commission errors predicted by executive control measures. Giovannetti, Schwartz, and Buxbaum (2007) recognized that while the NAT has been useful in characterizing everyday action impairments across a wide range of clinical populations, it does not sufficiently simulate conditions in which healthy or high-functioning individuals are prone to error. These authors claim to have observed that, on multiple occasions, healthy or high-functioning participants (AD patients) performed at or near the maximum possible score on this test.

There is actually a lack of valid experimental techniques for eliciting and observing errors on familiar and well-practiced tasks (Giovannetti et al., 2007). Hence, the objective of our study was to evaluate the benefits of virtual reality (VR) techniques for examining impairments in everyday activities in AD patients using tasks that parallel requirements of daily life (Plancher, Tirarda, Gyselincka, Nicolas, & Piolino, 2012). Computer-based VR techniques enable users to be immersed in simulated interactive environments that are similar to real world situations (Josman, Milika Ben-Chaim, Friedrich, & Weiss, 2008). VR can be used to objectively measure behavior in ecologically-valid environments, while maintaining strict experimental control over stimulus delivery and measurement (Rizzo, Schultheis, Kerns, & Mateer, 2004). Additionally, according to Nolin, Stipanicic, Henry, Joyal, and Allain (2012), automatic scoring reduces the number of errors that might skew the results. Furthermore, these authors suggest that by its very nature, VR testing detects subtle deficits, which are often imperceptible using traditional assessment approaches. These advantages make the use of VR techniques a promising avenue for improved identification of action deficits following AD (Yamaguchi, Foloppe, Richard, Richard and Allain, 2012). Studies that have used VR in cognitive evaluation of brain-damaged patients found that the ability to perform the assessment in a meaningful virtual environment enhanced prediction of the patients' daily life functioning (Zhang et al., 2003; Lee et al., 2003).

To date, only a few studies have used VR to test pathological aging, and they mainly focused on navigational or memory processes. Cushman, Stein, and Duffy (2008) found a close relationship between impairments in virtual and real navigation environments in MCI and AD patients. Widmann, Beinhoffa, and Riepea (2012) concluded that memory testing in virtual environments more realistically portrays everyday memory impairments in AD patients. These authors found that a photorealistic virtual environment was able to unmask memory deficits that are relevant to everyday life, which were not detectable with a classical list learning approach (see also Plancher et al., 2012). Concerning IADL, Werner, Rabinowitz, Klinger, Korcyn, and Josman (2009) supported the feasibility and the validity of a virtual action-planning supermarket for the diagnosis of patients with a MCI. Bialystok, Craik, & Stefurak (2008) demonstrated the value of a virtual breakfast preparation task to differentiate the pattern of performance of patients with Parkinson's disease from that of healthy controls.

In summary, VR technologies have considerable potential to detect functional limitations in IADL performance in AD patients, beyond that of current neuropsychological measures. Relative to self- and caregiver-report questionnaires, VR tasks may enhance specificity and ecological validity by presenting the patient with functional situations that resemble daily life. Moreover, VR tasks may be easier to implement and more sensitive to mild impairment. They may provide greater control over task parameters. Thus, the primary objective of this study was to present a novel laboratory method—The Non-Immersive Virtual Coffee Task (NI-VCT)—for the study of action errors in AD patients. We selected a non-immersive desktop computer program for two main reasons. First, it is considerably more portable than a three-dimensional environment (or the NAT), which could facilitate assessment in clinical settings. Second, the non-immersive environment limits the risk of simulation sickness, a well-known problem in virtual assessment of elderly subjects (Kawano et al., 2012). The simulated coffee preparation task was chosen because it constitutes a familiar task that requires multiple cognitive processes, such as serial ordering of several steps (open the drawer, etc.), object selection (filter, etc.), and technical manipulations (coffee machine) to achieve a practical goal. The coffee-making task has also been successfully used in studies of both patients and controls (Giovannetti et al., 2007) and in computational simulations (Botvinick & Plaut, 2004; Cooper & Shallice, 2000). The second objective of our study was to differentiate the IADL profiles of AD patients from those of healthy older adults, using the NI-VCT. In line with previous findings, we expected that NI-VCT performance would be inferior in AD patients compared to HE controls. Our third aim was to examine whether the NI-VCT was a better (e.g., potentially more sensitive) assessment of everyday action deficits than using a real condition (Real Coffee Task: RCT). Consistent with past findings, lower performance were expected in NI-VCT. The fourth objective was to verify the ecological validity of the NI-VCT. Close relationships between errors made in virtual and real environments were expected. Our fifth aim was to determine whether NI-VCT and RCT performance are each related to specific cognitive functioning. In agreement with the model of Giovannetti, Bettcher, Brennan, Libon, Kessler, et al. (2008), specific relations between action errors, global efficiency, and executive functioning were expected for both performance-based tasks. Finally, using regression analysis with the NI-VCT, RCT and informant-rated measures of everyday function, we sought to demonstrate that the NI-VCT might reflect everyday action deficits in AD patients more accurately than informant-rated measures.

METHOD

The study was performed according to institutional approval and guidelines for procedures concerning human subjects and according to the Declaration of Helsinki.

Participants and Experimental Design

Our patient sample consisted of 24 individuals with a diagnosis of probable AD according to NINCDS-ADRDA criteria (McKhann et al., 1984). Their medical history, neurological examination, brain imaging, and laboratory tests provided assurance that their dementia symptoms could not be attributed to an illness other than AD. The AD group consisted of individuals with a mild to moderate level of dementia severity [see Mini Mental State Examination (MMSE; Folstein, Folstein, & McHugh, 1975) in Table 1].

Thirty-two HE subjects served as healthy controls. All control participants were free of known serious medical, neurological and psychiatric illness and were living independently. None of them presented signs of global cognitive deterioration (MMSE, see Table 1).

Analyses of variance (ANOVAs) revealed no difference between the groups for educational level [$F(1,54) = 0.31$; $p = .57$] or age [$F(1,54) = 3.076$; $p = .08$]. AD and HE groups were comparable with respect to gender ($\chi^2 = 2.54$; $df = 1$; $p = .19$).

The present study manipulated two variables: population (AD patients and HE controls), coffee making condition (NI-VCT and RCT) within participants. The order in which participants encountered each coffee making condition was counterbalanced. After complete description of the study to the subjects, written informed consent was obtained.

Brief Cognitive Assessement

Each participant, with normal or corrected-to-normal vision, completed a short cognitive assessment including measures of global efficiency and executive functions. These measures were selected because they have been shown to be the best predictors of action errors in natural settings (Giovannetti, Bettcher, Brennan, Libon, Burke, et al., 2008; Giovannetti, Bettcher, Brennan, Libon, Kessler, et al., 2008). Global efficiency was evaluated using the MMSE. Executive functions assessment was performed using the Frontal Assessment Battery (FAB; Dubois, Slachevsky, Litvan, & Pillon, 2000), which consists of 6 subtests exploring conceptualization, mental flexibility, motor programming, sensitivity to interference, inhibitory control, and environmental autonomy. It takes approximately 10 min to administer. Neuropsychological scores are presented in Table 1. Scores of 25–30 of 30 on the MMSE and 16–18 of 18 on the FAB were considered normal.

None of the AD patients presented signs of aphasia ("Protocole Montréal-Toulouse d'Examen Linguistique de l'Aphasie"; Beland & Giroud, 1992), or signs of visual perception deficits ("Protocole Montréal-Toulouse d'Evaluation des Gnosies Visuelles"; Agniel, Joanette, Doyon, & Duchein, 1992).

Table 1. Demographic and neuropsychological characteristics of participants

	AD patients (n = 24)			HE controls (n = 32)			Statistical testing	
Variables	*M*	*SD*	Range	*M*	*SD*	Range	F	*P*
Age	76.96	6.05	63–87	74.13	5.93	65–88	3.07	.08
Gender								
Male	10	/	/	7	/	/	/	/
Female	14	/	/	25	/	/	/	/
Years of School*	9.29	2.90	7–15	9.63	1.45	7–15	0.31	.57
MMSE (max = 30)	21.80	2.54	18–26	29.06	1.08	26–30	216.96	.0001
FAB (max = 18)	12.25	2.42	8–17	17.25	0.92	15–18	115.35	.0001
IADL* (max = 4)	2.42	1.26	1–4	3.87	0.33	3–4	38.43	.0001

Note. Years of school are calculated since the first grade.
*Data were missing for five AD patients.

Functional Assessment

Informant-based functional assessment was completed by the four-item version of the Lawton-Brody IADL scale (Lawton & Brody, 1969), through an interview with a relative or a caregiver, to screen for IADL independence in the four following functional domains: using a telephone, shopping, cooking and managing finances. The total score is from 0 to 4, higher scores denoting more intact functional abilities. Scores for each group appears in Table 1. Data were missing for five patients because of various reasons (i.e., time constraints, scheduling conflicts, etc.).

MATERIALS

Virtual Apparatus

Our virtual system consisted of a laptop computer with a large screen size (17-inch), a mouse, and the NI-VCT application, which runs on the PC. The virtual environment was created with Virtools Dev 3.0 software (www.virtools.com). This system was designed to be as simple as possible to make it portable. The application was visually implemented in a three-dimensional environment to improve reality of the system. The environment was explored with the computer mouse.

Virtual Environment

The virtual environment simulated a fully textured, medium-sized kitchen with many closet shelves and drawers in the background. A hob and an oven were also present in the background. In the foreground, there was a work plane with all the objects needed to prepare a cup of coffee with milk and sugar (see Figure 1). In addition, there were three distractors, which were objects visually and/or semantically similar to the useful objects/utensils (a bottle of wine, a box of cocoa powder, and a fork). Placement of objects/utensils/distractors was standardized. Our virtual system included auditory events (e.g., water noise, etc.) to foster the sense of presence.

Virtual Interactions

Participants were seated in front of the screen of the laptop. They controlled the 2D cursor using the computer mouse. They could only move the cursor in the horizontal–vertical plane. To select a virtual object/utensil, participants had to put the 2D cursor over the targeted object/utensil and press the left button. To move a virtual object/utensil, they had to select it and move the cursor while continuously pressing the left button. All objects/utensils were automatically rotated or moved in depth, so that the participants only needed to control the vertical and horizontal position of the objects. Hence, interactions were facilitated to help our participants, who generally have limited experience in computer interaction (Dickinson, Arnott, & Prior, 2007), and to focus on the cognitive aspect of the task performance rather than on the sensory–motor coordination or spatial control aspects. Thus, we assessed action functioning while minimizing the impact of computer interaction difficulties.

Virtual Measures

All participants' actions were recorded in real-time and were saved as lists in .xls files. Following recent findings by Giovannetti, Bettcher, Brennan, Libon, Kessler, et al. (2008), 5 main outcome measures were calculated by the computer: the total time in seconds to complete the task, the percentage of task steps completed with or without error (Accomplishment score: range 0–100%), the total number of errors made on the NI-VCT [including omissions: steps not performed; commissions: sequence (anticipation-omission and reversal), perseveration, substitution, addition], and the number of omission and commission errors. Omission and commission errors were also analyzed separately using the model developed by Giovannetti, Bettcher, Brennan, Libon, Kessler, et al. (2008). In this model, the authors proposed that individuals with greater executive control deficits would show more commission errors in everyday tasks. Individuals with greater global efficiency/memory deficits would show more omission errors. We were interested to see if our NI-VCT task would distinguish between such 2 patterns of everyday action deficits.

Real Environment

The participants were seated at the center of a table so they could easily view and access all objects/utensils necessary to prepare a cup of coffee with milk and sugar. Three distractor

Fig. 1. Screen shot of the virtual kitchen and steps (numbers in squares) of the VCT. Note. To perform the task, subjects had to follow 14 steps: (1) open the coffee machine drawer, (2) put the filter inside the machine, (3) put the coffee powder on the filter, (4) close the coffee machine drawer, (5) open the water recipient, (6) put some water in the machine, (7) close the water recipient, (8) put the coffee recipient on machine, (9) turn on the coffee machine, (10) wait until the coffee is done, (11) put the coffee in the cup, (12) put back the coffee recipient, (13) put a sugar in the coffee cup, (14) put some milk in the coffee's cup.

objects/utensils were put in the array, so that the participants were presented with the same objects/utensils/distracters that were in the virtual environment. Placement of objects was also standardized. Performance was coded using the same measures described for the NI-VCT.

To verify that NI-VCT and RCT performance measures were reliably coded, error scores (omission and commission) and task steps completed with or without error were coded by two independent judges in both tasks. Consistent with previous research using these scores (Giovannetti et al., 2007; Giovannetti, Bettcher, Brennan, Libon, Kessler, et al., 2008), interrater reliability was very high (98%) or perfect.

PROCEDURE

Virtual Training Sessions

All participants underwent two virtual training sessions before the test session. The first session focused on the interaction technique using the computer mouse. Participants were asked to select and manipulate virtual objects in space by making them move with the movement of the mouse. The kitchen environment was similar to that used in the NI-VCT, but the objects on the kitchen work plane were different (e.g., a sphere, a cube). The second training session was designed to familiarize the participants with the virtual coffee machine (how to open the coffee machine drawer, etc.). The training sessions lasted until the participants felt familiar with the equipment (maximum 15 min). Participants who were unable to interact with the computer were excluded. Indeed, we decided that the participant was comfortable with the environment when he was able to move 8 of the 10 virtual objects properly with the computer mouse. Eight AD patients and one HE control did not complete the training sessions successfully.

Test Session

The participants were instructed to prepare a cup of coffee with milk and sugar in real (RCT) and virtual (NI-VCT) conditions. In both conditions, participants were seated in front of the real/virtual kitchen work plan and received general verbal information about the task. To perform the task, they had to follow 14 steps (see note of the Figure 1).

Statistics

For the cognitive (MMSE and FAB scores) and functional tasks (IADL score), one-way ANOVAs were used to compare group performance. The scores obtained on the NI-VCT and RCT conditions (time to completion, accomplishment score, total errors score, omission errors score, and commission errors score) were compared between groups using factorial mixed-design ANOVAs with *post hoc* paired *t* tests. Pearson correlations and regression analyses were also used to analyze the relationships between the NI-VCT and RCT scores and cognitive and functional measures. The significance threshold was set at $p < .05$. Nonparametric statistical analyses yielded similar results.

RESULTS

Neuropsychological Tasks and Functional Assessment

The scores on the MMSE, the FAB, and the IADL scale were all significantly lower for the AD patients, suggesting cognitive and functional impairments in these patients (Table 1).

NI-VCT and RCT (see Table 2)

Time to completion

The 2 × 2 ANOVA with population as the between-subjects factor and coffee-making condition as the within-subjects factor revealed a group main effect [$F(1,54) = 38.05$; $p < .0001$], indicating that AD patients took significantly longer to complete the tasks (mean: 1027 s) than HE subjects (mean: 613.81 s). There was also a main effect of coffee-making condition [$F(1,54) = 36.01$; $p < .0001$]. Time needed to complete the task was significantly longer in the virtual (NI-VCT mean: 1021.39 s) than in the real condition (RCT mean: 616.41 s). The interaction was also significant [$F(2,108) = 3.96$; $p = .04$], indicating that the mean difference between completion times for the coffee-making tasks was greater for AD patients (mean: 427.37 s) than HE controls (mean: 294.31 s), with significance for both groups (both *p*s < .0001 on paired *t* tests).

Accomplishment score

A main effect of group on the accomplishment score [$F(1,54) = 31.51$; $p < .0001$] was observed, revealing that

Table 2. Results of VCT and RCT tasks for AD patients and HE controls

	AD patients ($n = 24$)		HE controls ($n = 32$)	
Condition Dependant variables	*M*	*SD*	*M*	*SD*
VCT				
Time to completion in seconds	647.33	253.35	374.06	162.18
Accomplishment score (%)	81.45	19.44	97.59	5.79
Total errors (number)	5.25	3.85	0.62	0.87
Omission errors (number)	2.50	2.68	0.31	0.69
Commission errors (number)	2.75	1.70	0.31	0.53
RCT				
Time to completion in seconds	379.66	203.45	239.75	50.35
Accomplishment score (%)	92.71	10.62	99.56	1.72
Total errors (number)	1.71	1.92	0.15	0.51
Omission errors (number)	0.79	0.83	0.06	0.24
Commission errors (number)	0.91	1.01	0.09	0.29

AD patients (mean: 87.08%) performed significantly fewer steps than HE controls (mean: 97.57%). There was also a coffee-making condition main effect [$F(1,54) = 10.41$; $p = .0017$], and a significant interaction [$F(2,108) = 5.13$; $p < .0254$]. In general, participants performed more poorly in the NI-VCT condition than in the RCT condition. The interaction reflected a greater decline in accomplishment score for AD patients (mean: 11.26%) than for HE controls (mean: 1.97%). This decline was statistically significant for AD patients ($p = .008$) but not for HE controls ($p = .08$).

Total errors

The group main effect was significant [$F(1,54) = 62.35$; $p < .0001$], due to more errors for AD patients (mean: 6.95) than for HE controls (mean: 0.77). The main effect of coffee-making condition [$F(1,54) = 26.28$; $p < .0001$] and the interaction effect [$F(2,108) = 15.43$; $p = .0002$] were also significant. Errors were more frequent in the virtual condition (NI-VCT mean: 5.87) than in the real condition (RCT mean: 1.85), with a higher difference for AD patients (mean: 3.55) than controls (mean: 0.47; both *post hoc* paired t tests $< .009$).

Omission errors score

A main effect of group was observed [$F(1,54) = 29.89$; $p < .0001$], with more omission errors seen in AD patients (mean: 3.29) than in HE controls (mean: 0.37). The main effect of condition [$F(1,54) = 13.47$; $p = .0004$] and the interaction effect [$F(2,108) = 7.47$; $p = .0073$] were also significant. Omission errors were more frequent in the virtual (NI-VCT mean: 2.81) than in the real (RCT mean: 0.85) condition. The difference in omission errors between coffee-making tasks was higher for patients (mean: 2.19) than for controls (mean: 0.25), with a statistically significant difference seen only for patients ($p = .007$; for controls $p = .09$).

Commission errors score

The group main effect for commission errors was significant [$F(1,54) = 77.23$; $p < .0001$], due to more commission errors in the AD group (mean: 3.66) compared to HE controls (mean: 0.40). There was also a significant main effect of condition [$F(1,54) = 30.59$; $p < .0001$], indicating that commission errors were more frequent in the virtual condition (mean: 3.06) than in the real condition (mean: 1.01). Finally, the group × condition interaction was also significant [$F(2,108) = 18.94$; $p < .0001$]. Again the difference in commission errors between coffee-making conditions was greater for AD patients (mean: 1.83) than for HE controls (mean: 0.22). This difference was statistically significant only for AD patients ($p < .0001$), but not for HE controls ($p = .56$).

Relations among NI-VCT and RCT Variables and Neuropsychological Test Scores

Correlations among neuropsychological measures and NI-VCT and RCT scores are shown in Table 3. In AD patients, omissions, commissions and total errors from the NI-VCT and RCT were significantly correlated with both MMSE and FAB scores. All correlations indicated that more action errors were associated with more impaired scores on neuropsychological tests. In HE subjects, action error scores did not correlate with any cognitive measure. Finally, in both groups, a significant correlation was observed between NI-VCT time to completion and MMSE score, indicating that longer time to completion was associated with poorer performance on this global cognitive functioning measure.

To test the model of Giovannetti, Bettcher, Brennan, Libon, Kessler, et al. (2008), two stepwise multiple regressions with MMSE and FAB scores as the predictor variables and total (NI-VCT + RCT) omission and commission errors scores as the dependant variables were performed for the AD patients. The regression for omission errors accounted for 38.5% of the variance ($F = 13.80$; $p = .001$), and has only MMSE score as a significant predictor. The regression for commission errors accounted for 39% of the variance ($F = 14.08$; $p = .001$), and has only FAB score as a significant predictor.

Relation between NI-VCT and RCT performance

Our fourth objective was to verify whether the NI-VCT could objectively measure behavior in an ecologically valid environment. Therefore, we investigated the relationships between the time to completion, accomplishment, omissions, commissions and total error scores on the NI-VCT and RCT for both groups.

As expected, significant Pearson product-moment correlations were observed. More specifically, in the AD patients, there were significant correlations between the following pairs of NI-VCT and RCT scores: time to completion ($r = .57$; $p = .002$), accomplishment score ($r = .46$; $p = .01$), total error score ($r = .54$; $p = .005$), and commission error score

Table 3. Pearson correlations between VCT, RCT, and neuropsychological measures in both groups

	AD patients		HE controls	
	MMSE	FAB	MMSE	FAB
VCT				
Time to completion	−.48**	−.12	−.40*	−.10
Accomplishment score	.31	.25	−.19	−.13
Total errors	−.58***	−.62***	.06	.20
Omission errors	−.49**	−.46**	.18	.12
Commission errors	−.41*	−.54***	−.14	.16
RCT				
Time to completion	−.26	−.15	−.20	−.06
Accomplishment score	.09	.25	.25	.07
Total errors	−.51**	−.55***	−.16	−.33
Omission errors	−.54***	−.39*	−.25	−.07
Commission errors	−.40*	−.61***	−.22	−.32

Note. MMSE = *$p \leq .05$, **$p \leq .01$, ***$p \leq .005$.

($r = .44$; $p = .03$). No significant correlation emerged for the omission error score ($r = .23$; $p = .26$). In the HE controls, times to completion in NI-VCT and RCT correlated significantly ($r = .61$; $p < .0001$); however, no other significant correlations emerged from the analyses in this group (accomplishment score: $r = -.10$; $p = .55$; total error score: $r = -.13$; $p = .47$; omission error score: $r = -.11$; $p = .52$; commission error score: $r = .01$; $p = .94$).

To identify whether the RCT performance could be explained better by NI-VCT performance than by overall severity dementia, several stepwise multiple regression analyses were also performed for the AD patients, using neuropsychological test scores (MMSE and FAB) and NI-VCT measures as the predictor variables. Each regression used a RCT measure as a dependent variable. The regression for RCT total error score was significant accounted for 48% of the variance ($F = 20.33$; $p = .0002$), and has only NI-VCT commission error score as a significant predictor.

Relations among NI-VCT Variables RCT Scores and AD Patients' Caregiver Reports

Correlation analyses were performed between relative/caregiver IADL reports on the Lawton and Brody scale (1969) and NI-VCT variables in the AD group. There were significant correlations between three NI-VCT measures and the caregiver IADL rating: time to complete ($r = -.50$; $p = .01$), total errors ($r = -.44$; $p = .04$) and commission error ($r = -.47$; $p = .02$). More action errors and longer time to completion were associated with poorer autonomy on the IADL scale. To identify whether the IADL score could be better explained by NI-VCT performance than by overall dementia severity or RCT measures, a stepwise multiple regression analysis was performed with IADL total score as the dependent variable and neuropsychological scores, NI-VCT and RCT measures as the independent variables. The stepwise regression for IADL total score accounted for 29% of the variance ($F = 8.68$; $p = .009$), and had only NI-VCT total error score, as a significant predictor.

DISCUSSION

The principal aim of this study was to determine if a non-immersive virtual coffee-making task, the NI-VCT, could detect everyday action impairments in AD patients. The task was able to identify these impairments in these patients. We also wanted to investigate whether the NI-VCT measures everyday action deficits in an ecologically valid manner. We found significant relations between virtual and real coffee-making scores, and between virtual scores and IADL scale score, thereby supporting the ecological validity of the NI-VCT.

The NI-VCT specifically revealed significant differences between AD patients and HE on several measures (time to completion, accomplishment and error scores), with clearly higher performance on the NI-VCT for HE subjects. The data are consistent with previous studies showing everyday action impairments in AD (Giovannetti et al., 2002, 2006; Giovannetti, Bettcher, Brennan, Libon, Burke, et al. 2008; Giovannetti, Bettcher, Brennan, Libon, Kessler, et al. 2008; Ramsden, Kinsella, Ong, & Storey, 2008). However, this study is the first to demonstrate these differences in a non-immersive virtual kitchen. Werner et al. (2009) have already used a VR paradigm to assess subjects with normal and pathological aging in a complex, simulated real-life situation, but the task condition as well as the population were both different. In light of those results, we can conclude that VR environments are a promising alternative that can be used for the detection of action impairments in AD and in its prodromal form.

Furthermore, the NI-VCT reliably simulates real world coffee-making conditions. Therefore, similar cognitive mechanisms appear to be engaged under VR and real conditions. This assertion is supported by the observed correlations between scores in the NI-VCT and the RCT conditions, particularly those scores in the AD group. Our results parallel those of Cushman et al. (2008), who obtained similar findings with spatial cognition in normal and pathological aging. Our work is the first study to benchmark a non-immersive virtual task against a closely corresponding real-world task for the assessment of ability to perform everyday actions. It provides the same clear picture: virtual testing of everyday actions offers a valid assessment of these skills. This picture is also strongly supported by the results of our correlation and regression analyses, which demonstrated that several scores of our NI-VCT were predictors of independence in the ability to perform daily life activities.

It is particularly important to benchmark VR assessment against "real world" task performance because VR testing typically yields lower scores than classical testing. For example, Cushman et al. (2008) demonstrated that navigational performance was lower under VR conditions than in actual settings in all groups. In our study, virtual kitchen assessment also yielded significantly higher numbers of action errors relative to the real world coffee-making condition in both groups. Therefore, we can propose that our NI-VCT is a sensitive and ecologically valid test of everyday action impairments, because we had taken care to ensure that the NI-VCT performance was a good indicator of RCT performance and functional autonomy.

Indeed, the fact that the VR task seems to be harder than its real world counterpart could be interpreted in two directions. On the one hand, this discrepancy suggests that VR assessment is inherently more sensitive to action impairments. On the other hand, it calls into question the ecological validity of desktop virtual environment tasks for at least three reasons. First, interaction with the VR environment could increase cognitive load given that it requires participants to process more information simultaneously. A common effect of aging is a decline in attentional deployment (e.g., Staub, Doignon-Camus, Després, & Bonnefond, 2013). These deficits are particularly severe for patients with AD (Festa, Heindel, & Ott, 2010). Therefore, we can speculate that there could be more attentional requirements in the virtual kitchen, which would also affect the performance scores of both groups. Second, the virtual coffee-making task requires different motor manipulations, which could be more difficult for the HE subjects and the AD patients because of age-related decreases in dexterity and fine motor performance (Kluger et al., 1997; Yan, Rountree, Massman, Smith Doody, & Li, 2008). In fact, age-related studies have shown that interacting with a computer using traditional computer input devices can cause problems for many older adults, particularly for novices in this age group (Dickinson et al., 2007), and for those with age-related disabilities (Wood, Willoughby, Rushing, Bechtel, & Gilbert, 2005). Third, in our virtual environment, we attempted to facilitate interactions with objects/utensils using a computer mouse (i.e., using 2D manipulations). In doing so, object manipulations in the computer program were minimal or nonexistent, which may have deprived participants from useful tactile feedback to help them to produce the appropriate actions.

At the very least, it appears that our NI-VCT is sensitive for detection of everyday action impairments in AD patients. Our NI-VCT is not necessarily a more sensitive tool than its real world counterpart, but it may simply be a more complex, yet ecologically comparable task given that it requires participants to process more information simultaneously and to produce more fine motor movements. Thus, we cannot rule out the possibility that the increased difficulties observed for our HE subjects and AD patients during VR assessment were also related to attentional or motor limitations. However, because all of our participants underwent two training sessions before testing, we believe that only a part of the increased difficulties observed in our samples during virtual assessment could be attributable to additional attentional and motor problems. However, it will be important to compare NI-VCT and RCT performance using other types of interactive input approaches with older adults in future studies. For example, touchscreen interaction has the potential to offer a more intuitive approach for interaction with the computer than using a mouse, because it relies on existing skills, operating directly with objects (Wood et al., 2005). Based on subjective data obtained from questionnaires administered to our participants, it appears that while the objects, sounds, colors and actions in the VR condition were quite similar to the real situation, the use of the mouse and the coffee machine was difficult. Most participants believed that additional training sessions were necessary.

From a theoretical perspective, our findings that, in AD patients, omission and commission error scores were correlated in both tasks, and that both these scores correlated with global cognitive and executive impairments are compatible with the resource theory model of everyday action impairment developed by Schwartz et al. (1998). This model posits that everyday action places significant demands on cognitive resources and that performance errors arise when cognitive resources are limited as a consequence of brain damage. Following Schwartz et al. (1998), co-occurrence of high rates of omission and commission errors on everyday tasks and significant positive relations among different error types in patients may be explained by a unitary construct, namely resource limitations. This profile seems to characterize our AD patients. However, we also found that omissions and commissions errors were predicted by different neuropsychological test scores, respectively, by global efficiency (MMSE) and by executive control (FAB) measure. These results are more consistent with the proposal by Giovannetti, Bettcher, Brennan, Libon, Kessler, et al. (2008) that omission and commission errors would each independently account for overall cognitive impairment and overall executive impairment, reflecting dissociable aspects of everyday action performance. It is difficult to decide between these two models in this work. We must remain cautious in our interpretation since the models by Schwartz et al. (1998) and Giovannetti, Bettcher, Brennan, Libon, Kessler, et al. (2008) were based on regression and correlation analyses of a larger number of action and neuropsychological variables.

Despite the strengths of this study, we acknowledge several limitations. First, cohorts were relatively small. Second, patients are unlikely to be representative of the entire population with AD. Indeed, their MMSE score was between 18 and 26, and it would be interesting to test patients across a broader range of MMSE scores. Additional research is needed to examine whether our findings hold for individuals with other types of dementia, or for individuals at different stages of AD severity. Second, our neuropsychological protocol was brief and only assessed a limited number of cognitive processes for everyday action performance. For instance, measures of episodic memory, task familiarity and task knowledge were not administered. However, deficits in these cognitive domains could contribute to impaired performance, particularly episodic memory deficits that are primary in AD. Third, we limited the VR and real world assessments to one familiar everyday task (preparing coffee) that is comparable to one portion of the easiest of the three NAT tasks (preparing coffee and toast; Schwartz et al., 2003). The results should be replicated with various activities, including more complicated tasks. These limitations notwithstanding, we offer evidence that our NI-VCT task can be used to evaluate everyday action impairments in a way that is relevant to the patients' real deficits in AD. Overall, our study clearly demonstrated the feasibility of using VR technology to study deficits of AD patients in ecologically valid, controlled and safe environments.

ACKNOWLEDGMENTS

This study was supported by a Grant from the "Conseil Scientifique" of the University of Angers in France and from France Alzheimer. The authors have no conflicts of interest to declare. We thank Dr. John L. Woodard for his editing assistance, under the auspices of the Research and Editing Consultant Program of the International Neuropsychological Society International Liaison Committee. Doctors Allain and Foloppe have equally contributed to the preparation of this manuscript.

REFERENCES

Adam, S., Van der Linden, M., Juillerat, A.C., & Salmon, E. (2000). The cognitive management of daily life activities in patients with mild to moderate Alzheimer's disease in a day care center: A case report. *Neuropsychological Rehabilitation*, *10*, 485–509.

Agniel, A., Joanette, Y., Doyon, B., & Duchein, C. (1992). *Protocole Montréal-Toulouse d'évaluation des gnosies visuelles et auditives*. Isbergues: L'Ortho-Edition.

Barberger-Gateau, P., Fabrigoule, C., Rouch, I., Letenneur, L., & Dartigues, J.F. (1999). Neuropsychological correlates of self-reported performance in instrumental activities of daily living and prediction of dementia. *Journals of Gerontolology. Series B. Psychological Sciences and Social Sciences*, *54*, 293–303.

Beland, R., & Giroud, F. (1992). *Protocole Montréal-Toulouse d'examen linguistique de l'aphasie (Version révisée)*. Isbergues: L'Ortho-Edition.

Bialystok, E., Craik, F.I., & Stefurak, T. (2008). Planning and task management in Parkinson's disease: Differential emphasis in dual-task performance. *Journal of the International Neuropsychological Society*, *14*, 257–265.

Botvinick, M., & Plaut, D. (2004). Doing without schema hierarchies: A recurrent connectionist approach to routine sequential action and its pathologies. *Psychological Review*, *111*, 395–429.

Boyle, P.A., Malloy, P.F., Salloway, S., Cahn-Weiner, D.A., Cohen, R., & Cummings, J.L. (2003). Executive dysfunction and apathy predict functional impairment in Alzheimer disease. *American Journal of Geriatric Psychiatry*, *11*, 214–221.

Cahn-Weiner, D.A., Ready, R.E., & Malloy, P.F. (2003). Neuropsychological predictors of everyday memory and everyday functioning in patients with mild Alzheimer's disease. *Journal of Geriatric and Psychiatric Neurology*, *16*, 84–89.

Chaytor, N.S., Schmitter-Edgecombe, M., & Burr, R. (2006). Improving the ecological validity of executive functioning tests: Environmental demands and compensatory strategies. *Archives of Clinical Neuropsychology*, *21*, 217–227.

Cooper, R., & Shallice, T. (2000). Contention scheduling and the control of routine activities. *Cognitive Neuropsychology*, *17*, 297–338.

Cushman, L.A., Stein, K., & Duffy, C.J. (2008). Detecting navigational deficits in cognitive aging and Alzheimer disease using virtual reality. *Neurology*, *71*, 888–895.

De Bettignies, B.H., Mahurin, R.K., & Pirozzolo, F.J. (1993). Functional status in Alzheimer's disease and multi-infarct dementia: A comparison of patient performance and caregiver report. *Clinical Gerontologist*, *12*, 31–49.

Dickinson, A., Arnott, J., & Prior, S. (2007). Methods for human-computer interaction research with older people. *Behaviour & Information Technology*, *26*, 343–352.

Dubois, B., Slachevsky, A., Litvan, I., & Pillon, B. (2000). The FAB: A Frontal Assessment Battery at bedside. *Neurology*, *55*, 1621–1626.

Festa, E.K., Heindel, W.C., & Ott, B.R. (2010). Dual-task conditions modulate the efficiency of selective attention mechanisms in Alzheimer's disease. *Neuropsychologia*, *48*, 3252–3261.

Folstein, M., Folstein, S., & McHugh, P. (1975). Mini-mental state: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*, 189–198.

Giovannetti, T., Bettcher, B.M., Brennan, L., Libon, D.J., Burke, M., Duey, K., … Wambach, D. (2008). Characterization of everyday functioning in mild cognitive impairment: A direct assessment approach. *Dementia and Geriatric Cognitive Disorders*, *25*, 359–365.

Giovannetti, T., Bettcher, B.M., Brennan, L., Libon, D.J., Kessler, R.K., & Duey, K. (2008). Coffee with jelly or unbuttered toast: Omissions and commissions are dissociable aspects of everyday action impairment in Alzheimer's disease. *Neuropsychology*, *22*, 235–245.

Giovannetti, T., Schmidt, K., Gallo, J., Sestito, N., & Libon, D.J. (2006). Everyday action in dementia: Evidence for differential deficit in Alzheimer's disease versus subcortical vascular dementia. *Journal of the International Neuropsychological Society*, *12*, 45–53.

Giovannetti, T., Schwartz, M.F., & Buxbaum, L.J. (2007). The coffee challenge: A new method for the study of everyday action errors. *Journal of Clinical and Experimental Neuropsychology*, *29*, 690–705.

Josman, N., Milika Ben-Chaim, N., Friedrich, H., & Weiss, P.L. (2008). Effectiveness of virtual reality for teaching street-crossing skills to children and adolescents with autism. *Journal on Disability and Human Development*, *7*, 49–56.

Kawano, N., Iwamoto, K., Ebe, K., Aleksic, B., Noda, A., Umegaki, H., … Ozaki, N. (2012). Slower adaptation to driving simulator and simulator sickness in older adults. *Aging Clinical and Experimental Research*, *24*, 285–289.

Kluger, A., Gianutsos, J.G., Golomb, J., Ferris, S.H., George, A.E., Franssen, E., & Reisberg, B. (1997). Patterns of motor impairement in normal aging, mild cognitive decline, and early Alzheimer's disease. *Journal of Gerontology Series B: Psychological Sciences and Social Sciences*, *52*, 28–39.

Lawton, M.P., & Brody, E.M. (1969). Assessment of older people: Self-maintaining and instrumental activities of daily life. *The Gerontologist*, *9*, 179–186.

Lechowski, L., Dieudonne, B., Tortrat, D., Teillet, L., Robert, P.H., Benoit, M., … Vellas, B. (2003). Role of behavioural disturbance in the loss of autonomy for activities of daily living in Alzheimer patients. *International Journal of Geriatric Psychiatry*, *18*, 977–982.

Lee, J.H., Ku, J., Kim, H., Cho, W., Kim, B., Kim, I.Y., … Wiederhold, M.D. (2003). A virtual reality system for the assessment and rehabilitation of the activities of daily living. *Cyberpsychology and Behavior*, *6*, 383–388.

Mast, B.T., Azar, A.R., MacNeill, S.E., & Lichtenberg, P.A. (2004). Depression and activities of daily living predict rehospitalization within 6 months of discharge from geriatric rehabilitation. *Rehabilitation Psychology*, *49*, 219–223.

McKhann, G., Drachmann, D., Folstein, M., Katzman, R., Price, D., & Stadan, E.M. (1984). Clinical diagnosis of Alzheimer's disease: Report of the NINCDS-ADRDA work group under the auspices of the department of health and human services task force on Alzheimer's disease. *Neurology*, *34*, 939–943.

Millán-Calenti, J.C., Tubio, J., Pita-Fernandez, S., Rochette, S., Lorenzo, T., & Maseda, A. (2012). Cognitive impairment as predictor of functional dependence in an elderly sample. *Archives of Gerontology and Geriatrics*, *54*, 197–201.

Mitchell, M.B., Miller, L.S., Woodard, J.L., Davey, A., Martin, P., Burgess, M., & Poon, L.W. (2011). Regression-based estimates of observed functional status in centenarians. *The Gerontologist*, *51*(2), 179–189.

Noale, M., Maggi, S., Minicuci, N., Marzari, C., Destro, C., Farchi, G., … Crepaldi, G. (2003). Dementia and disability: Impact on mortality. *Dementia & Geriatric Cognitive Disorders*, *16*, 7–14.

Nolin, P., Stipanici, A., Henry, M., Joyal, C.C., & Allain, P. (2012). Virtual reality as a screening tool for sports concussion in adolescents. *Brain Injury*, *26*, 1564–1573.

Plancher, G., Tirarda, A., Gyselincka, V., Nicolas, S., & Piolino, P. (2012). Using virtual reality to characterize episodic memory profiles in amnestic mild cognitive impairment and Alzheimer's disease: Influence of active and passive encoding. *Neuropsychologia*, *50*, 592–602.

Ramsden, C.M., Kinsella, G.J., Ong, B., & Storey, E. (2008). Performance of everyday actions in mild Alzheimer's Disease. *Neuropsychology*, *22*, 17–26.

Rizzo, A.A., Schultheis, M.T., Kerns, K.A., & Mateer, C. (2004). Analysis of assets for virtual reality applications in neuropsychology. *Neuropsychological Rehabilitation*, *14*, 207–239.

Rusted, J.M., & Sheppard, L.M. (2002). Action-based memory in people with dementia: A longitudinal look at tea making. *Neurocase*, *8*, 111–126.

Schwartz, M.F., Buxbaum, L.J., Ferraro, M., Veramonti, T., & Segal, M. (2003). *Naturalistic action test*. Suffolk, England: Pearson Assessment.

Schwartz, M.F., Montgomery, M., Buxbaum, L.J., Lee, S., Carew, T.G., Coslett, H.B., … Mayer, N. (1998). Naturalistic action impairment in closed head injury. *Neuropsychology*, *12*(1), 13–28.

Schwartz, M.F., Segal, M.E., Veramonti, T., Ferraro, M., & Buxbaum, L.J. (2002). The Naturalistic Action Test: A standardised assessment for everyday-action impairment. *Neuropsychological Rehabilitation*, *12*, 311–339.

Senanarong, V., Poungvarin, N., Jamjumras, P., Sriboonroung, A., Danchaivijit, C., Udomphanthuruk, S., & Cummings, J.L. (2005). Neuropsychiatric symptoms, functional impairment and executive ability in Thai patients with Alzheimer's disease. *International Psychogeriatry*, *17*, 81–90.

Starkstein, S.E., Brockman, S., Bruce, D., & Petracca, G. (2010). Anosognosia is a significant predictor of apathy in Alzheimer's disease. *The Journal of Neuropsychiatry and Clinical Neuroscience*, *22*, 378–383.

Staub, B., Doignon-Camus, N., Després, O., & Bonnefond, A. (2013). Sustained attention in the elderly: What do we know and what does it tell us about cognitive aging? *Ageing Research Review*, *12*, 459–468.

Tekin, S., Fairbanks, L.A., O'Connor, S., Rosenberg, S., & Cummings, J.L. (2001). Activities of daily living in Alzheimer's disease: Neuropsychiatric, cognitive, and medical illness influences. *American Journal of Geriatric Psychiatry*, *9*, 81–86.

Werner, P., Rabinowitz, S., Klinger, E., Korcyn, A.S., & Josman, N. (2009). The use of the virtual action planning supermarket for the diagnosis of mild cognitive impairment. *Dementia and Geriatric Cognitive Disorders*, *27*, 301–309.

Widmann, C.N., Beinhoffa, U., & Riepea, M.W. (2012). Everyday memory deficits in very mild Alzheimer's disease. *Neurobiology of Aging*, *33*, 297–303.

Wood, E., Willoughby, T., Rushing, A., Bechtel, L., & Gilbert, J. (2005). Use of computer input devices by older adults. *Journal of Applied Gerontology*, *24*, 419–438.

Yamaguchi, T., Foloppe, D.A., Richard, P., Richard, E., & Allain, P. (2012). A dual-modal virtual reality kitchen for (re)learning of everyday cooking activities In Alzheimer's disease. *Presence: Teleoperators and Virtual Environments*, *21*, 43–57.

Yan, J.H., Rountree, S., Massman, P., Smith Doody, R., & Li, H. (2008). Alzheimer's disease and mild cognitive impairment deteriorate fine movement control. *Journal of Psychiatric Research*, *42*, 1203–1212.

Zhang, L., Abreu, B.C., Seale, G.S., Masel, B., Christiansen, C.H., & Ottenbacher, K.J. (2003). A virtual reality environment for evaluation of a daily living skill in brain injury rehabilitation: Reliability and validity. *Archives of Physical Medicine and Rehabilitation*, *84*, 1118–1124.