Feasibility and validity of mobile cognitive testing in the investigation of age-related cognitive decline

PIERRE SCHWEITZER,^{1,2} MATHILDE HUSKY,³ MICHÈLE ALLARD,^{1,2,4,5} HÉLÈNE AMIEVA,¹
KARINE PÉRÈS,^{6,7} ALEXANDRA FOUBERT-SAMIER,^{6,7,8} JEAN-FRANÇOIS DARTIGUES^{6,7,8}
& JOEL SWENDSEN^{1,2,4}

1 University of Bordeaux, Bordeaux, France

2 CNRS, INCIA, UMR 5287, Bordeaux, France

3 Institut de Psychologie, Université Paris Descartes, Sorbonne Paris Cité, Institut Universitaire de France, Boulogne Billancourt, France

4 EPHE, Bordeaux, France

5 CHU de Bordeaux, France

6 University of Bordeaux, ISPED, Centre INSERM U897-Epidémiologie-Biostatistique, Bordeaux, France

7 INSERM, ISPED, Centre INSERM U897-Epidémiologie-Biostatistique, Bordeaux, France

8 University Hospital, Memory Consultation, CMRR, Bordeaux, France

---

Key words

ecological momentary assessment, EMA, experience sampling method, ESM, mobile cognitive test, mobile neuropsychological test

Correspondence

Joel Swendsen, University of Bordeaux, Bordeaux, France.
Telephone (+33) 557 820 455
Fax (+33) 557 574 597
Email: Joel.Swendsen@u-bordeaux2.fr

Received 1 November 2015;
revised 3 June 2016;
accepted 20 June 2016

Abstract

Mobile cognitive testing may be used to help characterize subtle deficits at the earliest stages of cognitive decline. Despite growing interest in this approach, comprehensive information concerning its feasibility and validity has been lacking in elderly samples. Over a one-week period, this study applied mobile cognitive tests of semantic memory, episodic memory and executive functioning in a cohort of 114 elderly non-demented community residents. While the study acceptance rate was moderate (66%), the majority of recruited individuals met minimal compliance thresholds and responded to an average of 82% of the repeated daily assessments. Missing data did not increase over the course of the study, but practice effects were observed for several test scores. However, even when controlling for practice effects, traditional neuropsychological tests were significantly associated with mobile cognitive test scores. In particular, the Isaacs Set Test was associated with mobile assessments of semantic memory ($\gamma = 0.084$, $t = 5.598$, $p < 0.001$), the Grober and Buschke with mobile assessments of episodic memory ($\gamma = 0.069$, $t = 3.156$, $p < 0.01$), and the Weschler symbol coding with mobile assessments of executive functioning ($\gamma = 0.168$, $t = 4.562$, $p < 0.001$). Mobile cognitive testing in the elderly may provide complementary and potentially more sensitive data relative to traditional neuropsychological assessment. Copyright © 2016 John Wiley & Sons, Ltd.

Introduction

It is estimated worldwide that 6% to 8% of the population over 60 years of age has developed dementia, and Alzheimer's Disease (AD) represents approximately 60–70% of these cases (World Health Organization, 2015). The insidious onset of AD implies that patients experience subtle cognitive difficulties that increase in severity over a number of years, with the first decline in any cognitive functions occurring over a decade before clinical diagnosis is possible (Amieva *et al.*, 2008). From a methodological point of view, the detection of the earliest stages of cognitive decline leading to AD is not only complicated by the difficulty in separating pathological decline from that observed in healthy aging, but also by the relatively large margin of error associated with standard neuropsychological tests (Schmidt *et al.*, 2007; Metternich *et al.*, 2009; Tollenaar *et al.*, 2009; Hess *et al.*, 2012). In particular, the scores derived through punctual neuropsychological testing may vary as a function of diverse context-specific, patient-specific or clinician-specific influences, and therefore they often cannot differentiate variance attributable to measurement error from the subtle cognitive decline that occurs in patients who will eventually develop AD.

Methods relying on mobile technologies such as Ecological Momentary Assessment (EMA) hold promise for overcoming these barriers by providing repeated assessments multiple times a day and over several consecutive days in ecologically-valid contexts. EMA has been extensively validated in a range of normal and clinical populations (Johnson *et al.*, 2009a), and it has been used successfully in elderly samples (Cain *et al.*, 2009). Importantly, the expansion of EMA to provide repeated and *in vivo* assessments of cognitive functions may reduce the margin of error typically associated with neuropsychological testing and thereby more reliably characterize early cognitive decline in this population. In one recent study of non-demented elderly participants, intellectual activities of daily life were shown to improve momentary EMA measures of semantic memory, and these mobile memory scores were shown to be more sensitive than traditional neuropsychological measures in the identification of hippocampal atrophy (Allard *et al.*, 2015). The repetition of such tests through EMA can also provide information concerning variability of cognitive functions (distinct from average functioning scores) which has been shown to be strongly predictive of cognitive decline (Holtzer *et al.*, 2008).

Despite these encouraging findings, no published studies to date have examined the basic feasibility or

validity of mobile cognitive testing through EMA in the elderly. The present investigation responds to these concerns by examining a new EMA application designed to assess cognitive functioning in daily life using a smartphone. In light of their relevance to cognitive decline and dementia, mobile tests were developed to assess semantic memory, episodic memory, and executive functioning in daily life in a cohort of dementia-free elderly individuals. The objectives of this investigation are to: (1) examine the general acceptability and feasibility of EMA-based cognitive testing in this population and the compliance rate with the multiple tests provided per day; (2) assess potential biases associated with the repeated-assessment methodology over time; (3) estimate the convergent validity of mobile test scores with traditional neuropsychological assessments.

Method

Sample

The sample was drawn from the AMI cohort, a population-based cohort of 1002 French elderly agricultural workers in rural areas. The AMI cohort randomly selected participants aged 65 or older from the Farmers' Social Security Registry and its methodology has been fully described elsewhere (Pérès *et al.*, 2012). Study procedures were approved by the regional human research review board and all participants provided written informed consent. Following baseline inclusion in the full cohort from 2007 to 2009, participants were administered a neuropsychological test battery (see details later) and, in a sub-sample of the AMI cohort, participants accepted a neuroimaging examination that was administered approximately every two years. A first follow-up assessment of the imaging subgroup occurred between 2009 and 2012. At the second follow-up examination (between 2012 and 2014), a random sub-sample of these participants were also invited to participate in a week-long EMA study using an Android smartphone (Samsung Galaxy S with a 10.6 cm screen, default font size set to 12 point) programmed to administer electronic assessments. Inclusion criteria for the present study were having a sixth grade reading level, no visual or motor impairments, and no significant cognitive impairment or diagnosis of dementia (based on a neuropsychological evaluation by a psychologist, a clinical examination by a geriatrician and confirmation in a case consensus conference by three dementia specialists) and participation in the ancillary neuroimaging project. A total of 172 participants were contacted for participation

in the present EMA study (44% female) with a mean age of 72 years (standard deviation [SD] = 4.6).

Procedure

The electronic interview questioned participants regarding their physical environment, social interactions, and specific behaviors derived from previous EMA research (Johnson *et al.*, 2009b). A brief cognitive test was also administered at the end of a portion of electronic interviews to assess semantic memory and verbal fluency, episodic memory or executive functioning. These cognitive functions were selected given their frequent decline among elderly individuals (Buczyłowska and Petermann, 2016; Harada *et al.*, 2013) and the corresponding mobile cognitive tests were developed specifically for self-administration by the participant (see later). In order to avoid biases associated with time of day or test repetition, each test was administered five times during the week in an order that was counterbalanced across the day and a unique item content was developed for the five versions of each test. All functions of the device were deactivated so that it could only be used for the purposes of the study. All participants were trained in how to operate the device and two practice assessments were administered, one guided by the research staff and one completed independently by the individual. Participants experiencing difficulty in understanding or completing the interviews were provided additional training. The interviews occurred five times a day for one week and were administered at fixed intervals, randomized across individuals. The first and last interviews of the day were adjusted to correspond to the typical sleep schedule of the participants. Cognitive assessments in EMA used the voice-recording function of the smartphone (for semantic memory, free-recall for episodic memory, and for executive functioning), or through manual selection on the smartphone screen (for semantic memory recognition). Recorded verbal responses were coded by trained members of the research team. Interrater reliability was computed for a subsample of 30 participants and ranged from 0.90 to 0.98 for all tests.

AMI neuropsychological instruments

Semantic memory

The Isaacs Set Test (Isaacs and Kennie, 1973) was used to assess semantic memory through the presentation of one semantic category, such as “Colors” or “Towns”. Participants were asked to list as many words

belonging to a given category as possible in a 60-second span. The number of correct answers was used to score the test.

Episodic memory

Subjects were administered the Free and Cued Selective Reminding Test (Grober and Buschke, 1987). This test starts with a study phase where four sheets of paper each displaying four words were successively presented. The 16 words belong to 16 distinct semantic categories. For each list, the participant is asked to match the category with the corresponding item (e.g. “Among these words, can you tell me what is the fish?”). Participants are then tested for immediate recall: each category is presented and they are asked to give the corresponding item; when they fail to remember, they are given the correct answer. Once the study phase has been completed, the recall phase begins using free recall followed by cued recall for the items that were not remembered. Three free and cued trials are successively completed, separated by an interference task (counting backwards). After a 20-minute delay, a delayed free and cued recall trial is administered. The sum of correct answers for the three recall tests and the number of correct answers for the delayed recall period were both used to score the test.

Executive functioning

The digit-symbol test from the Wechsler Adult Intelligence Scale – Fourth Edition (WAIS-IV; Wechsler, 2008) was used to assess executive functioning. In this test, participants are presented a list of nine symbols paired with digits from one to nine. Participants are given a matrix of digits which they have to complete with the corresponding symbol in a 90-second span. The number of correct symbols was used to score the test.

EMA cognitive assessments

Semantic memory

A mobile test of semantic verbal fluency was developed to present participants with a semantic category, such as “Animals” or “Vegetables”, and then they were asked to say as many words as possible belonging to this category within 60 seconds. Verbal responses were recorded on the smartphone and later coded by research staff. The total number of correct words was used as the primary performance score.

Episodic memory

A mobile list-learning test was developed to assess immediate free recall and recognition as well as delayed recall and recognition. Lists of words representing material objects were selected from the BRULEX database (Content *et al.*, 1990) and that had a frequency of appearance between 200 and 1000 words among one million French words appearing in written text. A 10-word list was displayed for 30 seconds, after which participants were immediately asked to freely recall all words within one minute. Following the free-recall task, participants were presented a list of 20 words which included the 10 words in the previously-presented list as well as 10 additional words with matched frequency of appearance in French language. Participants were then asked to identify the words they recognized as belonging to the original list by selecting each word directly on the device. The number of correct words for both the free-recall task (recorded verbally) and the recognition task (recorded by selection on the EMA device screen) served as the primary scores for immediate episodic memory performance. At the following EMA assessment, approximately three hours later, participants were asked to complete the same tasks (free-recall and recognition). The number of correct words for both tasks served as the primary scores for delayed episodic memory performance.

Executive functions

A mobile letter-word generation test was used to assess executive functions. Participants were presented a letter such as “P” or “F”, and asked to name as many words beginning with that letter as possible within a one-minute span and without naming proper nouns. Verbal responses were recorded by the EMA device and the total number of correct words served as the primary performance score.

Data analysis

EMA data were analyzed using Hierarchical Linear and Non-linear Modeling Version 7 (Raudenbush *et al.*, 2004). Analyses of reactivity to the repeated testing methodology were performed in order to identify significant changes in the frequency or intensity of cognitive variables as a function of study duration. Convergent validity was examined by using neuropsychological test scores acquired at the baseline assessment as predictors of mobile cognitive performance at the second follow-up assessment (approximately two years later). Analyses of reactivity and convergent validity were performed with participants meeting minimum compliance with the EMA

methodology, defined as having completed at least one-third of programmed assessments (Johnson *et al.*, 2009a). Means-as-outcomes models were used for continuous outcome variables and Bernoulli models for dichotomous outcomes.

Results

Acceptance and compliance

Overall, 114 (66%) of the 172 eligible participants agreed to take part in the EMA portion of the study. Patients who refused participation were older, $t(170)=1.971$, $p<0.05$, and had lower Mini-Mental Status Examination (MMSE) scores, $t(170)=-4.417$, $p<0.001$, than those who accepted. In comparison with the overall cohort, the 114 participants enrolled in study were more often female, $\chi^2(1)=4.41$, $p<0.05$, younger (71.71 years versus 76.85 years) $t(1000)=7.96$, $p<0.001$, and had higher global MMSE scores (27.40 versus 24.33) $t(984)=-6.83$, $p<0.001$. Seventy-five of these participants completed at least one-third of programmed assessments and were considered as “minimally compliant” with the procedures. In comparison with the overall cohort, participants meeting the minimum compliance criterion were more often female, $\chi^2(1)=7.20$, $p<0.05$, younger (71.05 years versus 76.90 years) $t(961)=7.37$, $p<0.001$, and had higher global MMSE scores (27.70 versus 24.33) $t(945)=-6.10$, $p<0.001$.

The socio-demographic characteristics and cognitive performance scores for the sample of 75 participants meeting minimal compliance criteria are presented in Table 1. Compliance with the self-report EMA interviews by these participants was high, with 82% of the possible assessments being completed by participants in the context of their daily lives (resulting in 2158 observations). Examination of compliance with mobile cognitive tests that were recorded verbally by the device indicated that a small portion of participants may have received help from a spouse or other individual to complete the assessments. When these assessments were considered as void, 72% of the EMA cognitive assessments were completed.

Effects associated with duration of ambulatory monitoring

No fatigue effect was observed considering that the number of missing observations decreased as a function of study duration (Table 2). Concerning practice effects, semantic memory scores, immediate recognition and delayed free-recall improved as a function of study duration. However, performance on the delayed recognition condition of list-learning decreased over time. Scores

Table 1. Demographic and clinical characteristics of the final sample ($N=75$)

	Percentage	Mean	Standard deviation
<i>Demographic variables</i>			
Age		76.85	4.25
Gender			
Percentage female	57		
Education			
Less than elementary school	19.4		
Elementary school	31.3		
More than Elementary school	49.3		
<i>Baseline neuropsychological testing</i>			
MMSE		28.09	1.70
Semantic memory			
IST-60		64.87	12.32
Episodic memory			
Free recall		26.41	6.51
Delayed free recall		10.77	2.95
Executive functioning			
Digit-symbol test		31.95	8.41
<i>Mobile cognitive assessments</i>			
Semantic memory			
Verbal fluency		9.77	1.94
Category generation		1.70	0.37
Episodic memory			
Auto-biographic memory		7.11	0.54
Immediate free recall		4.82	1.22
Immediate recognition		7.18	1.78
Delayed free recall		2.50	2.05
Delayed recognition		3.07	1.51
Executive functioning			
Letter-word generation		8.78	2.93

Note: MMSE, Mini-Mental Status Examination; IST-60, Isaacs Set Test score at 60 seconds.

obtained from the other mobile cognitive tests were not significantly affected by study duration.

Concordance between baseline AMI neuropsychological tests and mobile cognitive tests

Scores for mobile cognitive tests were significantly correlated with baseline neuropsychological test scores, including when adjusted for practice effects associated with the number of mobile tests administered (Table 3). This was true for the semantic memory test ($\gamma=0.085$, $t=5.598$, $p<0.001$), the letter-word generation test ($\gamma=0.168$, $t=4.562$, $p<0.001$), the word list free recall ($\gamma=0.069$, $t=3.156$, $p<0.01$) and delayed recall ($\gamma=0.248$, $t=2.547$, $p<0.05$) tests, and the word list recognition test ($\gamma=0.091$, $t=2.609$, $p<0.05$).

Discussion

Mobile technologies have been applied in clinical research for nearly three decades and its use is rapidly expanding among elderly populations (Cain *et al.*, 2009). Very recent findings have confirmed the value of mobile cognitive testing in this population as a means of reducing the margin of error associated with scores derived from traditional neuropsychological assessments, thus providing more sensitive tools for the detection of corresponding brain markers as well as for the identification of the daily life activities that may improve cognitive functioning (Allard *et al.*, 2015). Despite this progress, the feasibility and validity of mobile cognitive testing has been examined by very few investigations (Brouillette *et al.*, 2013; Oliveira *et al.*, 2014; Scholey *et al.*, 2012; Schuster *et al.*, 2015; Timmers *et al.*, 2014) and no study to date has estimated fatigue

Table 2. Fatigue and practice effects associated with time in the study or repetition of mobile cognitive tests

	Coefficient	<i>t</i> Ratio	Significance
Fatigue effect	−0.360	−9.872	0.000
Practice effects			
Semantic memory and verbal fluency	0.420	2.851	0.005
Semantic category generation	0.430	1.425	0.155
List-learning immediate free-recall	0.090	1.285	0.200
List-learning immediate recognition	0.266	2.331	0.021
List-learning delayed free-recall	0.328	2.339	0.020
List-learning delayed recognition	−0.958	−10.064	0.000
Letter-word generation	−0.094	−0.592	0.554

Note. Fatigue is the effect of study duration on the number of missing observations. Practice effects are the effect of the number of tests completed on test scores.

Table 3. Association of neuropsychological tests with mobile cognitive test scores

Neuropsychological test	Mobile cognitive test	Covariates	Coefficient	<i>t</i> Ratio	<i>p</i> Value
Isaacs Set Test	Semantic memory and verbal fluency	no adjustment	0.086	5.832	0.000
		test number	0.085	5.598	0.000
		age, sex	0.090	5.977	0.000
		test number, age, sex	0.088	5.673	0.000
Grober and Buschke free recall	List-learning free recall	no adjustment	0.069	3.120	0.003
		test number	0.069	3.156	0.002
		age, sex	0.059	2.496	0.015
		test number, age, sex	0.059	2.524	0.014
Grober and Buschke free recall	List-learning recognition	no adjustment	0.092	2.550	0.013
		test number	0.091	2.609	0.011
		age, sex	0.085	2.267	0.027
		test number, age, sex	0.082	2.275	0.026
Grober and Buschke delayed free recall	List-learning delayed free recall	no adjustment	0.245	2.459	0.017
		test number	0.248	2.547	0.014
		age, sex	0.224	2.232	0.030
		test number, age, sex	0.226	2.316	0.024
Wechsler symbol coding	Letter-word generation	no adjustment	0.167	4.535	0.000
		test number	0.168	4.562	0.000
		age, sex	0.181	4.819	0.000
		test number, age, sex	0.182	4.834	0.000

or reactive effects in the elderly. In a cohort of elderly community residents, this investigation examined the acceptability and feasibility of mobile cognitive testing through EMA, as well as compliance with its multiple daily assessments, potential biases associated with repeated testing, and its convergent validity with traditional neuropsychological assessments.

Concerning the basic feasibility of mobile cognitive testing in the elderly, the findings indicate that the most significant barrier to its use is likely to be initial acceptance and compliance with study procedures. In this sample of

non-demented elderly persons, we found that only a moderate majority of participants (66%) accepted to use the mobile assessments. Those who refused participation were older and had lower scores for general cognitive functioning. These acceptance rates are lower than those found in either psychiatric or neurologic samples when EMA is used without cognitive testing (Johnson *et al.*, 2009a, 2009b), but may be lower due to the acceptance by this sub-sample of additional assessments involved in the larger research protocol (including magnetic resonance imaging [MRI]). It is also important to note that approximately one-third

(34%) of individuals who accepted to participate did not respond to enough EMA methods to be considered as minimally compliant with the study procedures. It is possible that these compliance rates could have been improved if specific procedures were implemented to encourage participation (such as financial compensation). Taken together, however, these acceptance and compliance rates indicate that the use of mobile cognitive assessments in the elderly may be limited to a subpopulation of higher-functioning individuals, but that such samples may still include individuals with very mild cognitive deficits.

While mobile cognitive testing may be feasible in the majority of elderly individuals, evaluation of participant compliance with the cognitive assessments requires attention to essential limitations associated with the fact that no clinician was present to administer the tests or to interpret the context in which the test was completed. In particular, a minority of the verbal responses recorded by participants provided potential evidence of unauthorized assistance (such as a spouse who reminds the participant of the words on a list-learning task). When these items were excluded from analyses, the sample nonetheless responded to 72% of all programmed cognitive tests. This response rate is only moderately lower than the overall response rate (82%) for other EMA questions in this same sample and which are again highly similar to those observed in healthy or clinical populations (Cain *et al.*, 2009; Johnson *et al.*, 2009a).

The validity of the data collected was also examined for potential biases associated with the repeated-testing methodology and for expected associations between mobile cognitive performance scores with traditional neuropsychological tests measuring similar cognitive functions. No evidence was found for a fatigue effect whereby individuals may increasingly miss or skip assessments as a function of time in the study, but rather participants became even more compliant over time in completing the mobile cognitive tests. However, the scores for some tests were found to vary as a function of the number of tests administered and therefore indicate potential practice effects. As the actual item content of each test was unique for each administration, these effects are likely to be attributable to the participant's learning of how to respond rather than memorizing previously-presented content. In any case, such practice effects can be readily controlled by researchers adjusting for the number of times in which a particular test had been administered to participants. Finally, the significant correlations observed with traditional neuropsychological tests of semantic memory, episodic memory and executive functions provide support for the convergent validity of the mobile cognitive assessments

developed for this study. While their use may not be feasible in all elderly individuals, the resulting data may nonetheless provide more broadly generalizable data for the identification of daily life activities associated with improved cognitive health (Allard *et al.*, 2015) as well as for the examination of clinical markers associated with initial cognitive decline.

The legitimate enthusiasm for mobile cognitive assessments in the elderly must be balanced with knowledge of its limitations and comparative value relative to other instruments and measures of cognitive functioning. The present study provides support for the feasibility and validity of these novel assessment tools, but the findings should be interpreted in light of several methodological issues. Perhaps most importantly, the present sample of elderly individuals is characterized by relatively low education levels that, along with other study procedures (such as MRI), may have affected study acceptance or compliance rates. The development of a wider range of mobile tests would also allow for more comprehensive assessments of cognitive functions and their examination relative to a range of socio-demographic characteristics should reinforce knowledge for the value of mobile testing. It is also important to note that mobile cognitive testing generates data that may be qualitatively different from that obtained from traditional neuropsychological testing, and therefore should be considered only as complementary information to such clinical information. As the neuropsychological tests used in the larger AMI study were not always adapted to administration through a mobile phone, the mobile tests were developed to broadly assess the same cognitive function as the neuropsychological test but they used content and administration procedures that were different. Moreover, as the study procedures excluded individuals with dementia, we were not able to examine the sensitivity or specificity of the mobile tests for detecting cases. Despite these limitations, the challenge of detecting subtle cognitive decline at the earliest stages of dementia requires the development of new methods and more sensitive assessment tools. Mobile technologies are an important contribution to this effort.

Acknowledgments

This investigation was supported by funding from the French Ministry of Health (PHRC "AmiMage"), the Agence Nationale de la Recherche as well as the Association France Alzheimer AAP SHS 2012.

Declaration of interest statement

The authors have no competing interests.

References

- Allard M., Husky M., Catheline G., Pelletier A., Dilharreguy B., Amieva H., Pérès K., Foubert-Samier A., Dartigues J.F., Swendsen J. (2015) Mobile technologies in the early detection of cognitive decline. *PLoS One*, **9**(12), e112197. DOI:10.1371/journal.pone.0112197.
- Amieva H., Le Goff M., Millet X., Orgogozo J.M., Peres K., Barberger-Gateau P., Jacqmin-Gadda H., Dartigues J.F. (2008) Prodromal Alzheimer's Disease: successive emergence of the clinical symptoms. *Annals of Neurology*, **64**(5), 492–498.
- Buczyłowska D., Petermann F. (2016) Age-related differences and heterogeneity in executive functions: analysis of NAB executive functions module scores. *Archives of Clinical Neuropsychology*, **31**(3), 254–262.
- Brouillette R.M., Foil H., Fontenot S., Corroero A., Allen R., Martin C.K., Bruce-Keller A.J., Keller J.N. (2013) Feasibility, reliability, and validity of a smartphone based application for the assessment of cognitive function in the elderly. *PLoS One*, **8**(6), e65925.
- Cain A.E., Depp C.A., Jeste D.V. (2009) Ecological momentary assessment in aging research: a critical review. *Journal of Psychiatric Research*, **43**(11), 987–996.
- Content A., Mousty P., Radeaux M. (1990) BRULEX: une base de données lexicale informatisée pour le français écrit et parlé. *L'Année Psychologique*, **90**, 551–566.
- Grober E., Buschke H. (1987) Genuine memory deficits in dementia. *Developmental Neuropsychology*, **3**(1), 13–36.
- Harada C.N., Natelson Love M.C., Triebel K.L. (2013) Normal cognitive aging. *Clinics in Geriatric Medicine*, **29**(4), 737–752.
- Hess T.M., Popham L.E., Emery L., Elliott T. (2012) Mood, motivation, and misinformation: aging and affective state influences on memory. *Neuropsychology, Development, and Cognition. Section B, Aging, Neuropsychology and Cognition*, **19**(1–2), 13–34. DOI:10.1080/13825585.2011.622740.
- Holtzer R., Verghese J., Wang C., Hall C.B., Lipton R.B. (2008) Within-person across-neuropsychological test variability and incident dementia. *JAMA*, **300**(7), 823–830.
- Isaacs B., Kennie A.T. (1973) The Set test as an aid to the detection of dementia in old people. *British Journal of Psychiatry*, **123**(575), 467–470.
- Johnson E.I., Grondin O., Barrault M., Faytout M., Helbig S., Husky M., Granholm E.L., Loh C., Nadeau L., Wittchen H.U., Swendsen J. (2009a) Ambulatory monitoring in psychiatry: a multi-site collaborative study of acceptability, compliance, and reactivity. *International Journal of Methods in Psychiatric Research*, **18**(1), 48–57.
- Johnson E.I., Sibon I., Renou P., Rouanet F., Allard M., Swendsen J. (2009b) Feasibility and validity of computerized ambulatory monitoring in stroke patients. *Neurology*, **73**(19), 1579–1583.
- Metternich B., Schmidtke K., Hüll M. (2009) How are memory complaints in functional memory disorder related to measures of affect, metamemory and cognition? *Journal of Psychosomatic Research*, **66**(5), 435–444. DOI:10.1016/j.jpsychores.2008.07.005.
- Oliveira J., Gamito P., Morais D., Brito R., Lopes P., Norberto L. (2014) Cognitive assessment of stroke patients with mobile apps: a controlled study. *Studies in Health Technology and Informatics*, **199**, 103–107.
- Pérès K., Matharan F., Allard M., Amieva H., Baldi I., Barberger-Gateau P., Bergua V., Bourdel-Marchasson I., Delcourt C., Foubert-Samier A., Fourrier-Réglat A., Gaimard M., Laberon S., Maubaret C., Postal V., Chantal C., Rainfray M., Rasle N., Dartigues J.F. (2012) Health and aging in elderly farmers: the AMI cohort. *BMC Public Health*, **12**, 558.
- Raudenbush S.W., Bryk A.S., Cheong Y.F., Congdon R.T. (2004) HLM 6: Hierarchical Linear and Nonlinear Modeling, Chicago, IL: Scientific Software International.
- Schmidt C., Collette F., Cajochen C., Peigneux P. (2007) A time to think: circadian rhythms in human cognition. *Cognitive Neuropsychology*, **24**(7), 755–789. DOI:10.1080/02643290701754158.
- Scholey A.B., Benson S., Neale C., Owen L., Tiplady B. (2012) Neurocognitive and mood effects of alcohol in a naturalistic setting. *Human Psychopharmacology*, **27**(5), 514–516.
- Schuster R.M., Mermelstein R.J., Hedeker D. (2015) Acceptability and feasibility of a visual working memory task in an ecological momentary assessment paradigm. *Psychological Assessment*, **27**(4), 1463–1470.
- Timmers C., Maeghs A., Vestjens M., Bonnemayer C., Hamers H., Blokland A. (2014) Ambulant cognitive assessment using a smartphone. *Applied Neuropsychology – Adult*, **21**(2), 136–142.
- Tollenaar M.S., Elzinga B.M., Spinhoven P., Everaerd W. (2009) Immediate and prolonged effects of cortisol, but not propranolol, on memory retrieval in healthy young men. *Neurobiology of Learning and Memory*, **91**(1), 23–31. DOI:10.1016/j.nlm.2008.08.002.
- Wechsler D. (2008) Wechsler Adult Intelligence Scale- Fourth Edition: Technical and Interpretive Manual, TX, Pearson: San Antonio.
- World Health Organization (2015) Dementia, Fact sheet n°362. <http://www.who.int/mediacentre/factsheets/fs362/en/> [July 2015].