International Psychogeriatrics (2015), 27:2, 231–242 © International Psychogeriatric Association 2014
doi:10.1017/S1041610214002038

The utility of a non-verbal prospective memory measure as a sensitive marker for early-stage Alzheimer's disease in Hong Kong

C. S. Tse,[1] J. F. Chang,[2] Ada W. T. Fung,[3] Linda C. W. Lam,[3] K. T. Hau,[1] Grace T. Y. Leung[3] and D. A. Balota[4]

[1]*Department of Educational Psychology, The Chinese University of Hong Kong, Hong Kong*
[2]*Department of Psychology, Guangdong University of Education, China*
[3]*Department of Psychiatry, The Chinese University of Hong Kong, Hong Kong*
[4]*Department of Psychology and Neurology, Washington University in St. Louis, USA*

ABSTRACT

Background: With the proportion of older adults in Hong Kong projected to double in size in the next 30 years, it is important to develop measures for detecting individuals in the earliest stage of Alzheimer's disease (AD, 0.5 in Clinical Dementia Rating, CDR). We tested the utility of a non-verbal prospective memory task (PM, ability to remember what one has to do when a specific event occurs in the future) as an early marker for AD in Hong Kong Chinese.

Methods: A large community dwelling sample of older adults who are healthy controls (CDR 0, $N = 125$), in the earliest stage of AD (CDR 0.5, $N = 125$), or with mild AD (CDR 1, $N = 30$) participated in this study. Their reaction time/accuracy data were analyzed by mixed-factor analyses of variance to compare the performance of the three CDR groups. Logistic regression analyses were performed to test the discriminative power of these measures for CDR 0 versus 0.5 participants.

Results: Prospective memory performance declined as a function of AD severity: CDR 0 > CDR 0.5 > CDR 1, suggesting the effects of early-stage AD and AD progression on PM. After partialling out the variance explained by psychometric measures (e.g., ADAS-Cog), reaction time/accuracy measures that reflected the PM still significantly discriminated between CDR 0 versus 0.5 participants in most of the cases.

Conclusion: The effectiveness of PM measures in discriminating individuals in the earliest stage of AD from healthy older adults suggests that these measures should be further developed as tools for early-stage AD discrimination.

Key words: Alzheimer's disease, cognitive testing, memory

Introduction

Given the high prevalence of dementia worldwide and that there is currently no therapeutic intervention that can reverse the course of most dementias (e.g., Jalbert *et al.*, 2008), there has been considerable interest in developing measures that discriminate healthy aging from individuals who are in the earliest stage of Alzheimer's disease (AD), as defined by 0.5 in Clinical Dementia Rating (CDR, e.g., Morris *et al.*, 1997; Lam *et al.*, 2008). The purpose of the current study was to test whether a nonverbal event-based prospective memory (PM) task (Burgess *et al.*, 2001) could discriminate individuals who are in the earliest stage of AD from healthy older adults in Hong Kong Chinese population.

Many cognitive tasks tap older adults' memory abilities but focus almost exclusively on their retrospective memory, e.g., studying a list of unrelated words and then recalling them (e.g., Lam *et al.*, 2005; Albert *et al.*, 2007). In contrast, PM, the ability to remember what one has to do in the future, has received much less attention in Hong Kong, despite its clear connection with older adults' daily-life activities. At the beginning of a PM task, participants are told to spontaneously perform an intended action (PM instruction) when a specific event occurs (event-based PM) or at a specific time (time-based PM), while doing a concurrent attention-demanding task. This task

Correspondence should be addressed to: C. S. Tse, Department of Educational Psychology, The Chinese University of Hong Kong, Shatin, N. T., Hong Kong. Phone: +852-39436751; fax: +852-26036921. E-mail: cstse@cuhk.edu.hk.
Received 23 Jun 2014; revision requested 1 Aug 2014; revised version received 17 Aug 2014; accepted 22 Aug 2014. First published online 23 September 2014.

design is analogous to daily-life situations, for example, remembering to take medicine before sleep (event-based PM) or after every four hours (time-based PM). PM performance can be affected by a retrospective component (encoding, storage, and retrieval of the intended action and the target event/time) and a prospective component (ability to initiate the intended action at the right moment without being given any explicit prompt to recall) (McDaniel and Einstein, 2007). In the current study, we controlled for the retrospective component (by including only participants who could recall the PM instruction at the end of the task) and investigated the effect of AD on the prospective component.

Previous studies reported that older adults with AD often demonstrated PM failures in daily-life activities, which pose a great challenge for their caretakers (Camp *et al.*, 1996). Healthy older adults with genetic risk factor for AD (apolipoprotein E allele) showed poorer event-based PM performance than those without this factor (e.g., Duchek *et al.*, 2006). PM performance is also a good discriminator between CDR 0 (healthy older adults) and CDR 0.5 individuals (older adults in the earliest stage of AD) (e.g., Duchek *et al.*, 2006; McDaniel *et al.*, 2011, see also similar results for individuals with mild cognitive impairment (MCI) in, e.g., Troyer and Murphy, 2007; Karantzoulis *et al.*, 2009; Costa *et al.*, 2010) and between CDR 0 and CDR 1 individuals (older adults with mild AD) (Martins and Damasceno, 2008, 2012). In Hong Kong, to our knowledge only one published study (Gao *et al.*, 2013) examined the effect of AD on event-based PM performance, although it focused on the effect of mild AD, but not early-stage AD. They found that mild AD individuals showed lower PM performance than healthy controls and the interference produced by holding a PM instruction on concurrent task performance was larger for mild AD individuals than for healthy controls.

It is important to note that older adults in Hong Kong received much less education than those in Western population. For example, about 80% of older adults finish their high-school education in the United States (U.S. Department of Health and Human Services, 2011), whereas in Hong Kong the estimate is only 31% (Hong Kong Census and Statistics Department, 2011). Previous studies reported that some items in psychometric tasks have differential discriminative powers of CDR 0.5 versus 0 groups for individuals who have high versus low education levels (e.g., Chang *et al.*, 2014). Hence, it is important to determine if a PM measure could discriminate the earliest detectable stage of AD (CDR 0.5), compared to healthy older adults in a less well-educated Hong Kong Chinese population. We also examined whether the PM measures would still discriminate the performance between CDR 0.5 versus 0 individuals after taking into account their general cognitive functioning, as reflected by standard psychometric measures. Furthermore, we included a group of CDR 1 participants to test whether the results of Gao *et al.* that mild AD individuals showed worse PM performance than healthy older adults could be conceptually replicated in another PM paradigm.

The time-based PM tasks are more difficult than event-based PM tasks for older adults, who often have difficulty monitoring time continuously. In order to avoid floor effects (which would decrease sensitivity), we adopted an event-based PM task from Burgess *et al.* (2001) in the current study. As shown in Appendix A (available as supplementary material attached to the electronic version of this paper at www.journals.cambridge.org/jid_IPG), on a non-PM trial, participants responded to a highlighted arrow by pressing left or right key. On some infrequent trials (*PM trials*—the specific event that demands a different response—in this case, the two color bars are in the same color), participants pressed an alternative key, instead of judging the arrow direction. Because the task does not involve verbal materials, it is suitable for our participants whose education levels are generally low.

To recapitulate, we investigated the utility of a nonverbal prospective memory measure as an early marker for Alzheimer's disease in a community sample of Hong Kong participants. We focused on PM trials as they should reflect the extent to which participants remembered to perform a certain action in response to the occurrence of a specific event. Two hypotheses were tested. First, individuals who were in the earliest stage of AD (CDR 0.5) would respond more slowly and less accurately to the PM trials than healthy older adults (CDR 0), and participants who received a CDR of 1 would produce even worse performance than CDR 0.5 participants. Second, PM-trial measures would be able to discriminate the CDR 0.5 versus 0 participants after taking into account the standardized measures of general cognitive functioning, as reflected by their psychometric performance.

Methods

Participants

A total of 125 CDR 0, 125 CDR 0.5, and 30 CDR 1 Hong Kong community-dwelling older adults aged 60 or above participated in this study (see Table 1). A progression to CDR 1 indicates conversion to mild AD and CDR 0.5 participants

Table 1. Demographic information of participants as a function of CDR status

	CDR 0		CDR 0.5		CDR 1	
	M	*SD*	*M*	*SD*	*M*	*SD*
Sample size	125		125		30	
Gender (male: female)	63:62		56:69		9:21	
Age	75.06	5.95	78.66	5.81	80.20	7.28
Number of years of education	7.10	4.91	4.74	4.59	4.10	4.30
Cantonese version of MMSE scores (Tse *et al.*, 2013)	27.72	1.84	24.79	2.94	21.43	4.47

Note. M = mean; SD = standard deviation. CDR 0 = healthy older adults. CDR 0.5 = early-stage AD individuals. CDR 1 = mild AD individuals. The C-MMSE scores was available for 114 CDR 0 participants and all CDR 0.5 and CDR 1 participants. Across the three CDR groups, age and number of years of education (education level) were significantly different, $F(2,277) = 15.20$, $MSE = 36.49$, $p < 0.01$, $\eta_p^2 = 0.10$ and $F(2,277) = 9.87$, $MSE = 22.15$, $p < 0.01$, $\eta_p^2 = 0.07$, whereas proportion of gender was not ($\chi^2(2) = 3.18$, $p = 0.20$).

are regarded as those individuals who are in the earliest stage of AD (i.e., early-stage AD). The samples were randomly chosen from those who participate in a prospective study of cognitive function led by one of the coauthors. Research assistants contacted all potential participants via phone invitation and arranged the appointments with consented participants and/or their caregivers. Participants provided their informed consents at the beginning of the study. They were tested individually by a research assistant on a laptop computer at their place of residence or regional social centers. This study was approved by The Joint Chinese University of Hong Kong—New Territories East Cluster Clinical Research Ethics Committee.

The level of global cognitive impairment was assessed by CDR (Morris *et al.*, 1997) based on a semi-structured clinical interview by trained and certified CDR raters. They assessed participants and obtained collateral information from their informants (e.g., spouses) on the current status of their functional abilities on memory, orientation, personal care, community affairs, judgment/problem solving, and home/hobbies, without considering their status in previous interviews or psychometric performance. The CDR has been shown to be highly predictive of pathology consistent with AD based on autopsy (e.g., Storandt *et al.*, 2006; Storandt, 2008). The reliability and validity of the CDR in identifying significant early cognitive deficit in older adults have been well-documented in Hong Kong population (e.g., Lam *et al.*, 2008, 2010). As indicated in Table 2, significant differences between CDR 0 and CDR 0.5 participants' performance in various psychometric measures also showed that CDR could discriminate participants with different dementia severity. Diagnosis of clinical AD followed the Diagnostic and Statistical Manual for Mental Disorder – IV edition (DSM-IV TR).

Participants' psychometric data were adapted from the aforementioned prospective study (see Table 2). We made use of these data in the current study because we wanted to test whether the new PM measure would still discriminate CDR 0.5 participants from CDR 0 participants after controlling for their performance in these psychometric tasks. This analytic strategy was similar to those used in previous studies (e.g., Aschenbrenner *et al.*, in press; Duchek *et al.*, 2009; Tse *et al.*, 2010a). To insure that the psychometric measures reflect current cognitive functioning of our participants, we used the psychometric data that were collected as closely to the time when they performed the PM testing as possible. The mean signed interval between the time of psychometric testing (first) and PM testing (second) sessions (i.e., session lag) was 53.40 days (SD = 108.96). Given that only 12 (40%) CDR 1 participants had their recent psychometric data available and there was a large imbalance in the sample size of this group with the other two (12 vs. ~100), we focus on the more important comparison of the CDR 0 and 0.5 participants. Participants were screened for depression, untreated hypertension, reversible dementia, and other disorders that could potentially produce cognitive impairment. Research assistants checked to make sure that they did not have color-blindness.

Design and procedures

The PM task was closely adapted from Burgess *et al.* (2001) and conducted via an E-Prime program. It was participant-paced: stimuli would stay on the screen until the participant responded. From trial to trial (see Appendix A available as supplementary material attached to the electronic version of this paper at www.journals.cambridge.org/jid_IPG), either the left or right arrow appeared randomly in dark grey and the color of the top and

Table 2. CDR 0 and CDR 0.5 participants' psychometric performance and their partial correlation with measures in PM trials

	CDR 0 PARTICIPANTS			CDR 0.5 PARTICIPANTS			COMPARISON		PARTIAL CORRELATION	
	N	*M*	*SD*	*N*	*M*	*SD*	*F*	η_p^2	PROPORTION CORRECT	RT FOR CORRECT RESPONSES
ADAS-Cog total error score	97	7.49	2.27	124	12.31	3.68	**88.99***	0.29	**− 0.32***	**0.22***
ADAS-Cog command errors	97	0.18	0.38	125	0.60	0.88	**15.36***	0.07	− 0.06	0.10
ADAS-Cog comprehension errors	97	0.02	0.14	125	0.12	0.41	3.81	0.02	− 0.05	0.15
ADAS-Cog constructional praxis errors	96	0.64	0.51	124	1.10	0.67	**12.85***	0.06	**− 0.24***	**0.17***
ADAS-Cog delayed recall scores	97	5.75	1.92	125	3.03	1.79	**88.57***	0.29	**0.21***	− 0.14
ADAS-Cog ideational praxis errors	97	0.05	0.22	125	0.17	0.42	1.58	0.01	− 0.13	0.11
ADAS-Cog word recall errors	97	4.62	1.27	125	6.06	1.29	**45.35***	0.17	**− 0.16***	**0.16***
ADAS-Cog language disability	97	0.01	0.10	125	0.01	0.09	0.50	0.002	–	–
ADAS-Cog naming objects & fingers errors	97	0.04	0.20	125	0.16	0.37	**4.07***	0.02	− 0.12	0.14
ADAS-Cog orientation errors	97	0.04	0.20	125	0.55	0.82	**19.33***	0.08	**− 0.15***	0.11
ADAS-Cog remembering instruction errors	97	0.01	0.10	125	0.10	0.30	2.62	0.01	**− 0.20***	0.02
ADAS-Cog spoken language errors	97	0.00	0.00	125	0.01	0.09	0.23	0.001	–	–
ADAS-Cog word finding difficulty	97	0.00	0.00	125	0.00	0.00	–	–	–	–
ADAS-Cog word recognition errors	97	1.90	1.21	124	3.50	1.98	**27.74***	0.11	**− 0.22***	0.07
Verbal fluency item generation in 30 s	97	27.92	5.85	125	24.11	5.82	**8.96***	0.04	0.09	**− 0.23***
Verbal fluency item generation in 60 s	97	39.60	8.36	125	33.56	8.28	**12.88***	0.06	0.09	**− 0.21***
Verbal fluency intrusion errors	97	0.16	0.49	125	0.16	0.46	0.28	0.001	**− 0.18***	**0.36***
Forward/backward digit span	97	7.55	1.13	125	6.82	1.30	**12.77***	0.06	0.07	**− 0.15***
Forward/backward visual span	96	3.99	0.88	122	3.67	0.91	2.10	0.01	**0.19***	0.03

Note. $^*p < 0.05$ (two-tailed). N = number of participants who have available psychometric data; M = mean; SD = standard deviation. The F and η_p^2 statistics of the significance test for a difference between the CDR 0 and 0.5 groups in each of the psychometric measures (after partialling out their age and education level) were presented in the "Comparison" column. The session lag (i.e., the signed interval between psychometric testing session and PM testing session), participants' age and education level were partialed out in the partial correlation analyses. Because (a) participants' scores were all zero in ADAS-Cog word finding difficulty, (b) only one CDR 0.5 participant scored 1 (with all other participants scored 0) in the ADAS-Cog spoken language errors, (c) only one CDR 0 and one CDR 0.5 participants scored 1 (with all other participants scored 0) in the ADAS-Cog language disability, no statistical test was performed for these three variables in the following analyses.

bottom bars randomly appeared in different color (red, blue, yellow or green).

Participants were positioned with forefinger, middle finger, and ring finger of their right hand resting on the 3 horizontally arranged keys. They were asked to press with their forefinger if the arrow appeared on the left and with their ring finger if it appeared on the right (i.e., the *non-PM trials*).

When the top and bottom bars were in the same color (i.e., the *PM trials*), they needed to respond with their middle finger instead. Prior to the task, the research assistant read aloud the instructions to the participants and asked them to repeat back the instructions to ensure that they fully understood the task. They were told it is important to keep this PM instruction in mind because they would be required to repeat that later. At the beginning of the task, 10 practice trials were given to familiarize the participants with the task setting, but there was no PM trial to interfere with the novelty of the trials. One hundred and twenty trials were then presented with 24 PM trials being randomly distributed in the sequence. There were 4 self-paced breaks (i.e., every 30 trials). Both accuracy and RT for PM and non-PM trials were recorded. Participants were not reminded to respond with their middle finger if they forgot to do so in any PM trials. At the end, participants were asked to recall the task instruction in a retrospective memory report (including PM trials, e.g., Costa *et al.*, 2010; Zhou *et al.*, 2012). The task took about 10–15 minutes to complete.

Data analyses

The significance level was set at $p < 0.05$, two-tailed. Given the significant differences in age and education level among the three CDR groups (see Table 1), we partialed out these variables in all of the following analyses, via analyses of covariance. Thirteen (10%) CDR 0, 15 (12%) CDR 0.5, and 14 (47%) CDR 1 participants failed to recall the task instruction in the retrospective memory report. Their data were excluded to ensure that participants' PM performance in the following analyses reflected the prospective component of their PM. We conducted three sets of analyses. First, to investigate the overall PM performance for the three CDR groups, we separately submitted participants' median RT for correct responses and mean proportion correct measures to a 3 (CDR: 0, 0.5, or 1) × 2 (Trial Type: PM or non-PM) mixed-factor ANCOVA. Second, we conducted correlation analyses to investigate the relationship between participants' performance in the PM task and in typical psychometric tasks. Third, we conducted logistic regression analyses to test whether the performance in the PM trials would discriminate CDR 0.5 versus 0 participants even after taking their psychometric performance into account in the regression analyses. Only the measures that significantly discriminated the two groups of participants, as shown in the ANCOVA, were examined in the regression analyses. As more than half of CDR 1 participants did not have their recent psychometric data available, we focus on the CDR 0 versus 0.5 comparisons in these analyses. Although we used the psychometric data that were collected as closely to the time when participants' performed the PM testing as possible, we still controlled for the lag between the time they did the psychometric tasks and the time they did the PM task by entering the signed interval between psychometric testing (first) and PM testing (second) sessions (*session lag*) in the first step of all regression analyses. We also controlled for participants' age and education level by entering these variables in this step.

Results

Overall performance of CDR 0, 0.5, and 1 groups

Figure 1 and Appendix B (available as supplementary material attached to the electronic version of this paper at www.journals.cambridge.org/jid_IPG) show the cell means of the findings. (The degrees of freedom are different in RT and proportion correct analyses due to some participants' zero accuracy in PM trials (i.e., missing cells) in the RT analyses.) The main effect of CDR was significant (RT: $F(2,203) = 13.27$, $MSE = 234283$, $\eta_p^2 = 0.12$; Proportion correct: $F(2,233) = 9.43$, $MSE = 0.04$, $\eta_p^2 = 0.08$), but the main effect of Trial Type was not (RT: $F(1,203) = 1.18$, $MSE = 142{,}803$, $\eta_p^2 = 0.01$; Proportion correct: $F(1,233) = 0.03$, $MSE = 0.04$, $\eta_p^2 = < 0.001$). More importantly, the CDR × Trial Type interaction was significant (RT: $F(2,203) = 3.86$, $MSE = 142{,}803$, $\eta_p^2 = 0.04$; Proportion correct: $F(2,233) = 3.94$, $MSE = 0.04$, $\eta_p^2 = 0.03$). Follow-up analyses were conducted to test the CDR 0 versus 0.5 difference and CDR 0.5 versus 1 difference. The median RT for correct responses and mean proportion correct measures of PM trials and non-PM trials were separately submitted to a 2 (CDR: 0 or 0.5) between-group ANCOVA and a 2 (CDR: 0.5 or 1) between-group ANCOVA.

The CDR 0 versus 0.5 group differences were significant in PM trials (RT: $F(1,191) = 7.32$, $MSE = 278295$, $\eta_p^2 = 0.04$; Proportion correct: $F(1,218) = 4.70$, $MSE = 0.08$, $\eta_p^2 = 0.02$) and non-PM trials (RT: $F(1,218) = 13.25$, $MSE = 74{,}908$, $\eta_p^2 = 0.06$; Proportion correct: $F(1,218) = 7.01$, $MSE = 0.004$, $\eta_p^2 = 0.03$). The CDR 0.5 versus 1 group differences were significant in PM trials (RT: $F(1,101) = 5.08$, $MSE = 476{,}885$, $\eta_p^2 = 0.05$; Proportion correct: $F(1,122) = 4.77$, $MSE = 0.09$, $\eta_p^2 = 0.04$), but not in non-PM trials (RT: $F(1,122) = 0.69$, $MSE = 107{,}270$, $\eta_p^2 = 0.01$; Proportion correct: $F(1,122)$

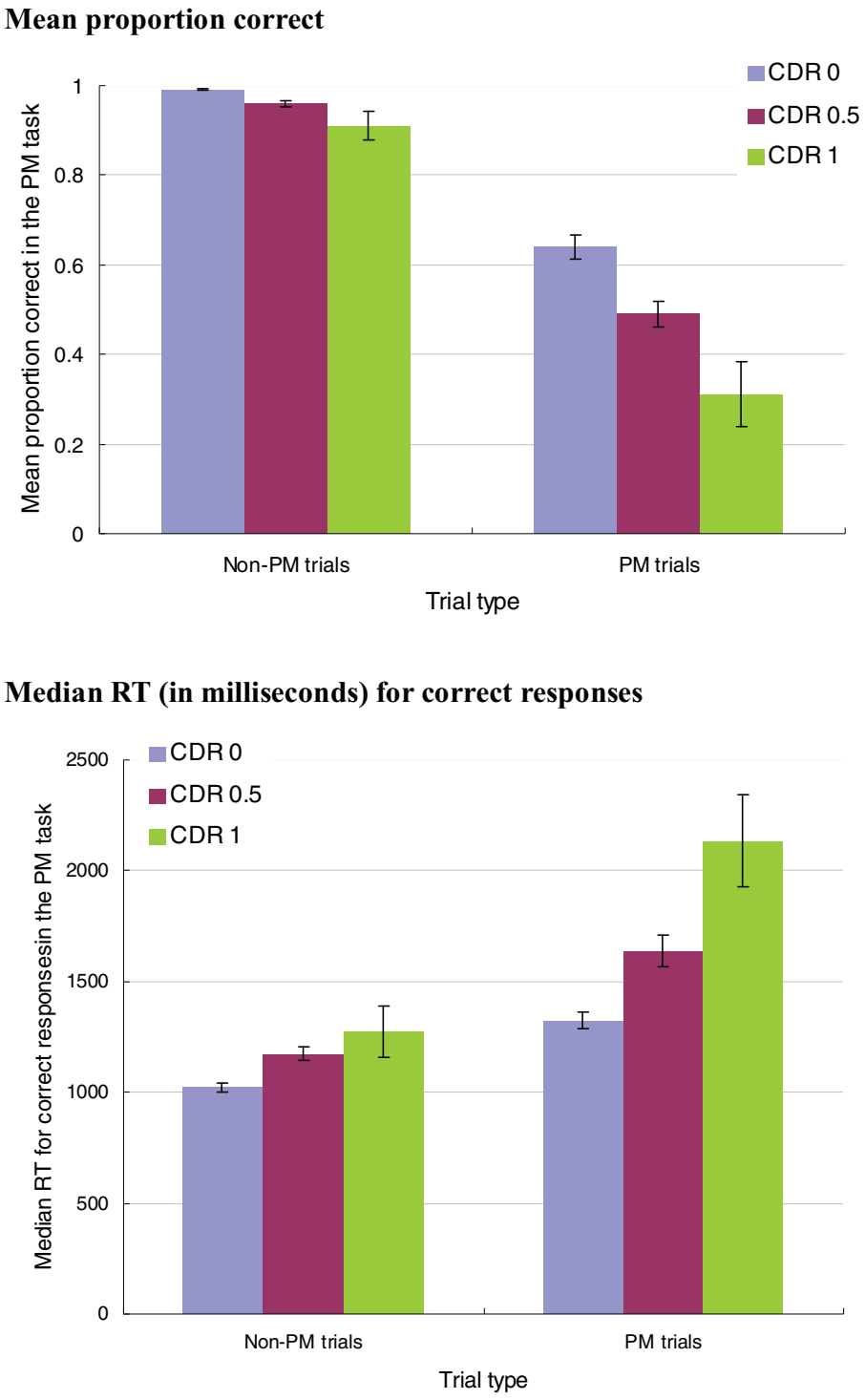

Figure 1. (Colour online) Performance in the prospective memory task as a function of CDR status. Mean proportion correct. Median RT (in milliseconds) for correct responses.
Note. Median RTs are in millisecond. Error bars indicate the standard errors of means.

$= 3.53$, $MSE = 0.01$, $\eta_p^2 = 0.03$). Overall, in PM trials that are of our most interest, the proportion correct significantly decreased, whereas the RT significantly increased, as a function of AD severity, suggesting that the prospective components of PM performance were significantly affected by AD even in the earliest detectable stage.

Predictive utility of PM measures in discriminating individuals in the earliest detectable stage of AD from healthy older adults

Table 2 presents the partial correlation between psychometric performance and the measures in

Table 3. Logistic regression analyses of proprtion correct in PM trials on predicting CDR status (0 vs. 0.5)

	FOR X, AFTER PARTIALLING OUT SESSION LAG, PARTICIPANTS' AGE, AND EDUCATION LEVEL (I.E., IN THE SECOND STEP)			FOR PROPORTION CORRECT IN PM TRIALS, AFTER PARTIALLING OUT X (I.E., IN THE THIRD STEP)		
VARIABLE X	$\chi^2(1)$	ODDS RATIO	NAGELKERKE'S R^2	$\chi^2(1)$	ODDS RATIO	NAGELKERKE'S R^2
Proportion correct in PM trials	**7.92***	**0.20***	0.21			
ADAS-Cog total error score	**75.85***	**4.17***	0.53	0.60	0.57	0.54
ADAS-Cog command errors	**13.26***	**1.45***	0.24	**6.06***	**0.23***	0.27
ADAS-Cog comprehension errors	2.53	1.14	0.17	**7.69***	**0.20***	0.22
ADAS-Cog constructional praxis errors	**8.77***	**1.60***	0.22	**5.28***	**0.25***	0.25
ADAS-Cog delayed recall scores	**72.82***	**0.22***	0.52	2.14	0.35	0.53
ADAS-Cog ideational praxis errors	1.77	1.17	0.17	**7.23***	**0.21***	0.21
ADAS-Cog word recall errors	**34.13***	**2.58***	0.35	**4.67***	**0.26***	0.37
ADAS-Cog naming objects & fingers errors	2.90	1.23	0.18	**6.79***	**0.22***	0.22
ADAS-Cog orientation errors	**23.55***	**1.52***	0.29	**4.98***	**0.27***	0.32
ADAS-Cog remembering instruction errors	0.85	1.10	0.16	**7.34***	**0.21***	0.21
ADAS-Cog word recognition errors	**32.65***	**2.19***	0.33	3.54	0.31	0.35
Verbal fluency item generation in 30 s	**6.99***	**0.65***	0.20	**6.97***	**0.22***	0.24
Verbal fluency item generation in 60 s	**10.50***	**0.59***	0.22	**6.65***	**0.22***	0.26
Verbal fluency intrusion errors	0.03	0.97	0.16	**8.46***	**0.18***	0.21
Forward/backward digit span	**10.91***	**0.63***	0.22	**6.94***	**0.21***	0.26
Forward/backward visual span	1.40	0.83	0.18	**7.84***	**0.19***	0.22

Note. *$p < 0.05$ (two-tailed). Following the procedures of previous studies (e.g., Tse *et al.*, 2010a), the psychometric measures entered in these models were first standardized using the mean performance of CDR 0 participants. Participants' age, education level, and session lag (i.e., the signed interval between psychometric testing session and PM testing session) were entered in the first step. Each of the psychometric measures was entered in the second step, prior to the entry of the PM measure in the third step. The predicted variable was the CDR status (0 vs. 0.5). The χ^2 column indicates the χ^2 statistics for the specific step of the regression model and the odds ratio column indicates the odds ratio for the proportion correct in PM trials as a predictor variable in the third step of the model. Nagelkerke's R^2 indicates the effect size.

the PM task, after controlling for session lag, participants' age, and education level. As shown, PM measures were weakly associated with episodic memory and attentional control measures (e.g., ADAS-Cog word recall errors and verbal fluency intrusion errors). This pattern suggests that the constructs measured in the PM trials are not highly overlapping with the other psychometric measures. Tables 3 and 4 summarize the findings of logistic regression analyses. When considered alone, the performance in the PM trials significantly discriminated CDR 0 versus 0.5 participants. After controlling for the variance explained by the psychometric measures, performance in the

Table 4. Logistic regression analyses of RT of correct responses in PM trials on predicting CDR status (0 vs. 0.5)

	FOR X, AFTER PARTIALLING OUT SESSION LAG, PARTICIPANTS' AGE, AND EDUCATION LEVEL (I.E., IN THE SECOND STEP)			FOR RT OF CORRECT RESPONSES IN PM TRIALS, AFTER PARTIALLING OUT X (I.E., IN THE THIRD STEP)		
VARIABLE X	$\chi^2(1)$	ODDS RATIO	NAGELKERKE'S R^2	$\chi^2(1)$	ODDS RATIO	NAGELKERKE'S R^2
RT of correct responses in PM trials	**6.56***	**1.001***	0.19			
ADAS-Cog total error score	**63.73***	**4.10***	0.51	1.24	1.001	0.51
ADAS-Cog command errors	**9.58***	**1.41***	0.21	**4.75***	1.001	0.24
ADAS-Cog comprehension errors	2.29	1.13	0.17	**5.85***	**1.001***	0.21
ADAS-Cog constructional praxis errors	**7.23***	**1.56***	0.21	**8.26***	**1.002***	0.26
ADAS-Cog delayed recall scores	**64.65***	**0.22***	0.52	2.33	1.001	0.53
ADAS-Cog ideational praxis errors	3.45	1.30	0.17	**5.71***	**1.001***	0.21
ADAS-Cog word recall errors	**26.77***	**2.40***	0.32	**4.07***	1.001	0.34
ADAS-Cog naming objects & fingers errors	3.50	1.29	0.17	**5.38***	**1.001***	0.21
ADAS-Cog orientation errors	**19.03***	**1.48***	0.27	**4.05***	1.001	0.30
ADAS-Cog remembering instruction errors	3.18	7.82	0.17	**6.56***	**1.001***	0.22
ADAS-Cog word recognition errors	**23.56***	**2.02***	0.30	**5.58***	**1.001***	0.33
Verbal fluency item generation in 30 s	**4.96***	**0.68***	0.18	**5.10***	**1.001***	0.22
Verbal fluency item generation in 60 s	**9.20***	**0.60***	0.21	**4.64***	1.001	0.24
Verbal fluency intrusion errors	0.18	0.93	0.15	**7.92***	**1.001***	0.20
Forward/backward digit span	**9.57***	**0.63***	0.21	**4.83***	1.001	0.24
Forward/backward visual span	0.14	0.94	0.16	**10.89***	**1.002***	0.23

Note. *$p < 0.05$ (two-tailed). Following the procedures of previous studies (e.g., Tse *et al.*, 2010a), the psychometric measures entered in these models were first standardized using the mean performance of CDR 0 participants. Participants' age, education level, and session lag (i.e., the signed interval between psychometric testing session and PM testing session) were entered in the first step. Each of the psychometric measures was entered in the second step, prior to the entry of the PM measure in the third step. The predicted variable was the CDR status (0 vs. 0.5). The χ^2 column indicates the χ^2 statistics for the specific step of the regression model and the odds ratio column indicates the odds ratio for the RT of correct responses in PM trials as a predictor variable in the third step of the model. Nagelkerke's R^2 indicates the effect size. Some participants had missing cells in median RT of correct responses in PM trials, so the statistics in the second step were not the same as those in Table 3.

PM trials (in particular proportion correct) still significantly predicted the CDR status in most cases, except the episodic memory measures (e.g., ADAS-Cog delayed recall). These were in line with those obtained in the correlation analyses that the constructs measured by the PM trials did not completely overlap with the psychometric measures.

Discussion

The goal of the current study was to test whether measures in the PM task could discriminate the earliest detectable stage of AD (CDR 0.5) from healthy older adults (CDR 0) in Hong Kong Chinese population. We also tested the

discriminative power of these measures after taking into account participants' psychometric measures. The current findings generally confirmed our hypotheses. Relative to CDR 0 participants, CDR 0.5 participants showed slower correct-response RT and more errors in both PM and non-PM trials. Relative to CDR 0.5 participants, CDR 1 participants showed slower correct-response RT and more errors in PM trials. This suggests that the PM task was sensitive to the changes in PM functioning due to early-stage AD (from healthy old to early-stage AD) and due to AD progression (from early-stage AD to mild AD). After partialling out the variance explained by the psychometric measures, proportion correct could still discriminate between CDR 0 and 0.5 groups on all but three of the psychometric measures (see Table 3). This was also the case for RT in PM trials (see Table 4), although it was not as robust as the proportion correct. These suggest that the PM measures are quite robust in the early-stage AD discrimination. It is noteworthy that the AD discriminative power was present even though the education level (as reflected by mean number of years of education) was lower in the current sample (~6) than those in previous studies (~14, e.g., McDaniel *et al.*, 2011). This suggests that the current nonverbal PM task is useful to discriminate the AD severity for older adults with lower education level.

Consistent with previous studies (e.g., Duchek *et al.*, 2006), CDR 0.5 participants performed worse than CDR 0 participants even though they remembered the PM instruction. This suggests their failure to track the target events and spontaneously trigger the intention, despite having intact episodic memory of the intention itself. In other words, the AD-related deficit lies in the prospective, rather than retrospective, component of PM (e.g., McDaniel and Einstein, 2007). However, it should be noted that after controlling for episodic memory abilities (e.g., ADAS-Cog delayed recall), PM measures no longer reliably discriminated CDR 0.5 versus 0 participants, although the pattern was in the predicted direction. This suggests that the difference in PM between these two groups could still be partially attributed to their differences in episodic memory. The contribution of participants' retrospective memory abilities to their PM performance should be further investigated in the future studies.

The decline in PM performance as a function of AD progression (CDR 1 vs. 0.5) was also in line with previous studies (e.g., Huppert *et al.*, 2000; Blanco-Campal *et al.*, 2009; Thompson *et al.*, 2010; McDaniel *et al.*, 2011; Gao *et al.*, 2013). Our study is limited that not all AD participants had undergone standard neuroimaging protocol for diagnosis. The neuroimaging investigation was determined at the clinical setting and had not constituted part of the present research protocol. In the future studies, it is important to increase the sample size of CDR 1 participants with standard neuroimaging measures and include the full psychometric data in order to test whether the PM measure could discriminate the early-stage versus mild AD after taking into account general cognitive abilities.

Although there are a number of important observations in these data regarding the utility of PM measures in the early detection of AD, there are also some important alternative interpretations that should be considered.

First, one could attribute the effects of early-stage and mild AD on PM performance to the high task demand. Specifically, participants might have failed to pay attention to the color of the peripheral bars, while judging the arrow direction in each trial. However, we found that group differences in PM trials remained significant after partialling out their performance in non-PM trials (RT: $F(2,202) = 6.17$, $MSE = 273{,}137$, $\eta_p^2 = 0.06$; proportion correct: $F(2,232) = 4.87$, $MSE = .08$, $\eta_p^2 = 0.04$). This shows that the CDR 0.5 and CDR 1 participants' poorer performance in PM trials was unlikely due to their being distracted more by non-PM trials.

Second, because the PM trials were randomly distributed in the PM task, it is possible that CDR 0.5 and CDR 1 participants would have relatively accurate memory for the PM instruction at the beginning of the task but then lost the intention over time as they would be fatigued by the task demands. To examine this possibility, we separately submitted the mean proportion correct and median RT for correct response per 8 PM trials to a 3 (CDR) × 3 (Trial Order: first, second and third 8) mixed-factor ANCOVA. The main effect of CDR was significant (RT: $F(2,184) = 14.77$, $MSE = 1{,}389{,}213$, $\eta_p^2 = 0.14$; Proportion correct: $F(2,233) = 6.82$, $MSE = 0.24$, $\eta_p^2 = 0.06$). The CDR × Trial Order interaction was significant in RT, $F(4,368) = 11.67$, $MSE = 503{,}669$, $\eta_p^2 = 0.11$. None of the other main effects or interactions was significant, all Fs < 1. The absence of an interaction effect including Trial Order in proportion correct suggested that the AD-related deficit in PM was not due to the fatigue induced by the task demands. The significant interaction in RT showed that individuals with higher AD severity showed greater improvement in their RTs to the PM trials across trial orders, which might merely be due to the floor effect for CDR 0 participants in their RTs. That is, CDR 0 participants might engage in the task more readily than the other groups of participants (see Appendix C available as supplementary material attached

to the electronic version of this paper at www.journals.cambridge.org/jid_IPG). In short, the AD effects that we obtained in the PM trials could not be explained by CDR 0.5/1 participants' loss of memory for their intention over time in the task.

Third, because of the number of PM trials (in one-fifth of all trials), one could argue that the current PM task might be like a switching task; that is, participants switched between the two task sets (i.e., to judge the arrow direction and to press the alternative key when the two bars were in same color). Given that previous studies showed significant differences in CDR 0.5 versus 0 individuals' task switching performance (e.g., Hutchison *et al.*, 2010; Tse *et al.*, 2010b), the current results might also possibly reflect the effects of AD on switching performance. While this possibility could not be completely ruled out in the current findings, we wanted to clarify that we used this amount of PM trials in order to increase the number of observations and in turn, the potential variance in the PM performance across participants, which could make the task more sensitive to detect any differences among the three CDR groups. Future studies should include fewer PM trials (e.g., one-eighth of trials) in the task to further investigate the effect of AD on PM performance.

To conclude, in the current study we have shown that measures from a nonverbal PM task can discriminate the individual who are in the earliest detectable stage of AD (CDR 0.5) from healthy older adults (CDR 0) in Hong Kong Chinese population. These measures showed the discriminative power for CDR 0 versus 0.5 even after taking into account the performance in various psychometric tasks. Previous studies showed that cognitive function based on both MCI and CDR criteria could identify participants who are potentially at-risk for further decline, although they did exhibit some differences in the detection profiles (e.g., Lam *et al.*, 2010). Hence, future studies should test whether the current results could be generalized when a different criterion is used to classify the participants (e.g., amnestic MCI, see Petersen *et al.*, 2001, for a review).

Given that PM is highly relevant to older adults' maintaining independence in the early stage of AD in everyday life, future studies should evaluate whether the current PM measures would predict older adults' everyday functioning, as indicated by, e.g., the Chinese version of Comprehensive Assessment of Prospective Memory Questionnaire (e.g., Chan *et al.*, 2010). Such a measure should provide information about older adults' abilities to engage in basic daily care successfully such as door locking and to achieve certain activities such as bill paying. Even though our task may not be as ecologically valid as the questionnaire data (e.g., Chan *et al.*, 2010), it is more objective and relies less on participants' self-reported data, which should be interpreted with caution because clinical populations like AD might not be able to report their problems accurately (e.g., Crawford *et al.*, 2006). Based on the research assistants' reports, older adults were generally positive about this short and easy to administer PM task, even though it was conducted in a computer platform that not all of our older-adult participants were familiar. Nevertheless, the possibility of using this computer task in clinic and community settings (e.g., multiple testing of the patients during consecutive visits) should be further investigated in the future studies. Overall, the current study serves as the first step in the development and validation of the utility of PM measures for detecting the earliest stage of AD in Hong Kong Chinese population.

Conflict of interests

None.

Description of authors' roles

C. S. Tse designed the study, conducted the statistical analyses, and drafted the paper. J. Chang and K. T. Hau contributed to the design of study and statistical analyses. G. T. Y. Leung and A. W. T. Fung undertook the interviews. D. A. Balota and L. C. W. Lam helped the formulation of the study, designed and supervised the data collection, and commented on the drafts of paper.

Acknowledgments

The project is supported by the Health and Health Services Research Fund (#08090041), Hong Kong Special Administrative Region Government, awarded to the first author. We thank Janet Kwan and Kaver Hui for their assistance in data collection and all older adults for their participation.

Supplementary material

To view supplementary material for this article, please visit http://dx.doi.org/10.1017/S1041610214002038

References

Albert, M., Moss, M. B., Blacker, D., Tanzi, R. and McArdle, J. J. (2007). Longitudinal change in cognitive

performance among individuals with mild cognitive impairment. *Neuropsychology*, 21, 158–169.

Aschenbrenner, A. J. *et al.* (in press). Alzheimer disease biomarkers, attentional control and semantic memory retrieval: synergistic and mediational effects of biomarkers on a sensitive cognitive measure in non-demented older adults. *Neuropsychology*.

Blanco-Campal, A., Coen, R. F., Lawlor, B. A., Walsh, J. B. and Burke, T. E. (2009). Detection of prospective memory deficits in mild cognitive impairment of suspected Alzheimer's disease etiology using a novel event-based prospective memory task. *Journal of International Neuropsychological Society*, 15, 154–159.

Burgess, P. W., Quayle, A. and Frith, C. D. (2001). Brain regions involved in prospective memory as determined by positron emission tomography. *Neuropsychologia*, 39, 545–555.

Camp, C. J., Foss, J. W., Stevens, A. B. and O'Hanlon, A. M. (1996). Improving prospective memory task performance in persons with Alzheimer's disease. In M. Brandimonte, G. O. Einstein and M. A. McDaniel (eds.), *Prospective Memory: Theory and Applications* (pp. 351–367). New Jersey: Lawrence Erlbaum Associates.

Chan, R. C., Qing, Y., Wu, Q. and Shum, D. (2010). Prospective memory in healthy Chinese people: the latent structure of the comprehensive assessment of prospective memory questionnaire. *Neuropsychological Rehabilitation*, 20, 459–470.

Chang, J. F. *et al.* (2014). Bias in discriminating very mild dementia for older adults with different levels of education in Hong Kong. *International Psychogeriatrics*, 26, 995–1010.

Costa, A. *et al.* (2010). Prospective memory functioning in mild cognitive impairment. *Neuropsychology*, 24, 327–335.

Crawford, J. R., Henry, J. D., Ward, A. L. and Blake, J. (2006). The prospective and retrospective memory questionnaire (PRMQ): latent structure, normative data and discrepancy analysis for proxy-ratings. *British Journal of Clinical Psychology*, 45, 83–104.

Duchek, J. M., Balota, D. A. and Cortese, M. (2006). Prospective memory and apolipoprotein E in healthy aging and early stage Alzheimer's disease. *Neuropsychology*, 20, 633–644.

Duchek, J. M., Balota, D. A., Tse, C.-S., Holtzman, D. M., Fagan, A. M. and Goate, A. M. (2009). The utility of intraindividual variability in selective attention tasks as an early marker for Alzheimer's disease. *Neuropsychology*, 23, 746–758.

Gao, J. L., Cheung, R. T. F., Chan, Y. S., Chu, L. W. and Lee, T. M. C. (2013). Increased prospective memory interference in normal and pathological aging: different roles of motor and verbal processing speed. *Aging, Neuropsychology, and Cognition*, 20, 80–100.

Hong Kong Census and Statistics Department. (2011). *Population Census 2011*. Retrieved from http://www.census2011.gov.hk/en/main-table.html.

Huppert, F. A., Johnson, T. and Nickson, J. (2000). High prevalence of prospective memory impairment in the elderly and in early-stage dementia: findings from a population-based study. *Applied Cognitive Psychology*, 14, 63–81.

Hutchison, K. A., Balota, D. A. and Duchek, J. M. (2010). The utility of stroop task switching as a marker for early stage Alzheimer's disease. *Psychology and Aging*, 25, 545–559.

Jalbert, J. J., Daeillo, L. A. and Lapane, K. L. (2008). Dementia of the Alzheimer type. *Epidemiologic Reviews*, 30, 15–34.

Karantzoulis, S., Troyer, A. K. and Rich, J. B. (2009). Prospective memory in amnestic mild cognitive impairment. *Journal of International Neuropsychological Society*, 15, 407–415.

Lam, L. C. W., Lui, V. W. C., Tam, C. W. C. and Chiu, H. F. (2005). Subjective memory complaints in Chinese subjects with mild cognitive impairment and early Alzheimer's disease. *International Journal of Geriatric Psychiatry*, 20, 876–882.

Lam, L. C. W. *et al.* (2008). Use of clinical dementia rating in detecting early cognitive deficits in a community-based sample of Chinese older persons in Hong Kong. *Alzheimer Disease and Associated Disorder*, 22, 153–157.

Lam, L. C. W. *et al.* (2010).Combined clinical and cognitive criteria to identify mild cognitive impairment in a southern chinese community. *Alzheimer Disease and Associated Disorders*, 24, 343–347.

Martins, S. P. and Damasceno, B. P. (2008). Prospective and retrospective memory in mild Alzheimer's disease. *Arquivos de Neuro Psiquiatria*, 66, 318–322.

Martins, S. P. and Damasceno, B. P. (2012). Accuracy of prospective memory tests in mild Alzheimer's disease. *Arquivos Neuropsiquiatria*, 70, 17–21.

McDaniel, M. A. and Einstein, G. O. (2007). *Prospective Memory: An Overview and Synthesis of an Emerging Field*, Thousand Oaks, CA: Sage Publications

McDaniel, M. A., Shelton, J. T., Breneiser, J. E., Moynan, S. and Balota, D. A. (2011). Focal and nonfocal prospective memory performance in very mild dementia: a signature decline. *Neuropsychology*, 25, 387–396.

Morris, J. C. *et al.* (1997). Clinical dementia rating (CDR) training and reliability protocol: the Alzheimer's disease cooperative study unit experience. *Neurology*, 48, 1508–1510.

Petersen, R. C., Stevens, J. C., Ganguli, M., Tangalos, E. G., Cummings, J. L. and DeKosky, S. T. (2001). Practice parameter: early detection of dementia: mild cognitive impairment (an evidence-based review). *Neurology*, 56, 1133–1142.

Storandt, M. (2008). Cognitive deficits in the early stages of Alzheimer's disease. *Current Directions in Psychological Science*, 17, 198–202.

Storandt, M., Grant, E. A., Miller, J. P. and Morris, J. C. (2006). Longitudinal course and neuropathologic outcomes in original versus revised MCI and in pre-MCI. *Neurology*, 67, 467–473.

Thompson, C., Henry, J. D., Rendell, P. G., Withall, A. and Brodaty, H. (2010). Prospective memory function in mild cognitive impairment and early dementia. *Journal of International Neuropsychological Society*, 16, 318–325.

Troyer, A. K. and Murphy, K. J. (2007). Memory for intentions in amnestic mild cognitive impairment: time- and event-based prospective memory. *Journal of International Neuropsychological Society*, 13, 365–369.

Tse, C.-S., Balota, D. A., Moynan, S. C., Duchek, J. M. and Jacoby, L. L. (2010a). The utility of placing

recollection in opposition to familiarity in early discrimination of healthy aging and very mild dementia of the Alzheimer's type. *Neuropsychology*, 24, 49–67.

Tse, C.-S., Balota, D. A., Yap, M. J., Duchek, J. M. and McCabe, D. P. (2010b). Effects of healthy aging and early-stage dementia of the Alzheimer's type on components of response time distributions in three attention tasks. *Neuropsychology*, 24, 300–315.

Tse, C.-S. *et al.* (2013). Effects of education on very mild dementia among Chinese people in Hong Kong: potential mediators in the cantonese mini-mental state examination (C-MMSE) tasks. *Aging and Mental Health*, 17, 310–318.

U.S. Department of Health and Human Services. (2011). *A Profile of Older Americans: 2011,* Retrieved from http://www.aoa.gov/aoaroot/aging_statistics/Profile/2011/docs/2011profile.pdf

Zhou, T., Broster, L. S., Jiang, Y., Bao, F., Wang, H. and Li, J. (2012). Deficits in retrospective and prospective components underlying prospective memory tasks in amnestic mild cognitive impairment. *Behavioral and Brain Functions*, 8, 39.