Research Paper: Preliminary Reliability and Validity of the Iranian Computerized Version of Memory Tasks of the Betula Study

Javad Hatami[1,2] , Babak Hemmatian Borujeni[3], Ehsan Abdekhodaie[1*] , Reza Kormi-Nouri[4], Zahra Mayeli[1], Atieh Sadat Mottaghi Ghamsari[1], Maryam Taghavi[1]

1. Department of Psychology, Faculty of Psychology and Education, University of Tehran, Tehran, Iran.
2. Department of Cognitive Psychology, Institute for Cognitive Science Studies, Tehran, Iran.
3. Department of Cognitive, Linguistic, Psychological Sciences, Brown University, RI,US.
4. Center for Health & Medical Psychology, Örebro University, Sweden.

ABSTRACT

Introduction: This study aimed to evaluate the reliability and validity of an Iranian computerized memory battery modeled after the Betula study.

Methods: This study aimed to evaluate the reliability and validity of an Iranian computerized memory battery modeled after the Betula study (Nilsson et al., 1997). The researchers developed this battery as an assessment tool in the Sepidar prospective cohort study. One hundred and ninety-nine participants aged 19-83 years were tested extensively on different aspects of memory. Exploratory factor analysis of the data demonstrated factors similar to those reported by the Betula study.

Results: The authors succeeded to converge the cross-sectional findings of the study and the data from longitudinal studies of memory aging by correcting possible cohort effects. Investigating age differences in episodic and semantic memory factor scores corrected by education and socioeconomic status revealed no significant difference between younger and older adults before ages 53 to 60, though linear age-related declines existed thereafter.

Conclusion: The results support the reliability and construct validity of this computerized battery for memory assessment in Iranian adults.

* Corresponding Author:
Ehsan Abdekhodaie, MSc.
Address: Department of Psychology, Faculty of Psychology and Education, University of Tehran, Tehran, Iran.
Tel: +98 (915) 1191332
E-mail: abdekhodaie@ut.ac.ir

Highlights

● Researchers tried to discover developmental trends of memory processes during adulthood.

● There is still much controversy about the shape, pace, and meaning of age-related cognitive changes in adulthood.

● It is essential to assess memory functions accurately.

● Battery of memory tasks was used for memory assessment.

Plain Language Summary

Researchers focussed during the last few decades on discovering developmental trends of memory processes during adulthood. Some key questions were developed: When the age-related cognitive decline in healthy aging begins? How age trends are different for various memory tasks? Are changes in adults' performance in psychometric tests truly reflects cognitive decline? To answer, studies on memory development through adulthood remained to be important and challenging. It is essential to assess memory functions accurately. We aimed at investigating the reliability and validity of the scores of an Iranian computerized battery of memory assessment tasks. A comprehensive battery of memory tasks was used for memory assessment. The battery was based on the work of some authors composed of tasks exploring a wide variety of processes and hypothetical memory systems.

1. Introduction

Researchers have strived during the last few decades to discover developmental trends of memory processes during adulthood and senility in normal and clinical populations. Although current studies offer valuable insights into the developmental trends of memory and factors that moderate these trends, there is still much controversy about shape, pace, and meaning of age-related cognitive changes in adulthood, particularly concerning answering some questions. For example, the following questions can be considered: when the age-related cognitive decline in healthy aging (if there is any) begins (Nilsson, Sternäng, Rönnlund, & Nyberg, 2009; Salthouse, 2009; Singh-Manoux et al., 2012), how age trends are different for various memory tasks (e.g. episodic and semantic memory, repetition priming, etc.) (Fleischman, 2007; Rönnlund, Nyberg, Bäckman, & Nilsson, 2005; Salthouse, 2005; Schaie, 2012), or whether changes in adults' performance in psychometric tests truly reflects cognitive decline (Ramscar, Hendrix, Shaoul, Milin, & Baayen, 2014; Ratcliff & McKoon, 2015). Thus, studies on memory development through adulthood remains to be important and challenging. It is, in turn, essential to assess memory functions accurately; particularly considering putative cultural differences in memory performance (Hedden et al., 2002; Gutchess et al., 2006) in the context of non-western societies (Park & Gutchess, 2006).

In the current study, we aimed at investigating the reliability and validity of the scores of an Iranian computerized battery of memory assessment tasks adapted from the Betula study (Nilsson et al., 1997) and reporting some of the preliminary findings from the data obtained in an Iranian sample, and comparing the results to those of other cross-sectional and longitudinal studies on the effect of aging on memory performance. Also, some suggestions were provided on how to adapt the current computerized memory assessment tasks for use in other languages/cultures in order to accelerate the conducting relevant cross-cultural studies. It should be noted that the battery was developed for memory assessment in a larger prospective cohort study called Sepidar study.

2. Materials and Methods

2.1. Participants

Two hundred and five participants were recruited via advertisements in social networks and flyers distributed in the faculty's vicinity, by which it was indicated that those aged 25 to 70 years living in Tehran could participate in a study at the University of Tehran involving free assessment of memory and they would receive reports on their performance and a brief consultation. The participants first completed a series of demographic questionnaires and then, were tested using a comprehensive battery of memory tasks at the University of Tehran's Center for Cognitive Rehabilitation. Six participants reported

536

Hatami, J., et al. (2020). Reliability and Validity of the Iranian Computerized Version of Memory Tasks. *BCN, 11*(4), 535-548.

computer failures and the final sample consisted of 199 participants (152 women, 42 men, 5 missing genders) who were tested between September 2014 and March 2015. The Ethical Committee of Iran University of Medical Sciences and Health Services approved the study.

2.2. Measures

2.2.1. Memory assessment

A comprehensive battery of memory tasks was used for memory assessment. The battery was based on the seminal work of Nilsson et al. (1997), which was composed of tasks exploring a wide variety of processes and hypothetical memory systems (see also Nilsson et al., 2002). The battery included tasks associated with short-term and long-term memory processing, recall of general knowledge and word fluency (semantic memory), as well as recall of studied events (episodic memory). Specific tasks related to the conscious and non-conscious recollection of information, focused versus divided attention encoding, visual, auditory, motoric, and spatial aspects of memory, retrospective versus prospective memory, and intentional versus non-intentional learning were considered. Every care was taken to tailor the items for Iranian populations without compromising the theoretical aims of the tasks.

This included some changes to the content of items (e.g. for newly-acquired facts and verbal fluency measures), as well as changes in presentation procedures for parts where Iranian participants had a particularly hard time to grasp the tasks. In addition, considerable effort was made to computerize the aforementioned tasks through the development of a computer package in PsychoPy (Peirce, 2008) so that all of the experiments could be done on personal computers and the process of data collection and analysis could be improved in terms of accuracy and ease. Table 1 presents the memory tasks according to the order, in which they were carried out.

Undergraduate and graduate psychology students were carefully trained to conduct memory testing. A detailed manual was provided to the testers. After studying the manual, the participants carried out several tests before conducting memory tests in the actual setting and some of them were under the supervision of one of the authors to ensure they precisely followed the testing protocol.

Since the tasks were very similar to those described in the original study by Nilsson et al. (1997), we will only mention the differences or specific methods pertaining to

the current study and refer the readers to the Nilsson et al. research for further information.

At the beginning of the testing session, the participants were instructed to remind the examiner at the end of the session to sign a piece of paper. This instruction was used to engage the participant with a prospective memory task.

For recognition of faces, 32 color pictures of faces were obtained from Iranian children aged 8 to 12 years. Half of the pictures belonged to girls and the other half to boys. Participants were presented with 16 pictures chosen randomly from the available pictures by the computer program. The photographed children lived in a city far from Tehran to reduce the chance of coincidental familiarity. The computer program chose 16 made-up names randomly from 32 given names and 32 family names and showed them beneath the pictures on the screen. Thus, the pictures, names, and their order were decided randomly for all the participants.

For the tasks referred to as free and cued recall of sentences with or without enactment, two sets of imperative sentences were designed for enactment and non-enactment tasks. The sentences were constructed with regard to daily activities and cultural practices of the Iranian community. The sentences were chosen from a pool of sentences, which have been studied in the Iranian population (Jafarian-Namini, Kormi-Nouri, & Usefi-Louyeh, 2001). The nouns of the sentences in each list were members of four different categories with four instances each. Although no category norms were available for the Iranian adults in the literature, according to a study by Kormi-Nouri, et al. (2012), the mean number of correct responses for monolingual Iranian children in these enactment and non-enactment lists was 9.28 and 8, respectively; indicating the effect of enactment in line with previous studies (Engelkamp, 1998; Hald, van den Hurk, & Bekkering, 2015). The sentences were presented in random order.

Regarding the word stem completion task (Greene, 1986), 32 two-letter stems were introduced to the participants and they were asked to tell the examiner the first family name beginning with those letters that came to the mind. Sixteen stems of names presented in the face-name recognition task, and the remaining 16 were distractors. The mean number of common family names with the 32 word stems was 9.09 (ranged: 7-19).

In the word fluency task, the participants were asked to say as many words as they could according to five cri-

Hatami, J., et al. (2020). Reliability and Validity of the Iranian Computerized Version of Memory Tasks. *BCN, 11*(4), 535-548.

537

Table 1. Memory tasks used in the study modeled after Nilsson et al. (1997)

Task	Outcome variables
1. Instruction about a later test of prospective memory.	-
2. Study episode for a later face/name recognition test (16 items).	-
3. Study of the list of short sentences with or without enactment (16 items).	-
4. Immediate free recall of the list of sentences presented in 3.	FR non-enacted/enacted sentences
5. Study of the list of short sentences with or without enactment (16 items).	-
6. Immediate free recall of the list of sentences presented in task 5.	FR non-enacted /enacted sentences
7. Cued recall of nouns in sentences presented in tasks 3 and 5 (32 items, category cue).	CR non-enacted/enacted sentences noun with category cue
8. Stem completion of family names presented in task 2.	Priming
9. Word fluency: Initial letter Alef without Madda.	WF A
10. Word fluency: Initial letter Alef with Madda.	WF Aa
11. Word fluency: Initial letter Teh for five-letter words.	WF 5 letters T
12. Word fluency: initial letter Mim for names of professions.	WF professions M
13. Word fluency: initial letter Gaf for four-letter names of animals.	WF animals 4 letters G
14. Free-choice recognition of faces presented in task 2 (12 of 32 items with 12 distractors, yes/no).	RN faces
15. Forced-choice recognition of names presented in task 2 (16 items, four alternative forced choices).	RN family name RN given name
16. Free-choice recognition of nouns in sentences presented in tasks 3 and 5 (16 of 32 items presented in 3 and 5, 16 distractors, yes/no).	RN non-enacted/enacted sentences noun
17. Cued recall of nouns in sentences presented in tasks 3 and 5 (32 items) and source recall of the sentences (32 items, enactment/no enactment).	CR non-enacted/enacted noun with verb cue RE non-enacted/enacted sentence source
Break	-
18. Study of the word list along with concurrent task (12 items).	-
19. Immediate free recall of the word list presented in task 18 with the concurrent task (12 items).	FR words divided attention at study test
20. Study of the word list with the concurrent task (12 items).	-
21. Immediate free recall of the word list from task 20 without concurrent task (12 items).	FR words divided attention at study
22. Study of the word list without the concurrent task (12 items).	-
23. Immediate free recall of the word list from 22 with the concurrent task (12 items).	FR words divided attention at the test
24. Study of the word list without the concurrent task (12 items).	-
25. Immediate free recall of the word list from task 24 without concurrent task (12 items).	FR words no divided attention
26. Study of the facts presented on yellow/red screen or by male/female voice.	-
27. Recall of the facts presented in task 26, source recall of the facts, and recall of the general knowledge and general knowledge source.	RE newly acquired facts RE newly acquired facts source RE general knowledge RE general knowledge source
28. Reminding activities in the whole session.	FR activities
29. Prospective memory test with or without cue.	Prospective memory

NEUR★SCIENCE

Hatami, J., et al. (2020). Reliability and Validity of the Iranian Computerized Version of Memory Tasks. *BCN, 11*(4), 535-548.

teria, each within one minute. The first and second criteria, initial letter Alef with/without Madda, were based on the fact that in the Persian lexicon, more words begin with Alef than other letters (together with Hebrew Aleph, Greek Alpha, and Latin A) and it is descended from Phoenician āleph. A modern Persian corpus was used to compute the number of words with different criteria (Ale Ahmad et al., 2009; Darrudi, Hejazi, & Oroumchian, 2004). In addition, research on Persian monolinguals indicated that this letter has the highest frequency in the letter fluency task (Kormi-Nouri et al., 2012; Kormi-Nouri et al., 2008). The third and fourth criteria, based on the combination of a letter fluency task and either a category fluency task or an n-letter words fluency task, were developed so that the number of actual words with the criteria were large enough to allow for enough variance (AleAhmad et al. 2009; Darrudi et al., 2004). The last criterion was developed so that the number of fulfilling words was very few (only three words), similar to the fourth criterion in the word fluency task in the original study.

After this task, the participants were given a free-choice recognition of faces task, which included 12 target faces randomly chosen from the 16 presented faces in the study task and 12 distractor faces, and a forced-choice recognition of names of all the 16 faces that had been presented.

The next task was the recognition of nouns in imperative sentences. Accordingly, 16 nouns were chosen randomly from the study lists (8 in sentences with enactment and 8 in sentences without enactment) and 16 distractor nouns appeared in a randomized order. The participants were asked to tell whether the noun was presented in the imperative sentence tasks or not. After recognition of the nouns, the participants did a cued recall task, in which they were asked to recall the nouns in the imperative sentences with verb phrases served as the cues. They were also asked to tell for each noun whether they enacted or not enacted the corresponding sentence (external source recall).

After a short break, the participants performed the free recall of lists of words under conditions of focused and divided attention tasks. Eight 12-item word lists were constructed. Twelve categories of the words were chosen and eight common words from each category were included, with the difference between mean word frequencies of the lists being negligible (M=52.99 per million words, 95% CIs [52.97, 53.00]; the frequencies obtained from Darrudi et al. data, 2004). For each participant, four lists were randomly selected from the word lists and presented in a predetermined order of conditions (Table 1). The concurrent task was to sort a deck of cards into two piles based on the symbols on them. In the first condition, this task was done during both study and test. In the second condition, the card sorting task was done at study only. In the third condition, the card sorting was done at the test but not at the study. In the last condition, there was no simultaneous task during either study or test.

The next task in the sequence involved the study and recall of facts about famous and unknown Iranian people, along with source recall of these facts. Twenty different statements were presented to each participant (e.g. Jamshid Mashayekhi is a pioneer actor who is 83 years old). Each statement was presented in one of four different ways: visually on a yellow or red rectangle on the computer screen, or auditory by playing a female or male voice. At the test, the participants were presented with 20 questions corresponding to the studied statements (e.g. how old is Jamshid Mashayekhi?), as well as another 10 questions related to generally known information (e.g. where did Hafiz live?), and also another 10 questions, which were made up and therefore were impossible to answer (distractors). The experimenter also asked the participants to tell whether the information related to the question was presented in the experiment. In case the participants said they learned the information within the experiment, they were asked to indicate whether it had been presented on a yellow or red card, or via a female or male voice. If they reported the information was not learned within the experiment, they were asked to indicate whether they had learned this information in school, in papers or books, on radio or TV, or through some other means (source recall).

In the end, the participants were asked to remember all the activities done in the session. The examiner then told the participant that the testing session was completed and thanked them for their participation; at this moment, the participant was supposed to remember the task of signing a piece of paper (prospective memory task). The computer program was designed to wait for 15 s and then show a window. If the participant failed to remember the task within this period, they were cued by the examiner; firstly by asking them whether there was something else to be done. If the participant did not remember the task, they were cued by asking them whether they remember they should remind the examiner something. Finally, if the participant failed to remember the task by this cue, they were asked whether they remember the specific task of signing a piece of paper. The participant was rated from 4 to 0 according to when they remembered the prospective task in the procedure mentioned above.

Hatami, J., et al. (2020). Reliability and Validity of the Iranian Computerized Version of Memory Tasks. *BCN, 11*(4), 535-548.

539

2.3. Demographics

The demographic data, such as date of birth, gender, job, family expenses (as a measure of Socio-Economic Status (SES)), years of education, marriage status, age of (first) marriage, postal address, and residence status were obtained along with other data using a questionnaire designed by the authors. The demographic data were obtained before the memory assessment began.

3. Results

3.1. Descriptive data

Table 2 presents the number and percentage of participants for different age groups in this study. The Mean±SD of education of the participants were 14.9±3.0 years). The participants were categorized into five SES categories (1-5 values) based on family expenses they reported. The first category belonged to expenses less than approximately 140 $/month and the last category pertained to expenses more than about 1,700 $/month. The Mean±SD of category value was 2.69±0.75. Participants aged 19-83 years old (Mean±SD of age: 46.5±14.4).

Men and women were not different in terms of their age or SES (P>0.45) but men had 1.4 years more schooling than did women (Mann-Whitney U=2171, P=0.01).

3.2. Exploratory factor analysis

Exploratory factor analysis was performed on the Spearman correlation matrix of the outcome variables using Principal Component Analysis (PCA). The factors were extracted by the Kaiser criterion (i.e. Eigenvalue>1) and the Varimax rotation was used.
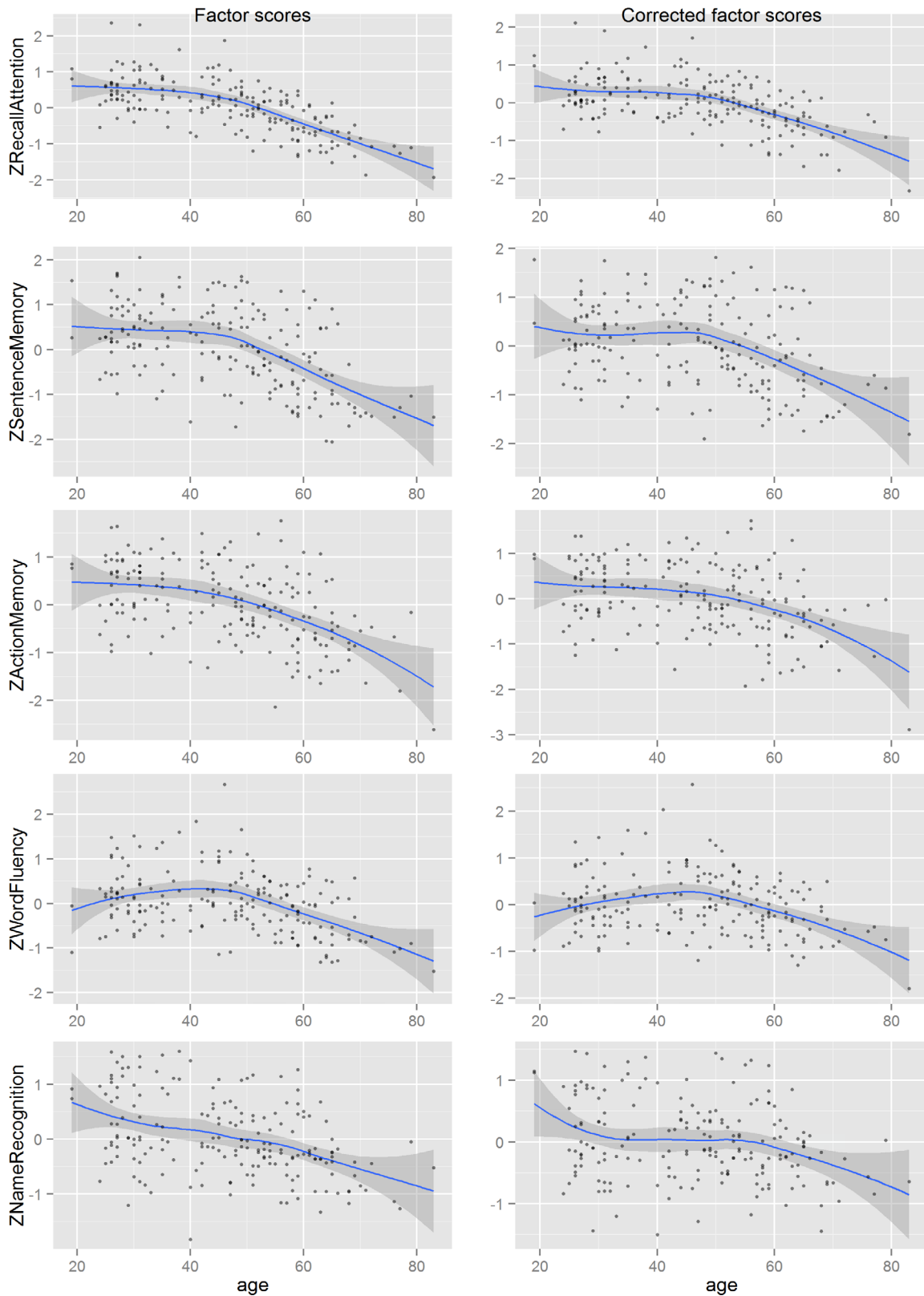
Firstly, the correlation matrix was computed for all variables other than "RE general knowledge source". Because it was not feasible to identify the truth of the sources, by which participants claimed to have learned general knowledge items, we did not analyze this variable. The factor analysis led to an eight-factor structure. Then, we dropped the variables "Prospective memory" and "RE general knowledge" due to their low communalities (<0.45). Lastly, we also removed the variable "WF animals 4 letters G" from the correlation matrix because it caused the appearance of a single factor independent of other word fluency measures. The aforementioned procedure resulted in a 25-variable correlation matrix, on which we carried out the final factor analysis. A six-factor solution is deemed appropriate based on the Kaiser criterion. The derived factors were similar to those reported by Nilsson et al. (1997). The factors explained 58.3% of the variance in the correlation matrix. The labels of the factors and rotated loadings of different tasks belonging to each factor are presented in Table 3.

To compute the composite factor scores of the data, the raw scores obtained in various tasks were transformed into z scores and the factor scores were computed by averaging the z scores of tasks related to each factor. However, the tasks "RN non-enacted sentences noun", "RE non-enacted sentence source", "RN enacted sentences noun", "RE enacted sentence source" were not included in the calculation of the composite scores because

Table 2. The age of the participants

Age Group (y)	No. (%)
Born after 1989 (<25)	7 (3.5)
Born between 1984 and 1988 (25-29)	25 (12.6)
Born between 1979 and 1983 (30-34)	26 (13.1)
Born between 1974 and 1978 (35-39)	6 (3.0)
Born between 1969 and 1973 (40-44)	21 (10.6)
Born between 1964 and 1968 (45-49)	18 (9.0)
Born between 1959 and 1963 (50-54)	26 (13.1)
Born between 1954 and 1958(55-59)	21 (10.6)
Born between 1949 and 1953 (60-64)	24 (12.1)
Born between 1944 and 1948 (65-69)	7 (3.5)
Born in or before 1943 (>69)	7 (3.5)
Missing age	11 (5.5)

y: year.

NEURSCIENCE

Hatami, J., et al. (2020). Reliability and Validity of the Iranian Computerized Version of Memory Tasks. *BCN, 11*(4), 535-548.

NEUR⚙SCIENCE

Figure 1. Scatter plots of the corrected vs. non-corrected memory factor scores against age with LOESS fitting line

The shadows indicate 95% confidence intervals.

Hatami, J., et al. (2020). Reliability and Validity of the Iranian Computerized Version of Memory Tasks. *BCN, 11*(4), 535-548.

541

Table 3. List of the tasks included in the Principal Component Analysis with Varimax rotation, their rotated loadings, and the chosen labels for the different components

Characteristics	Task	1	2	3	4	5	6	% of the total explained variance
1. Recall-Attention	FR words with divided attention at the test	0.68						
	FR words with divided attention at the study	0.64	0.34					
	RE newly acquired facts source	0.62			0.34			
	FR words with divided attention at the test	0.59	0.34	0.33				16.8
	RE non-enacted sentence source*	0.57						
	FR words without divided attention	0.56	0.44					
	FR activities	0.55		0.31				
2. Sentence Memory	RN non-enacted sentence (noun)*		0.75					
	FR non-enacted sentence (verb and noun)	0.36	0.74					11.7
	CR non-enacted sentence (category cue)	0.40	0.70					
	CR non-enacted sentence (verb cue)	0.34	0.70	0.30				
3. Action Memory	RE enacted sentence source*			0.71				
	CR enacted sentence (verb cue)	0.32		0.66				
	RN enacted sentence (noun)*			0.59				9.3
	FR enacted sentence (verb and noun)**	0.62		0.45				
	CR enacted sentence (category cue)**	0.62		0.49				
4. Word Fluency	WF 5 letters T				0.72			
	WF A				0.72			9.0
	WF Aa				0.57			
	WF professions M			0.36	0.42		-0.30	
5. Name Recognition	RN family name					0.73		
	RN faces					0.66		6.6
	RE newly acquired facts***	0.46				0.46		
	RN first name					0.35		
6. Priming	Priming						0.83	4.9

NEUR☼SCIENCE

* Not used to compute composite scores because of non-normal distribution (very high skewness); ** Assigned to the second most loaded factor (Action Memory) to be more interpretable; *** Not used to compute composite scores due to ambiguous factors. Factor loadings <0.3 are not shown in the table.

of their large negative skewness (ranged from -0.91 to -2.57). Also, the task "RE newly acquired facts" was not included because it was almost equally loaded on two factors.

3.3. Reliability

We computed the reliability of factor scores (except Priming that consisted of only one item) of the memory assessment tasks in terms of internal consistency (Table 4). The results show that Recall-Attention, Sentence Memory,

and Action Memory factor scores demonstrated satisfactory reliability (Cronbach's α ranged from .76 to 0.91) but internal consistencies were lower for factor scores of Word Fluency and Name Recognition (α<0.7).

3.4. Age trends, gender differences, and education

To explore construct validity (Cronbach & Meehl, 1955; Nunnally & Bernstein, 1994) of the memory assessment battery, we examined the age trends of memory function. The factor scores, Recall-Attention, Sentence Memory,

Hatami, J., et al. (2020). Reliability and Validity of the Iranian Computerized Version of Memory Tasks. *BCN, 11*(4), 535-548.

Table 4. Cronbach's alphas for memory factor scores

Factor Score	Cronbach's alpha
1. Recall-Attention	0.84
2. Sentence Memory	0.91
3. Action Memory	0.76
4. Word Fluency	0.62
5. Name Recognition	0.51

NEUR⊗SCIENCE

Table 5. Age at onset of a decline in memory performance (years) for factor scores and corrected factor scores (by regressing on socio-economic status and education)

Factor Score	Onset for Memory Factor Scores	Onset for Corrected Factor Scores
1. Recall-Attention	46	53
2. Sentence Memory	52	58
3. Action Memory	50	56
4. Word Fluency	55	58
5. Name Recognition	47	60
6. Priming	N/A	N/A

Age at onset was calculated based on "loess" fit 95% confidence intervals.

NEUR⊗SCIENCE

Action Memory, and Name Recognition were used to investigate episodic memory, Priming factor scores (only one item) were used to study non-declarative memory, and Word Fluency scores were analyzed as measures of semantic memory since the general knowledge measure had a low commonality and a very skewed distribution.

To examine the age trends in different memory factor scores, we used loess regression fit, which is a nonparametric method, in which the regression is carried out by fitting parametric functions locally (Cleveland & Grosse, 1991; Cleveland, Grosse, & Shyu, 1992), with 95% confidence interval (except for Priming scores because of non-normal distribution; Figure 1). The episodic and semantic memory factor scores followed nonlinear trends, each with a steady (no significant change) stage followed by a typically linear decline after a particular age (decline onset). We detected the age when the upper band of the fitted line became less than the maximum of the lower band as the onset of a decline in performance for each variable. To control for cohort effects, due to differences in SES or education, we corrected the factor scores by regressing them on education and SES and also using residuals. This resulted in an average of seven years passed from the onset of cognitive decline but little change in the general shapes of the curves.

Table 5 represents the onset of decline pertaining to different memory scores. Correction of the education and SES also resulted in closer onset ages (difference in SD=1.0 year) suggesting that the dispersion in onset age data before correction might be due to cohort effects. The average age of onset of decline in corrected episodic and semantic memory measures was 57 years (ranged: 53-60, SD=2.6). We calculated the rate of decline in the average of standardized episodic memory and semantic memory scores after 60 years of age. There was a moderate rate of decline; however, the rates of decline were not significantly different between episodic and semantic memory (rs= -0.52, P<0.001 and rs=-0.41, p=-0.011, respectively, 95% CIs [-0.72, -0.25], and [-0.64, -0.10]). Also, episodic and semantic memory performance was equally correlated with education (rs=-0.47, P<0.001 and rs=-0.31, P<0.001, respectively, 95% CIs [0.35, -0.57], and [0.18, .43]).

Priming scores were not different between the young (younger than 35 years, Mdyoung=2) and middle-aged adults (36 to 55 years old, Mdmiddle-aged=2, Mann-Whitney U=1181, P=0.38), but scores of the older adults (older than 55 years, Mdolder=1) were worse than those of the middle-aged adults (Mann-Whitney U=1181, P<0.001). Priming scores were not related to education (rs= -0.10, P=0.18).

Hatami, J., et al. (2020). Reliability and Validity of the Iranian Computerized Version of Memory Tasks. *BCN, 11*(4), 535-548.

543

There were not any gender differences in factor scores even after adjusting for education (Welch's t= -1.8-0.2, df'=56.2-64.0, P=0.08-0.86). However, considering the small effect size of gender differences observed in previous studies, e.g. Habib, Nyberg, & Nilsson (2007) and the small number of men participating in the study, our sample did not have enough power to show gender differences that might have existed in memory functions.

4. Discussion

In this study, we examined the factor structure and the preliminary evidence regarding the reliability and validity of the Iranian computerized version of the memory battery developed by Nilsson et al. (1997), (the Betula study) in a sample of Iranian adults aged 19-83 years. We reached a factor structure similar to that proposed by Nilsson et al. (1997) and explored the age trends of memory measures scores using cross-sectional data.

The results of exploratory factor analysis were similar to those reported by Nilsson et al. (1997), yielding six main factors for the data: Recall-Attention, Sentence Memory, Action Memory, Word Fluency, Name Recognition, and Priming. Even the portion of variance explained by these factors was comparable between the current study and that of Nilsson et al. (1997). These results seem to show that at least as far as the latent structure of memory functions are concerned, there is no powerful cultural effect causing a difference between the memory factor structure in the Swedish and Iranian samples. Because the number of items pertaining to each factor was fairly small and we did not measure test-retest reliability, factor scores cannot be evaluated completely by means of only internal consistencies. However, the scores of the three factors, including Recall-Attention, Sentence Memory, and Action Memory had acceptable internal consistencies and their resultant scores were quite reliable for assessment of memory performance in an Iranian population.

Considering age differences in memory performance, this study presented intriguing findings. The age trends of episodic and semantic memory measures, though from cross-sectional data, were quite consistent with those of several longitudinal studies (Rönnlund et al., 2005; Schaie, 2012; Singh-Manoux et al., 2012; Zelinski & Burnight, 1997), when memory measures have been corrected by education and SES; that is, steady performance before 50-60 years of age with a fairly linear decline thereafter. These results support the construct validity of the current memory battery. Furthermore, this contrasts with the existing divergence in the literature between

cross-sectional studies, which typically have shown a gradual decline in episodic performance beginning from the first years of adulthood (Li et al., 2004; Nilsson et al., 1997; Park et al., 1996) as well as longitudinal studies, which hardly reported any significant decline before 60 years of age (Rönnlund et al., 2005; Schaie, 2012; Singh-Manoux et al., 2012). The divergence often is attributed to cohort effects contaminating the results of cross-sectional studies or positive effects from prior test experiences in longitudinal studies (Kuhlen, 1940; Nilsson et al., 2009; Rönnlund et al., 2005; Salthouse, 2009), resulted in a recent author's assertion that only longitudinal studies can accurately assess age-related decline in cognitive function (Singh-Manoux et al., 2012). Our findings and those of (Rönnlund et al., 2005) regarding the effectiveness of education correction in converging cross-sectional and longitudinal data suggest that the assertion above may be a precarious one and this simple correction for cohort effects may be utilized to make the ubiquitous cross-sectional data regarding aging more meaningful.

In the current study, priming scores were weakly related to age. The literature on age-related decline in priming has been inconsistent. Some researchers have found evidence for such declines (Chiarello & Hoyer, 1988; Hultsch, Masson, & Small, 1991; Karlsson, Adolfsson, Börjesson, & Nilsson, 2003), whereas others have failed to demonstrate any significant drop in performance for older adults (Light & Singh, 1987; Light, Singh, & Capps, 1986; Mitchel & Bruss, 2003). The inconsistent findings regarding priming in the aging have been attributed to task characteristics, such as explicit contamination or difference between latency and accuracy measures, or participant characteristics, like the presence of people with mild cognitive impairment in samples, the age profile of the samples, or small sample sizes (Fleischman, 2007; Mitchel & Bruss, 2003). However, given the fact that we did not put rigor in monitoring explicit contamination, it should be cautiously concluded that priming function was impaired in older adults. Younger adults might intentionally have used the data they were primed on in the priming task possibly due to their greater familiarity with computerized tests/games.

Because of the insufficient power of our sample to detect gender differences in memory scores (the small number of men), we cannot elaborate on the absence of these differences in our data. However, the participants' performance in episodic memory and word fluency was significantly correlated with the years of education they had undergone. This variable accounted for 21.7% of the variance in episodic memory and 9.4% of the variance in

word fluency. Such results can be related to the evidence put forward by Habib et al. (2007) who found education to be the most powerful predictor for healthy memory in senility. Other cross-sectional and longitudinal studies also mostly reported significant contribution of education to cognitive performance in general and memory performance in particular (Anstey & Christensen, 2000; Arenaza-Urquijo et al., 2013). This suggests that using educationally diverse samples and correcting education while analysis of the data of an aging study should become standard practices.

The limitations of the current study should be considered. Our sample consisted of adults mainly residing in the North or West Tehran. Given the diversity of cultures and ethnic groups of Iran, an assumption that the results of the current study are easily generalizable to other Iranian populations would be flawed. As older people had to come to the University for testing, our sampling method might have placed high demands on them and older adults in the sample may over-represent their performance as older adults (Hedden & Gabrieli, 2004). In terms of measurement of memory functions, although the battery in the current study included a range of different tasks to assess episodic memory, it included only word fluency measures for the assessment of semantic memory and we did not reliably measure general knowledge in this study. Also, we did not carry out test-retest reliability analyses for the measures and also did not investigate concurrent or discriminant validity of the scores of memory assessment tasks using other memory measures or by assessing how people with, e.g. mild cognitive impairment perform on the battery.

Given the dearth of cross-cultural studies of memory aging (Park & Gutchess, 2006), especially in Asia and the Middle East, and the particular characteristics of the computerized memory assessment battery used in the current study, we suggest that this battery can be easily adapted to other cultures and is applicable to accelerate conducting relevant cross-cultural studies. The used adaptation process involved fairly easy tasks: we obtained color pictures of children, made-up given names and family names, using sentences related to daily activities in Iranian culture (Jafarian-Namini et al., 2001), and use of tentative category norms (Kormi-Nouri et al., 2008; Kormi-Nouri et al., 2012), and a Farsi corpus to determine frequencies of the words (AleAhmad et al., 2009; Darrudi et al, 2004). These could effortlessly be carried out in various cultures and the new content could readily be replaced by the one in the current testing software. The results obtained from this pilot study were largely comparable with other large studies done in the Western

societies; therefore, it is hoped that this outcome could possibly be replicated in other countries without the exhaustive design of memory assessment batteries and this would contribute to the emergence of the new cognitive neuroscience of culture and aging (Park, 2002; Park & Gutchess, 2006).

To conclude, we examined the factor structure of the Iranian computerized version of the memory task battery developed by the Betula, computed reliability in terms of internal consistencies, and compared the age trends of scores of episodic memory and word fluency factors with the findings of previous longitudinal and cross-sectional research. Overall, the present study provided preliminary evidence showing the reliability and construct validity of this battery for memory assessment in the Iranian population.

Ethical Considerations

Compliance with ethical guidelines

All ethical principles were considered in this article.

Funding

Authors' contributions

All authors contributed equally in preparing all parts of the research.

Conflict of interest

The authors declared no conflict of interest.

Acknowledgments

Reference

AleAhmad, A., Amiri, H., Darrudi, E., Rahgozar, M., & Oroumchian, F. (2009). Hamshahri: A standard Persian text collection. *Knowledge-Based Systems, 22*(5), 382-7. [DOI:10.1016/j.knosys.2009.05.002]

Anstey, K., & Christensen, H. (2000). Education, activity, health, blood pressure and apolipoprotein E as predictors of cognitive change in old age: A review. *Gerontology, 46*(3), 163-77. [DOI:10.1159/000022153] [PMID]

Hatami, J., et al. (2020). Reliability and Validity of the Iranian Computerized Version of Memory Tasks. *BCN, 11*(4), 535-548.

545

Basic and Clinical
NEUROSCIENCE

Arenaza-Urquijo, E. M., Landeau, B., La Joie, R., Mevel, K., Mézenge, F., & Perrotin, A., et al. (2013). Relationships between years of education and gray matter volume, metabolism and functional connectivity in healthy elders. *NeuroImage, 83,* 450-7. [DOI:10.1016/j.neuroimage.2013.06.053] [PMID]

Chiarello, C., & Hoyer, W. J. (1988). Adult age differences in implicit and explicit memory: Time course and encoding effects. *Psychology and Aging, 3*(4), 358. [DOI:10.1037/0882-7974.3.4.358] [PMID]

Cleveland, W. S., & Grosse, E. (1991). Computational methods for local regression. *Statistics and Computing, 1*(1), 47-62. [DOI:10.1007/BF01890836]

Cleveland, W. S., Grosse, E., & Shyu, W. M. (1992). Local regression models. In J. M. Chambers & T. J. Hastie (Eds.), Statistical Models in S. Pacific Grove, California: Wadsworth & Brooks/Cole.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52*(4), 281-302. [DOI:10.1037/h0040957] [PMID]

Darrudi, E., Hejazi, M. R., & Oroumchian, F. (2004). Assessment of a modern Farsi corpus. *In Proceedings of The 2nd Workshop on Information Technology & its Disciplines (WITID), ITRC.* http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.129.2092

Engelkamp, J. (1998). Memory for actions. Hove, UK: Psychology Press. [DOI:10.1080/741942360] [PMID]

Fleischman, D. A. (2007). Repetition priming in aging and Alzheimer's disease: An integrative review and future directions. *Cortex, 43*(7), 889-97. [DOI:10.1016/S0010-9452(08)70688-9]

Greene, R. L. (1986). Word stems as cues in recall and completion tasks. *The Quarterly Journal of Experimental Psychology Section A, 38*(4), 663-673. [DOI:10.1080/14640748608401619]

Gutchess, A. H., Yoon, C., Luo, T., Feinberg, F., Hedden, T., & Jing, Q., et al . (2006). Categorical organization in free recall across culture and age. *Gerontology, 52*(5), 314-23. [DOI:10.1159/000094613] [PMID]

Habib, R., Nyberg, L., & Nilsson, L. G. (2007). Cognitive and non-cognitive factors contributing to the longitudinal identification of successful older adults in the Betula study. *Aging, Neuropsychology, and Cognition, 14*(3), 257-73. [DOI:10.1080/13825580600582412] [PMID]

Hald, L. A., van den Hurk, M., & Bekkering, H. (2015). Learning verbs more effectively through meaning congruent action animations. *Learning and Instruction, 39,* 107-22. [DOI:10.1016/j.learninstruc.2015.05.010]

Hedden, T., & Gabrieli, J. D. E. (2004). Insights into the ageing mind: A view from cognitive neuroscience. *Nature Reviews Neuroscience, 5*(2), 87-96. [DOI:10.1038/nrn1323] [PMID]

Hedden, T., Park, D. C., Nisbett, R., Ji, L.-J., Jing, Q., & Jiao, S. (2002). Cultural variation in verbal versus spatial neuropsychological function across the life span. *Neuropsychology, 16*(1), 65-73. [DOI:10.1037//0894-4105.16.1.65] [PMID]

Hultsch, D. F., Masson, M. E. J., & Small, B. J. (1991). Adult age differences in direct and indirect tests of memory. *Journal of Gerontology, 46*(1), P22-P30. [DOI:10.1093/geronj/46.1.P22] [PMID]

Jafarian-Namini, F., Kormi-Nouri, R., Usefi-Louyeh, M. (2001). The effect of both divided and concentrated attention on verbal and action memory in dyslexic and normal children (Persian)]. *Advances in Cognitive Science, 4*(1), 25-36.

Karlsson, T., Adolfsson, R., Börjesson, A. & Nilsson, L.G. (2003). Primed word-fragment completion and successive memory test performance in normal aging. *Scandinavian Journal of Psychology, 44,* 355-61. [DOI:10.1111/1467-9450.00355] [PMID]

Kormi-Nouri, R., Moradi, A. R., Moradi, S., Akbari-Zardkhaneh, S., & Zahedian, H. (2012). The effect of bilingualism on letter and category fluency tasks in primary school children: Advantage or disadvantage? *Bilingualism: Language and Cognition, 15*(02), 351-64. [DOI:10.1017/S1366728910000192]

Kormi-Nouri, R., Shojaei, R., Moniri, S., Gholami, A., Moradi, A., & Akbari-Zardkhaneh, S., et al. (2008). The effect of childhood bilingualism on episodic and semantic memory tasks. *Scandinavian Journal of Psychology, 49*(2), 93. [DOI:10.1111/j.1467-9450.2008.00633.x] [PMID]

Kuhlen, R.G. (1940). Social change: A neglected factor in psychological studies of the life span. *School & Society, 52,* 14-16.

Li, S. C., Lindenberger, U., Hommel, B., Aschersleben, G., Prinz, W., & Baltes, P. B. (2004). Transformations in the couplings among intellectual abilities and constituent cognitive processes across the life span. *Psychological Science, 15*(3), 155-63. [DOI:10.1111/j.0956-7976.2004.01503003.x] [PMID]

Light, L. L., & Singh, A. (1987). Implicit and explicit memory in young and older adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*(4), 531. [DOI:10.1037/0278-7393.13.4.531]

Light, L. L., Singh, A., & Capps, J. L. (1986). Dissociation of memory and awareness in young and older adults. Journal of Clinical and Experimental Neuropsychology, 8(1), 62-74. [DOI:10.1080/01688638608401297] [PMID]

Mitchell, D. B., & Bruss, P. J. (2003). Age differences in implicit memory: Conceptual, perceptual, or methodological? *Psychology and Aging, 18*(4), 807-22. [DOI:10.1037/0882-7974.18.4.807] [PMID]

Nilsson, L. G., Adolfsson, R., Bäckman, L., Cruts, M., Edvardsson, H., Nyberg, L., & Van Broeckhoven, C. (2002). Memory development in adulthood and old age: The Betula prospective-cohort study. In P. Graf & N. Ohta (Eds.), Lifespan development of human memory (pp. 158-204). Cambridge, Mass: MIT Press.

Nilsson, L. G., Bäckman, L., Erngrund, K., Nyberg, L., Adolfsson, R., & Bucht, G., et al. (1997). The Betula prospective cohort study: Memory, health, and aging. *Aging, Neuropsychology, and Cognition, 4*(1), 1-32. [DOI:10.1080/13825589708256633]

Nilsson, L. G., Sternäng, O., Rönnlund, M., & Nyberg, L. (2009). Challenging the notion of an early-onset of cognitive decline. *Neurobiology of Aging, 30*(4), 521-4. [DOI:10.1016/j.neurobiolaging.2008.11.013] [PMID]

Nunnally, J. C., & Bernstein, I. H. (1994). Psychometric theory (3rd ed). New York: McGraw-Hill.

Park, D. C. (2002). Aging, cognition, and culture: A neuroscientific perspective. *Neuroscience & Biobehavioral Reviews, 26*(7), 859-67. [DOI:10.1016/S0149-7634(02)00072-6]

Park, D. C., & Gutchess, A. (2006). The cognitive neuroscience of aging and culture. *Current Directions in Psychological Science, 15*(3), 105-8. [DOI:10.1111/j.0963-7214.2006.00416.x]

Park, D. C., Smith, A. D., Lautenschlager, G., Earles, J. L., & et al. (1996). Mediators of long-term memory performance across the life span. *Psychology and Aging, 11*(4), 621-637. [DOI:10.1037/0882-7974.11.4.621] [PMID]

Peirce, J. W. (2008). Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics, 2.* [DOI:10.3389/neuro.11.010.2008] [PMID] [PMCID]

Ratcliff, R., & McKoon, G. (2015). Aging effects in item and associative recognition memory for pictures and words. *Psychology and Aging, 30*(3), 669-74. [DOI:10.1037/pag0000030] [PMID] [PMCID]

Ramscar, M., Hendrix, P., Shaoul, C., Milin, P., & Baayen, H. (2014). The myth of cognitive decline: Non-Linear dynamics of lifelong learning. *Topics in Cognitive Science, 6*(1), 5-42. [DOI:10.1111/tops.12078] [PMID]

Rönnlund, M., Nyberg, L., Bäckman, L., & Nilsson, L. G. (2005). Stability, growth, and decline in adult life span development of declarative memory: Cross-sectional and longitudinal data from a population-based study. *Psychology and Aging, 20*(1), 3. [DOI:10.1037/0882-7974.20.1.3] [PMID]

Salthouse, T. A. (2005). Relations between cognitive abilities and measures of executive functioning. *Neuropsychology, 19*(4), 532. [DOI:10.1037/0894-4105.19.4.532] [PMID]

Salthouse, T. A. (2009). When does age-related cognitive decline begin? *Neurobiology of Aging, 30*(4), 507-514. [DOI:10.1016/j.neurobiolaging.2008.09.023] [PMID] [PMCID]

Schaie, K. W. (2013). Developmental influences on adult intelligence: The seattle longitudinal study (2nd ed). New York, NY: Oxford University Press. [DOI:10.1093/acprof:osobl/9780195386134.001.0001]

Singh-Manoux, A., Kivimaki, M., Glymour, M. M., Elbaz, A., Berr, C., & Ebmeier, K. P., et al. (2012). Timing of onset of cognitive decline: Results from Whitehall II prospective cohort study. *BMJ, 344*(jan04 4), d7622-d7622. [DOI:10.1136/bmj.d7622] [PMID] [PMCID]

Zelinski, E. M., & Burnight, K. P. (1997). Sixteen-year longitudinal and time lag changes in memory and cognition in older adults. *Psychology and Aging, 12*(3), 503-513. [DOI:10.1037/0882-7974.12.3.503] [PMID]

Hatami, J., et al. (2020). Reliability and Validity of the Iranian Computerized Version of Memory Tasks. *BCN, 11*(4), 535-548.

547

This Page Intentionally Left Blank