Journal of Studies in Learning
and Teaching English
Volume. 6, Issue. 2, Ser. 12, (2017), 125-147

Heterogeneous Test-takers, Homogeneous Test Construct! Foundations of L2 Assessment Revisited: A Study of Cognition-related Assessment

Behzad Nasirpour*
Ph.D. candidate of TEFL
Department of English Language Teaching
Shiraz Branch, Islamic Azad University
Shiraz, Iran.
Email: behzadnasirpour@gmail.com

Mohammad Sadegh Bagheri
Associate Professor in TEFL
Department of English Language
Larestan Shiraz, Islamic Azad University
Shiraz, Iran
Email: bagheries@gmail.com

Mortaza Yamini
Associate Professor in TEFL
Zand Institute of Higher Education
Shiraz, Iran
E-mail: mortazayamini@gmail.com

Abstract. Researchers have tried to increase knowledge about how to test language ability.Still, important elements of cognition have not been perceived in traditional tests, and assessments are not designed to assess differences in learners' levels of understanding. The question of whether Classical Test Theory (CTT), Analysis of Variance (ANOVA), Generalizability (G) Theory, and Psychometrics can justify the heterogeneity of EFL learners remains unanswered. This study investigatedthe

Received: May 2017; Accepted: July 2017
*Corresponding author

extent heterogeneous EFL learners,who have been different in their L2 memory pathways, cognitive control and type-of-test preferences, have similarly been assessed. This study was conducted on 101 adult participants (50 males; mean age 34.63). The participants' EFL levels fell within the scope of 49 basic and 52 independent users. The researchers gathered data usingSelf-Report Measure of Phonological Memory Questionnaire (SMPM) to tap into their memory pathways and type-of-test preferences. Both descriptive statistics (mean, standard deviations, frequencies) and inferential statistics (T-test) were run on the data. The analyses indicated that 50.5% of the EFL test-takers were eager to take multiple-choice/true-false test formats. Thus, they needed prompts to repair their chain of sequential events in their explicit knowledge. 8% of the participants had mature L2 cognitive control, and 14% of themwere not mature in this regard. Most EFL learners benefited implicit knowledgein retrieving information. It clarified that inputs which were acquired by explicit memory could ordinarily be supported by implicit knowledge. Cognition-related assessment/discipline should be taken into accounton both learning and performance. One type of assessment does not fit all.

Keywords: Assessment, type-of-test preferences, L2 memory pathways, L2 cognitive control

1. Introduction

In today's world of interdisciplinary fields, neuro-education applications have tried to make a significant contribution in how cognitive science can affect learning because the teaching methodology by itself did not seem to make a major breakthrough in how/what should an individual learn and retrieve (e.g., Hardiman, 2003; Mehta, 2009). It seems that L2 methodologies/assessment should be concerned with different multidisciplinary issues Should the L2 learning/assessment come within the purview of cognition and memory? To what extent do different groups of test-takershave the same repertoires, knowledge, and retrieval to take a similar test? Should the design of assessment and its interpretation fit the model of L2 cognition and learning?

It is widely believed that the performance of individuals on a test should be estimated in terms of their abilities/knowledge on the psychological construct. Since assessment is the process of quantifying the physical and mental characteristics of individuals, it is the abilities/attributes test developers want to measure but not persons (Bachman, 1995). Hence, learning is not an overnight process. Accordingly, Brooks and Kempe (2013)claimed that linguistic unit is perceived and processed in the sequence. It is worth noting that sequence learning is linked to individual differences in L2 learning (Kaufman, et al., 2010).

National Research Council (NRC) committees have been exploring equally compelling issues related to human cognition and learning. The NRC Committee on Developments in the Science of Learning examined findings from cognitive science that have advanced understanding of how people think and learn (Pellegrino, Chudowsky & Glaser, 2001). NRC along with its Board on Testing and Assessment (BOTA) pursue both assessment and the understanding of human learning; each of these disciplines could enrich the other (Pellegrino et al., 2001). NRC committee has reviewed advances in the cognitive and measurement sciences to find a joint approach between the two aforementioned disciplines. The committee members worked on topics e.g., assessment practices in terms of cognitive principles, new statistical models in assessing cognitive performances,and cognitively based instructional intervention programs(Pellegrino et al., 2001).

Bachman and Palmer (2010) believe that one cannot assume a test to be the best since a test can be inappropriate at least for some of the test takers. Therefore, a test could not meet all of the needs of the test users. They also mention that there is no just one best test for any situations, and test developers should not focus on a single quality of the test (Bachman & Palmer, 2010). All in all, advances in the study of thinking and learning (cognitive science) and measurement (psychometrics) make us think about how learners learn and what they know, what is worth assessing, and how to obtain useful information about student competencies (Pellegrino et al., 2001). In a nutshell, a joint approach to assessment, curriculum and instruction seems to be inevitable because assessment does not exist in isolation; it must be interacted with curriculum and instruction to support learning. Therefore, assessments in L2 contexts should entail learning/competence, the current knowledge about learner'scognitive control(which is affected by memory) and learning. Although considerable research has been assigned to language testing, little substantial research considers different aspects of L2 cognitive processes and the relevant knowledge. This study focused on EFL memory pathways and cognitive control that might mediate the performance of EFL individuals.

2. Literature Review

2.1. Can g-theory and classical test theory justify the heterogeneity of EFL learners?

Generalizability (G) theory is about the dependability of behavioral measurement; it just refers to the accuracy of generalizing from a person's observed score on a test to the average score that a person would have received under all the possible conditions. In other words, the individuals' knowledge, attitude, skill or other measured attributes are in steady states (Shavelson & Webb, 1991). Both Generalizability (G) theory and Classical Test Theory (CTT) set realistic goals to replace the single true score by the notion of a universe score, which can be thought of as an average score for a person across all the conditions of measurement (Bachman & Palmer, 1996). However, G-theory draws on test takers heterogeneity, certain individual cognitive abilities in performance (e.g., retrieval) might not be captured.

Moreover, in the selection and use of any psychological test, it is essential that a test be fair to all applicants, and not be biased against a segment of the applicant population. Bias can result in systematic errors that distort the inferences made in selection and classification (Zumbo, 1999). Test bias has been the subject of a great deal of recent research, and a technique called Differential Item Functioning (DIF) analysis has become the new standard in the psychometric bias analysis (Zumbo, 1999). DIF statistical techniques are based on the principle that if different groups of test-takers (e.g., males and females) have roughly the same level of something (e.g., knowledge), then they should perform similarly on individual test items regardless of group membership.

The question of whether the relationship between "the probability of passing an item" and "the level of ability" can justify the heterogeneity of EFL learners remains unclear. It can be concluded that other sources of variance in learning and performance(many aspects of cognitive processes) remain recondite. Probably, the item facet mechanism might set another optimal goal to elicit abilities/knowledge of individuals who seem to be different in decoding, retrieval, memory pathways and cognitive control. How can a test extract information, with one test format, from different knowledge pathways of individuals? And how can an assessment interpret and decide on abilities/attributes through one test format? It seems that the source of variance should be revisited. However, the assessment is based on area of content such as the nature of language ability and need analysis (Bachman & Palmer, 2010), and it seems that the nature of decoding, cognitive control, memory pathways and retrieval should be taken into account.

2.2. Cognition accounts for individual differences

The differential perspective focuses on the nature of individual differences in what they know, in their potential for learning and in their performance. Assessment practices need to move beyond a focus on component skills and discrete knowledge to entail the more complex aspects of learners' achievement (Pellegrino et al., 2001). Assessments should evaluate what schemas an individual has (Pellegrino et al., 2001) and under what circumstances he/she uses a certain memory pathway to retrieve from (Nasirpour, 2013, 2014). Cognitive control is the ensemble or set of mechanisms that confine one's thoughts and responses according to his/her goals; besides, cognitive control mechanisms permit an individual to access the internal representation in a goal directed manner (Wagner, Bunge & Badre, 2004). To them, cognitive control entails retrieval and online-maintenance of knowledge, and the functional/neuroanatomical basis of cognitive control has an influence on long term memory such as working memory, semantic memory, and priming. The prefrontal cortex (PFC) supports cognitive control and guides mnemonic processing to overcome interference and the uncertainty of the ineffective retrievals (Race, Kuhl, Badre & Wagner, 2009). Functional brain imaging studies prove the neuroanatomical basis of cognitive control which affects and is affected by long term memory (Wagner et al., 2004).

To Squire and Zola-Morgan (1991), explicit memory entails conscious recollection of facts and events. Explicit memories are available to conscious recollection. Implicit memory relates to a collection of abilities wherein performance changes due to experience, but without affording access to conscious memory of the original experience. Understanding the contents of long-term memory is especially critical for determining what people know; how they know it; and how they can use that knowledge to answer questions, solve problems, and engage in additional learning (Pellegrino et al., 2001). To Squire (1992) and Paller and Squire (2007), explicit memory resides in hippocampus in the medial temporal lobes. It is the memory we can describe, elucidate, and write about. The written exam, essay or any exam-type recall used in schools and colleges are one of the types. Jensen (1998) states that explicit memory depends on medial temporal diencephalon and extensive engrams of neocortex. Explicit memory has many forms, including word-based semantic memory and the event-type episodic memory. To Paradis (2004), knowledge of language is the internalized implicit knowledge and explicit knowledge; the first is unavailable to introspection and the second corresponds for example, to the declarative component of skills.

Working memory capacity (WMC) is strongly related to executive functions (e.g., McCabe, Roediger, McDaniel, Balota & Ham-brick, 2010). Individual differences in WMC have been implicated in social psychological phenomena such as stereotype threat, emotion regulation, and intrusive thought suppression (Redicket al., 2012). Kane, Conway, Hambrick and Engle (2007) elucidate that performance on Operation, Symmetry, and Reading Span reflect individual differences in executive attention. Nasirpour (2018) observed that working memory (WM) in phonological loop was correlated to automatic semantic memory and WM-priming interactions. His investigation also revealed that there are differences between the means of semantic memory, episodic memory, and declarative and non-declarative memories among the EFL learners (Nasirpour, 2013, 2014). Although the subjects were rather good at episodic retrieval, they did not seem to be good at semantic retrieval (Nasirpour, 2013, 2014).

In France, Binet and Simon (1980) designed mental tests to identify the differences among learners. Their approach revealed that individuals differed in their mental capacities and that these differences defined stable mental traits, namely, aspects of knowledge, skill, and intellectual competence that could be measured. The differential perspective was assigned to evaluate the intelligence or cognitive ability that were separate from the processes and content of academic learning.

Modularity implies domain specific, innate learning constraints. However, normal individuals may differ in how they recruit regions into networks (Ovsiew, 2007). The methods used in studying groups of subjects in functional imaging experiments may obscure such individual differences (Ovsiew, 2007). For example, robust individual differences in patterns of activation emerged in a memory task, differences putatively reflecting different strategies in performing the task. Individual differences in the organization of language cortex are clinically evident in the unusual occurrence of crossed aphasia (aphasia due to right hemisphere injury in a dextral), crossed non-aphasia (lack of aphasia with a left hemisphere injury causes aphasia in a dextral), and aphasic deficits in both dextrals and sinistrals (Ovsiew, 2007).

Recent neuroimaging studies have revealed potentially important differences in the timing of PFC development across typical and atypical individuals (Thompson-Schill et al., 2009). Thompson-Schill et al. (2009) suggested researchers to account for individual differences in cognitive abilities among different developmental groups (i.e., infants, toddlers, adolescents) since different stages of prefrontal maturation are coupled with different learning opportunities. Cognition without control (i.e. initially underdeveloped PFC) may allow for other subcortical networks (e.g., hippocampus) to facilitate certain types of learning at different developmental stages (Aston-Jones & Cohen, 2005).

2.3. Cognitively based assessment

Most of the educational tests do a reasonable job with certain functions of testing in measuring knowledge of basic facts/procedures and producing overall estimates of proficiency within the curriculum (Pellegrino et al., 2001). Thus, the test strengths are a product of their adherence to theories of learning and measurement that fail to capture the breadth and richness of knowledge and cognition (Pellegrino et al., 2001). To what extent these assessments capture learners' complex knowledge and skills that are essential for success in the information-based economy (Resnick & Resnick, 1992). Traditional tests do not consider many aspects of cognition; therefore, they are not structured to capture critical differences in students' levels of understanding (Pellegrino et al., 2001). Thus, the validity of the inferences drawn from these kinds of results might be questioned.

Three key elements underlying any assessment are cognition, observations, and interpretation which are portrayed as the assessment triangle; for an assessment to be effective, the three elements must be in synchrony (Pellegrino et al., 2001). Cognition indicates how students represent knowledge and develop competence in a subject domain. The question arises as to whether tasks are linked to the cognitive model of learning and individuals' knowledge pathways. This raises another question whether a method, which is optimized for performance, can be optimal for learning. To Pellegrino et al. (2001), tasks must be carefully designed to be linked to the cognitive learning and to support the types of decisions that are going to be based on the assessment results. Firstly, L2 pedagogy should consider sequence learning in any individual since sequence learning is linked to individual differences in L2 learning (Kaufman et al., 2010). Secondly, L2 pedagogy should highlight the learning-performance discrepancy since a system which is optimized for performance may not be optimal for learning (Thompson-Schill, Ramscar & Chrysikou, 2009). Observations require the assessment task to elicit illuminating responses from students; furthermore, any assessment is based on interpreting the evidence collected from observations (Pellegrino et al., 2001). The NRC committee claimed that assessment, curriculum, and instruction could be better achieved if they were derived from a shared knowledge base about cognition and learning (NRC, 1999). The central core, namely cognition, is the scientific understanding of how individuals learn.

3. Purpose of the Study

This study was carried out to investigate whether the theories of assessment entail the breadth and richness of cognition among EFL individuals. Is there an absolute sure-fire assessment for different cognitive beings (EFL individuals)? To Pellegrino et al. (2001), assessment must be linked to the cognitive learning. Thus, the authors designed a scale to measure L2 cognitive control and memory pathways, namely SMPM. Accordingly, the following questions were addressed to answer the objectives of this study:

1. Are there any significant differences in literal memory pathways between EFL basic and independent test-takers taking the same test?

2. Which type of test has a greater degree of preference among EFL test-takers?

3. Is there any relation between EFL learners' test preference and their types of memory?

4. Method

4.1. Participants

Five different universities and public/private sectors where English is the medium or studied as a foreign language were negotiated in Shiraz, Iran, in the early 2018. The participants of the study comprised 101 (50 males; 52 EFL independent users; 49 EFL basic users; mean age 34.63) graduate/post graduate students of universities, EFL learners and non-students, namely Agriculture College of Shiraz University (ACSU); Islamic Azad University, Shiraz Branch (IAU); South Industrial Management Institute (SIMI); Fars Regional Water Company (FRWC), and Pooräb Fars Engineering Consulting Company (PFECC). Since English program is a major course in the Iranian university and school systems, students take EFL classes as part of the university/school curriculum. The reason for selecting these students and staff members was that they were mature enough to perform on the SMPM questionnaire and cognitive tests. All the participants from each stratum received the questionnaire (i.e., SMPM).

Table 1: Universities, Institutes, and Public/Private Sectors and Number of Participants

Sector	Specialism	Participant groups	Number of participants
A	English Dpt., Islamic Azad University, Shiraz Branch	2	27
B	R & D Dpt., Fars Regional Water Company, Shiraz	1	26
C	English Dpt., South Industrial Management Inst.	2	23
D	Department of Agricultural Extension and Education, Shiraz University	2	15
E	International Trade Dpt., Pooräb Fars Engineering Consulting Co.	1	10
	Total	8	101

Note: Universities, institutes and public/private sectors are presented in chronological data collection order

4.2. Instrument

For the purpose of data collection in this study the Self-Report Measure of Phonological Memory Questionnaire-SMPM (Nasirpour, 2013, 2014, 2018) was used. Self-Report Measure of Phonological Memory Questionnaire (SMPM) was based on a Likert scale where a set of ordered responses was arranged on a scale of 1 to 5. The numbers in front of the choices are the values they carry that manifested the degree of the preference or tendency of the participants towards the items of the questionnaire, on a five-point Likert scale. The SMPM questionnaire combines 41 items identifying different memory questions: explicit/semantic memory questions; episodic memory types; implicit memory; priming/recognition memory questions; working memory questions, and implicit/cell memory. The respondents were supposed to read the items and select choices. Accordingly, items 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, and 23 were indicative of explicit/episodic memory; items 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, and 24 were identified explicit/semantic memory. Items 2, 4, 6 and 8 determined controlled semantic retrieval and items 10, 12, 14, 18 and 22 determined selection semantic retrieval. Moreover, items 25, 27, 29, 31, and 33 were working memory; items 24, 30, 32, and 34 indicated whether learners used different types of memory when they used different learning skills. Besides, items 26 and 28 were the non-declarative/procedural memory; items 35, 36, 37, and 38 were brain potentials associated with priming/explicit (item verbal memory); and items 39, 40 and 41 indicated implicit/cell memory. The total index of cognitive control was assessed calculating the sum of the means of semantic memory,priming and working memory. The reliability index for the SMPM obtained through Cronbach's Alpha was 0.81.

4.3. Data collection and analysis procedures

The data were collected in Shiraz in two successive months in the winter 2018 from diverse strata serving the purpose of quota sampling. All the data collection was carried out by the researchers themselves and their two assistants who have been trained. However, the informed consent was already obtained, the nature and purpose of the research were again explained to the participants by the researchers. The participants were assured of the confidentiality of the results and the advantages of the study. They took the Persian versions of questionnaires for the ease of reading and answering.

This project aimed at two EFL levels, namely basic users (A1 & A2) and independent users (B1 & B2) based on the CEFR description (in English). Identification of the proficiency level of the participants was based on 1) actual IELTS Test reports which include CEFR level, and 2) IELTS Mock Test (2018) which was administered in different aforementioned sectors. As for Mock Test, their band scores closely aligned with the levels of the Common European Framework of Reference (CEFR). Most of the participants had taken the test between two to four weeks before the study while few of them took the Mock Test one to three weeks after the study. The participants' overall CEFR levels fell within the scope of EFL basic users (49 subjects) and EFL independent users (52 subjects).

The descriptive statistics such as mean, frequency, valid percentage, cumulative percentage, and standard deviation were calculated for the data to determine the types of memory pathways and test preferences the participants were inclined to use. Later, inferential statistical procedures such as T-test were run on the data. Analysis was performed using SPSS software version 25.0.

5. Results

The results of data analysis for each research question are provided below.

Question One:
To observe the difference between L2 memory pathways (semantic/episodicmemory,explicit/implicit memory; cognitive control),the mean and standard deviation were run. The mean score of the memory pathways in SMPM fell at high and medium level, which showed discrepancies among EFL learners in tilting toward a certain type of test or memory pathway.

As it is evident in Table 2, the mean and standard deviation of the participants' L2 memory pathways are displayed. The highest frequently memory used in Table 2 was implicit memory with a high mean of 3.84 (SD= 0.83). With regard to implicit memory, the mean of explicit memory was running at about 3.30 (SD= 0.34).

Table 2: Mean and SD of the Memory Pathways

	N	Min.	Max.	Mean	Std. Deviation
Explicit memory	100	2.50	4.21	3.30	.34
Implicit memory	101	1.50	5.00	3.84	.83
Semantic memory	100	2.33	4.08	3.26	.33
Episodic memory	101	2.25	4.67	3.33	.56
Working memory	101	1.00	4.40	2.72	.64
Choosing MC/TF in preference	101	1.00	5.00	3.40	1.20
Cognitive control	101	1.71	4.07	3.00	.42
Valid N (list wise)	100				

It is worth noting that the lowest frequently memory used in Table 2 was working memory with means of 2.72 (SD= 0.64).

Following this, it was decided to observe differences between those participants who were mature, mediocre or immature in terms of the frequency of L2 cognitive control (i.e., the sum of semantic memory, WM and priming).

As it is evident in Table 3, almost 8% of the participants in this study hadmatureL2 cognitive control. 78% of the participants were moderate, and nearly 14% of the participants were not mature in L2 cognitive control.

Table 3: The Percentage and Frequency of L2 Cognitive Control

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Mature	8	7.9	8.0	8.0
	Moderate	78	77.2	78.0	86.0
	Immature	14	13.9	14.0	100.0
	Total	101	100.0		

It was also decided to assess the significant difference between EFL basic and independent users in using specific L2 memory pathways. Thus, information in Descriptive Statistics and Independent Samples Test was summarized in Table 4.

As it is evident in Table 4, the difference between the means of EFL basic and independent users in semantic memory (total) was significant at 0.05. The mean score of EFL independent users was greater than that of EFL basic users regarding semantic memory. The mean score of L2 cognitive control of the EFL independent users was greater than that of the EFL basic users.

Table 4: Descriptive Statistics and T-test Results Combined on Explicit/Implicit Memory and Cognitive Control Used by EFL Basic and Independent Users

		N	Mean	SD	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference
Explicit	basic users	47	3.26	.349	.123	.727	-1.02	97	.309	-.07
memory	indep. users	52	3.33	.336			-1.02	95	.310	-.07
Implicit	basic users	48	3.90	.726	5.29	.024	.58	98	.558	.09
memory	indep. users	52	3.80	.929			.59	95	.555	.09
Semantic	basic users	47	3.12	.238	5.22	.024	-4.32	97	.000	-.26
memory	indep. users	52	3.39	.362			-4.41	88	.000	-.26
Episodic	basic users	48	3.40	.585	.237	.628	1.13	98	.260	.12
memory	indep. users	52	3.27	.537			1.12	95	.262	.12
Cognitive	basic users	47	2.85	.34	1.52	.220	-3.40	97	.001	-.280
control	indep. users	52	3.13	.45			-3.45	94	.001	-.280

Question Two:

The researchers assessed significant differences between participants (EFL basic and independent users) in terms of the frequency of the individual preference for multiple-choice/true-false (MC/TF)exam (see Table 5).

As illustrated in Table 5, the frequency distribution and percentage present the data in order to indicate that 50.5% of the participants in this study were inclined to take multiple-choice/true-false (MC/TF) exam rather than the written one. Almost 23% of the participants were neutral and 26.7% of the participants did not prefer to take MC/TF exam.

Table 5: The Percentage and Frequency of MC/TF Preference

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	MC/TF were preferred	51	50.5	50.5	50.5
	No difference	23	22.8	22.8	73.3
	MC/TF were not preferred	27	26.7	26.7	100.0
	Total	101	100.0	100.0	

Question Three:

To see whether there was any relation between EFL learners' test preference and their types of memory Pearson Correlation was run. As illustrated in Table 6, the correlation coefficient between MC/TF preference and episodic memory was .560 and the p-value was 0.000. It can be concluded that the correlation coefficient between MC/TF preference and episodic memory was significant.

Table 6: Pearson Correlation Between MC/TF Preference and Episodic/Semantic Memory

		Episodic memory	Semantic memory
MC/TF preference	Pearson Correlation	.560**	.034
	Sig. (2-tailed)	.000	.738
	N	101	100

**. Correlation is significant at the 0.01 level (2-tailed).

6. Discussion

To offer a comprehensive depiction of cognitively based assessment, this study is premised on the idea that heterogeneity of test-takers should be explored in terms of many aspects of L2 cognitive control and L2 memory pathways. As for the first research question, significant differences in literal memory pathways were observed. According to the data, it was detected that objective functioning (implicit memory) was greater than subjective functioning (explicit memory). As it is evident in Table 2, the highest frequently memory used was implicit memory with a high mean of 3.84 (SD= 0.83). By contrast, explicit memory (the combination of semantic and episodic memory) had a moderate mean of 3.30 (SD= 0.34). For the purpose of the current study, the researchers juxtaposed the results with psycho-philosophical notions of Chalmers (2009) who found out that science of consciousness integrates two key classes of data into a scientific framework: 1) third-person data, or data about behavior and brain processes (such as perceptual discrimination of external stimuli, levels of access to internally represented information, etc.), and 2) first-person data, or data about subjective experience (such as visual/bodily experience, emotional experience, etc.) (Chalmers, 2009). From the third person point of view, it seems that unconscious perception of visually presented linguistic/nonlinguistic stimuli were stronger (deciphering implicit questions). Subjects might benefit the distinctive quality of subjective/objective experience associated with the theme of the implicit memory questions which explained the various third-person data. Based on subjective/objective experience, different individuals may have different internal speech in pre-motor coding. Comparing the means of explicit (conscious) and implicit (unconscious) memory, the researchers suggested that the objective functioning of the phonological system was greater than subjective functioning. Thus, their first-person data-the data of subjective experience-might not be data about objective functioning.

The analysis also revealed that inputs which were once acquired by explicit memory could ordinarily be supported by implicit memory. That's why implicit memory was the highly frequently used memory.

Besides, the capacity for episodic recollection and synthesis of episodic memory (i.e., semantic memory) depends on the same input structures and sequences. It seems some of the participants did not benefit from the frequency of occurrence, categorical relations, overlapping, and timeless routine experiences. From the third person point of view, it seems that unconscious (implicit) perception of visually presented linguistic/non-linguistic stimuli were stronger than that of explicit (conscious) memory. Thus, inputs which were once acquired by explicit memory could be supported by implicit knowledge. This corroborates the evidence found by Thompson-Schill et al. (2009) who broached that a system which is optimized for performance may not be optimal for learning. These findings revealed that the individual's performance in an assessment would be different from his/her previous learning. Thus, cognitive-minded assessment would entaila psycho-cognitive construct which could take account of changes in learning and test performance.

To Wagner et al. (2004), functional basis of cognitive control has an influence on long term memory, such as working memory, semantic memory, and priming. Thus the mean of these three memories can entail cognitive control. As illustrated in Table 3, 8% of the participants had mature cognitive control, and 78% of them were moderate. But nearly 14% of the participants did not mature in cognitive control. This clarifies that certain types of learning are facilitated at different developmental stages (Aston-Jones & Cohen, 2005). It is also in agreement with Thompson-Schill et al. (2009) who claimed that different stages of maturation are coupled with different learning opportunities. Moreover, individual differences were observed in WMC (Redick et al., 2012) and executive attention (Kane et al., 2007). Due to the late prefrontal cortex development, what is optimized for learning might not be optimal for performance(Thompson-Schill et al., 2009). This reveals that an assessment system (as the individual's performance) seems to be different from a learning system since different stages of cognitive maturation are connected with different learning opportunities (Thompson-Schill et al., 2009). It answers the first and third research questions.

As it is evident in Table 4, The mean score of EFL independent users was greater than that of basic users regarding semantic memory. However, there was no significant difference between EFL basic and independent users in implicit and episodic memory. The findings in Table 4 revealed that EFL basic users would process semantic memory as part of a sequence. The analysis also revealed that language retrieval is sequential and sensitive to morphological, semantic and syntactic regularities. The findings of this study was in line with Eichenbaum (2004) who broached that the episodic memory task can come first since episodic memory is the gateway through which we encode events and semantic memory. In other words, EFL learners might need to shift their attention towards the salient memory pathways one at a time (e.g., episodic memory and perceptual priming); later they may use them concomitantly (e.g., semantic unification). Accordingly, Brooks and Kempe (2013) observed that each linguistic unit is perceived or produced as part of a sequence, with one unit processed at a time. Thus, sequence learning/retrieval were linked to participants' differences in L2 learning. This was in agreement with the findings of Kaufman et al. (2010).

As it is evident in Table 4, the participants in this study had different stages of cognitive maturation based on different learning opportunities. The results were in line with Huttenlocher and Dabholkar (1997) and Thompson-Schill et al. (2009) who broached that due to the late prefrontal cortex development and heterochronicity, children exhibit impaired behavioral and cognitive control for years. Appropriate uses of assessment in standards-based reform (e.g., cognitive-minded assessment, new psychometric model), a movement that can reshape curriculum, syllabus and education throughout the world seems to be inevitable.

Considering the second and third research questions, our findings showed that participants who chose "MC/TF in preference to written exam" had a mean of 3.40 (SD=1.20), indicating that the participants needed episodic sequence of information to recall (Table 2). It answers both the second and third research questions. The contamination of episodic/explicit processing often happens to students who really do know their material but lack the specific "hooks", "mental file names" (Jensen, 1998), "inkling", "whisper", and "indication" to retrieve all their learning. Now we find out why 50.5% of the EFL test-takers like multiple-choice/true-false (MC/TF) test formats; they provide the prompts that their brain needs. As it is evident in Table 6, there was a high correlation between MC/TF preference and episodic memory. Therefore, at least 50.5% of the subjects needed prompts to repair their chain of sequential events in their episodic knowledge (see Table 5 and Table 6). Forgetting occurs because such cues are rarely present when the recall is needed. It answers the second research question.

Since assessment is the process of quantifying the mental characteristics of individuals, and different fields of knowledge should be defined by different individuals, universe score variance in G-theory does not seem to be tenable due to individuals' L2 cognitive maturation. Why language learners were different in retrieving materials was partly influenced by part of their previous individual experiences; besides, their explicit learning was individually dependent. Since semantic memory is the synthesis of episodic memory, which is associatively networked and stored in a distributed network, it links the conceptual representations. As a result, the association varies between individuals in strength, their overlap in features and their categorical relations (Nasirpour, 2012, 2013). Therefore, if a psychological test construct is based on semantic dimension, test-takers, who are not mature enough in linking conceptual representations, will be the losers.

DIF statistical techniques whose proponents advocate that test-takers should perform similarly on individual test items regardless of group membership might reconsider their notion because based on heterochronicity (in cognitive maturation) and different memory pathways of individuals, no two classmates should take the same test (psychological test construct). As a result, cognitive-minded assessment seems to be called for.

7. Conclusion

The findings suggested that cognitive maturity in L2 learning/retrieval may be attributable to the individual differences in using mnemonics, detecting underlying rules and processing semantic unification. The processes learning and performance entail at different levels of competence(Pellegrino et al., 2001). Thus, it does not seem that one type of assessment can fit all. Nowadays, a single assessment is used for multiple purposes/knowledge in general. However, the more purposes a single assessment aims to serve, the more sources of variance (heterochronicity, different memory pathways) should be compromised (e.g., Pellegrino et al., 2001). The developmental time course of L2 cognitive processes testifies the theory of plasticity of connectomes observed among EFL independent users who seemed to benefited more from experience-dependent plasticity and analogical reasoning than those of EFL basic users who struggled with miss-wiring. The findings of L2 cognitive control (e.g., evolutionary compromises between L2 memories) open up perspectives in how EFL learners should be cognitively individuated. The level of L2 learning/retrieval is highly dependent onthe learners'different cognitive stages, the extent to which they are able to use linguistic cues sequentially or concomitantly. Therefore, cognition-related discipline/assessment should be taken into account in both L2 learning and performance. This model should be based on the best available understanding of how different individuals retrieve, perform and develop competence in EFL context. It should serve as the centerpiece of the assessment design process as well.

8. Implications

The results of the present study highlight the need for merging L2 methodologies in cognition and learning and methods of L2 measurement. L2 assessments should be able to assess how cognitively mature learners retrieve linguistically dependent cues, in comparison with L2 learners who are less mature in this regard. Such information can help L2 educational systems dealing with L2 learning and assessment to differentiate among many levels of plasticity and competence. To sum up, to design better assessments, test designers should integrate cognitive and measurement principles.

References

Aston-Jones, G. and Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: Adaptive gain and optimal performance. *Journal of Annual Review of Neuroscience, 28*, 403-450.

Bachman, L. (1995). *Fundamental considerations in languagetesting* (3rd ed.) Cambridge: Cambridge University Press.

Bachman, L. and Palmer, A. (1996). *Language testing in practice: Designing and developing useful language tests.* Oxford: Oxford University Press.

Bachman, L. and Palmer, A. (2010). *Languageassessment in practice.* New York: Oxford University Press.

Binet, A. and Simon, T. (1980). *The development of intelligence in children.* Nashville, TN: Williams.

Brooks, P. J. and Kempe, V. (2013). Individual differences in adult foreign language learning: The mediating effect of metalinguistic awareness. *Journal of Memory and Cognition, 41,* 281-296.

Chalmers, D. J. (2009). How can we construct a science of consciousness? In M. S. Gazzaniga (Ed.), *The cognitive neuroscience* (4th ed.). Cambridge, Mass.: MIT Press.

Eichenbaum, H. (2004). An information processing framework for memory representation by the hippocampus. In M. S. Gazzaniga, *The cognitive neuroscience.* London: A Bradford book.

Hardiman, M. M. (2003). *Connecting brain research with effective teaching: The brain-targeted teaching model.* Rowman & Littlefield Education, Landham, MD.

Huttenlocher, P. R. and Dabholkar, A. S. (1997). Regional differences in synaptogenesis in human cerebral cortex. *Journal of Comparative Neurology, 387*(2), 167-178.

Jensen, E. (1998). *Teaching with the brain in mind.* Virginia: Association for Supervision and Curriculum Development, Alexandria.

Kane, M. J., Conway, A. R. A., Hambrick, D. Z., and Engle, R. W. (2007). Variation in working memory capacity as variation in executive attention and control. In A. R. A. Conway, C. Jarrold, M. J. Kane, A. Miyake, & J. N. Towse (Eds.), *Variation in working memory.* New York: Oxford University Press.

Kaufman, S. B., DeYoung, C. G., Gray, J. R., Jimnez, L., Brown, J., and Macintosh, N. (2010). Implicit learning as an ability. *Journal of Cognition, 116,* 321-340.

McCabe, D. P., Roediger, I, H. L., McDaniel, M. A., Balota, D. A., and Hambrick, D. Z. (2010). The relationship between working memory capacity and executive functioning: Evidence for a common executive attention construct. *Journal of Neuropsychology, 24,* 222-243.

Mehta, A. (2009). Neuroeducation emerges as insights into the brain and learning grow. *The Dana Foundation, 19*(3), 1-3.

Nasirpour, B. (2012). *Investigating different types of memory L2 learners are inclined to use: innately endowed or non-innately enriched.* Unpublished MA Thesis, IAU, Shiraz Branch.

Nasirpour, B. (2013). Investigating different memory pathways L2 learners use: A practical application to the brain compatible learning. *Journal ofStudies in Learning and Teaching English, 1*(3), 77-103.

Nasirpour, B. (2014). *Brainy brain: Neuropsychological linguistics.* Saarbrucken: LAP Lambert Academic Publishing.

Nasirpour, B. (2018). The assessment of organizational intelligence among experts and engineers in Fars Regional Water Company: Implementation of meta-knowledge management (Report No. 110/14061). Iran's Ministry of Energy, Practical Research Committee of Fars Regional Water Co.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics.* Reston, VA: Author.

National Research Council. (1996). National science education standards. National Committee on Science Education Standards and Assessment. Coordinating Council for Education. Washington, DC: National Academy Press.

National Research Council. (1999). *High stakes: Testing for tracking, promotion, and graduation.* Committee on Appropriate Test Use. In J.P. Heubert and R.M. Hauser, (Eds.), *Commission on behavioral and social sciences and education.* Washington, DC: National Academy Press.

Ovsiew, F. (2007). Neuropsychiatry and behavioral neurology. In B. J. Sadock & V. A. Sadock (Eds.), *Comprehensive textbook of psychiatry.* Philadelphia, PA: Lippincott. Williams & Wilkins.

Paller, K. A. and Squire, L. R. (2007). Biology of memory. In B. J. Sadock & V. A. Sadock (Eds.), *Comprehensive textbook of psychiatry.* Philadelphia, PA: Lippincott. Williams & Wilkins.

Paradis, M. (2004). *A neurolinguistic theory of bilingualism.* Amsterdam: John Benjamins.

Pellegrino, J. W., Chudowsky, N., and Glaser, R. (2001). *Knowing what students know: The science and design of educational assessment.* National Academy of Sciences. Retrieved
http://www.nap.edu/catalog/10019.html.

Race, E. A., Kuhl, B. A., Badre, D., and Wagner, A. D. (2009). The dynamic interplay between cognitive control and memory. In M. S. Gazzaniga (Ed.), *The cognitive neuroscience* (4th ed.). Cambridge, Mass.: MIT Press.

Redick, T. S., Broadway, J. M., Meier, M. E., Kuriakose, P. S., Unsworth, N., Kane, M. J., et al. (2012). Measuring working memory capacity with automated complex span tasks. *European Journal of Psychological Assessment, 28*(3), 164-171.

Resnick, L. B. and Resnick, D. P. (1992). Assessing the thinking curriculum: New tools for educational reform. In B.R. Gifford and M.C. O'Connor (Eds.), *Changing assessments:Alternative views of aptitude, achievement, and instruction.* Boston: Kluwer.

Shavelson, R. J. and Webb, N. M. (1991). *Generalizability theory: A primer.* Newbury Park: Sage Publications Inc.

Squire, L. R. (1992). Memory and the hippocampus: A synthesis from findings with rats, monkeys, and humans. *Journal of Psychological Review, 99*(2), 195-231.

Squire, L. R. and Zola-Morgan, S. (1991). The medial temporal lobe memory system. *Journal of Science, 235,* 1380-1386.

Thompson-Schill, S. L., Ramscar, M., and Chrysikou, E. G. (2009). Cognition without control: When a little frontal lobe goes a long way. *Journal of Current Directions in Psychological Science, 18*(5), 259-263.

Wagner, A. D., Bunge, S. A., and Badre, D. (2004). Cognitive control, semantic memory, and priming: Contribution from the prefrontal cortex. In M. S. Gazzaniga (Ed.), *The cognitive neuroscience* (3rd ed.). Cambridge, Mass.: MIT Press.

Zumbo, B. D. (1999). *A Handbook on the theory and methods of differential item functioning (DIF): Logistic regression modeling as a unitary framework for binary and Likert-type (ordinal) item scores.* Ottawa, ON: Directorate of Human ResourcesResearch and Evaluation, Department of National Defense.