On the relationship among Iranian EFL Leaners' Crystallized Intelligence, Fluid Intelligence, Creativity, and their Performances on Multiple-Choice and Constructed-Response Items

Nilufar Ebrahimpur, Khalil Motallebzadeh*, Mitra Zeraatpishe

Department of English Language Teaching, Mashhad Branch, Islamic Azad University, Mashhad, Iran

Corresponding Author: Khalil Motallebzadeh, E-mail: k.motalleb@iautorbat.ac.ir

ARTICLE INFO

ABSTRACT

Studies have been done on finding construct-irrelevant factors and cognitive processes involved in test taking. Previous studies have explored correlations between reading comprehension (RC) ability and psychological variables such as creativity and intelligence reporting significant relations. Many researchers trying to figure out the difference between diverse formats performance. The present study takes interest in investigating if each test format performance is affected by cognitive traits of test takers. It investigates the effects of three psychological variables including Fluid intelligence (Gf), Crystallized intelligence (gc), and creativity (C) on reading comprehension (RC) performance where Multiple-Choice (MC) and Constructed-Response (CR) formats are involved for English as a Foreign Language (EFL) Learners. The relations among all five variables are examined applying Structural Equation Modeling (SEM) by hypothesizing a model related to the previous researches. The model goes through modifications twice and the final revision reports the relationships. Path analysis demonstrates direct significant effects in paths for Gc-MC, Gc-CR, Gf-CR and indirect significant effects in Gf-MC, creativity-MC, and creativity-CR. Therefore, CR items are the most affected format by those cognitive variables. The results are further discussed and concluded in more details.

Key words: Fluid Intelligence (Gf), Crystallized Intelligence (Gc), Creativity (C), Reading Comprehension (RC), Multiple-Choice (MC) Items, Constructed-Response (CR) Items

INTRODUCTION

Test validity has been the concern of the researchers and teachers as an important factor in designing tests in language testing. Tests should not measure the learners' abilities which are certain to result in success on the part of the learner. According to Bachman (1990), communicative language ability, test method facets, personal attributes, and random factors affect test performance favoring one person over another. Kunnan (2004) argues that certain test formats make some students with specific abilities more successful in test performance than the others who do not possess those abilities. This threatens the validity of the tests.

The investigations for the effects of test method facets have received significant attention in the field of language testing or other fields of study. Many studies have indicated the differences among different formats specially multiple choice (MC) and constructed response (CR) question types in language learning, mathematics, and economics (Ackerman & Smith, 1988; Heim & Watts, 1966; Kastner & Stangl, 2011; Katz, Bennett, & Berger, 2000; Kennedy & Walstad, 1997; Martinez, 1991; Rodriguez, 2003). As stated by In'nami and Koizumi (2009), these two formats (MC & CR) in the literature have been the most explored comparison with

each study reporting a different result. This leads the current study to consider only those formats in English reading comprehension (RC) for analysis.

Individuals' characteristics, i.e, personality and cognitive styles interacting with different aspects of test method result in different performances (Bachman, 1990). The relations among test takers' characteristics and test or item formats need closer attention and investigations. The correct response to different types of items can be affected by many different cognitive traits. There are researches that have investigated the effects of cognitive and personal characteristics of the learners on their performance on different test formats of English language tests (Barati, Ravand, & Ghasemi, 2013; Birenbaum & Pinku, 1997; Gamer & Engelhard, 1999; Hansen, 1984; Kordjazi & Ghonsooly, 2015; Pishghadam & Tabataba'ian, 2011a, 2011b).

Among many of the test takers' characteristics, the two major components of intelligence, crystallized and fluid intelligence in Cattell-Horn-Carroll (CHC) model of intelligence seem to have not been investigated for the effects they have on the test format performance; however, general intelligence (g) effects on reading comprehension (having MC items) have been measured and proved in the literature

(Ekstrand, 1977; Ghabanchi & Rastegar, 2014; Lohnes & Gray, 1972). Messick (1987) highlights that MC questions emphasize memory and convergent thinking which according to Chamorro-Premuzic and Reichenbacher (2008) is required by intelligence. Solving well-defined and rational problems with one correct answer requires convergent thinking (Csikszentmihalyi, 1996). As long as Gc includes lexical knowledge, general information, and information about culture (Flanagan, 2008), it seems that it is also effective on CR items performance as well. RC with CR items is still a test calling for a response from a test taker to produce an answer in his or her language after comprehending a text. This is a reason which rationally explains why Gc is more effective in answering CR format of RC tests. Previously, researches have explored the relationship between Gc and RC (Benson, 2008; Evens, Floyed, & McGrew, 2002; Reynolds & Turek, 2012; Taub & Benson, 2013).

Gf, as another factor of intelligence, involves inductive, deductive, and quantitative reasoning needed for working with new processes and materials While Gc is the acquired knowledge for answering questions and solving problems that are presented at familiar processes and materials (Sternberg & Kauffman, 2011). The relationship between Gf and RC ability is investigated (Benson, 2008; Evans, Floyed, & McGrew, 2002; Kassaian, 2008; McGrew, 1993; Motallebzadeh & Tabatabaee, 2016). Its relation with RC test including CR item has not been investigated in the literature. This research hypothesizes that Gf also affects CR items performance because it requires the test takers to write the answers themselves which is not similar to MC items including options for the test taker to pick from as the answers. There is a lack of research in exploring the relationship between Gf and the performance on CR items of RC tests.

Another influential factor in tests might be creativity. Creativity is considered as a psychological trait of the foreign language learners and test takers, and according to Boden (1998), is a fundamental activity of human information processing. Creativity is said to be almost under-researched in foreign language learning investigations (Dornyei, 2005) and also there seems to be no research about the relationships between this construct and performance on different test formats in language testing. Despite this fact, there have been researches regarding the relations between creativity and other variables such as language achievement (Pishghadam, khodaday, Zabihi, 2011; Sutrisno, 2007) and foreign language reading comprehension (Gould, 1972; Mousavi, Maghsoudi, Yarahmadi, 2013). Messick, (1993) emphasized that divergent production requires some form of CR for their direct examination.

Considering test item formats in assessing foreign language learning, most of the studies have paid attention to comparing the performance on different forms of the tests. As mentioned previously, few studies seem to have explored the interactions between the characteristics of the learners and test method facets. The literature of language testing, specially the works on the correlations between RC and personal characteristics of the test takers, seem to have mostly evaluated the relationships by administering MC items in

reading comprehensions. It is also important to understand the relations between other test formats and those psychological traits in order to find the most affected test format and the extent to which each format is construct-irrelevant. The three variables have not been measured for the effects they have on other test formats (eg.CR) by comparing the amount of effect on each form (MC & CR). It seems that the psychological variables in this study and their relationships specifically with MC and CR test formats have not been investigated to the researchers' knowledge.

This is an important investigation for the researchers and test designers who intend to provide the test takers with sound assessment method specially by knowing the strength and significance of the relationships among the variables. Teachers also benefit from the results of this study as they will be able to use the appropriate form of a test by understanding their students' personal traits first. It can also prompt other researchers to study the relations among the creativity, intelligence, and other test formats so that the least affected form of the test is found.

There is a lack of research that provides information on the relations among these variables, all in one established model. The links between the variables are analyzed by employing multiple regression, path, and factor analysis as building blocks of SEM (Schumacker & Lomax, 2010). In accordance with the related literature, a theoretical model in SEM is examined to determine if it is supported by the sample data. This has not been done in the past researches to show the interrelationships among the variables mentioned in this study.

Intelligence

As Kline notes (1991, p.1) "Intelligence is popularly defined as the ability to learn, understand and deal with novel situations". This study applies Cattell-Horn-Carroll (CHC) theory of intelligence the construction of which started by Cattell (1940) designing culture-free tests with items of "common knowledge" that could be applicable to all cultures (as cited in Carroll, 1984). This theory is similar to Hebb's (1949) classification of intelligence into type A and B based on clinical investigations. According to Hebb (1949) A type of intelligence is an innate ability and hereditary which was not alone responsible for intelligence development. Type B intelligence is experiences that affect mature behavior.

In the factor analysis done by Cattle in 1963 and 1967, two separate general factors are extracted (gf & gc) using culture fair intelligence tests, Thurstone's primaries, and High School Personality Questionnaire. Culture-fair, relation-perceiving tests load on gf with verbal, spatial, numerical ability, trained reasoning, and other scholastic abilities loading on the gc factor (as cited in Cattell, 1987).

In 1990s, Horn added more abilities in Gf-Gc model including visual perception or processing, short-term memory, long-term storage and retrieval, and speed of processing. Later auditory processing ability reaction time and decision speed also became a part of the model. Finally based on Horn (e.g. 1991) and Woodcock (1994), quantitative and broad reading-writing factors expanded the model (as cited

On the relationship among Iranian EFL Leaners' Crystallized Intelligence, Fluid Intelligence,
Creativity, and their Performances on Multiple-Choice and Constructed-Response Items

165

in Flanagan, 2008).

Carroll (1993) proposes the three stratum theory in which a general factor (g) is considered in the stratum III, Gf-Gc domains as the broad abilities in stratum II, and the narrow abilities in stratum I. According to Flanagan (2008), there are also different ability classifications, for example Carroll places quantitative reasoning as a narrow ability subsumed by Gf rather than by quantitative knowledge in Cattell-Horn classification and views distinct broad reading/writing ability in Cattel-Horn as being subsumed by Gc. McGrew and Flanagan (1998) further combine the Cattell-Horn and Carroll's theories and create a single taxonomy. Finally they present an integrated model with some revisions (as cited in Flanagan, 2008).

CHC theory defines 16 broad cognitive abilities subsumed by 60 narrow abilities and the general intelligence is omitted from the model. These categories involve abilities named as reasoning, acquired knowledge, memory and efficiency, sensory, motor, and speed and efficiency which are suggested by Schneider and McGrew (2012). The reasoning ability involving Gf (fluid intelligence) is included out of the other groupings of intelligence in this study. Gc (crystallized intelligence) as one of the categories under the ability of acquired knowledge is too chosen for the purpose of this study as they have been investigated in the previous literature to find relationships with other variables such as creativity or reading comprehension. These two types of intelligence are also mentioned to be the two main components in the intelligence area (Cattell, 1987).

This study applies Gc definition as stated in CHC theory. Sternberg and Kaufman (2011, p.29) define Gc according to the research Horn and Cattell (1966) conducted: " Crystallized intelligence entails the application of consolidated knowledge typically acquired in academic settings". Fluid intelligence (Gf): It is defined by Sternberg and Kaufman (2011, p.29) as stated in accordance with the research Horn and Cattell (1966) conducted: " Fluid intelligence involves the processing of new information and the solution of novel types of problems.

Creativity

Rhodes (1961) states that all the definitions of creativity (C) share the same content that form four strands referred to as the four P's of creativity: person, process, press, and products (the four Ps of creative thinking). "Person" includes information about personality, intellect, temperament, physique, traits, habits, attitudes, self-concept, value systems, defense mechanisms, and behavior. One of the questions asked in this approach is about the coefficient of correlation between intelligence test scores and creativity (Rhodes, 1961).

The next term, "process", refers to motivation, perception, learning, thinking, and communicating. It questions the processes involved in creative thinking and according to Wallas (1926) who formulates Helmholtz's ideas, there are four stages or processes for creative thinking: reparation, incubation, inspiration, and verification. In the preparation process, individuals observe, listen, ask, read, collect, compare, contrast, analyze, and relate all kinds of objects and information. The incubation process being both conscious and unconscious is composed of thinking about parts and relationships, reasoning, and often a fallow period. Inspiration is the release of tension in order to be creative and the beginning of a new idea. Verification is a process of converting thoughts into articulated form or object (as cited in Rhodes, 1961).

"Press" as the relationship between an individual and environment, is a stage where creativity needs advanced stage of culture and a proper technical heritage to flourish. The communicated thought in the form of words, paint, or other materials define the term "idea" (Rhodes, 1961). Simonton (1990) suggests another term as "persuasion" which means that creative achievements change the way other people think. Runco (2006) proposes an additional category concerned with the "potential" in children and students which may or may not result in a creative behavior. Those people performing at very high levels still have the capacity to grow and make improvements which shows the existence of unused potential. The next part deals with the related theories that consider some of the P's mentioned earlier.

Cognitive theories of creativity put emphasis on creative process, being the cognitive mechanisms and person, being the individual differences in creative thought. One of the related theories is Guilford's distinction between convergent and divergent thinking processes which are both involved in creative thinking (Kaufman & Sternberg, 2010).

Convergent thinking process needs manipulation of the existing knowledge that leads to a single correct answer to a question. Accordingly, Cropley (2006) states that recalling from stored knowledge and working out from what is known exist where a ready-made answer is present. As noted by Cropley (2006), divergent thinking in contrast to convergent thinking applies available information to produce alternative answers by the use of making unexpected combinations, recognizing links among remote associates, and transforming information into unexpected forms. Chamorro-Premuzic and Reichenbacher (2008) notes that intelligence needs convergent thinking while creativity requires divergent thinking.

The Psychometric theories focus on the measurement and thus the product of creativity. It is concerned with the reliability and validity of the assessments. Discriminant validity makes distinction between tests of non-creative talents like traditional intelligence, IQ, convergent thinking.

Ryhammer and Brolin (1999) highlight that within the cognitive approach, creativity is considered to be in relation to intelligence (for example Guilford, 1967; Gardner, 1983) and it is also known as an aspect of intelligence (for example Binet & Henri, 1896). Guilford (1967) does not find a perfect correlation between creativity and intelligence. Moreover, the relations become weaker with the intelligence going up. This proves that there is no correlation between the two constructs for more than average IQ. This is known as "threshold hypothesis" which implies that creativity and intelligence are related to a certain level (Sligh, Conners, Roskos – ewoldsen,2005). The research on this theory has had diverse results throughout the literature (Jauk, Benedek, Dunst, & Neubauer, 2013; Kim, 2005; Park, Lubinski, &

Benbow, 2008; Runco & Albert, 1987).

According to Ryhammer and Brolin (1999), Guilford defines creativity as divergent thinking and performed factor analysis to find the cognitive capacities for creative thinking. The results indicate 120 components for Guilford's structure of intellect model (1967), important parts of which are considered to be under a new concept of divergent production (as cited in Ryhammer & Brolin, 1999). As stated by Runco and Acar (2012), based on empirical studies some of the components are supported and they are included in most divergent thinking tests: fluency, originality, flexibility, and elaboration.

Kaufman and Sternberg (2010, p.52) define fluency as "the number of responses to a given stimuli", originality as "as the uniqueness of responses to a given stimuli", flexibility as "the number and/or uniqueness of categories of responses to a given stimuli", and elaboration as "the extension of ideas within a specific category of responses to a given stimuli". Most recently, Auzmendi, Villa, and Abedi (1996) developed a multiple-choice measure of divergent thinking (Runco & Acar, 2012).

The Relationship between Intelligence and Creativity

Based on the investment theory of creativity, Sternberg and Lubart (1991) state that a combination of factors lead to creative production among which intelligence plays a role. The six resources of creativity according to this theory include: process of intelligence, knowledge, intellectual styles, personality, motivation, and environmental context. Traditionally, Guilford (1950) views creativity and intelligence as separate constructs; however, some studies report the correlations between these two to range from. 05 to .73 (as cited in Olive, 1973). Many studies have proved that intelligence and creativity have a correlation of a small value (e.g. Furnham & Bachtiar, 2008; Chamorro- Premuzic & Furnham, 2006; Getzels & Jackson, 1958). In a meta-analysis, kim (2005), confirms the fact that correlations between intelligence and creativity are not significant which means that these are two separate constructs. Apparently, studies on investigating the relations between intelligence and creativity lead to controversial results. Further research in this area, needs to be done.

In Cattell's view (1971), creativity is first determined by Gf and then personality factors. Kaufman (2009) also states that in the CHC theory, not exactly specified, creativity seems to be a component of Gf (as cited in Kaufman & Sternberg, 2010). Mendic (1962) suggests that individuals with high Gc and good cognitive organization, structure, or processes can combine new elements better (as cited in Batey & Furnham, 2006). Carroll (1993) considers intelligence and creativity to be different and at the same time, creativity to be dependent on general mental abilities: Speed of retrieval (in timed tests) to access stored knowledge (Gc), Gf to manipulate the retrieved information, and motor skills (as cited in Batey & Furnham, 2006). It is worth mentioning that in the system model of creativity, Csikszentmihalyi (2014) emphasizes that an existing body of knowledge, developed set of skills and abilities, mastery of the rules, sym-

bols, skills, values, and practices of a domain to accomplish something creative.

Factors affecting Test Performance

Bachman's (1990) four categories reveal the factors that affect test performance which include: communicative language ability, test method facets, personal attributes, and random factors. Test method facets according to Bachman (1990) have five major categories: Facets of the testing environment, facets of the test rubrics, facets of the input, Facets of the expected response, and the relationship between input and response. As mentioned previously, the difference between test methods (Mc & CR formats) have been done in the literature. Some other studies have considered the interaction among these formats and personal attributes of the test takers (Messick, 1987; Pishghadam & Tabataba'ian, 2011a & 2011b).

To the researchers' knowledge, there are a few studies on the relations between personal characteristics and diverse test method performance. This study, attempts to include Intelligence (Gf & Gc) and creativity as personal characteristics of the test takers. More elaborated in the next section, the relationship between reading comprehension performance in general and the two intelligence constructs will be discussed.

Intelligence and Reading Comprehension Ability

Throughout research, it has been proved that intelligence predicts learning in general (Chamorro-Premuzic, 2007). As mentioned by Oller (1981), language ability has a deep relationship with intelligence as a result of intelligence tests having a deep language component in them. Jensen (1998) maintains that intelligence is a predictor of academic attainment and that reading comprehension is highly g loaded.

Evans et al (2002) conducted a study in which participants with 2 to 95 range of age took basic comprehension skills and reading comprehension clusters from Woodcock and Mather (2001) along with CHC factor clusters from Woodcock-Johnson III Tests of Cognitive Abilities (WJ III COG). Analyzing the results, no significant correlation is reported between reading comprehension and Gf. On the other hand, Gc is the strongest predictor of reading skills and increased with age.

Taub and Benson (2013), administered WJ III COG, WJ III Tests of Achievement, and WJ III Diagnostic Supplement (Woodcock, McGrew, Mather, & Schrank, 2003) to college students. Gc, after visual-spatial thinking is the strongest predictor of reading comprehension and the g factor indirectly predicts RC.

Learning French as a second language, students in 4, 7 and 11 grade took a battery of language test which consisted of reading, language usage, listening comprehension and interpersonal communication. This study done by Genesee (1976), indicates a highly correlated relationship between reading as well as usage skills with intelligence. Ekstrand (1977) also has found a significant correlation between reading comprehension and intelligence for immigrant children receiving instruction in Swedish.

On the relationship among Iranian EFL Leaners' Crystallized Intelligence, Fluid Intelligence, Creativity, and their Performances on Multiple-Choice and Constructed-Response Items

167

In a study by Ghabanchi and Rastegar (2014), undergraduate students took Raven's Advanced Progressive Matrices as a measure of intelligence and TOEFL RC test. Correlational analysis indicates that intelligence (in this case, Gf) significantly and strongly predicts reading comprehension. Therefore, intelligence is considered as a facilitator of RC. Kassaian (2008) employed Raven's Progressive Matrices (a measure of Gf) and concludes that intelligence significantly correlates with language proficiency.

Creativity and Reading Comprehension Ability

Creativity is an effective factor on language learning and academic achievement. McCabe (1991) indicates a strong correlation between creativity and academic achievement (as cited in Meera & Remya, 2010). Sutrisno (2007) reports a significant positive correlation between creativity and academic achievement for tenth grade students. Meera and Remya (2010) used ANOVA which shows a significant main effect of creativity on English achievement for secondary students.

Various researches have also investigated the role of creativity in reading comprehension. Mousavie et al (2013) correlated RC and creativity (measured by Abedi's creativity questionnaire) which reveals a weak significant correlation in a sample of male university students. Applying UNIANOVA, high creativity group has higher RC scores than the low creativity group.

Reading Comprehension Test Format and Cognition

A multiple choice question requires the test taker to select an option from a set of possible answers. The open ended items on the other hand, involve the test takers to create their own responses which may ask the learners to produce one-word answers or more challenging essay type constructed responses. Explaining the process for answering MC questions, Messick (1987) states that this type of item emphasizes convergent thinking and memory. By contrast, answering CR questions involves divergent thinking. The effects of these cognitive processes on item response predict the relationships in this study.

Convergent thinking, according to Messick (1987) influences MC items performance. Intelligence needs convergent thinking (Chamorro-Premuzic & Reichenbacher, 2008) and Guilford (1967) consideres divergent thinking as a component of creativity. In CHC theory, intelligence consists of Gf and Gc. Gc as one of the intelligence components has a primary influence on MC questions. This acquired knowledge being mostly verbal, seems to be necessary for CR items that require individuals come up with their own formulated answers.

Messick (1987) notes that MC items require intelligence. Gf is another component of intelligence in CHC theory. Some researches have studied the relations between MC RC tests and Gf component reporting effective significant results (Kassaian, 2008; Qapanchi & Rastegar, 2014). Gf involves inductive, deductive, and quantitative reasoning needed for working with new processes and materials (Sternberg & Kauffman, 2011). If Gf correlates with MC RC test

scores, it is supposed to correlate with CR items scores as the test takers try to comprehend the texts and formulate the answers themselves. Therefore Gf and Gc both are effective in answering MC and CR items. Current literature has found correlations between MC reading comprehension tests and intelligence, Gf, and Gc and not with CR question types.

According to Pishghadam and Tabataba'ian (2011a), IQ and its subscales have impacts on the scores of some test formats. IQ has no relations with MC item scores; on the other hand, it predicted the performance on cloze test, c-test, and summary writing (Pishghadam, 2011a). The study includes subscales of intelligence such as information, Digit Span, Vocabulary Knowledge, Arithmetic,Comprehension, and Similarities. Among all the subscales, information with cloze test, Comprehension with cloze test and c-test, and Vocabulary knowledge with summary writing have significant relationships.

For the relationship between intelligence and creativity, investment theory postulates intelligence as one of the factors that leads to creative thinking (Sternberg & Lubart; 1991). Gf component predicts creativity (Cattell, 1987) and for incorporating Gc in forming creative thinking, there is an argument. Cropley (2006) and Csikszentmihalyi (2014) highlight that both convergent and divergent thinking are crucial for creativity and emphasize that relevant existing knowledge make the creative thought.

In conclusion, Gf, Gc, and creativity are all effective on MC and CR formats of RC performance. Crystallized and fluid intelligence are exogenous (independent) variables that affect MC items, CR items, and creativity as the endogenous (dependent) variables. Creativity has an exogenous trait in relation to MC and CR items: however, it becomes an endogenous variable predicted by Gf and Gc. Figure 1 demonstrated the hypothesized model.

This study aims at finding answers to the following research questions:
1. Is there any significant relationship between Iranian EFL learners' fluid intelligence and their performances on multiple-choice reading comprehension items?
2. Is there any significant relationship between Iranian EFL learners' crystallized intelligence and their performances on multiple-choice reading comprehension items?
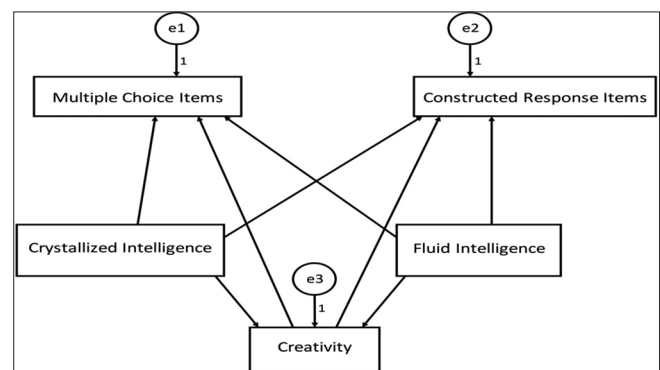

Figure 1. The schematic representation of the hypothesized for relationships among Crystallized intelligence, Fluid intelligence, Creativity, Multiple choice items and Constructed response items

3. Is there any significant relationship between Iranian EFL learners' creativity and their performances on multiple-choice reading comprehension items?

4. Is there any significant relationship between Iranian EFL learners' fluid intelligence and their performances on constructed response reading comprehension items?

5. Is there any significant relationship between Iranian EFL learners' crystallized intelligence and their performances on constructed response reading comprehension items?

6. Is there any significant relationship between Iranian EFL learners' creativity and their performances on constructed response reading comprehension items?

METHODOLOGY

Participants

Convenience type of nonprobability sampling (Griffee, 2012) is used and 300 participants took all the tests in the study, 120 of which are valid for the analysis. The number of the participants comprises 89 females and 31 males who are either Master of Arts (MA) students of English, engineering, biology, and law or English teachers of diverse institutes in Mashhad, Iran. Their age ranges from 23 to 40 and their first language is Farsi. They are considered as experienced test takers and upper intermediate to advanced English learners because students need to be at this level of knowledge to pass the entrance examination of universities for MA acceptance.

As stated by Stevens (1996), the minimum sample size for structural equation modeling needs to be 15 cases per observed variable. With 5 observed variables in this study, 75 participants are required. Schumacker and Lomax (2010) emphasize that sample size varies from 100, 200, 500 or more depending on model complexity and cross-validation requirements.

Instruments

Raven progressive matrices

Arthure and Day (1994) developed a short form of Raven Progressive Matrices (RPM) including 12 items which demonstrates the psychometric properties similar to the long form. It measures fluid intelligence of the participants using a set of figures and requires the examinee to choose a piece to complete the figure designs. According to Arthur and Day (1994) obtained items 1. 4, 8, 11, 15, 18, 21, 23, 25, 30, 31, make the short form of the original 36-item test. This study adopts only these 12 items to shorten the administration time. This test is culture fair as it has low level of culture loading. The reported Cronbach alpha for this short form test is.72 which has a correlation of .90 with the original form (Arthure & Day, 1994).

C-test

This study employs a C-test of two texts with 50 blanks. Tabataba'ian and Pishghadam (2011b) designed and used this C-test in their study under the guidance of C-Test construc-

tion principles proposed by Klein-Braley and Raatz (1984) and Klein-Braley (1985). This test has a reliability of.70. C-test is used as a measure of crystallized intelligence according to research results over the years. In a factor model by Vernon (1950), crystallized intelligence correlates with verbal abilities and it is measured by tests of vocabulary comprehension, reading comprehension, grammar, and syntax matching proverbs, etc. C-test has been used as a general language proficiency measure and in the literature it is also reported that C-tests can assess vocabulary knowledge (Chapelle & Abraham, 1990; Karimi, 2011). C-tests containing the desired cultural and factual knowledge can assess the knowledge component of Gc (Baghaei & Tabatabaee, 2015).

Creativity test

Abedi (1983) designed the creativity test (CT) based on Torrance (1974) test of creative thinking constructs (as cited in Abedi, 1993) and includes four subscales: Fluency (22 items), Flexibility (11 items), Originality (16 items), and Elaboration (11 items). Each item has three options ranging from the least to the most creative response. The Farsi version of the test is selected for this study developed by Abedi in 1983 (Auzmendi, Villa, & Abedi, 1996). Reliability for each factor is measured:.85 (fluency)., 84 (flexibility)., 82(originality), and .80 (Elaboration) as stated by Abedi (1993). Using Villa and Auzmendi (1996) creativity test as a criteria for concurrent validity in a latent variable modeling (Abedi, 2002), its subscales highly correlate with CT subscales (.80 for Fluency., 53 for Flexibility., 53 for Originality, and .64 for Elaboration).

Reading comprehension test

Master of Art (MA) university students and EFL teachers in different institutes participate in this study. The participants at this level of education in universities and as teachers in institutes are at upper intermediate or higher levels of language learning. The two reading texts are taken from the Test of English as a Foreign Language (TOEFL) which has 10 multiple choice questions with 0.52 of Cronbach's alpha (Pishghadam & Tabataba'ian, 2011b). The MC items are then turned into constructed response items by removing the item options which results into stem equivalent items of MC and CR for each text. An expert of the field judges the items in order to see if they represent characteristics of appropriate items (Rel:.67). The tests are piloted with 11 students of the target population to make sure that the tests are appropriate for their proficiency level.

In this study, two or three possible answers are suggested for some of the items in each RC test considering the correct choice in MC format as well. Two raters rate CR items of the two texts and the inter-rater reliability using Cohen's Kappa, indicated .87 of agreement for the first text and .84 for the second. Cohen interprets a value of 0.81–1.00 as almost perfect agreement for inter-rater reliability (McHugh,2012). The two raters resolve the issues regarding disagreements and come up with a final decision for rating.

To examine the reliability of the scales, Cronbach's alpha is used. The reliability coefficient is.95 for Crystallized Intelligence scale with 47 items, which shows that the scale has high reliability. The reliability of the Fluid Intelligence Scale is.82 indicating acceptable reliability (See Table 1).

Table 2 shows the reliability coefficient of creativity questionnaire and its subscales. The reliability coefficient is.91 for Creativity Scale with 51 items, which shows that the scale is highly reliable.

Procedure

Due to a high time restriction in university classes, the administration of the tests were organized in two parts and the second part was given to be done at home and brought to the class for the next session. Accordingly, measures of intelligence (crystallized & fluid) were taken in the class as the first part and the reading texts and the creativity test were selected into the second part of the tests. Clear and complete instructions were given at the beginning of each test. For each item of the creativity test, there is no correct answer and it simply elicits test takers' opinion about themselves. Therefore, any missing data for those items causes the removal of that participant from data analysis. For all the other tests, any blank answer sheet is considered unacceptable for data analysis. Accordingly, the study does not contain any missing data.

As the present study investigates the stem equivalent items, the design used in a similar study by Barati et al (2013) is applied by giving MC and CR item formats of different texts to the test takers. A participant takes the test of MC items based on text A, and CR based on text B. On the other hand, another student takes MC items of text B and CR items of text A. This design solves the problems caused by other methods of presenting MC and CR items. For example, in a within subject design in which the test takers take the two formats of test based on the same text, they may get familiar with the text after reading it for the second time(Shohamy, 1984). Another example is using a between subject design in which only

one format is given to each individual based on the same text. Therefore, this study, applies none of the two designs and intends to follow the method used by Barati et al (2013).

DATA ANALYSIS AND RESULTS

In this research, the relationship among observed variables in a "Path" model is investigated. Path analysis as a member of SEM family is a structural model for observed variables (Kline, 2011). It describes the directed dependencies among a set of variables in which only single indicators are employed for each of the variables in the causal model.

The variables in the model analyzed through SEM are dependent observed variables (Performance on MC & CR test formats) and independent latent variables (Gc & Gf). Creativity is placed as being either the dependent (influenced by Gc & Gf) and independent (affecting MC & CR test performances) variable in SEM. The relationship among these variables and the extent to which independent variables predict the dependent ones is studied by testing a model after providing all the tests with their factor models and appropriate items through the use of confirmatory factor analysis (CFA) or exploratory factor analysis (EFA). The hypothesized model goes through changes to finally present a model of the best fit. The direct or indirect effects of the latent variables on the observed variables are specified.

To examine the structural relations, the proposed model is tested using the Amos 22 statistical package. A number of fit indices are examined to evaluate the model fit: the chi-square magnitude which shouldn't be significant, Chi-square/df ratio which should be lower than 2 or 3, the normed fit index (NFI), the good fit index (GFI), and the comparative fit index (CFI) with the cut value greater than .90, and the Root Mean Square Error of Approximation (RMSEA) of about .06 or.07 (Schreiber, Nora, Stage, Barlow, & King, 2006).

Exploratory factor analysis (EFA) tests the instruments in this study for validity, the results of which suggest revisions on the three questionnaires for the items of less than .30 loading. Three items are removed from Gc (6,7,25) and Gf (1,2,5) tests. Creativity test also has items 3, 13, 14, 30, 31, 36, 41, 51, and 53 excluded after running EFA.

Table 3, indicates the path analysis of crystallized intelligence, fluid intelligence, creativity, multiple choice items and constructed response items. The table shows, the chi-square value (31.635), the chi-square/df ratio (15.817), RMSEA (.35), GFI (.91), and CFI (.75). Only one fit index (GFI) lies within the acceptable fit thresholds. Table 3 shows goodness of fit indices.

The standardized estimates on each path in the model indicate the strength of the causal relations among the variables in Figure 2. This standardized estimate is the standardized coefficient or beta coefficients (β) resulted from an analysis carried out on independent variables. It explains the predictive power of the independent variable and the effect size. The closer the magnitude to 1.0, the higher the correlation and the greater the predictive power of the variable is.

The results demonstrate that Crystallized intelligence is positive predictor of learners' performance on Multiple

Table 1. The reliability indices of crystallized intelligence and fluid intelligence questionnaires

Subscale	Number of items	Cronbach alpha
Crystallized intelligence	47 items	0.95
Fluid intelligence	9 items	0.82

Table 2. The reliability indices of creativity questionnaire and its subscales

Subscale	Number of items	Cronbach alpha
Fluency	19 items	0.82
Elaboration	9 items	0.81
Originality	14 items	0.80
Flexibility	9 items	0.81
Creativity scale	51 items	0.91

choice items (β = .44, P=.00) and their performance on Constructed response items (β = .41, P=.00). However, there is not a significant relationship between learners' Crystallized intelligence and their Creativity (β = .03, P = .733).

Moreover, the results show that Fluid intelligence is positive predictor of learners' performance on Constructed response items (β=.32, P=.00) and their creativity (β=.40, P=.00). However, the power of the path from Fluid intelligence to Multiple choice items (β=.05, P>.05) is not significant. Lastly, the findings indicate that the paths from creativity to Multiple choice items (β=.11, P>.05) and Constructed response items (β= -.01, P>.05) are not significant.

These non-significant paths are deleted from the model according to their t-values (<1.96). The first removed path is C-CR (t= -.064) followed by Gc-C (t=.341), Gf-MC (t=.680), and C-MC (t= 1.631). Table 4 shows the deleted paths and fitness changes. Figure 3 demonstrates the modified model.

Figure 3, indicates the modified version of the path analysis of crystallized intelligence, fluid intelligence, creativity, multiple choice items and constructed response items. As mentioned previously, non-significant paths are removed which results in having an endogenous creativity variable that have no direct relations with the MC and CR items. In other words, creativity is left out in the model. Therefore, on the basis of the previous research, the path direction is inverted in order to find the indirect relationship with CR items. Following investment theory of creativity (Sternberg & Lubart, 1991) as well as Carroll's view (1993) considering creativity to be dependent on intelligence, the model is generated. However, Kaufman (2009) notes that in CHC theory, not exactly specified, creativity is

a component of Gf (as cited in Kaufman & Sternberg, 2010). Accordingly, the path is inverted in the model.

Modification indices in Amos also suggest adding bidirectional relationships for Gc-e3 showing that what is not shared by Gf construct correlates with Gc. Another change suggested a bidirectional relation between e1 and e2 indicating a correlation between construct irrelevant characteristics of stem equivalent MC and CR items. Goodness of fit indices are indicated in Table 5.

As demonstrated by Table 5, the chi-square value (5.913), the chi-square/df ratio (1.478), RMSEA (.06), GFI (.98), and CFI (.98), lie within the acceptable fit thresholds. Therefore, the model is fit.

This investigation further aims at finding indirect relationship between Gf and MC items performance as well as creativity and MC items scores. Previously Cattell (1943; 1963) divides intelligence into separate factors as Gf and Gc (as cited in Sternberg & Kaufman, 2011). Cattell (1987/1971) in investment theory of intelligence notes that Gf provides the basis for Gc development (as cited in Rindermann, Flores-Mendoza, & Mansur-Alves, 2010). Therefore a path from Gf to Gc is added, so that indirect effects can be estimated (see Figure 4).

Figure 4 demonstrates the addition of a direct path from fluid intelligence to crystallized intelligence. This path helps finding the indirect relationship between Gf and MC items scores showing Gc as a mediating variable. In addition to the indirect path from creativity to CR items, an indirect path is

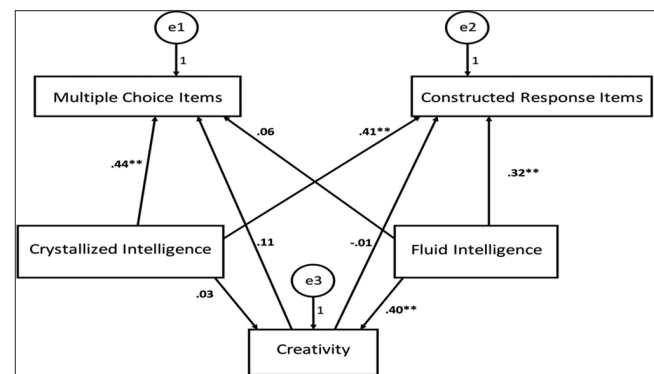**Table 3.** Goodness of fit indices

	X^2	df	X^2/df	GFI	CFI	RMSEA
Acceptable fit			<3	>0.90	>0.90	<0.08
Model fit	31.635	2	15.817	0.91	0.75	0.35

Table 4. Goodness of fit indices for the received models

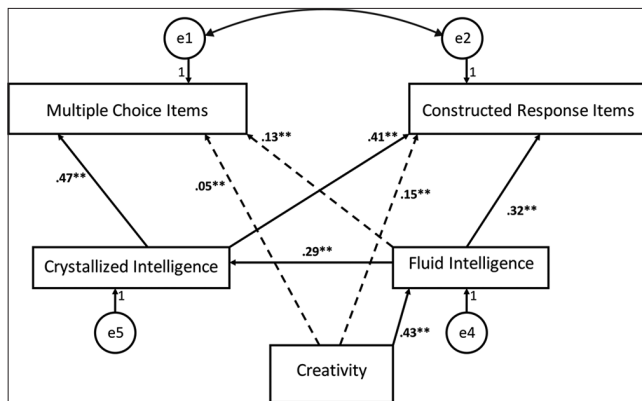|	X^2	df	X^2/df	GFI	CFI	RMSEA
Acceptable fit			<3	>0.90	>0.90	>0.08
Model: C-CR Path removed	31.639	3	10.546	0.91	0.76	0.28
Model: Gc-C Path removed	31.745	4	7.936	0.91	0.77	0.24
Model: Gf-MC Path removed	32.175	5	6.435	0.90	0.77	0.21
Model: C-MC Path removed	34.753	6	5.792	0.90	0.76	0.20

Figure 2. The Schematic Representation of the Relationships among Crystallized Intelligence, Fluid Intelligence, Creativity, Multiple Choice Items and Constructed Response Items
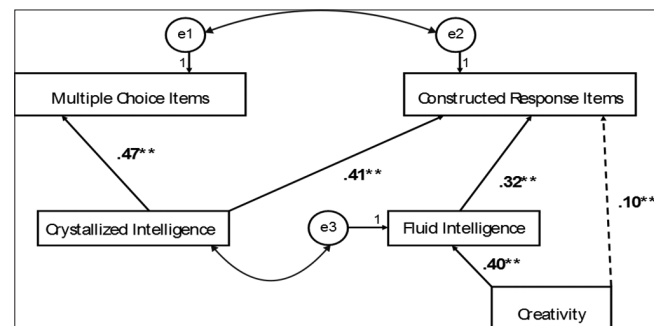
Figure 3. The Modified version of Path Analysis of Crystallized Intelligence, Fluid Intelligence, Creativity, Multiple Choice Items and Constructed Response Items

Table 5. Goodness of fit indices

	X^2	df	X^2/df	GFI	CFI	RMSEA
Acceptable fit			<3	>0.90	0>0.90	<0.08
Model	5.913	4	1.478	0.98	0.98	0.06

On the relationship among Iranian EFL Leaners' Crystallized Intelligence, Fluid Intelligence,
Creativity, and their Performances on Multiple-Choice and Constructed-Response Items

171

Figure 4. The Second Modified Version of Path Analysis of Crystallized Intelligence, Fluid Intelligence, Creativity, Multiple Choice Items and Constructed Response Items

Table 6. Goodness of fit indices

	X^2	df	X^2/df	GFI	CFI	RMSEA
Acceptable fit			<3	>0.90	>0.90	<0.08
Model	3.486	4	0.87	0.98	1.00	0.00

placed between creativity and MC items scores. This path indicates Gf and Gc are considered as mediating variables.

s demonstrated by Table 6, the chi-square value (3.486), the chi-square/df ratio (.871), RMSEA (.00), GFI (.98), and CFI (1.00), lie within the acceptable fit thresholds and they indicate improvement in this second modified version. Goodness of fit indices in Table 6, indicate improvement. The results are more acceptable considering fitness in compare to the previous model.

This study, assesses RC by conducting SEM. As indicated in Figure 2, the path from Fluid intelligence to Multiple choice items (β=.06, P>.05) is not significant in the first version of the model. Bootstrap in Amos, calculates the indirect effect of Gf on MC items performance which is significant (β=.13, P<.05). Figure 2 shows that, EFL learners' performances on multiple-choice reading comprehension items is influenced positively and significantly by their crystallized intelligence (β=.47, p < 0.05). The path from creativity to Multiple choice items (β = .11, P>.05) is not significant. However, bootstrap method as shown in figure 4, shows a small, but significant indirect effect of creativity on MC items performance (β=.05, P=003).

As it can be seen in figure 2, Fluid intelligence is a positive predictor of learners' performance on Constructed response items (β=.32, P=.00). Crystallized intelligence is a positive predictor of learners' performance on Constructed response items (β=.41, P=.00). The path from creativity to learners' performance on Constructed response items (β= -.01, P>.05) is not significant (figure 2). The indirect effect of creativity on CR items performance reports a significant impact (β=.15, P=008).

DISCUSSION AND CONCLUSION

The indirect relation between Gf and MC items performance is significant (β=.13, P<.05). The direct effect reported insig-

nificant result (β=.06,P>05) in the first hypothesized model. When Gf goes up by 1 standard deviation, MC goes up by 0.137 standard deviations. This indicates the same result obtained with multiple regression by Evans et al (2002) in which direct effect from Gf on reading comprehension is not reported. Administering the same tests, Benson (2008) does not report a meaningful relation between Gf and reading achievement. The study involves SEM analysis for kindergarten children as participants.

McGrew (1993) finds that Gf predicts reading achievement using multiple regression. Adult test takers participated in the study while using Woodcock-Johnson Psycho-Educational Battery-Revised (WJ-R). Kassaian (2008) measured intelligence with Raven Progressive Matrices (a measure of Gf) for which Pearson correlation results in a significant relationship between Gf and Reading comprehension. Another study reporting different results conducted by Motallebzadeh and Tabatabaee (2016). That study applied multiple regression analysis confirming Gf as a significant predictor of RC ability.

The results do not agree with Messick's (1987) view who believes that MC items require intelligence. Different results may be due to the statistical methods used in each study. For the current study, it seems that MC items does not activate reasoning abilities of the test takers to work with new situations. This is mostly due to the present information in the options of each item.

The current study applies SEM for EFL university students who took RPM test and MC reading comprehension. The results present similarities with those studies that used analysis methods other than Pearson correlation. SEM gives validity and reliability of observed scores explicitly taking measurement into account. Benson (2008) who did not find a significant relationship in SEM, claims that the extent to which cognitive processes affect reading comprehension performance will be either underestimated or overestimated when they are included in isolation. Accordingly, the model in the current study only counts Gc and Gf from CHC theory. There are other factors even out of CHC cognitive components that can predict MC reading performance.

Gc is a good predictor of MC reading comprehension performance (β=.47, p<.05). In a factor analysis by Cattell (1987), verbal ability loads on Gc factor. Gc is all about the application of acquired knowledge for answering questions specially the verbal ones (Sternberg & Kaufman, 2011). That is why it is supposed to have relations with reading comprehension test performance. The result is in line with Messick's (1987) view on the need for intelligence to answer MC questions. Gc is the access to stored knowledge and Gf is the manipulation of that knowledge to new materials (Batey&Furnham, 2006).

Previously, Gf and MC had no significant relationship and it seems that answering MC items does not require concept formation and fluid reasoning in addition to the manipulation of the knowledge test takers have. It can be explained by the presence of options in MC questions without which students needed to try to reorganize and make concepts to come up with their own answers. Therefore, Gc has the only significant effect on MC reading performance as it is expect-

ed because of the nature of Gc.

The results are in line with the work of Benson (2008) who finds a direct significant relation between Gc and reading comprehension and predicts the relation to go up as age increases. The current study has included graduate EFL university students who seem to have high knowledge specially in English. Evans et al (2002) conducted a study using WJ-III test with participants including adults as well. They report Gc as a predictor of reading ability which increases by age. Reynolds and Turek (2012) postulate a unidirectional relation between Gc and RC for participants to 15 years of age. Taub and Benson (2013) studied participants of an age range from 20 to 39 and report Gc as a strong predictor of RC after Visual-Spatial thinking. All these four studies applied SEM method and WJ-III tests.

Motallebzadeh and Tabatabaee (2016) does not find Gc as a meaningful predictor of RC performance using MC items. In this study, a Farsi measure of c-test does not correlate with RC ability of university students. The results may be due to the fact that they applied Farsi C-tests and correlated the scores with English RC tests. Using multiple regression method for analyzing the results can be another reason for the opposing results to the current study.

Significant indirect relationship is found between MC reading test performance and creativity (β=.05, P>.05) having Gf and Gc as mediator variables. Negligible direct effect is indicated in the first hypothesized model (β=.11, P>.05). Creative thinking gives an individual the ability to produce a novel and useful product (Plucker, Beghetto, & Dow, 2004). Creativity process according to Wallas (1926), involves preparation, incubation, inspiration, and verification as described in literature review. The first three stages seem to help RC process (as cited in Rhodes, 1961).

According to Rhodes' citation (1961), Wallas (1926) states that verification process is where the reader needs to answer some questions (articulating thoughts) and in this case, they try to formulate MC question answers. In accordance with what Messick (1987) notes, MC items activate convergent thinking (intelligence measures) and CR items don't. Test takers have to choose from multiple items rather than producing their own response.

As stated by Runco and Acar (2012), divergent thinking components including fluency, originality, flexibility, and elaboration, form the factors measured by the creativity test which is used in this study. These factors involve giving a number of unique responses to stimuli with extension of ideas within a specific category (Kaufman and Sternberg, 2010). It appears that MC items do not tap into the participants' creative thinking by limiting their options and creation process. Despite this fact, previous researches have found significant correlations between these two variables. Mousavie et al (2013) applying the same creativity test and Pearson correlation method, report a small significant relation between them. The statistical method of this study may have brought limitations on accepting the significance of relation between the variables. Having its primary effect on Gf development which in turn influences Gc, finally, creativity exerts its effects on MC items performance in small amount.

A significant relationship is found between test takers Gf and CR items scores (β=.32, P=.00). Gf involves inductive, deductive, and quantitative reasoning needed for working with new processes and materials (Sternberg & Kauffman, 2011). It seems that formulating their own answers, test takers activate Gf which involves " forming and recognizing concepts, perceiving relationships among patterns, drawing inferences, comprehending implications, problem solving, extrapolating, and reorganizing or transforming information" (Flanagan, 2008, para.17). Gf helps manipulating the retrieved information, and motor skills (Batey & Furnham, 2006) which are applied when RC items have their options removed and that is why MC items require Gc without Gf's participation.

Researches to the author's knowledge have not investigated the relationship between CR items scores and Gf. The reports of this study indicate a relationship between RC when items of the test are made in a CR format. This rejects Messick's view on how convergent thinking affects MC items and not the CR ones. According to Messick (1987), CR items can be designed so that they need convergent thinking ability; however, this study uses stem equivalent items resulting in CR items affected only.

The relationship between Gc and CR items of RC is significant (β=.41, P=.00). Slightly smaller than the relation between Gc and MC items performance which can be explained by the existence of relationship between both intelligence variables (Gc & Gf) and CR items performance. As mentioned previously, Gc and RC relationship significance is not negligible throughout the literature which used MC items. CR items require test takers to read and produce their own responses. Therefore, they are nothing less than MC items in requiring cognitive abilities. Another measure of convergent thinking (Gc) has meaningful relation with CR items scores which does not agree with Messick's view.

In accordance with the processes involved in reading comprehension listed by Grabe (2009), reading as an interactive process activates readers' background knowledge to comprehend what the writer intends to say. One of those processes called linguistic process, emphasizes graphemic-phonemic connections, word and structural phrases recognition along with a store of linguistic knowledge to comprehend the language of a text. Grabe (2009) notes that it is not possible to read without these processes working. In Gf-Gc theory, Cattell (1987) considers reading comprehension to be predicted by Gc as stored linguistic and general background knowledge. Kintsch (1998) highlights schemas as messy networks of knowledge that are activated by active information in working memory and that they are not fixed generalized structures (as cited in Grabe, 2009).

In CR items, there are no options for the test taker to read and pick from. The test taker must formulate and write accurate and acceptable response. Considering this, when Gc is supposed to be in relation with MC reading comprehension ability, its relation with CR reading comprehension performance is expected as well.

Despite the negligible direct relation between CR items and creativity, the indirect effect indicates a significant one

On the relationship among Iranian EFL Leaners' Crystallized Intelligence, Fluid Intelligence,
Creativity, and their Performances on Multiple-Choice and Constructed-Response Items

173

(β=.15, P=.008) mediating Gf variable. Creativity directly predicts Gf (β=.43, P=.00) followed by a direct path from Gf to CR items scores. Investment theory of Sternberg and Lubart (1991) suggests intelligence being as a subset of creativity. Cattell (1971) states that Gf determines creativity and as Carroll (1993) notes that they are different while creativity is still dependent on intelligence (having overlap).

The result of the current research suggests that there can be a bidirectional relation between Gf and creativity with Gf having a more powerful effect on CR format of RC performance. That is because CR items do not seem to need an answer that is purely original and out of the context of the reading text. There is only one correct answer written in any form of language which is first related to the effects of Gf. This does not conform to what Messick (1987) claims about this effect which is probably due to the nature of the CR items used in this study. They seem to be designed in a way that divergent thinking components have only an small indirect effect on.

To sum up, SEM found model fit indicating the relationships among all the variables in this study. The model underwent modification twice and the third model aimed at finding indirect relationships along with the direct ones. Direct relations were found in paths including Gc-MC, Gc-CR, Gf-CR. Indirect relations between variables including creativity-MC (Gf & Gc as mediators), Gf-MC (Gc mediating), creativity-CR (Gf mediating) demonstrated significant results.

Teachers and test designers seek valid tests that measure only the construct they aim to measure. The results of this study indicate MC items as a less affected RC format by cognitive variables such as Gc, Gf, and creativity. The only variable that predicts MC reading ability test directly is Gc. To a close degree, Gc also predicts CR reading comprehension test performance which shows that linguistic knowledge is primarily essential for RC performance. There are many cognitive processes working together while reading a text. Test designers ought to avoid testing RC ability through a heavy emphasis on linguistic resources or general knowledge specially when the students put effort in figuring out the unknown in a large amount.

According to Sternberg and Kaufman (2011), Horn considers Gf to be dependent on learning as much as Gc. This study recommends use of diverse techniques to motivate Gf ability in students such as working memory training and solving figural reasoning problem. Creativity, as concluded in this study, builds Gf and Gf predicts Gc ability. Improving creativity should be one of the goals in language teaching by providing students with creative tasks and help them produce novel linguistic outputs in classroom.

Test designers need to be careful with CR items to create those that include less measures of Gf and creativity or at least indirectly so that RC tests do not become biased. MC items also need to measure a more real life trait that does not limit test takers production while enjoying construct relevant factors included. Considering the results, teachers and test takers should know that Gc, Gf, and creativity all affect both formats; however, they had smaller and more indirect effect on MC items. They can use MC items as a less affected (by only the variables in this research) RC item format, but the real life nature of the CR items should be considered as well.

Further research may focus on the cognitive variables other than creativity and intelligence in a model to find the power of relations for each variable. Hypothesizing a bigger model can provide more precise results. Other test formats can replace with the variables in this study, such as Cloze test and summary writing. Considering many formats at a time will produce complete information on the relation between the use of a particular test and each different personal trait.

REFERENCES

Abedi, J., (1993). Creativity and new techniques to quantify it. *Psychological Research, 1 & 2, Tehran. Alghazo, I.M., 2006. Quality of Internet use by teachers in the United Arab Emirates. Education, 26*(4), 769-781

Abedi, J. (2002). A latent-Variable modeling approach to assessing reliability and validity of a creativity instrument. *Creativity Research Journal, 14 (2*), 267-276.

Ackerman, T.A., & Smith, P.L. (1988). A comparison of the information provided by essay, multiple-choice and free-response writing tests. *Applied Psychological Measurement*, 12, 117-128.

Arthur, W., & Day, D.V. (1994). Development of a Short form for the Raven Advanced Progressive Matrices Test. *Educational and Psychological Measurement, 54 (*2), pp. 394 – 403. doi:10.1177/0013164494054002013

Auzmendi,E. Villa, A. Abedi, J. (1996). Reliability and validity of a newly-constructed multiple-choice creativity instrument. *Creativity Research Journal, 9 (*1) (1996), pp. 89-95.

Bachman, L. (1990). *Fundamental considerations in language testing.* Oxford: Oxford University Press.

Baghaei, P., & Tabatabaee, M. (2015). The C-Test: An integrative measure of crystallized intelligence. *Journal of Intelligence, 3*, 46–58.

Barati, H., Ravand, H., & Ghasemi, V. (2013). Investigating Relationships among Test Takers' Characteristics and Response Formats in a Reading Comprehension Test: A Structural Equation Modeling Approach. *Iranian Journal of Language Testing, 3 (*2), October 2013.

Batey, M., & Furnham, A. (2006). Creativity, intelligence, and Personality: A Critical Review of the Scattered Literature. *Genetic, Social, and General Psychology Monographs, 132 (*4), 355−429.

Benson, N. (2008). Cattell–Horn–Carroll Cognitive Abilities and Reading Achievement. *Journal of Psychoeducational Assessment, 26*, 27-41.

Birenbaum, M. Pinku, P. (1997). Effects of test anxiety, information organization, & testing situation on performance on two test formats. *Contemporary Educational Psychology,22*, 23–38.

Boden, M. A. (1998). Creativity & artificial intelligence. *Artificial Intelligence, 103*, 347-356.

Carroll, J. B. (1984). Raymond B. Cattell's contributions to the theory of cognitive abilities. *Multivariate Behavioral Research, 19*, 300-306.

Carroll, J. B. (1993). *Human cognitive abilities: A Survey of Factor Analytic Studies*. New York: Cambridge University Press.

Cattell, R. B. (1971). *Abilities, their structure, growth & action*. Boston, MA: Houghton Mifflin.

Cattell, R. B. (1987). *Intelligence: Its Structure, Growth, and Action*. New York: Elsevier Science.

Chamorro-Premuzic, T. (2007). *Personality and individual differences*. Oxford: Blackwell

Chamorro-Premuzic, T., & Reichenbacher, L. (2008). Effects of Personality and Threat of Evaluation on Divergent and Convergent Thinking. *Journal of Research in Personality, 42*, 1095–1101.

Chamorro-Premuzic, T., & Furnham, A. (2006). Intellectual competence and the intelligent personality: A third way in differential psychology. *Review of General Psychology,10*, 251–267.

Chapelle, C., & Abraham, R. (1990). Cloze Method: " What difference does it make?". *Language Testing, 7 (*2), 121-145.

Cropley, A. J. (2006). In Praise of Convergent Thinking. *Creativity Research Journal, 18*, 391–404.

Csikszentmihalyi, M. (1996). *Creativity: Flow & the psychology of discovery & invention.* NewYork: Harper Perennial.

Csikszentmihalyi, M. (2014). *The Systems Model of Creativity: The Collected Works of Mihaly Csikszentmihalyi*. Dordrecht: Springer, 2014. ISBN 978-94-017-9084-0

Dornyei, Z. (2005). *The psychology of the language learner*. US: Lawrence Erlbaum Associates.

Ekstrand, L. H. (1977). Social and Individual Frame Factors in Second Language Learning. *In T. Skutnabb-Kangas (Ed.), Papers from the first Nordic conference on Bilingualism*, 40– 61.

Evans, J. J. Floyd, R. G., & McGrew, K. S. (2002). The Relations between Measures of Cattell-Horn- Carroll (CHC) Cognitive Abilities and Reading Achievement During Childhood and Adolescence. *School Psychology Review, 31 (*2), 246-262.

Flanagan, D. P. (2008). The Cattell-Horn-Carroll Theory of Cognitive Abilities. *Encyclopedia of Special Education.* 381:368–382.

Furnham, A., & Bachtiar, V. (2008). Personality and intelligence as predictors of creativity. *Personality and Individual Differences, 45 (*7), 613-617.

Gamer, M. Engelhard Jr., G. (1999). Gender differences in performance on Multiple-Choice & Constructed-Response mathematics items. *Applied Measurement in Education, 12 (*1), 29-51. doi:10.1207/s15324818ame1201_3

Genesee, F. (1976). The Role of Intelligence In Second Language Learning. *Language Learning, 26*, 267–280.

Getzels, J. W., & Jackson, P. W. (1958). The Meaning of "Giftedness"—An Examination of an Expanding Concept. *Phi Delta Kappan, 40*, 75–77.

Ghabanchi, Z., & Rastegar, R. (2014). The Correlation of IQ and Emotional Intelligence with Reading Comprehension. *An International Online Journal, 14 (*2), 135-144

Gould, K.L. (1972). Relationships of Creativity, Reading Comprehension, Intelligence, & Response to a Literature Selection for Fourth Grade Inner-City Children (Doctoral dissertation). Retrieved from http://eric.ed.gov/?id=ED072434

Grabe, W. (2009). *Reading in a second language: Moving from theory to practice*. New York: Cambridge University Press.

Griffee, D.T (2012). *An Introduction to Second Language Research Methods: Design and Data*. TESL-EJ Publications.

Guilford, J. P. (1967). *The Nature of Human Intelligence*. New York, NY, US: McGraw-Hill.

Hansen, L. (1984). Field Dependence-Independence & language testing: Evidence from six pacific island cultures. *TESOL Quarterly, 18*, 311–324. doi:10.2307/3586696

Hebb, D. O. (1949). *The Organization of Behavior: A Neuropsychological Theory.* New York: Wiley.

Heim, A. W., & Watts, K. P. (1966). An Experiment on Multiple-Choice Versus Open-Ended Answering in A Vocabulary Test. *British Journal of Educational Psychology, 37,*339-346.4

In'nami, Y., & Koizumi,R. (2009). A Meta-analysis of Test Format Effects on Reading and Listening Test Performance: Focus on Multiple-choice and Open-ended Formats. *Language Testing, 26 (*2), 219–244.

Jauk, E., Benedek, M., Dunst, B., & Neubauer, A. C. (2013). The relationship between intelligence and creativity: New Support for the Threshold Hypothesis by Means of Empirical Breakpoint Detection. *Intelligence, 41*, 212–221.

Jensen, A. R. (1998). *The g factor: The science of mental ability*. Westport, CT: Praeger.

Karimi, N. (2011). C-test and Vocabulary Knowledge. *Language Testing in Asia, 1 (*7). 7-38.

Kassaeian, Z., & Kassaeian, N. (2008). The Relationship between EFL Learners' Proficiency, Intelligence, and Creativity. *Journal of Literature and Language, 27*, 72-82.

Kastner,M., & Stangl, B. (2011). 'Multiple Choice and Constructed Response Tests: Do Test Format and Scoring Matter?'. *Procedia Social and Behavioral Sciences, 12*, 263–273.

Kaufman, J. C., & Sternberg, R. J. (Eds). (2010). *Cambridge Handbook of Creativity.* New York: Cambridge University Press.

Katz, I. R., Bennett, R. E., & Berger, A. E. (2000). Effects of Response Format on Difficulty of SAT- Mathematics Items: It's not the Strategy. *Journal of Educational Measurement, 37*, 39-57.

Kennedy, P., & Walstad, W. B. (1997). Combining Multiple-Choice and Constructed-Response Test Scores: An Economist's View. *Applied Measurement in Education, 10 (*4), 359-375.

Kim, K. H. (2005). "Can only Intelligent People be Creative?". *Journal of Secondary Gifted Education, 16*, 57– 66.

Kline, P. (1991). *Intelligence: The Psychometric View*. London: Routledge.

Kline, R. B. (2011). *Principles and Practices of Structural Equation Modeling (3rd ed.)*. New York: The Guilford Press.

On the relationship among Iranian EFL Leaners' Crystallized Intelligence, Fluid Intelligence, Creativity, and their Performances on Multiple-Choice and Constructed-Response Items

175

Kordjazi, Z. Ghonsooly, B. (2015). Brain dominance & test format: A Case of vocabulary. *Journal of Language Teaching and Research, 6 (*3), pp. 695-703. doi: 10.17507/jltr.0603.29

Kunnan, A.J. (2004). Test fairness. In M. Milanovic & C. Weir (Eds.). *European language testing in a global context: Proceeding of the ALTE Barcelona conference*, 27- 50.

Lohnes, P. R. Gray, M. M. (1972). Intelligence & the co-operative reading studies. *Reading Research Quarterly, 7 (*3), 466-476. doi: 10.2307/746994

Martinez, M. E. (1991). A Comparison of Multiple-Choice and Constructed Figural Response Items. *Journal of Educational Measurement Summer 1991, 28 (*2), 131-145.

McGrew, K. S. (1993). The Relationship between the WJ-R Gf-Gc Cognitive Clusters and Reading Achievement Across the Life-Span. *Journal of Psychoeducational Assessment, Monograph Series WJ-R*, 39-53.

McHugh M. (2012). Inter-rater reliability: The Kappa statistic. *Biochemia Medica, 22 (*3), 276-282.

Meera, KP., & Remya, P (2010). Effect of Extensive Reading and Creativity on Achievement in English Language. *E-journal of All India Association for Educational Research, 22 (*1),16-22.

Messick, S. (1987). VALIDITY. *ETS Research Report Series, 1987 (*2), i–208.

Messick, S. (1993). Trait equivalence as construct validity of score interpretation across multiple methods of measurement. *Construction versus choice in cognitive measurement*, 61-74.

Motallebzadeh, Kh., & Tabatabaee, M. (2016). The Relationship between EFL Learners' ReadingComprehension Ability and their Fluid Intelligence, Crystallized Intelligence, and Processing Speed. *Cogent Education (2016), 3*, 1-8.

Mousavi, M., Maghsoudi, M., & Yarahmadi, M. (2013). The Impact of Creativity on Iranian EFL Learner's Reading Comprehension Ability. *Indian Journal of Fundamental and Applied Life Sciences, 3*, 437-445.

Olive, H. (1973). The Relationship of Divergent Thinking to Intelligence, Social Class, and Achievement in High-School Students. *The Journal of Genetic Psychology: Research and Theory on Human Development, 122 (*2), 179-186.

Oller J. W. (1981). "Language as intelligence?". *Language Learning, 31*, 465-492.

Park, G., Lubinski, D., & Benbow, C. P. (2008). Ability differences among people who have commensurate degrees matter for scientific creativity. *Psychological Science, 19*, 957–961.

Pishghadam, R. Khodadady, E. Zabihi, R. (2011). Learner creativity in foreign language achievement. *European Journal of Educational Studies, 3 (*3), 465-472.

Pishghadam, R., & Tabataba'ian, M. (2011a). IQ and Test Format: A Study into Test Fairness. *Iranian Journal of Language Testing, 1 (*1), 17-29.

Pishghadam, R., & Tabataba'ian, M. (2011b). ' Emotional Intelligence: Can It Be a Predictor of Performance on Different Test Formats?'. *International Journal of Linguistics, 3 (*1), 1-21.

Plucker, J. A., Beghetto, R. A., & Dow, G. T. (2004). Why isn't creativity more important to educational psychologists? Potentials, pitfalls, and future directions in creativity research. *Educational Psychologist, 39*, 83–96.

Reynolds, R. M. Turek, J. J (2012). A dynamic developmental link between verbal Comprehension-Knowledge (Gc) and Reading Comprehension: Verbal Comprehension-Knowledge drives positive change in reading comprehension. *Journal of School Psychology, 50 (*6), 841–863. doi:10.1016/j.jsp.2012.07.002.

Rhodes, M. (1961). An analysis of creativity. *Phi Delta Kappan, 42*, 205-210.

Rindermann, H., Flores-Mendoza, C., & Mansur-Alves, M. (2010). Reciprocal effects between fluid and crystallized intelligence and their dependence on parents' socioeconomic status and education. *Learning and Individual Differences, 20*, 544–548.

Rodriguez,M. (2003), Construct Equivalence of Multiple-Choice and Constructed-Response Items: A Random Effects Synthesis of Correlations. *Journal of Educational Measurement Summer 2003, 40 (*2), 163-184.

Runco, M. A.(2006). Personal creativity and the uncertainty of creative potential. In J. Gonzelez (Ed.), *Proceedings of I Simposio International Sobre Altas Capacidades Intellectuales*. Las Palmas, Spain: Directorate for Educational Planning and Innovation.

Runco, M. A., & Acar, S. (2012). Divergent Thinking as an Indicator of Creative Potential. *Creativity Research Journal, 24 (*1), 1–10.

Runco M. A., & Albert R. S. (1987). The Threshold Hypothesis Regarding Creativity and Intelligence: an Empirical Test with Gifted and Non-Gifted Children. *Creat. Child Adult Q, 11*, 212–218

Ryhammar, L., & Brolin, C. (1999). Creativity Research: Historical Considerations and Main Lines of Development. *Scandinavian Journal of Educational Research, 43 (*3),259-273.

Schreiber, J. B., Nora, A., Stage, F. K., Barlow, E. A., & King, J. (2006). Reporting Structural Equation Modeling and Confirmatory Factor Analysis Results: A Review. *The Journal of Educational Research, 99 (*6), 323-337.

Schneider, W. J., & McGrew, K. (2012). The Cattell-Horn-Carroll model of intelligence. In, D. Flanagan & P. Harrison (Eds.), *Contemporary Intellectual Assessment: Theories, Tests, and Issues (3*rd *ed.)*, 99- 144. New York: Guilford.

Schumacker, R. E., & Lomax, R. G. (2010*). A beginner's guide to structural equation modeling (3*rd *ed.)*. New York: Routledge.

Shohamy, E. (1984). Does the testing method make a difference? The case of reading. *Language Testing, 1*, 147–170.

Simonton, D. K. (1990). History, Chemistry, Psychology, and Genius: An Intellectual Autobiography of Historiometry. *In M. A. Runco & R. S. Albert (Eds.), Theories of creativity*, 92–115. Newbury Park, CA: Sage.

Sligh, A. C., Conners, F. A., & Roskos – ewoldsen, B. (2005). Relation of Creativity to Fluid and Crystallized Intelligence. *Second Quarter, 39 (*2), 123-136.

Sternberg, R. J., & Kaufman, S. B. (2011). *The Cambridge Handbook of Intelligence*. Cambridge: Cambridge University Press.

Sternberg R.J., & Lubart T.I.(1991). An Investment Theory of Creativity and Its Development. *Human Development, 34*, 1–31.

Stevens, J. (1996). *Applied multivariate statistics for the social sciences*. Mahwah, NJ: Lawrence Erlbaum Publishers.

Sutrisno (2007). *Students' Creativity and Its Relation to English Learning Achievement: A Case Study of the Tenth Grade Students of SMA N 1 Brebes in the academic year 2006/2007* (dissertation). Semarang State University.

Taub, G. E., & Benson, N. (2013). Identifying the Effects of Specific CHC Factors on College Students' Reading Comprehension. *International Journal for the Scholarship of Teaching and Learning, 7 (*2), 1-13.

Vernon, P. E. (1950). *The structure of human abilities*. London, England: Methuen.