

Vulnerability of Multiple Intelligences Assessment Instrumentation to Gender Variance in EFL Context

¹Fariba Hooshyar

²Akbar Afghari*

IJEAP- 1808-1265

³Bahram Hadian

Abstract

Robust gender effect has been reported in Multiple Intelligences (MI)-EFL studies across cultures. Nevertheless, research results vary in effect size, Intelligence type, overall estimation etc. The variance is mainly explained in terms of socio-cultural factors. However, although manifold MI inventories and scales are used in studies, the role of the MI-test type in creating variance has not been studied thus far. This article sets out to comparatively examine two widely used MI tests, MMII (McKenzie's Multiple Intelligences Inventory) and the MIDAS (Multiple Intelligences Developmental Assessment Scales) in terms of gender effect. Kolmogorov-Smirnov test, paired-samples and independent-samples t-tests were applied to the data. The findings indicated non-normal distribution of MMII and normal distribution of MIDAS data. No male-gender effect was found in the MIDAS test results while in MMII results significant male effect was observed in Logical/Mathematical intelligence. While both tests showed significant female gender effects, the Intelligence types indicating the effect varied. In MMII, Spatial, Musical, Kinesthetic and Linguistic Intelligences signified female effect and in the MIDAS, Spatial, Linguistic and Musical. The study provides new insights into how MI assessment instrumentation can have significant impacts on MI-EFL research findings and MI-differentiated EFL teaching.

Keywords: Gender effect, Multiple Intelligences, MI test Instrumentation

1. Introduction

Among the theories that have had impacts on post-positivist education philosophy and practice is Gardner's (1983) MI theory which emphasizes multidimensionality of human intelligence as opposed to the positivist definition of general intelligence or g factor as a unitary reductionist concept (Spearman, 1904). MI theory has extended the horizons of (EFL) educational research and practice in terms of learners' abilities and intelligences spectrum and their perceptions of their MI. Research in MI-EFL fields indicates gender effect in learners' self-estimation of MI (Cheng, Hou, Hou & Chung, 2010; Loori, 2005; Zare-Ee, Mohd Don, Knowles & Tohidian, 2015). However, the results differ significantly regarding the effect size and the type of the intelligence/s that signify gender effect. The overall profiles also differ in favor of either sex. There are also differences in participants' general ratings of intelligences within male/female groups.

¹ PhD Candidate, hooshyar_fariba@yahoo.com; Department of English Language, Isfahan (Khorasgan) Branch, Islamic Azad University, Isfahan, Iran

² Associate Professor, afghary@yahoo.com; Department of English Language, Isfahan (Khorasgan) Branch, Islamic Azad University, Isfahan, Iran

³ Assistant Professor, bah.hadian@Khuisf.ac.ir; Department of English Language, Isfahan (Khorasgan) Branch, Islamic Azad University, Isfahan, Iran

Various explanations have been suggested for the variance in gender-effect which basically relate to socio-cultural diversity of the environments in which individuals' perceptions of their cognitive abilities and interests are formed (Bennet 2000, Bowles, 2012; Furnham, Shahidi & Baluch, 2002; Wigfield & Cambria, 2010). However, MI assessment instrumentation may also, in some measure, justify the discrepancy of MI-EFL research results. After a brief review of literature on Gardner's MI theory, MI assessment, and MI-gender effect, the present research comparatively studies the relationship between gender effect and MI instrumentation based on the findings obtained from an experiment conducted on Iranian students applying two widely used MI assessment inventories, the MIDAS (Multiple Intelligences Developmental Assessment Scales) and McKenzie Multiple Intelligences Inventory (MMII).

This article will provide EFL educationalists and investigators with a deeper and more accurate understanding of MI-fair assessment and will help differentiated, learner-oriented EFL education be more Intelligence-fair and efficient. Based on the significant differences between the results obtained from MMII and those from the MIDAS, it is suggested that MI assessment instrumentation can have a significant effect on the validity of MI-related research.

2. Literature Review

2.1. MI Theory

Defining intelligence as “the ability to solve problems, or to create products, that are valued within one or more cultural settings” (p. xxviii), Gardner expands the concept of intelligence to areas that the positivists psychometrists would either ignore or see as facets of g factor (Sternberg, 2015). He bases his theory on evidence obtained from neurological, evolutionary, anthropological and cross-cultural studies (Gardner, 1983, p. xii). Gardner (Gardner & Hatch, 1989) formulated a list of eight intelligences defined as follows:

Logical-mathematical: Sensitivity to, and capacity to discern, logical or numerical patterns; ability to handle long chains of reasoning.

Linguistic: Sensitivity to the sounds, rhythms, and meanings of words; sensitivity to the different functions of language.

Musical: Abilities to produce and appreciate rhythm, pitch, and timbre; appreciation of the forms of musical expressiveness.

Spatial: Capacities to perceive the visual-spatial world accurately and to perform transformations on one's initial perceptions.

Bodily-kinesthetic: Abilities to control one's body movements and to handle objects skillfully.

Interpersonal: Capacities to discern and respond appropriately to the moods, temperaments, motivations, and desires of other people.

Intrapersonal: Access to one's own feelings and to discriminate among them and draw upon them to guide behavior; knowledge of one's own strengths, weaknesses, desires, and intelligences. (p.6)

Naturalist: “Expertise in the recognition and classification of the numerous species—the flora and fauna—of an individual's environment” (Armstrong, 2009, p. 7).

Gardner (2011, p. 4) argues that intellect is “distinctly pluralistic” and human beings possess all the intelligence types in varying degrees. This makes their personal “cognitive profile” (Gardner, 2017, p.2) a unique composition of strengths and weaknesses. He further states that his choice of the word “intelligences” (p.6) was deliberate but expresses his willingness to refer to MI as “abilities, talents or mental skills” provided verbal and logical/mathematical capacities are considered talents too. In line with Bronfenbrenner (1986, 1977), Gardner (1999) emphasizes the impact of cultural and ecological-contextual factors on the development of intelligence. As such, MI are not static products but interactive

and dynamic operations with each intelligence entailing their own “set of psychological processes” (Gardner & Hatch, 1989, p.6). Gardner (1998) criticizes reductionist intelligence theories for casting socio-cultural and environmental considerations aside.

2.2. *MI Assessment*

Assessing MI has always been a great challenge. Making a distinction between testing and assessment, Gardner (1999, p.138) expresses concerns about “abusing” MI theory by inappropriate MI testing methods instead of adopting multi-dimensional, contextualized assessment. Gardner (2004) casts doubts on the reliability and ecological validity of extant MI tests within the existing paradigms:

There are several batteries of short tests that claim to measure the intelligences, but these tend to be strongly linguistic and often confound an interest in an intelligence with a demonstrated skill in that intelligence. These tests simply multiply by seven or eight the sins of original intelligence tests (Gardner, 1999, p. 138)

Gardner (2016) emphasizes that MI assessment must be done in a triangulation framework using more than one source of information to avoid bias and unfairness. He calls for “intelligence-fair assessment” in a “contextually appropriate” setting (Gardner, 1993, p.22). Such “intelligence-fair assessment” seeks to “respect the different modes of thinking and performance that distinguish each intelligence” (Gardner & Hatch, 1989, p.6). However, aside from Project Zero, Gardner (2013) has never got seriously involved in MI assessments.

In a similar vein, Shearer (2001) warns that wrong configuration of learners’ Intelligences will result in irrelevant MI-based educational programs and hence their failure. Appreciating Gardner’s theory enthusiastically, Sternberg (1991, p.257) fears “bad” MI tests which may create a “psychometric nightmare”. Baum, Viens and Slatin (2007) posit that MI cannot be assessed in a linear fashion using article-and-pencil tests since MI assessment is multifocal and “tightly contextualized” (p.43). Armstrong (2009) maintains that “no test can accurately determine the nature or quality of a person’s intelligences” (p.21) and that there is no “mega-test” (p. 33) on the market that can give a complete account of a persons’ MI configuration. In line with Armstrong, Tirri, Nokelainen and Komulainen (2013) express concerns about the soundness of MI assessment as a means of operationalizing MI theory. Chen (2005) and Tomlinson (2001) also call for sound and reliable MI assessment in EFL education. Tomlinson emphasizes that “differentiated instruction is rooted in assessment” (p. 4). McMahon and Rose (2004) noticed that MI tests lack validity: they do not measure what they are claimed to and do not allow predictions.

Among the widely used MI tests are MMII (McKenzie, 1999/2017) and the MIDAS (Shearer, 1994) and their different adaptations. McKenzie developed his inventory in 1999. MMII is an agree-disagree questionnaire with nine subscales, representing the nine types of intelligences (i.e. Naturalist, Musical, Logical, Existential, Interpersonal, Kinesthetic, Verbal, Intrapersonal and Visual), each with 10 questions. Although McKenzie stressed that his questionnaire is not a test but “a snapshot in time of an individual's perceived MI preferences” (p.1), MMII and its adaptations are extensively used for research and educational purposes (Hoerr, 2000; McKenzie, 2009; Nicholson-Nelson, 1998).

DeVellis (1991) found a reliability coefficient of 0.86 for MMII, which he considers acceptable for many studies and evaluations. According to Al-Balhan (2006), MMII had a reliability coefficient of 0.85. The overall internal consistency of the subscales of MMII was reported to be about 0.85 - 0.90 (Al-Balhan, 2006; Hajhashemi & Bee Eng, 2010; Hajhashemi, Cook & Bee Eng, 2012; Razmjoo, 2008). However, MMII validity is not believed to be as high. O’Grady (2017) criticizes MMII for lacking the capacity to assess what it claims to. His findings indicated “no evidence of any research supporting the validity of the measure” (para. 31), nor of its predictive validity. He seriously questions the findings of

the research which applied MMII. Similarly, Armstrong (2009) suggests using standard tests such as the MIDAS rather than the non-standard MMII.

The MIDAS was developed by Shearer (2012) through a period of more than six years starting in 1987. It is a five-point Likert scale self-discovery questionnaire which, Shearer (2017) claims, is “significantly different from the brief MI checklists” (para.2). The test includes 8 main sections, 119 descriptive questions and 23 qualitative subscales. There are five versions of the MIDAS for different age groups including the MIDAS-ADULTS, MIDAS_ TEENS, and three different versions of MIDAS-KIDS.

Shearer (1997, 2001, 2009, 2013, & 2017) has constantly been making attempts to validate the MIDAS. The alpha reliability coefficient for the test scales are reported to “range from .79 - .89” (Shearer, 2012, p.137). According to Shearer and Jones (1994, p.12), “accumulated evidence supports its [MIDAS] validity as a tool to gather useful and meaningful data regarding an individual’s profile in seven areas of everyday intellectual functioning”. The MIDAS is claimed to provide both a qualitative and quantitative configuration of a person’s MI (Gardner, 2016; Oliver, 2010; Shearer, 1997, 2012). From among various MI tests, Gardner (2016, p. 9) appears to favor the MIDAS as a standard test positing that “performance on the MIDAS correlates with abilities in certain areas, as determined by other indexes”.

Although both tests are claimed to assess MI, there are noticeable differences between MMII and the MIDAS. The subscales of the MIDAS (Shearer, 1994, p.3) and MMII (McKenzie, 2017/1999, section 2) vary greatly in terms of item type, content, the number of items, the concepts each test considers as components of intelligences and the proportional weight given to each component. Answer choices of each test also vary significantly. While the MIDAS is a six-scale test (A= No, B= A little, C= Fair, D= Good, E= Excellent, F= I don't know), MMII is an inventory with only two answer choices (A. Agree, B. Disagree). Moreover, the content of answer choices in the MIDAS are not fixed across items whereas the answer choices in MMII are fixed.

2.3. Gender Effect in MI Assessment

During the past 30 years, research into the learners’ perceptions of their MI has revealed robust gender effects. Tirri and Nokelainen, (2008) found men rated their Logical/Mathematical intelligence significantly higher than women while Verbal/Linguistic intelligence was ranked higher by women. The overall males’ MI self-estimates were higher than females’ in Tirri and Nokelainen’s study. Examining research on MI self-estimation across twelve countries (Australia, Austria, Brazil, France, Iran, Israel, Malaysia, South Africa, Spain, Turkey, UK and US), von Stumm, Chamorro-Premuzic and Furnham (2009, p. 429) basically observed significant sex effects in favor of males. They noticed Logical/Mathematical and Visual/Spatial intelligences, which they call “male-normative” (p. 439), were rated the highest by men in almost all their studies. However, Kaur’s (2014) findings indicated significant gender differences in Natural, Logical-Mathematical, Musical and Bodily-Kinesthetic intelligences in women’s group while no significant difference was observed in men’s group. Concerning within-group ratings, females rated their Linguistic, Visual and Interpersonal intelligences the highest while males rated their Natural, Musical and Logical intelligences the highest. Table 1 gives an account of a number of studies on gender effect in MI self-estimation processes.

Table 1: A Comparison of Gender Effect in MI Research Results

Research by	MI Test Type	Gender Effect
		Male Female
Bennet (2000)	Not Given	Kin. Ling., Spat.
Workman (2003)	Not Given	
Furnham, et al. (2002)	FCMIQ	Inter.
Furnham and Chamorro-Premuzic (2005)	FCMIQ	Math-Spat.-Music
Yuen and Furnham (2005)	FCMIQ	Log./Math-Kin.- Spat. - Music Exist.-Intra. Nat.
Tirri and Nokelainen, (2008)	MIPQIII	Log. /Math. Ling.
Bowels (2012)	Not Given	Kin.-Ling.-Inter.
Al-Onizat (2014)	MIDAS	Nat. Music-Spat.
Kaur (2014)	RIMI	Nat.-Log./Math.- Music-Kin.
Joshi (2016)	RMMIQ	Nat. Music
Kang and Furnham (2016)	FCMIQ	Log./math.-Spat.

Note. RIMI=Roger’s Indicator of Multiple Intelligences; FCMIQ: Furnham Constructed MI Questionnaire; MIPQIII: Multiple Intelligences Profiling Questionnaire; RIMI: Roger’s Indicator of Multiple Intelligences; RMMIQ: Researcher Made Multiple Intelligences Inventory; Log=Logical; Kin. =Kinesthetic; Spat. =Spatial; Inter. =Interpersonal; Intra. =Intrapersonal; Nat. =Natural; Exist. =Existential, Music=Musical; Ling. =Linguistic; Verb. =Verbal.

MI-EFL studies also report vigorous sex effects. Conducting his investigation with ESL learners in the USA, Loori (2005) found significant differences in Logical/Mathematical intelligence in favor of males and Intrapersonal intelligence in favor of females. Zare-Ee, et al. (2015) observed Iranian male learners differed significantly in their estimation of Intrapersonal intelligence. Concerning WGHR, both groups rated their Existential and Logical-Mathematical intelligences the highest. Their findings implied women had a higher overall estimation of their MI. Esmaeili, Behnam and Esmaeili (2014) conducted an investigation on EFL learners in Azerbaijan, Iran, where people’s first language is Turkish. Results revealed gender effect for men in Intrapersonal intelligence. No significant difference between males and females’ overall MI assessments was reported. Hajhashemi, Akef, and Anderson (2012) found female EFL learners were significantly different in Bodily/Kinesthetic intelligence while men did not rate any of their intelligences significantly higher. The results of their study exhibited an overall higher MI self-estimation for females. Table 2 gives a summary of a number of MI-EFL studies results.

Table 2: A Comparison of Gender Effect in MI-EFL Research Results

	MI Test Type	Gender Effect	
		Male	Female
Loori (2005)	Teele Inventory	Log. /Math.	Intra.
Al-Faoury, Khataybeh and Al-Sheikh (2011)	RMMIQ	Inter.	Intra.-Ling
Hajhashemi, Akef, and Anderson (2012)	MMII		Kin.
Sadeghi (2013)	Christion MII	Spat.	Inter.
Biria, Boshrabadi and Nikbakht (2014)	MMII		Ling.
Esmaeili, et al. (2014)	Armstrong's MII	Intra.	
Zare-Ee, et al. (2015)	MMII	Inter.	

Note. MMII: McKenzie Multiple Intelligences Inventory; RMMIQ: Researcher Made Multiple Intelligences Inventory; Log=Logical; Kin. =Kinesthetic; Spat. =Spatial; Inter. =Interpersonal; Intra. =Intrapersonal; Nat. =Natural; Exist. =Existential, Music=Musical; Ling. =Linguistic; Verb. =Verbal.

Although gender effect in MI research is widely acknowledged, findings vary-even within the same country-in the effect size, intelligence type, gender category and overall self-estimation (Tables 1 and 2). As an example, the findings of Hajhashemi, et al. (2012), Zare-Ee, et al. (2015) and Sadeghi (2013) do not concord with those of Furnham, et al. (2002) although they were all conducted in Iranian context. Gender-effect variance is often explained as a function of socio-ecological and cultural contexts (Bowles, 2012; Furnham, et al., 2002; Shearer, 2009; Tirri & Nokelainen, 2008; Zare-Ee, et al., 2015). Wigfield & Cambria (2009, p. 4) argue that “self -schema” and “perceptions of competence” impact an individual’s expectancies, values and interests, hence *self*-estimations of MI.

While socio-cultural factors may to some extent explain gender effect, part of the variance may be attributed to MI assessment instrumentation. More than ten different MI test methods were used in the eighteen studies presented in Tables 1 and 2. Some were researcher-made and some were commonly used large-scale tests like the MIDAS. They differ significantly in content, number of items and components, the proportional weight given to each component, modality (written, pictorial), item type (question, statement) and answer choices. The present research sets to investigate the relationship between MI test instrumentation and gender effect in Iranian context.

3. Method

3.1. Participants

A cohort of 369 (149 male and 220 female) high school students aged between 15 to18 were selected from twenty schools in the cities of Tehran and Isfahan through random clustering. The schools of three districts of Isfahan and Tehran were divided into clusters. Then, simple random samples were selected from the population. The study focused on this age group since they make the highest population of secondary school students in Iran (Danakhabar, 2018) and English is an important subject at this level.

3.2. Instruments

Two MI test inventories, the Persian version of the MIDAS (P-MIDAS) and the Persian version of McKenzie’s MII (P- MMII), were used for this comparative study. To reduce language and culture-relevant problems, the adapted and validated versions of the tests were applied. The P-MIDAST and P-MMII were translated (and back-translated), adapted and validated by Saeidi, Ostvar, Shearer and Asghari Jafarabadi (2014) and Hajhashemi and BeeEng (2010) respectively. The item content validity index (I-CVI) and scale content validity index (S-CVI) of the P-MIDAS were from good (between 0.60-

0.74) to excellent (above 0.74) (Saeidi et. al, 2014, p. 124). The reliability of the P-MIDAS also fell in “the high-moderate to high range with alpha coefficients ranging from 0.82 to 0.90 and a median of 0.86” (P. 125). As concerns P-MMII, the Cronbach alpha for the test was found to be 0.90 which indicates high reliability (Hajhashemi & BeeEng, 2010, p. 343). To make the comparison of the results balanced, the Existential Intelligence subscale of P-MMII was not included in the study since the MIDAS lacks this section.

3.3. Procedure

In order to prevent or reduce probable problems related to time management, participant involvement and appropriateness of the procedure, two pilot studies were conducted. A group of 25 female high school students were selected for the pilot study. Preliminary to the application of the tests, the students were told about the research objectives and ethical permission. They were then assigned to P-MIDAS first and, after a short break, to P-MMII. As the procedure took more than 120 minutes, the participants showed signs of extreme fatigue. The results of the tests revealed that they had not performed on P-MMII as carefully and completely as they did on the P-MIDAS. Cronbach alpha for P-MIDAS was acceptable ($\alpha = 0.74$) and for P-MMII was questionable ($\alpha = 0.61$).

In the next pilot study, another class of 28 students were assigned to each test with a time interval of one week. They responded positively and assessed their MI more carefully on both tests. Concerning the comprehensibility of the items and the allotted time, no serious problem was observed. Cronbach alphas for both P-MIDAS ($\alpha = 0.86$) and P-MMII ($\alpha = 0.84$ respectively) were good. For the main study, hence, the participants were assigned to P-MIII and P-MIDAS in two different sessions with a time interval of at least one week. To avoid question/test-order bias, half the students were assigned to the P-MIDAS first and half to the P-MMII first. The process took about six weeks.

4. Results

4.1. Descriptive Statistics

As Table 3 demonstrates, the highest mean score in the P-MIDAS results was found to be related to Intrapersonal intelligence with a value of 59.62 and the lowest was of Musical with a value of 45.62. In P-MMII, the highest mean score was for Intrapersonal intelligence with a value of 88.15 and the lowest was 57.46 for Interpersonal intelligence. Students’ scores on all Intelligences in MMII were higher than those on the MIDAS. Figure 1 gives a visual view of the data.

The results do not show similar SDs for any of the Intelligence subcategory pairs. The SD for Verbal/Linguistic Intelligence is the largest in P-MMIII (SD=21.8) while in the P-MIDAS, Musical intelligence has the largest SD (SD=20.19). The lowest SDs in P-MMII and in the P-MIDAS are for Intrapersonal intelligences (12.28 and 14.52 respectively). Maximum SD difference is between Verbal/Linguistic intelligences (21.81-17.35 = 4.46) and the smallest between Kinesthetic intelligences SDs (18.24-17.91 =0.33).

Table 3: P-MIDAS and P-MMII Variables Descriptive Data

P-MMII			P_MIDAS		
Variables	Mean	SD	Variables	Mean	SD
Naturalist	70.19	20.75	Naturalist	46.20	19.31
Musical	75.07	18.86	Musical	45.62	*20.19
Logical/Math	67.24	16.70	Logical/Math	57.15	15.40
Interpersonal	57.46	18.73	Interpersonal	54.13	15.70
Kinesthetic	72.44	18.24	Kinesthetic	48.42	17.91
Verbal/Linguistic	61.61	*21.81	Verbal/Linguistic	49.73	17.35
Intrapersonal	88.15	*12.28	Intrapersonal	59.62	*14.52
Visual/Spatial	78.83	15.56	Visual/Spatial	51.02	19.18

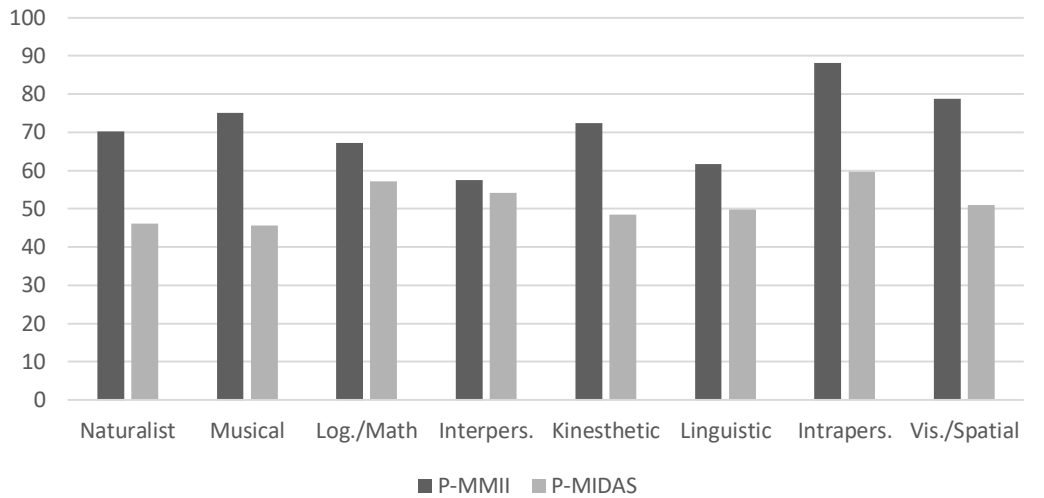


Figure 1: A Comparison of Variables Means of P-MIDAST and P-MMII

4.2. Inferential Statistics

As Table 4 illustrates, the results of Kolmogorov-Smirnov Test indicate P-MMII data were not normally distributed ($p < 0.001$ ($p < 0.05$)), but the P-MIDAS results had normal distribution ($p > 0.05$). Tests of paired-samples and independent-samples t-test results indicated normal distribution of all P-MIDAST subcategories data (Asymp. Sig. (2-tailed) > 0.05), while the results show non-normality of all P-MMII subcategories data (Asymp. Sig. (2-tailed) < 0.05).

Table 4: Kolmogorov-Smirnov Test Results for the MIDAS and MMII

MIDAS	Mus.	Kin.	Log./Math	Spat./Vis	Ling/Verb	Inter.	Intra.	Nat.
K-ST	0.78	0.99	0.90	0.96	1.03	0.85	0.94	0.92
Asymp. Sig. (2-tailed)	0.57	0.27	0.39	0.30	0.23	0.45	0.33	0.36
MMII								
K-ST Z	3.17	3.308	2.49	2.45	2.03	2.05	4.35	2.34
Asymp. Sig. (2-tailed)	0.00	0.000	0.000	0.000	0.001	0.000	0.000	0.000

Overall, on both MI tests, females estimated their MI higher than did males. For the MIDAS, the overall means for males and females were 50.77 and 52.52 respectively. However, females rated themselves lower than males in logical/mathematical ($M=56.79$, $M=57.39$ respectively), Interpersonal ($M=53.84$, $M=54.32$ respectively) and Natural ($M=44.44$, $M=47.30$ respectively) Intelligences. Kinesthetic Intelligence was rather equally estimated by both genders (females: $M=48.39$, Males= 48.44). For other intelligences, females estimated their MI higher than did males. The results of Independent Student t-test analyses revealed female gender effect for Spatial ($t(340) = -2.16$, $p = .031$), Linguistic/Verbal ($t(360) = -2.30$, $p = .02$) and Musical ($t(354) = -5.29$, $p < .001$) intelligences in P-MIDAST (Table 5). No significant gender effect was observed in Logical/Mathematical intelligence for men.

Table 5: Gender effect in P-MIDAS T-test results

	<i>t</i>	<i>df</i>	<i>Sig. (2-tailed)</i>
Musical	-5.29	354	0.000*
Kinesthetic	0.02	361	0.980
Math/Logical	0.35	356	0.722
Spatial/Visual	-2.16	340	0.031*
Ling./Verbal	-2.30	360	0.022*
Interpersonal	0.27	343	0.781
Intrapersonal	1.26	345	0.206
Natural	1.33	344	0.182

*= significant difference

Concerning P-MMII, the overall mean for females was 73.25 and for males 69.98. As on the P-MIDAS, females estimated their Logical/Mathematical Intelligence lower than did males on P-MMII (females: M=64.33, Males: M=69.41). Interpersonal Intelligence was rather equally estimated by both genders (females: M=57.96, Males=57.08). For all other Intelligences, females had a greater estimation of their capacities than did males.

Independent Student t-test analyses indicate gender effect for Musical, Logical, Kinesthetic, Verbal and Visual intelligences (Musical: $t(302) = -3.65, P = 0.000$, Logical: $t(295) = 2.62, p = .009$, Kinesthetic: $t(314) = -2.00, p = .046$, Verbal: $t(315) = -4.59, p = 0.000$, Visual: $t(313) = -2.59, p = .010$). The female group assessed their Musical, Kinesthetic, Verbal and Visual intelligences significantly higher than males, while males estimated their Logical/Mathematical intelligence significantly higher (Table 6).

Table 6: Gender effect in P-MMII T-test Results

	<i>t</i>	<i>df</i>	<i>Sig. (2-tailed)</i>
Naturalist	-0.30	308	0.765
Musical	-3.65	302	0.000*
Logical/Math.	2.62	295	0.009*
Interpersonal	-0.41	313	0.681
Kinesthetic	-2.00	314	0.046*
Ling./Verbal	-4.59	315	0.000*
Intrapersonal	-1.53	322	0.126
Spatial/Visual	-2.59	313	0.010*

*= significant difference

It follows that the data are not comparable in terms of normal distribution. The P-MIDAS data fall on a symmetrical curve while MMII data don't. Moreover, the results are not in accordance in terms of MI-type/s exhibiting gender effect, nor are they comparable outright in their effect sizes. Aside from Intrapersonal intelligence rated the first in both tests, other Intelligence rankings on P-MIDAST and P-MMII do not accord (Table 7).

Table 7: A Comparison of Gender Effect in the Results of P-MIDAST and P-MMII

MI Test	Male WGHR	Female WGHR	Gender Effect	
MMII	Intra.- Spat.- Music	Intra.- Spat. Music	*Male	*Female
MIDAS	Intra.-Log.-Inter.	Intra.-Log.- Inter.	Log. /Math.	Spat. -Music- Kin.- Ling.
				Spat.- Ling.-Music

Note. WGHR= Within Group Highest Ratings

According to Table 7, results relatively concord in that in both P-MIDAST and P-MMII, Spatial/Visual, Musical and Linguistic/Verbal intelligences exhibited female effect. However, Bodily/Kinesthetic and Logical/mathematical intelligences signified gender difference in P-MMII but not in P-MIDAST. Concerning the effect size and MI rating, excluding Spatial/Visual Intelligence, other effect sizes are significantly different, i.e., Musical intelligence is the second in P-MMII but the last in P-MIDAST. Moreover, the P-MMII data indicated four Intelligences creating female effect whereas for the P-MIDAST, there were only three intelligence types. There was no gender effect in favor of males in the P-MIDAST, whereas in P-MMII, male effect for Logical/Mathematical intelligence was observed. Concerning within- group highest ratings (WGHR), both tests yielded compatible data.

5. Discussion

The results of the study are in line with Zare-Ee, et al. (2015) and Hajhashemi, et al. (2012), with MMII for their instrumentation, and Sadeghi (2013) using Christian MII, which indicated Iranian female students had an overall higher MI self-estimation. However, Zare-Ee, et al. found Interpersonal Intelligence male effect whereas neither the P-MIDAST nor P-MMII results in this study revealed such gender effect. The results of P-MMII of the present study are partially in line with Hajhashemi, et al. in that both indicated Bodily/Kinesthetic female effect. But, the findings of P-MIDAST in this article are not in accordance with Hajhashemi, et al.

The findings of this article do not accord with the results of Furnham, et al. (2002)-conducted in Iran-with FCMIQ as their instrument. While the present article indicated an overall higher female MI self-estimation, Furnham, et al. observed an overall higher male MI self-estimation. As concerns gender effect, the results of P-MMII in this article are not in line with Furnham, et al. (2002) who found Interpersonal intelligence as the only Intelligence type to be significantly different in favor of females. This study indicated Logical/Mathematical male-effect size and Visual/Spatial, Musical, Bodily/Kinesthetic and Verbal/Linguistic female effect in P-MMII results. Furnham, et al. did not reveal any male-related gender difference while this study indicated Logical/Mathematical male effect in P-MMII. As for the P-MIDAST, the results are partly similar to those of Furnham, et al. in that neither indicated male gender effect and both indicated female effect. But the type and number of intelligences associated with the effect totally differed (Tables 1, 2 and 7).

Concerning WGHR, the P-MMII findings are partially in line with Esmaili, et al. (2014) and Hajhashemi, et al. (2012). Intrapersonal and Musical intelligences were rated the highest in both studies. The results related to the MIDAS data are partially comparable with Zare-Ee, et al. (2015) in that Logical-Mathematical intelligence was one of the WGHR intelligences for both genders.

The relative consistency in the findings of Furnham and his co-researchers' (e.g., Furnham and Chamorro-Premuzic, 2005; Furnham, et al., 2002; Szymanowicz & Furnham, 2013; Yuen & Furnham, 2005), who were involved in complementary replications of MI research across different cultures and countries, may in some measure be due to the application of a similar MI assessment instrument, FCMIQ, which was developed by Furnham (2000, p. 15) and applied in "all" their studies, while the variance in gender effect findings in a number of other relatively similar studies conducted in Iranian context, including the present research, may be attributed to variety in MI assessment instrumentation and method.

In line with the argument of this article, Saeidi and Karvandi (2014) noticed their findings were in accordance with some research results which used the same MI test as they did- the MIDAS- and different from the research findings obtained from MMII. They concluded that "One of the possible explanations for the discrepancies among the results of different studies concerning the relationship between Iranian EFL learners' MI and their writing skill can be the type of MI scaleused"(p.200).

The difference in the content, test item type (questions, statements, pictures, and actions), the number and components of each intelligence category, the weight each component is given and the type and number of answer choices of MI tests may cause variance in the results. As an instance, a close look at Musical Intelligence sections in the MIDAS (Shearer, 1994, p.3) and MMII (McKenzie, 2017/1999, section 2) reveals differences and similarities in approach. The MIDAS appears to consider the individual's personal response as an important indicator of Musical Intelligence, i.e., the term "like" is frequently used in the MIDAS while MMII is less concerned about liking. Dancing or "moving to the beat" is considered as an index of Musical Intelligence by MMII but ignored by the MIDAS. The MIDAS gives greater weight to composing and playing the music-even drumming fingers- whereas MMII considers picking patterns, beats and rhymes as constituents of Musical Intelligence. MMII is concerned with the sound of nature, poetry and lyrics while the MIDAS completely ignores them. Musical memory and remembering sounds and patterns are taken into consideration in MMII more than they are in the MIDAS. It follows that they not only differ in their definition of Musical Intelligence domain components but in the weight they give to each.

The problem of translation may be added to the list since it is rare when a translated test is in total agreement with the original one. As different MI tests are usually made to address a certain socio-culturally defined context, like the MIDAS which was originally constructed for the American society, the questions may not always fit new contexts. For example, the MIDAS emphasizes playing and making music as indicators of Musical Intelligence. However, in certain cultures listening to music is forbidden or is highly restricted especially for women and in some low-income communities, playing and composing music are far to imagine. Such restrictions may create false understanding of one's abilities as they may have a faint idea of what such concepts may signify. The present researchers noticed that some students in less privileged areas had trouble understanding what making music could possibly involve. Many of the participants did not know what "tune", "rhyme", "rhythm", "pattern" (Shearer, 1994, p.3) and "cadence of poetry" (McKenzie, 2017/1999, section 2) really meant.

As concerns answer choices, the present researchers noticed despite its lower number of questions and fewer answer choices (agree-disagree), MMII took a proportionately much longer time (between 1.25 to 1.5 times longer) for the participants to answer than did the MIDAS with a considerably larger number of questions and six answer choices. The participants said that they had completely been lost between "agree-disagree" choices because they fell on two extremes and there were no in-between choices.

6. Conclusion

A comparative study of the results of MMII and the MIDAS which were applied to the same students in the same Iranian context (Isfahan and Tehran), revealed the non-normality of the data obtained from MMII and the normal distribution of the MIDAS data. The overall scores on MMII were higher than on the MIDAS. There was significant gender effect differences between the results of MMII and the MIDAS. MMII data signified male gender effect but no male effect was observed in the MIDAS. Except for Visual/Spatial Intelligence which was of the same effect size for females other intelligences did not indicate the same effect size. Whereas MMII data indicated female effect for four Intelligences, there were three Intelligences creating female effect in the MIDAS. Intrapersonal intelligence excluded, there was a difference between intelligences in terms of WGHR (Table7).

The findings of the present research indicate the importance of instrumentation in MI-EFL research and pedagogy. Because MI test facets vary greatly in their contents, number of items and components, the proportional weight each component receives, modality, item type and answer choices, they may yield significantly different results. Hence, individuals' estimated MI profiles may vary depending on what test instrument have been used for assessing their MI profiles. As an example, an

individual may be found to have high Interpersonal intelligence if tested by the MIDAS but be assessed as being kinesthetically intelligent if tested by MMII. Such inconsistencies indicate the need for more comprehensive and valid MI-test methods.

Stahl (1999) expresses doubts about the positive impacts of some differential educational programs not because he is against MI-fair teaching methodology but because he fears the instrument won't assess individual differences as they are claimed to have the capacity to. He suspects the validity and reliability of "any measure [including MI tests] that asks subjects to report about themselves" (p. 4). According to Stahl, ambiguity of the questions will result in low construct validity. Gardner (2016, p.8) is also well aware of the problem of MI assessment. He postulates that unless multiple measures are developed for assessing individual's MI profile, "intelligence-fair" assessment won't be possible.

It is hoped that this article will open new fields of enquiry into MI-fair assessment and MI-EFL research and pedagogy. The results of this research may provide applied linguists and EFL investigators with a more profound understanding of MI assessment and will help MI-based differentiated EFL teaching be more Intelligence-fair, hence more effective.

We propose replication of this study in new contexts and conducting comparative research on other widely-used MI tests such as MIPQIII and FCMIQ (Furnham Constructed MI Questionnaire). Research into multimodal methods of MI assessment compared to 'mono-modal' instrumentation will definitely offer new insights into MI assessment and research.

References

- Alizadeh1, H., Saeidi, M., & Hadidi Tamjid, N. (2014). The relationship between Iranian EFL learners' Multiple Intelligences and their writing performance across different genders. *The Journal of Applied Linguistics*, 7(14), 1-22.
- Ahur, T., & Abdi, M. (2015). The Relationship between EFL Learners' Multiple Intelligences and Vocabulary Learning Strategies Use with a Focus on Gender. *Theory and Practice in Language Studies*, 5(4), 800-809. doi:<http://dx.doi.org/10.17507/tpls.0504.16>.
- Akbari, R., & Hosseini, K. (2008). Multiple intelligences and language learning strategies: Investigating possible relations. *System*, 36(2), 141-155. doi:10.1016/j.system.2007.09.008.
- Al-balham, E. M. (2006). Multiple intelligence styles in relation to improved academic performance in Kuwaiti middle school reading. *Digest of Middle East Studies*, 15(1), 18-34. Retrieved from <http://onlinelibrary.wiley.com/doi/10.1111/j.1949-3606.2006.tb00002.x/pdf>.
- Armstrong, T. (2009). *Multiple intelligences in the classroom*. Alexandria, USA: ASCD.
- Baum, S., Viens, J., & Slatin, B. (2007). *Multiple intelligences in the elementary classroom*. New York: Teachers College Press.
- Bennet, M. (2000). Self-estimates and population estimates of of ability in men and women. *Australian Journal of Psychology*, 52(1), 23-28.
- Biria, R., Boshraadi, A., & Nikbakht, E. (2014). The relationship between Multiple Intelligences and Iranian EFL learners' level of L2 lexical knowledge: The case of gender. *Advances in Language and Literary Studies*, 5(3), 9-17. doi:10.7575/aiac.all.v.5n.3p.9.
- Bowles, T. (2012). The longer they stay the less talented they perceive they are: Females' talent based on approaches to learning. *Australian Journal of Educational & Developmental Psychology*, 12, 12-20.
- Brualdi, A. C. (1996). Multiple intelligences: Gardner's theory. *Practical Assessment, Research & Evaluation*, 5(10). Retrieved from <http://PAREonline.net/getvn.asp?v=5&n=10>.
- Chen, S.-F. (2005). *Cooperative learning, multiple intelligences and proficiency: Application in college English language teaching and learning*. Unpublished doctoral dissertation, Australian Catholic University, Australia.
- Cheng , H. Y., Hou, Y. J., Hou, Y. A., & Chung , W. Y. (2010). Gender differences in Multiple Intelligences, learning behavior and English performance: A case study in Taiwan. *International Journal of the Humanities*, 8(2), 255-270. doi:10.18848/1447-9508/CGP/v08i02/42844.
- Christison, M. A. (1996). Teaching and learning languages through Multiple Intelligences. *TESOL Journal*, 6(1), 10-14.
- DeVellis, R. F. (1991). *Scale development: Theory and applications*. Newbury Park: Sage Publication.
- Esmaeili, F., Behnam, B., & Esmaeili, K. (2014). A study of relationship between Multiple Intelligences and writing ability of Iranian female and male students. *Mediterranean Journal of Social Sciences*, 5(20). doi:10.5901/mjss.2014.v5n20p2663.
- Gardner, H. (1983). *Frames of mind: The theory of Multiple Intelligences*. New York: Basic Books. Retrieved 2016.
- Gardner, H. (1999). *Intelligence reframed: Multiple intelligences for the 21st century*. New York: Basic Books.

- Gardner, H. (2004). Audiences for the theory of Multiple Intelligences. *Teachers College Record*, 106(1), 212-220. doi:10.1111/j.1467-9620.2004.00329.
- Gardner, H. (2011). *The theory of Multiple Intelligences: As psychology, as education, as social science*. Retrieved from Harvard Graduate School of Education, Project Zero Fifty Years: <http://www.pz.harvard.edu/resources/the-theory-of-multiple-intelligences-as-psychology-as-education-as-social-science>.
- Gardner, H. (2013). *Harvard Project Zero: A Personal History*. Retrieved from Harvard Graduate School of Education Project Zero Fifty Years: <http://www.pz.harvard.edu/resources/harvard-project-zero-a-personal-history>.
- Gardner, H. (2016). *Frequently Asked Questions—Multiple Intelligence and Related Educational Topics*. Retrieved from Howard Gardner: <https://howardgardner.com/faq/>.
- Gardner, H., & Hatch, T. (1989). Multiple intelligences go to school: Educational implications of the theory of Multiple Intelligence. *Educational Researcher*, 18(8), 4-10. Retrieved from <http://www.jstor.org/stable/1176460>.
- Guba, E. G., & Lincoln, E. S. (1994). Competing Paradigms in Qualitative Research. In N. K. Denzin, & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 105-117). Thousand Oaks, CA: Sage.
- Hajhashemi, K., & Bee Eng, W. (2010). A validation study of the Persian version of McKenzie's Multiple Intelligences Inventory to measure profiles of pre-university students. *Pertanika J. Soc. Sci. & Hum.*, 18(2), 343 - 355 .
- Hajhashemi, K., Akef, K., & Anderso, N. (2012). The relationship between multiple intelligences and reading proficiency of Iranian EFL students. *World Applied Sciences Journal*, 9(10), 1475-1483. doi:10.5829/idosi.wasj.2012.19.10.3134.
- Hajhashemi, K., Cook, J., & Bee Eng, W. (2012). MI as a predictor of students' performance in reading competency. *English Language Teaching*, 5(3), 240-251. doi:10.5539/elt.v5n3p240.
- Haley, M. H. (2004). Learner-Centered instruction and the theory of Multiple Intelligences with second language learners. *Teachers College Record*, 106(1), 163-180. doi:10.1111/j.1467-9620.2004.00326.x.
- Hoerr, T. R. (2000). *Becoming a multiple intelligence school*. Alexandria, USA: Association for Supervision and Curriculum Development.
- Joshi, J. (2016). A study of Multiple Intelligences of students in context to gender. *The International Journal of Indian Psychology*, 3(3), 47-53.
- Kaur, M. (2014). Gender differences in Multiple Intelligences with respect to grade level. *American International Journal of Research in Humanities, Arts and Social Sciences*, 8(2), 117-120.
- Kilgore, k., Griffin, C., Sindelar, P. T., & Webb, R. B. (2015). Restructuring for inclusion: Changing teaching practices (Part II). *Middle School Journal*, 33(3), 7-13.
- Liu, H.-j., & Chen, T.-h. (2013). Foreign language anxiety in young learners: How it relates to multiple intelligences, learner attitudes, and perceived competence. *Journal of Language Teaching and Research*, 4(5), 932-938. doi:10.4304/jltr.4.5.932-938.
- Loori, A. A. (2005). Multiple intelligences: A comparative study between the preferences of males and females. *Social Behavior and Personality: An international journal*, 33(1), 77-88. doi:<https://doi.org/10.2224/sbp.2005.33.1.77>.

- Myers, M. D. (1997). Qualitative Research in Information Systems. *MIS Quarterly*, 21(2), 241-242. Retrieved from <http://www.qual.auckland.ac.nz/>.
- Nicholson-Nelson, K. (1998). *Developing students' multiple intelligences*. New York: Scholastic Professional Books.
- O'Grady, M. (2017). *Critique of Teaching & Learning Unit Report*. Retrieved from Institute of Technology Tralee: http://www.ittralee.ie/en/InformationFor/Staff/Review/TLUHome/Journals/Journals20062007/Title_15442_en.html.
- Oliver, Z. M. (2010). *Falling but fulfilled: reflections on Multiple Intelligences*. Honolulu: Savants Books and Publications.
- Palmberg, R. (2011). Multiple intelligences revisited. Vaasa, Finland. Retrieved 2017, from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.455.949&rep=rep1&type=pdf>.
- Phakiti, A. (2003). A closer look at the relationship of cognitive and metacognitive. *Language Testing*, 20(1), 26-56.
- Razmjoo, S. A. (2008). On the relationship between multiple intelligences and language proficiency. *The Reading Matrix*, 8(2), 155-174. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.540.3303&rep=rep1&type=pdf>.
- Rostami, F., & Soleimani, H. (2015). The relationship between multiple intelligences and Iranian EFL learners' performance on writing different essay types. *Journal of Language and Literature*, 6(4), 77-87. doi: doi:10.7813/jll.2015/6-4/15.
- Sabah Hasan Al-Onizat, S. H. (2014). The psychometric properties of an Arabic version of Multiple Intelligence Development Assessment Scale for Adolescents (TEEN-MIDAS). *Creative Education*, 5, 590-605. Retrieved from <http://dx.doi.org/10.4236/ce.2014.58070>.
- Sadeghi, M. (2013). EFL male and female learning styles and Multiple Intelligences; A case of Iranian EFL university students. *IJRELT*, 1(2).
- Seliger, H. W., & Shohamy, E. (1989). *Second Language Research Methods*. Oxford: Oxford University Press.
- Shearer, B. (2012). Cross cultural factor analytic studies of a Multiple Intelligences self-assessment. *The International Journal of Educational and Psychological Assessment*, 12(1).
- Shearer, B. (2017). *About MIDAS™*. Retrieved from Multiple Intelligences Research and Consulting Inc.: <http://www.miresearch.org/midas/midas-2/>.
- Shearer, C. B. (1997). Reliability, Validity and Utility of a Multiple Intelligences Assessment for Career Planning. *105th Annual Convention of the American Psychological Association* (pp. 1-26). Chicago, Illinois: Eric. Retrieved from <http://files.eric.ed.gov/fulltext/ED415476.pdf>.
- Shearer, C. B. (2001). The Multiple Intelligences Developmental Assessment Scales for Adolescents. Retrieved from <http://www.miresearch.org/?s=The+Multiple+Intelligences+Developmental+Assessment+Scales+for+Adolescents&sa=Search>.
- Shearer, C. B. (2013, December 13). *About MIDAS*. Retrieved 2016, from Multiple Intelligences Research and Consulting, Inc.:

<http://www.miresearch.org/?s=Which+Multiple+Intelligences+Assessment+is+Appropriate+for+Y+our+Purposes%3F&sa=Search>.

- Shearer, C. B., & Jones, J. A. (1994). The validation of the Hillside Assessment of Perceived Intelligences (HAPI): A measure of Howard Gardner's theory of Multiple Intelligences. *The annual meeting of the American Educational Research Association* (pp. 1-20). New Orleans, LA: Eric. Retrieved from <http://files.eric.ed.gov/fulltext/ED372077.pdf>.
- Shearer, C. B., & Luzzo, D. A. (2009). Exploring the application of multiple intelligences theory to career counseling. *The Career Development Quarterly*, 3-13. doi:10.1002/j.2161-0045.2009.tb00169.x.
- Snider, D. P. (2001). *Multiple Intelligences Theory and foreign language teaching*. Unpublished Dissertation, University of Utah.
- Spearman, C. (1904). General intelligence, objectively determined and measured. *The American Journal of Psychology*, 15(2), 201-292.
- Sternberg, R. J. (2015). *Human intelligence psychology*. Retrieved from Encyclopedica Britannica: <https://www.britannica.com/topic/human-intelligence-psychology>.
- Szymanowicz, A., & Furnham, A. (2013). Gender and gender role differences in self- and other-estimates of Multiple Intelligences. *The Journal of Social Psychology*, 153:4, 399-423. doi:10.1080/00224545.2012.754397
- Tomlinson, C. A. (2001). *How to differentiate instruction in mixed-ability classrooms* (2nd ed.). Alexandria, USA: Association for Supervision and Curriculum Development.
- Yuen, M., & Furnham, A. (2005). Sex differences in self-estimation of Multiple Intelligences among Hong Kong Chinese adolescents. *High Ability Studies*. doi:10.1080/13598130600618009
- von Stumm, S., Chamorro-Premuzic, T., & Furnham, A. (2009). Decomposing self-estimates of intelligence: Structure and sex differences across 12 nations. *British Journal of Psychology*, 100, 429-442.
- Wigfield, A., & Cambria, J. (2010). Students' achievement values, goal orientations, and interest: Definitions, development, and relations to achievement outcomes. *Developmental Review*, 30, 1-35. doi:10.1016/j.dr.2009.12.001.
- Workman, L. (2003). Self-perception of intelligence in male and female undergraduates in old and new Welsh universities. *Psychology Learning and Teaching*, 4(1), 22-26.
- Zare-ee, A., Mohd Don, Z., Knowles, G., & Tohidian, I. (2015). Gender differences in self-estimates of multiple intelligences among learners of English. *International Online Journal of Education and Teaching*, 2(4), 249-264.