

An Inter-rater Reliability Study of a Self-assessment for the Multiple Intelligences

C. Branton Shearer¹

¹ MI Research and Consulting Inc, Kent, Ohio, USA

Correspondence: C. Branton Shearer, MI Research and Consulting Inc., 1316 South Lincoln Street, Kent, Ohio 44240, USA. Tel: 1-330-687-1735. E-mail: sbranton@kent.edu

Received: July 19, 2012 Accepted: August 8, 2012 Online Published: August 22, 2012

doi:10.5539/ijps.v4n3p131

URL: <http://dx.doi.org/10.5539/ijps.v4n3p131>

Abstract

This paper describes the results of an inter-rater reliability study of a self-assessment for the multiple intelligences (Gardner, 1983/1993). Previous studies found the *Multiple Intelligences Developmental Assessment Scales* (MIDAS; Shearer, 2007) demonstrate content, criterion group, concurrent and construct validity (Shearer, in press) but that respondent reliability needed further investigation. An original multi-informant study ($n=74$) reported moderate to high levels of categorical agreement among raters (40% exact and 80% within one category). This cross-cultural replication study found even higher levels of agreement (46% exact and 92% within one category) among primary and secondary raters ($n=173$) for respondents from five different countries. Implications for both theory and practice are discussed.

Keywords: multiple intelligences, assessment, self-report, inter-rater reliability

1. Introduction

The use of self-report questionnaires in education and psychology as a means to obtain valid information about adults and students has a long and checkered history. Before the advent of the first successful IQ test (Binet, 1916) as a measure of cognitive abilities, introspection was a widely accepted method of assessment. Standardized IQ tests developed in lockstep with the industrial revolution as objective, quantitative measures that gained respectability as “scientific instruments”. Once performance tests became the accepted standard for measuring various cognitive abilities researchers strongly criticized self-report methods of gathering data as being “merely subjective” and lacking validity due to respondent bias and distortion. (DeNisi & Shaw, 1977; Mabe & West, 1982)

The *Multiple Intelligences Developmental Assessment Scales* (MIDAS) were first developed in 1987 with the goal of gathering valid and clinically useful information regarding the multiple intelligences’ profiles of person’s undergoing cognitive rehabilitation following brain trauma. The theory of multiple intelligences (MI; Gardner, 1983/1993) was chosen as the basis for this assessment because of its broad, unique and practical approach of describing intelligence in everyday, observable terms. This was essential because family members needed to be interviewed regarding the pre-morbid intelligence of memory impaired patients as a means of creating strengths-based, client-centered treatment plans.

Multiple intelligences theory differs from IQ in a number of ways but perhaps most critically in its essential definition (Gardner, 1999): “*a biopsychological potential to process information that can be activated in a cultural setting to solve problems or create products that are of value in a culture*” (p. 34). Based on this definition and using eight criteria Gardner’s research identified seven (and later research added an eighth) candidate intelligences: linguistic, logical-mathematical, spatial, musical, kinesthetic, intrapersonal, interpersonal and naturalist.

Some critics view MI as being “anti-IQ” but this is not true because IQ-related skills are integral components of the linguistic and logical-mathematical intelligences. However, MI expands beyond academic skills to include creative (poetry, novels) and practical abilities (technical manuals, explanations) associated with each of eight intelligences.

The visual-spatial intelligence is likewise an essential part of many standard IQ-type tests as visual problem-solving but MI extends this set of abilities to include the creative arts and imagination. The same holds

true for the musical and kinesthetic intelligences where there are obvious academic aspects to each (ballet and classical music) as well as creative (choreography, jazz and improvisation) and pragmatic applications (handcrafts and social music). The naturalist intelligence is represented in the standard academic curriculum in science but it also includes understanding of living things and pattern recognition.

The intrapersonal and interpersonal intelligences have been the subject of a great deal of research in recent years in terms of metacognition and social-emotional abilities, respectively. (Goleman, 1995 and 2006; Sternberg, 1985) The core features of these two intelligences are the understanding of oneself (Intra) and the understanding of other people (Inter). As common sense would indicate, along with linguistic these two intelligences are strongly associated with success in school.

The eight intelligences are neither monolithic nor simplistic, but rather each is comprised of a set of specific skills that are all related to the core cognitive components. For example, musical intelligence is expressed in skills related to musical composition, vocal ability, instrumental skills and the understanding of music. Detailed definitions of each intelligence are presented in Appendix 1.

The MIDAS is unique because it is not a simplistic checklist, but instead was carefully designed as a structured interview where respondents could describe their skills and abilities in both quantitative and qualitative terms. Each question is written to describe specific behaviors associated with core cognitive components of the target intelligence. Response choices are lettered and not numbered and are uniquely written to match with the content of each question so that respondents are encouraged to answer in a thoughtful rather than superficial manner. See sample question in Appendix 3.

Each response choice is given a numerical value by the scoring program (0 to 4) so that the main and subscale scores are calculated as percentages from the total number of answers provided. *I don't know* or *Does Not Apply* choices are not included in the calculations. A majority of items score only on their primary designated scale, but a small number of items are scored on two, and in a few instances three, scales. These co-scored items were identified from factor analytic results and a qualitative analysis of item content. Scale scores are expressed as simple percentages ranging from 0 to 100%. Using both large scale data and criterion group statistics skill level categories are defined thus: 0 – 19%= Very Low; 20% - 39% = Low; 40 – 59%= Moderate; 60% - 79%= High and 80% - 100= Very High (Shearer, 2007).

Starting in 1987 a series of preliminary validity and reliability studies of the MIDAS were conducted to determine if, indeed, a respondent (or informant who knew the person well) could provide a realistic appraisal of one's abilities in each of the eight intelligences. Numerous studies around the world have investigated the validity of the MIDAS and many research results are summarized in detail in *The MIDAS Professional Manual* (Shearer, 2007). The MIDAS provides a profile of the respondent's "intellectual disposition" that has been favorably evaluated in Buros Mental Measurements Yearbook (Prackard & Trevisan, 1999) suggesting support for use within educational contexts. See summary in Appendix 2.

Early inter-rater reliability investigations were promising but additional studies are necessary to confirm their results. In the 25 years since those initial studies a number of investigations have provided consistent support for the scales' validity (Shearer, in press) but important questions remain regarding respondents' reliability.

The first inter-rater reliability study reported in the *Professional Manual* involved 74 self-reports and an additional 138 assessments by primary and secondary respondents. To describe rates of agreement in practical terms, scale scores are divided into five categories from Very High to Very Low. An exact agreement rate of 40% was obtained. An 80% rate of agreement that is plus-or-minus one category between self, primary and secondary respondents was also found. Given the difficulty of observing high rates of agreement among independent raters, these percentages were judged to be more than adequate.

This paper describes the results of an inter-rater reliability study of the MIDAS that addresses the question, *How well does a self-report correspond with the assessments provided by significant others who know him or her well?* The goal of this study is to replicate the original inter-rater reliability investigation that found high levels of agreement among self, primary and secondary respondents (Shearer, 2007).

2. Method

Educators in a variety of contexts (universities, community colleges and high schools) volunteered to recruit students and colleagues to participate in a multi-informant study of the MIDAS questionnaire. Students were offered extra credit for completing their own MIDAS assessment and then asking two people "who know you well" (designated as Primary and Secondary) to also complete the questionnaire online. All participants were voluntary and provided with their own profiles. Data was aggregated anonymously for statistical analysis.

2.1 Participants

The total sample of 173 participants included 65 self-reports, 62 Primary and 46 Secondary respondents. There are respondents from five different countries: Canada (21), United Kingdom (18), Germany (14), US (8) and Ireland (4). Fifty-nine percent are female (38) and 26 are male. The mean age is 26.3 and ranges from 14 through 59 years. Included in the sample are 26 adults, 21 university students and 13 teenagers. The types of primary and secondary respondents are predominantly parents (30), friends (18) and spouses (11). See Table 1 for more details.

Table 1. Type of primary and secondary respondents

	Respondents	
	Primary	Secondary
Child	1	0
Parents	15	15
Spouse	7	4
Family	6	5
Boss	0	0
Co-Worker	5	4
Friend	6	12
Boy/girlfriend	3	1
Classmate	1	1
Teacher	1	1
Counselor	0	0
Other	3	0
<i>Totals</i>	<i>47</i>	<i>43</i>

Self respondents were asked to describe how long the chosen Primary and Secondary respondents have known him or her. A majority knew the person for more than ten years (68%). See breakdown in Table 2.

Table 2. How long has primary and secondary respondent known you?

	Primary	Secondary
More than 10 years:	73%	63%
5 – 10 years	7%	15%
3 – 5 years	5%	20%
1 – 2 years	10%	3%
Less than 1 year	4%	

3. Results

Out of a total number of 742 paired comparisons (Self to Primary; Self to Secondary), there was a 46% rate of exact categorical agreement. When ratings are compared within plus-or-minus-one category the agreement rate increases to 92%. There are seven percent of ratings that are different by two categories and only one percent differs by three categories.

Table 3. Agreement rates between self and primary and secondary respondents

Respondent Scale	Primary		Secondary	
	± 1	Exact	Exact	± 1
Musical	94%	45%	39%	85%
Kinesthetic	93%	51%	50%	96%
Logical-math	92%	50%	41%	83%
Spatial	88%	39%	33%	91%
Linguistic	93%	46%	39%	95%
Interpersonal	93%	36%	35%	89%
Intrapersonal	93%	54%	60%	95%
Naturalist	92%	57%	51%	93%
Total mean	92%	47%	44%	91%
Grand means	92%	± 1 cat	46%	exact

Note. Italics = Primary agreement higher than Secondary ($n=8$)

Bold = Secondary agreement higher than Primary ($n= 5$)

The Primary informant agrees exactly with the self-rating more frequently than does the Secondary informant (five scales vs. one; 47% and 44% agreement). However, the Secondary informant agrees with the self-rating within one category more often than does the Primary information (four scales vs. three).

Both Primary and Secondary informants tend to provide higher ratings than does the person rating him or her self. This is most evident in a number of extreme ratings that are two or three categories higher than the self-rating (48 are higher vs. 22 lower).

The Naturalist and Intrapersonal scales have the highest percentage of exact agreement (57% and 60%, respectively) while the Spatial and Interpersonal scales have the lowest (33% and 35%, respectively). The overall highest rates of agreement ± 1 category is 96% for Kinesthetic and 95% for Intrapersonal.

4. Discussion

The results of this investigation confirm that respondents are able to provide “reasonable descriptions” of their multiple intelligences strengths and limitations as compared with knowledgeable informants. In fact, these data are surprisingly robust given two hurdles. The first hurdle was that participants of varying ages came from five different countries and participated under a variety of circumstances. The respondents in Germany completed the questionnaire in English, their non-native language. The second hurdle was that primary and secondary informants had to complete the online questionnaire about someone else when the questions were written in second person (e.g., *Do you ever...*). There was concern that informants would have difficulty making this mental translation and thus provide inaccurate reports. This does not appear to have been a problem and indicates the ease with which respondents are able to provide accurate ratings using the MIDAS questions.

Reliability is an important yet often overlooked attribute of assessments intended for research, classroom and clinical applications. The essential validity of the multiple intelligences construct is a matter of ongoing debate and investigation. Researchers typically use standard performance measures to examine MI validity (Gottfriedson, 1998; Herrnstein & Murray, 1994; Visser, et al, 2006; White, 1988) but these provide skewed results because each of the eight intelligences are comprised of more than the convergent problem-solving skills assessed by performance tests. The MIDAS includes the divergent thinking and practical tasks associated with each MI and it has empirical research support for its cross cultural validity (Kim, 2007; Pizarro, 2003; Shearer, in press; Wu, 2007; Yoong, 2001). The data reported here add crucial evidence of reliability as a solid basis for judging the essential validity of the MIDAS to assess the eight multiple intelligences. Taken together these investigations provide large scale empirical support for the idea that the human brain possesses at least eight distinct, relatively independent forms of intelligence that are evident across cultures.

The MIDAS “process approach” toward assessing the multiple intelligences is unique because it gathers both quantitative and qualitative information describing the intellectual disposition of the respondent—from his or her perspective. This phenomenological approach respects the person as an important source of information that will be useful for both educational and clinical purposes. However, it is equally important to be able to gauge the trustworthiness of the scores and descriptions generated from a particular respondent’s responses to the questionnaire.

This research provided two surprising results pertinent to practical applications. First, the highest rate of agreement was for the Intrapersonal scale and the Spatial scale was among the lowest. A common sense assumption is that the scales with the most observable behaviors would have the greatest rate of agreement between informants, and this is true for the Kinesthetic and Interpersonal but not for the Spatial and Intrapersonal scales. It is affirming, but a bit perplexing why the least tangible of intelligences-Intrapersonal-would be one of the scales with the greatest agreement among raters. This finding is particularly important because if informants were unable to agree on behaviors associated with Intrapersonal ability then the fundamental validity of a MIDAS self-assessment would be called into question. Instead, our confidence in the results of the questionnaire is strengthened by the knowledge that external raters agree 95% of the time with the Intrapersonal scale (plus-or-minus one category). Knowing that the Spatial and Interpersonal scales have a tendency to differ from external raters can help administrators during Profile verification and interpretation.

5. Conclusions

Three points are of particular note. First, the higher agreement rates among respondents obtained from these diverse participants as compared to the original inter-rater research indicate that respondents are generally able to be reliable self-reporters using the online MIDAS questionnaire. However, the results are not perfect and so profile verification strategies described in the *Manual* should be followed to enhance educational utility. Second, contrary to some expectations, self-ratings are rarely higher than the ratings provided by people who know the respondent well. Third, these strong reliability data from five countries support the cross-cultural validity of both the multiple intelligences construct and the unique design of the MIDAS “process approach” to assessment (Shearer, in press).

Critics of “introspection” as a method of data collection may be right in their negative appraisals but not because people cannot be accurate self-reporters in general. The problem may be due to the theoretical design and construction of the introspective methods and the questions employed. The MIDAS originated as a structured interview using the theory of multiple intelligences as a guide to the selection and construction of items that could be accurately responded to by an outside informant. It was then refined into a self-report through a series of investigations both in-depth and large scale so as to sharpen the focus of both questions and response choices that are uniquely written to match the content of each question. This research supports the common sense adage that it is important both *what* you ask about and *how* you go about asking if reliable information is to be obtained. These data indicate that MIDAS self-reports are generally reliable, but of course on a case-by-case basis distortions can occur. This is true for any form of cognitive measurement. There are no perfectly reliable tests and so all test administrators are wise to trust but verify the accuracy of any profile, if maximum benefit is to be obtained.

Finally, contrary to some research findings, these data support the idea that people know themselves well enough. The importance of Intrapersonal intelligence is a key human ability that has supported our survival in the face of adversity for millennia as well as the development of a complex civilization. We must be able to deploy our cognitive capacities to our best advantage and an accurate self-appraisal is an essential skill in this process.

References

- Binet, A., & Simon, T. (1916). *The development of intelligence in children*. Baltimore: Williams & Wilkens.
- DeNisi, A., & Shaw, J. (1977). Investigation of the uses of self-reports of abilities. *Journal of Applied Psychology*, 62(5), 641-644. <http://dx.doi.org/10.1037/0021-9010.62.5.641>
- Gardner, H. (1999). *Intelligence reframed: Multiple intelligences for the 21st century*. New York: Basic Books.
- Gardner, H. (1983/1993). *Frames of mind: The theory of multiple intelligences*. New York: Basic Books.
- Goleman, D. (2006). *Social intelligence*. New York: Bantam Dell.
- Gottfredson, L. (1998). The General Intelligence Factor, *Scientific American*, 9(Winter), 4.
- Goleman, D. (1995). *Emotional intelligences: The new science of human relationships*. New York: Bantam Dell.

- Herrnstein, R., & Murray, C. (1994). *The bell curve*. New York: Free Press.
- Kim, J. (2007). *Validation studies of the Korean translation of the multiple intelligences developmental assessment scales*. Multiple Intelligences Institute, Seoul, Korea
- Kim, M., & Kim, Y. (1996). Multiple intelligences analysis of secondary school students. *Hanyang Education Review*, 12, 151-183.
- Mabe, P., West, S. (1982). Validity of self-evaluation of ability: A review and meta-analysis. *Journal of Applied Psychology*, 67(3), 280-296. <http://dx.doi.org/10.1037/0021-9010.67.3.280>
- Prackard, A., & Trevisan, M. S. (1999). Review of the MIDAS: Multiple intelligences developmental assessment scales. In B. S. Plake & J. C. Impara (Eds.), *The supplement to the thirteenth mental measurements yearbook* (pp.154-157). Lincoln, NE: Buros Institute of Mental Measurements.
- Pizarro, S. R. (April, 2003). *Psychometric analyses of the multiple intelligences developmental assessment scales*. Paper presented at the annual conference of the American Educational Research Association (AERA), Chicago, Ill. Research Association, Montreal, CA.
- Shearer, C. B. (in press). Cross cultural factor analytic studies of a multiple intelligences self-assessment. *International Journal of Educational and Psychological Assessment*.
- Shearer, C. B. (2007). *The MIDAS: Professional manual*. (Rev.ed.) Kent, Ohio: MI Research and Consulting, Inc.
- Shearer, C. B. (2006a). *Math skill and the multiple intelligences: An investigation into the MI profiles of high school students with varying levels of math skill*. Kent, Ohio: MI Research and Consulting, Inc.
- Shearer, C. B. (2006b). *Reading skill and the multiple intelligences: An investigation into the MI profiles of high school students with varying levels of reading skill*. Kent, Ohio: MI Research and Consulting, Inc.
- Shearer, C. B. (2005). *Large scale factor analysis of the Multiple Intelligences Developmental Assessment Scales*. Paper presented at the annual meeting of the American Educational.
- Shearer, C. B. (2004). *Criterion related validity of the MIDAS scales*. Kent, Ohio: MI Research and Consulting, Inc.
- Sternberg, R. (1985). *Beyond IQ: A triarchic theory of human intelligence*. New York: Cambridge University Press.
- Visser, B., Ashton, M., & Vernon, P. (2006). Beyond g: Putting multiple intelligences theory to the test. *Intelligence*, 34, 487-502. <http://dx.doi.org/10.1016/j.intell.2006.02.004>
- Waterhouse, L. (2006). Multiple intelligences, the Mozart effect, and emotional intelligence: A critical review. *Educational Psychologist*, 41(4), 207-225. http://dx.doi.org/10.1207/s15326985ep4104_1
- White, J. (1988). *Do Howard Gardner's Multiple Intelligences Add Up?* London: Institute of Education, University of London.
- Wiswell, A., Hardy, C. R., & Reio, T. G. (2001). An examination of the Multiple Intelligences Developmental Assessment Scales (MIDAS). Paper presented at the annual meeting of the Academy of Human Resource Development, Tulsa, OK.
- Wu, W. (2007). *Manual for C-MIDAS Chinese translation*. Taipei, Taiwan: Psychological Publishing Company.
- Yoong, S. (2001). *Multiple intelligences: A construct validation of the MIDAS scale in Malaysia*. Paper presented at the International Conference on Measurement and Evaluation, Penang, Malaysia.

Appendix

Appendix 1. Description of the multiple intelligences

Intelligence	Description
Interpersonal	To think about and understand another person. To have empathy and recognize distinctions among people and to appreciate their perspectives with sensitivity to their motives, moods and intentions. It involves interacting effectively with one or more people in familiar, casual or working circumstances.
Intrapersonal	To think about and understand one's self. To be aware of one's strengths and weaknesses and to plan effectively to achieve personal goals. Reflecting on and monitoring one's thoughts and feelings and regulating them effectively. The ability to monitor one's self in interpersonal relationships and to act with personal efficacy.
Kinesthetic	To think in movements and to use the body in skilled and complicated ways for expressive and goal directed activities. A sense of timing, coordination for whole body movement and the use of hands for manipulating objects.
Linguistic	To think in words and to use language to express and understand complex meanings. Sensitivity to the meaning of words and the order among words, sounds, rhythms, inflections. To reflect on the use of language in everyday life.
Logical-Mathematical	To think of cause and effect connections and to understand relationships among actions, objects or ideas. To calculate, quantify or consider propositions and perform complex mathematical or logical operations. It involves inductive and deductive reasoning skills as well as critical and creative problem solving.
Musical	To think in sounds, rhythms, melodies and rhymes. To be sensitive to pitch, rhythm, timbre and tone. To recognize, create and reproduce music by using an instrument or voice. Active listening and a strong connection between music and emotions.
Naturalist	To understand the natural world including plants, animals and scientific studies. To recognize, name and classify individuals, species and ecological relationships. To interact effectively with living creatures and discern patterns of life & natural forces.
Visual-Spatial	To think in pictures and to perceive the visual world accurately. To think in three-dimensions and to transform one's perceptions and re-create aspects of one's visual experience via imagination. To work with objects effectively.

Appendix 2. Summary of reliability and validity research of the MIDAS*

Reference Citation of Study	Results of Investigation
Wiswell, Hardy, & Reio (2001)	Reliability coefficients for MIDAS scales range from .85-.90.
Shearer (2006a, 2006b)	Scores on the Mathematics MIDAS subscale are moderately correlated ($r = .58$) with scores on the Ohio State Math Achievement Test; scores on the Reading-Writing MIDAS subscale are moderately correlated with scores on the Ohio State Reading Test.
Pizarro (2003)	Seven factors identified in analysis of MIDAS Spanish translation on Chilean sample ($n = 429$). Mean Alpha = .81.
Shearer (2005)	Large scale exploratory and confirmatory analyses ($n = 23,000$) find nine factors that correspond with MIDAS main scales and two Spatial subscales. Alpha reliabilities for scales range from .79 - .89.

Yoong (2001)	Eight factors confirmed for MIDAS-BH Bahasa translation on Malaysian sample (n= 644). Alphas range .72 - .91. Concurrent validity results find significant correlations between Logical-math and math achievement; science with Naturalist; language achievement with Linguistic scales.
Wu (2007)	Initial reliability and validity studies of the C-MIDAS Chinese MIDAS translation correspond strongly with original English version data.
Shearer & Jones (1994)	A review of several concurrent validity studies concludes that MIDAS provides a "reasonable estimate" of the respondent's intellectual disposition.
Shearer (2007)	Test-retest correlations range from .77 - .92; inter-rater reliability ranges from 40% exact agreement to 80% \pm 1 category. Minimal differences between African-American and Caucasian university students observed. Validity studies find adequate rates of discrimination and convergence with matched tests.
Kim, J.(2007); Kim, H. (2004)	Criterion groups validity study of K-MIDAS Korean translation finds significant and appropriate pattern of correlations among MIDAS scales and student groups.
Shearer (2004)	Matched criterion group mean scale scores conform with theoretical predictions for matched student and adult groups.

*This chart was previously published in *Career Development Quarterly*, 2009, 58, 1, 3–13.

Appendix 3. Sample MIDAS question

Do you have a good voice for singing with other people in harmony?

A= A little bit.

B= Fair.

C= Good.

D= Very good.

E= Excellent.

F= I don't know.