

# IQ estimate smackdown: Comparing IQ proxy measures to the WAIS-III

RUTH SPINKS,<sup>1</sup> LOWELL W. MCKIRGAN,<sup>1</sup> STEPHAN ARNDT,<sup>1,2,3</sup> KRISTIN CASPERS,<sup>1</sup>  
REBECCA YUCUIS,<sup>1</sup> AND CHRISTOPHER J. PFALZGRAF<sup>1</sup>

<sup>1</sup>Department of Psychiatry, Carver College of Medicine, University of Iowa, Iowa City, Iowa

<sup>2</sup>Iowa Consortium for Substance Abuse Research and Evaluation, University of Iowa, Iowa City, Iowa

<sup>3</sup>Department of Biostatistics, College of Public Health, University of Iowa, Iowa City, Iowa

(RECEIVED February 22, 2008; FINAL REVISION March 5, 2009; ACCEPTED March 11, 2009)

## Abstract

Brief assessments of general cognitive ability are frequently needed by neuropsychologists, and many methods of estimating intelligence quotient (IQ) have been published. While these measures typically present overall correlations with the Wechsler Adult Intelligence Scale (WAIS) Full Scale IQ, it is tacitly acknowledged that these estimates are most accurate within 1 standard deviation of the mean and that accuracy diminishes moving toward the tails of the IQ distribution. However, little work has been done to systematically characterize proxy measures at the tails of the IQ distribution. Additionally, while these measures are all correlated with the WAIS, multiple proxy measures are rarely presented in one manuscript. The current article has two goals: (1) Examine various IQ proxies against Wechsler Adult Intelligence Scale (Third Version) scores, showing the overall accuracy of each measure against the gold standard IQ measure. This comparison will assist in selecting the best proxy measure for particular clinical constraints. (2) The sample is then divided into three groups (below, average, and above-average ability), and each group is analyzed separately to characterize proxy performance at the tails of the IQ distribution. Repeated measures multivariate analysis of variance compares the different proxy measures across ability levels. All IQ estimates are represented in tables so that they can be examined side by side. (*JINS*, 2009, *15*, 590–596.)

**Keywords:** Intelligence, Cognitive ability, Assessment, Neuropsychology, Psychological tests, Educational tests

## INTRODUCTION

Neuropsychologists and cognitive researchers often need quick estimates of global cognitive functioning [i.e., intelligence quotient (IQ)]. IQ is often estimated using various methods using limited testing and/or demographic variables. The use of demographic variables is particularly attractive when a patient has little or no tolerance for formal testing. The formality and complexity of demographic estimates vary a great deal. Informal estimates may simply be a crude judgment of level of functioning based solely on occupational status or years of formal education (Sattler, 2001). Formal estimating formulae vary in complexity and use an array of demographic variables (c.f., Barona et al., 1984; Crawford & Allan, 1997). Commonly used demographic variables include educational attainment, occupational sta-

tus, and age in a weighted regression formula (Crawford & Allan, 1997). The Barona estimate also incorporates race, region of the country in which the person is living, and whether they live in an urban or rural environment (Barona et al., 1984).

The most popular method of estimating IQ involves a shortened administration of the Wechsler scales. A variety of subtest combinations are used to estimate a Full Scale IQ (FSIQ) based on administering as few as one Wechsler Adult Intelligence Scale (WAIS) subtest and as many as seven (Axelrod et al., 2000; Engelhart et al., 1999; Jeyakumar et al., 2004; Mendella et al., 2000; Pilgrim et al., 1999; Schoenberg et al., 2002, 2004a, 2004b). Many of these subtest combinations are based on their correlations with FSIQ from the Wechsler Adult Intelligence Scale-Revised (WAIS-R) or Wechsler Adult Intelligence Scale (Third Edition) (WAIS-III) standardization sample. Psychologists wanting a brief measure will often use the subtest with the highest correlation with FSIQ. The advantages of using a shortened form of an established test include a relatively quick familiar method of

Correspondence and reprint requests to: Ruth Spinks, Psychiatry Research, Medical Examination Building, University of Iowa, Iowa City, Iowa 52242. E-mail: ruthspinks@gmail.com

testing that produces high correlations with the referent measure with both measures based on a very large, representative normative sample established by the original instrument (Wechsler, 1997). A drawback to partial test administration is interpolating the subtest(s) scores into a Full Scale estimate when the true Full Scale measure is computed using more subtests or questions.

A hybrid of these two types of proxy IQ measures combines demographic variables with limited testing using select subtests of the WAIS-III. The Oklahoma Premorbid Intelligence Estimate-3 (OPIE3; Schoenberg et al., 2002) provides five different formulae using between one and four WAIS-III subtests combined with demographic information to estimate FSIQ. The OPIE3-4 subtest (ST) uses the Vocabulary, Information, Matrix Reasoning, and Picture Completion subtests together with age, ethnicity, education level, and region of residence. The OPIE3-2ST FSIQ combines the Vocabulary (V) and Matrix Reasoning (MR) raw scores from the WAIS-III with age, education, ethnicity, and gender. Shorter OPIE3 uses the Vocabulary, Matrix Reasoning, and/or Picture Completion subtests with demographic variables.

The last form of proxy IQ measures examined here are original tests that provide an IQ estimate [the North American Adult Reading Test (NAART) and the Shipley Institute of Living Scales (SILS)] and school achievement testing. The NAART (Blair & Spreen, 1989) is an estimate of premorbid IQ and taps a relatively well-preserved function, pronouncing irregularly spelled words. The SILS (Zachary, 1986) is a two-subtest measure designed to produce an estimated FSIQ and two subscales. The conceptual quotient (CQ) is a measure of impairment, and the abstraction quotient adjusts the CQ for age and education (Zachary, 1986). Finally, we examine a prominent standardized school achievement test, the Iowa Test of Basic Skills (ITBS; Hoover et al., 2003).

The current article examined 11 proxy measures to determine their level of agreement with WAIS-III FSIQ across the entire sample. Two measures, the Barona and Crawford demographics formulae, were originally formulated for use with the WAIS-R. Since they are still in use, clinically there were examined to see how they related to the WAIS-III. The sample was also divided into three ability levels to determine how well the proxy measures perform at the tails of the IQ distribution.

## METHODS

All procedures were approved by the Internal Review Board on Human Research.

### Participants

Data for 313 participants from the Iowa Adoption Studies were used for the current study. The Iowa Adoption Studies is a series of studies examining Gene  $\times$  Environment risk for developing substance abuse or psychopathology. Study par-

ticipants underwent a complete neuropsychological test battery as part of the most recent follow-up. The average age was 43.89 years ( $SD = 6.78$ ) and ranged from 31 to 60 years. The sample was predominantly female (61.86%). Average education was 14.17 years ( $SD = 2.26$ ).

## Procedures

All participants in the current follow-up were given a neuropsychological test battery that included the WAIS-III, the NAART (Blair & Spreen, 1989), and the SILS (Zachary, 1986) as global measures of cognitive ability (IQ). All measures were administered by trained research assistants under standard conditions and double scored by two trained raters who had achieved very high interrater reliability (average reliability  $\geq 0.90$ ). Files were then reviewed by a neuropsychologist. The WAIS-III was always given first in the battery to minimize fatigue effects, and testing typically began in the morning. The order of all other tests in the test battery was varied according to a Latin square design. School achievement data were obtained from the participant's elementary and/or secondary school or from Iowa Testing Services at the University of Iowa after obtaining signed consent from the research participant.

## Measures Evaluated

Measures evaluated for this study included the Ward-7ST short form developed by Ward and modified for the WAIS-III by Pilgrim et al. (1999), the NAART, the SILS, ITBS, the Barona and Crawford demographic regression formulae, and the five OPIE3 hybrids combining demographic and WAIS-III subtest information. The final estimate examined was the ITBS (Hoover et al., 2003), a nationally recognized standardized school achievement test. School achievement is strongly related to IQ (Sattler, 2001), and the ITBS correlated .64 with WAIS-III FSIQ (Spinks et al., 2007). All proxy measures were computed per previously published guidelines. ITBS Iowa state percentile rank scores were converted to IQ scores using table 1.1 in Strauss et al., (2006). Only FSIQ measures were examined for the various proxy measures.

## ANALYSES

### Entire Sample

WAIS-III FSIQ was considered our referent measure. Means, standard deviations, minimum and maximum scores for the WAIS-III FSIQ, and all the proxy measures for the entire sample are shown in Table 1, Entire Sample 1. Pearson correlations and confidence intervals were used as a measure of agreement between the WAIS-III FSIQ and the various proxy measures. Spearman correlations were compared to the Pearson correlations to check for any nonlinearity in the data. Finally, intraclass correlations were calculated to examine the case-by-case correspondence of the WAIS-III

**Table 1.** Descriptive statistics for FSIQ estimates

Variable	Entire sample (N = 313)					Below-average FSIQ (n = 18)					Average FSIQ (n = 211)					Above-average FSIQ (n = 84)				
	Mean	SD	Minimum	Maximum	% Within 5 points FSIQ	Mean	SD	Minimum	Maximum	% Within 5 points FSIQ	Mean	SD	Minimum	Maximum	% Within 5 points FSIQ	Mean	SD	Minimum	Maximum	% Within 5 points FSIQ
WAIS-III	106.72	13.44	67.00	155.00	—	78.67	4.75	67.00	84.00	—	102.71	7.73	85.00	114.00	—	122.51	7.45	115.00	155.00	—
Ward-7ST	103.23	13.27	66.00	144.00	65.18	77.39	4.63	66.00	84.00	88.89	99.40	8.26	78.00	116.00	68.57	118.08	8.66	100.00	144.00	51.76
OPIE3-4ST	109.06	10.81	70.48	132.04	61.98	86.59	7.10	70.48	101.33	<b>22.22</b>	106.56	7.96	78.97	120.95	61.43	119.98	4.62	108.51	132.04	71.76
OPIE3-2ST	110.53	10.95	74.67	132.23	<b>48.56</b>	90.09	8.48	74.67	101.22	<b>16.67</b>	109.21	8.55	82.58	123.61	<b>44.29</b>	121.65	4.93	109.09	132.23	65.88
OPIE3-V	107.74	10.34	68.70	124.66	<b>48.88</b>	90.52	8.72	73.04	102.56	<b>16.67</b>	105.66	9.09	68.70	122.75	50.48	116.51	4.88	102.39	124.66	51.78
OPIE3-MR	107.42	9.65	75.08	126.20	<b>40.58</b>	91.37	9.05	75.08	105.23	<b>16.67</b>	105.77	8.33	79.67	122.95	<b>45.24</b>	114.90	6.13	96.48	126.20	<b>34.12</b>
OPIE3-PC	107.60	9.61	71.84	128.67	51.12	89.42	9.47	71.84	103.19	<b>33.33</b>	105.98	7.95	80.10	122.16	55.71	115.45	5.24	101.20	128.67	<b>43.53</b>
SILS	105.05	9.14	72.00	131.00	54.31	87.50	9.38	72.00	100.00	<b>38.89</b>	103.49	7.26	85.00	118.00	66.19	112.61	5.59	97.00	131.00	<b>28.24</b>
NAART	102.39	10.98	67.32	125.76	<b>40.26</b>	87.92	10.76	67.32	105.53	<b>33.33</b>	99.72	9.58	78.56	125.76	<b>48.57</b>	112.04	6.23	97.66	124.64	<b>21.18</b>
ITBS	103.66	12.76	64.00	136.00	<b>36.74</b>	86.09	9.85	64.00	106.00	<b>38.89</b>	101.15	11.00	75.50	136.00	<b>40.95</b>	113.74	10.00	91.00	136.00	<b>25.88</b>
Crawford	99.88	7.61	79.79	113.45	<b>32.27</b>	89.44	6.90	79.79	102.54	<b>22.22</b>	99.17	7.26	81.44	113.27	<b>44.76</b>	103.81	5.86	88.74	113.45	<b>3.53</b>
Barona	108.31	4.10	86.34	120.55	<b>37.38</b>	107.79	4.42	100.78	117.73	<b>0.00</b>	107.78	4.11	<b>86.34</b>	118.91	50.48	109.72	3.74	98.61	120.55	<b>12.94</b>

Note. The sample size, mean, SD, minimum and maximum values, and the percentage of each sample where the participant's proxy score was within 5 points of their WAIS-III FSIQ score are listed for the entire sample (N = 313) and for the three IQ groups determined from the WAIS-III FSIQ scores [e.g., below (n = 18), average (n = 211), and above average (n = 84)]. Values in italics refer to outside defined range and values in boldface refer to less than 50% of sample or group estimated within 5 points. ST, subtest; V, Vocabulary; PC, Picture Completion; MR, Matrix Reasoning.

and proxy measures. The Spearman and intraclass correlations were slightly lower than Pearson correlations, but all three correlation matrices were quite similar. Therefore, the Spearman and intraclass correlations are not reported here. Percent agreement (defined as ±5 IQ points) between the WAIS-III FSIQ and each proxy measure was also calculated.

Repeated measures multivariate analysis of variance (MANOVA) and *post hoc* comparisons examined the statistical difference between the proxy measures and the WAIS-III FSIQ.

### Different Ability Levels

To examine the relationship of the proxy measures and WAIS-III FSIQ at the tails of the IQ score distribution, the sample was divided into three groups according to WAIS-III FSIQ. Individuals with an FSIQ at or above 115 were classified “above average” (actual score range 115–155). The “average-ability” group had FSIQs ranging from 85 to 114. The “below-average” individuals were those with an FSIQ below 85 (actual score range 67–84). Analyses computed on the entire sample were also performed on the three ability groups to determine how each proxy measure performed at the tails of the IQ distribution.

## RESULTS

### Participants

As a group, WAIS-III FSIQ was slightly above average overall (mean IQ = 106.68, SD = 13.43, range 67–155), and many individuals had above-average IQs (n = 84) than below-average IQs (n = 18) (Table 1). The average level of formal education was 14.17 years (SD = 2.26), with a range of 8–17 years.

### Analyses on the Entire Sample

All the group means for the various IQ estimates produced were within 7 points of WAIS-III FSIQ. However, the range of IQs estimated by the proxy measures differed greatly from the referent measure (Table 1).

Table 1 shows the percent agreement of each proxy and FSIQ. The highest percent agreement was 65.18% by the Ward-7ST IQ estimate. The lowest percent agreement was 32.27% produced by the Crawford demographics equation.

The Pearson correlation and confidence interval between WAIS-III FSIQ and each proxy measure are shown in Table 2. Correlations ranged from  $r = .25$  for the Barona estimate to  $r = .95$  for the Ward-7ST short form. Three proxy measures, the ITBS, Barona, and Crawford, had Pearson correlations with WAIS-III FSIQ below  $r = .70$ , indicating they were not reliable enough for clinical use.

Repeated measures MANOVA tested all the proxy measures against the WAIS-III FSIQ. The main effect for proxy measure was highly significant,  $F(df = 1, 12) = 253.35, p < .0001$ .

**Table 2.** Correlations between WAIS-III and estimated FSIQ for each proxy measure

Variables	Entire sample ( <i>N</i> = 313)		Below-average FSIQ ( <i>n</i> = 18)		Average FSIQ ( <i>n</i> = 211)		Above-average FSIQ ( <i>n</i> = 84)	
	<i>r</i>	CI	<i>r</i>	CI	<i>r</i>	CI	<i>r</i>	CI
Ward-7ST	.95	0.93–0.96	.80	0.52–0.92	.87	0.83–0.90	.83	0.74–0.88
OPIE3-4ST	.92	0.90–0.94	.86	0.65–0.95	.84	0.79–0.88	.69	0.55–0.79
OPIE3-2ST	.87	0.84–0.90	.78	0.48–0.92	.79	0.73–0.84	.58	0.41–0.71
OPIE3-V	.77	0.72–0.81	.56	0.11–0.82	.64	0.55–0.72	.45	0.25–0.61
OPIE3-MR	.73	0.67–0.78	.70	0.33–0.88	.59	0.49–0.67	.35	0.14–0.53
OPIE3-PC	.79	0.74–0.83	.58	0.14–0.83	.68	0.60–0.75	.44	0.24–0.60
SILS	.78	0.73–0.82	.60	0.17–0.83	.68	0.59–0.74	.34	0.13–0.51
NAART	.71	0.65–0.76	.37	–0.14 to 0.72	.19	0.05–0.32	.11	–0.12 to 0.32
ITBS	.64	0.58–0.67	.04	–0.45 to 0.51	.47	0.35–0.57	.23	0.01–0.43
Barona	.25	0.14–0.36	.33	–0.18 to 0.70	.36	0.23–0.48	.01	–0.21 to 0.23
Crawford	.49	0.39–0.57	.48	–0.02 to 0.77	.74	0.67–0.80	.52	0.33–0.66

*Note.* Column groups of the table show the correlations of WAIS-III and proxy measures for (1) the entire sample, (2) individuals with WAIS-III FSIQs below 85 (below-average IQ), (3) individuals with IQs ranging from 85 to 115, and (4) IQs above 115. Values in italics refer to correlations below .70 (minimum accepted clinical correlation). ST, subtest; V, Vocabulary; PC, Picture Completion; MR, Matrix Reasoning; CI, confidence interval.

*Post hoc* comparisons between each proxy and the WAIS-III FSIQ are shown in Table 3. The OPIE3-V, OPIE3-MR, OPIE3-PC, SILS, and the Barona estimates did not differ significantly from the WAIS-III FSIQ.

### Different Ability Levels

The performance of the proxy measures across the different cognitive ability groups was examined next. The sample sizes of the groups ranged from 18 in the below-average group, 211 in the average-ability groups, and 84 in the above-average group. Means, standard deviations, minimum and maximum scores, and the percentage of group scoring within 5 points of the WAIS-III FSIQ score are listed in Table 1, Below-average FSIQ, Average FSIQ, Above-average FSIQ. Pearson correlations and confidence intervals are shown in Table 2. Note that restriction of range attenuated the correlations somewhat in the different ability levels.

The Ability Level  $\times$  Proxy Measure interaction of the repeated measures MANOVA was highly significant  $F(df = 1, 12) = 105.08$ ,  $p < .0001$  (Table 3). *Post hoc* contrasts between the WAIS-III FSIQ and each proxy measure are shown in Table 3.

### Average-IQ Group

The average-IQ group was the largest of the three ability groups ( $n = 211$ ). The mean WAIS-III FSIQ was 102.71 ( $SD = 7.73$ ). The percentage of group members scoring within 5 points of the WAIS-III FSIQ score for each proxy measure ranged from a high of 68.57% for Ward-7ST to a low of 40.95% for the ITBS (Table 1, Average FSIQ, percentage within 5 points FSIQ).

Correlation coefficients and confidence intervals between each proxy measure and the WAIS-III FSIQ are shown in Table 2. The Pearson correlation for the Ward-7ST, the Barona, and the Crawford was higher for the average-ability group

than for the entire sample (Table 2, Average FSIQ). The remaining correlation coefficients were reduced in the average-ability group. Some attenuation of the correlations was expected due to restriction of range, but the reduction of the correlation coefficient between FSIQ and NAART was unusually large. The correlation coefficient ranged from  $r = .71$  for the entire sample to  $r = .19$  for the average-ability group. The *post hoc* contrasts from the mixed-model MANOVA indicated that 9 of the 11 proxy measures were significantly different from the WAIS-III FSIQ. Only the SILS and ITBS estimates were not statistically different at the  $p < .0001$  level (Table 3).

### Below-Average Group

The mean WAIS-III FSIQ for the below-average group was 78.67 ( $SD = 4.75$ ). All the proxy measures except the Ward-7ST produced mean IQs above the upper cutoff of the low-ability group (i.e., above 85; Table 1, Below-average FSIQ). The overall percent agreement (within 5 points of FSIQ) was poor in the low-ability group. Only one proxy (Ward-7ST) came within 5 points of FSIQ for more than 40% of the low-ability group. Seven of the 11 proxy measures did not produce clinically reliable correlations with WAIS-III FSIQ (i.e.,  $r \geq .70$ ). *Post hoc* contrasts from the MANOVA indicated that only the Ward-7ST and ITBS estimates did not statistically differ from the WAIS-III FSIQ.

### Above-Average IQ Group

Six of the 11 proxy measures judged the mean IQ of the above-average group to be in the average-ability range (Table 1). The five measures producing above-average mean IQs all used WAIS-III subtest scores. The four measures requiring the greatest amount of formal testing were the only proxies to have more than 50% of group members estimated



**Table 3.** *Post hoc* comparisons of repeated measures MANOVA comparing all proxy measures to WAIS-III FSIQ

Proxy measure	Entire sample (N = 313)				Below-average ability (n = 18)				Average ability (n = 211)				Above-average ability (n = 84)			
	Mean	SD	t Value	p Value	Mean	SD	t Value	p Value	Mean	SD	t Value	p Value	Mean	SD	t Value	p Value
WAIS-III FSIQ	106.68	13.43	—	—	78.67	4.75	—	—	102.68	7.72	—	—	122.51	7.45	—	—
Ward-7ST	103.20	13.27	14.16	<.0001	77.39	4.63	1.83	.0852	99.36	8.26	11.51	<.0001	118.08	8.66	8.38	<.0001
OPIE3-4ST	109.06	10.81	7.49	<.0001	86.59	7.10	8.60	<.0001	106.56	7.96	12.25	<.0001	119.98	4.62	3.87	.0002
OPIE3-2ST	111.53	10.85	12.63	<.0001	90.09	8.48	8.50	<.0001	109.21	8.55	17.00	<.0001	121.65	4.93	0.97	.3340
OPIE3-V	107.74	10.34	2.24	.259	90.52	8.72	6.80	<.0001	105.66	9.09	5.73	<.0001	116.51	4.88	7.74	<.0001
OPIE3-MR	107.42	9.65	1.49	.1377	91.37	9.05	7.94	<.0001	105.77	8.33	5.95	<.0001	114.90	6.13	8.55	<.0001
OPIE3-PC	107.60	9.61	2.02	.0445	89.42	9.47	5.75	<.0001	105.98	7.95	7.34	<.0001	115.45	5.24	8.93	<.0001
SILS FSIQ	105.04	9.13	3.43	.0007	87.50	9.38	4.96	.0001	103.48	7.24	1.91	.0576	112.61	5.59	11.95	<.0001
NAART	102.32	11.03	8.09	<.0001	87.92	10.76	4.13	.0007	99.63	9.65	5.23	<.0001	112.04	6.23	12.52	<.0001
ITBS	103.66	12.76	4.42	<.0001	86.09	9.85	2.89	.0113	101.15	11.00	1.91	.0576	113.74	10.00	7.74	<.0001
Crawford	99.75	7.92	10.04	<.0001	89.44	6.90	6.35	<.0001	98.98	7.71	5.88	<.0001	103.81	5.86	17.93	<.0001
Barona	108.31	4.10	2.22	.0272	107.79	4.23	22.96	<.0001	107.78	4.10	8.83	<.0001	109.72	3.74	14.38	<.0001

*Note.* A significance value of  $p < .0001$  was adopted to account for multiple comparisons. Values in italics refer to no significant difference from WAIS-III FSIQ.

within 5 points of actual FSIQ IQ. The Ward-7ST estimate was the only proxy to correlate above  $r = .70$  for the high-ability group. *Post hoc* comparisons from the MANOVA indicated that only the OPIE3-4ST and OPIE3-2ST proxy measures did not differ from WAIS-III FSIQ.

### DISCUSSION

The most important finding of this article is how poorly the IQ proxy measures performed at the tails of the IQ distribution. The proxy measures consistently overestimated the IQ of low-functioning individuals and underestimated the IQs of high-functioning individuals. Inaccurate assessment of global ability could be very problematic for clinicians designing treatment programs and managing family expectations for recovery. Researchers using a poor measure of global functioning as a covariate may miss a significant theoretical finding because too much (or too little) variance is being attributed to global functioning and factored out of the model.

The poor performance at the tails of the IQ distribution is concealed in the current literature, which only reports correlations for the entire sample. This has given the neuropsychological community a false sense of security for many years. The proxy measures consistently produced higher percent agreement (Table 1) and correlations (Table 2) across the entire sample than for the low- or high-ability individuals. This pattern cannot be completely accounted for by restriction of range.

The proxy measures that consistently produced scores closest to the WAIS-III FSIQ were measures that required the greatest amount of formal testing and used WAIS-III subtests. Similar analyses of verbal IQ and performance IQ proxy measures, not published here, found the same pattern (i.e., the greater the amount of formal testing, the more accurate the IQ estimate).

The second major finding of this article is how the different proxy measures performed depending on the method of evaluation. The majority of proxy articles published only examined correlations. We presented Pearson correlations, percent agreement defined as  $\pm 5$  IQ points, and  $F$  and  $t$  tests from repeated measures MANOVA with *post hoc* comparisons. The Pearson correlations were the most consistent measures from group to group and from each group to the whole sample. The Ward-7ST and OPIE3-4ST produced the most consistently reliable correlations (i.e.,  $r > .70$ ) (Table 2). The Ward-7ST had the highest percent agreement index with WAIS-III FSIQ, with the whole sample and all three subgroups producing scores within 5 points of FSIQ at least 50% of the time. Conclusions determined from the MANOVA *post hoc* comparisons had very little overlap conclusions from the Pearson correlations or percent agreement. The OPIE3-Vocabulary, OPIE3-Matrix Reasoning, OPIE3-Picture Completion, SILS FSIQ, and Barona FSIQ did not differ from WAIS-III FSIQ statistically for the whole group. Overall, the proxy measures did not perform well in either the low- or the high-ability groups, regardless of the method of evaluation.

One interesting note on the various OPIE3, different formulae used different demographic variables. The only

demographic variables found in all the OPIE3 equations were educational attainment and ethnicity. Both gender and age were used in only four of five of the equations. Interestingly, when gender was used, it was a negative coefficient (males coded 0 and females coded 1), subtracting points from the overall equation. Females were also rated heavier than males (i.e., females had a coefficient of 2 where males = 1). An interesting point about using age in the regression formulae is that even raw subtest scores on the WAIS-III are age adjusted. It should be noted that the bulk of the OPIE equations needed additional age correction.

The NAART generally performed very poorly. In addition, the NAART had an obvious ceiling effect producing poor estimates for extremely bright individuals. However, the NAART is believed to represent a true measure of premorbid functioning in that the ability to pronounce irregularly spelled words is well preserved even in moderate dementia (Strauss et al., 2006). This may make the NAART the measure of choice when cognition suddenly becomes impaired due to injury or illness. Again, the clinician and researcher must consider the target sample in selecting a measure.

## CONCLUSIONS

The major emphasis of this article is that the IQ proxy measure of choice depends on the question being asked. Is the testing an estimate of current functioning, of premorbid functioning, or is the assessment for research? What are the basic patient characteristics that will influence testing (e.g., estimated ability level and the amount of testing time the patient will tolerate)? The administration of the full WAIS-III is recommended when possible. When the use of a proxy estimate is necessitated, consideration must be given to the examples shown here. The measures using the greatest amount of testing performed somewhat better than those using little or no formal testing. However, when the situation dictates, the use of a proxy, several proxy measures, produced marginal reliability and low correlations with WAIS-III FSIQ in individuals with cognitive functioning at the ends of the ability distribution.

## ACKNOWLEDGMENTS

This study was supported by the National Institute on Drug Abuse grant number R01 DA005821, Gene  $\times$  Environment Interactions in the Development of Drug Abuse. This study was presented at the Annual Conference of the International Society for Intelligence Research, December 2007, Amsterdam, The Netherlands. The authors have no financial or other relationships that could be interpreted as a conflict of interest affecting this article.

## REFERENCES

- Axelrod, B.N., Dingell, J.D., Ryan, J.J., & Ward, L.C. (2000). Estimation of the Wechsler Adult Intelligence Scale-III with the 7-subtest short form in a clinical sample. *Assessment*, 7, 157–161.
- Barona, A., Reynolds, C.R., & Chastain, R. (1984). A demographically based index of premorbid intelligence test for the WAIS-R. *Journal of Consulting and Clinical Psychology*, 52, 885–887.

- Blair, J.R. & Spreen, O. (1989). Predicting premorbid IQ: A revision of the National Adult Reading Test. *The Clinical Neuropsychologist*, 3, 129–136.
- Crawford, J.R. & Allan, K.M. (1997). Estimating premorbid WAIS-R IQ with demographic variables: Regression equation derived from a UK sample. *The Clinical Neuropsychologist*, 11, 192–197.
- Engelhart, C.I., Eisenstein, J., Johnson, V., & Losonczy, M. (1999). Comparison of linear equating and prorated short forms for estimating WAIS-R FSIQ in a neuropsychological population. *The Clinical Neuropsychologist*, 13, 95–99.
- Hoover, H.D., Dunbar, S.B., & Frisbie, D.A. (2003). *Iowa Tests of Basic Skills Complete/Core Battery Norms: Score conversions, student norms and school average norms, (Form A)*. Itasca, IL: Riverside Publishing.
- Jeyakumar, S.L.E., Warriner, E.M., Raval, V.V., & Ahmad, S.A. (2004). Balancing the need for reliability and time efficiency: Short forms of the Wechsler Adult Intelligence Scale-III. *Educational and Psychological Measurement*, 64, 71–87.
- Mendella, P.D., McFadden, L., Regan, J., & Medlock, L. (2000). Short-form prediction of WAIS-R scores in a sample of individuals diagnosed with multiple sclerosis. *Applied Neuropsychology*, 7, 102–107.
- Pilgrim, B.M., Meyers, J.E., Bayless, J., & Whetstone. (1999). Validity of the Ward seven-subtest WAIS-III short form in a neuropsychological population. *Applied Neuropsychology*, 6, 243–246.
- Sattler J.M. (Ed.). (2001). *Assessment of children: Cognitive applications* (4th ed.). San Diego, CA: Jerome M. Sattler Publishing, Inc.
- Schoenberg, M.R., Duff, K., Scott, J.G., & Adams, R.L. (2002). An evaluation of the clinical utility of the OPIE-3 as an estimate of premorbid WAIS-III FSIQ. *Clinical Neuropsychologist*, 17, 308–321.
- Schoenberg, M.R., Duff, K., Dorfman, K., & Adams, R.L. (2004a). Differential estimation of verbal intelligence and performance intelligence scores from combined performance and demographic variables: The OPIE-3 verbal and performance algorithms. *Clinical Neuropsychology*, 16, 266–276.
- Schoenberg, M.R., Scott, J.G., Ruwe, W., Patton, D., & Adams, R.L. (2004b). Assumptions that underlie predicting premorbid IQ: A comment on the "Evaluation of the accuracy of two regression-based methods for estimating premorbid IQ." *Archives of Clinical Neuropsychology*, 19, 1103–1106.
- Spinks, R., Arndt, S., Caspers, K., Yucuis, R., McKirgan, L.W., Pfalzgraf, C.J., & Waterman, E.E. (2007). School achievement strongly predicts midlife IQ. *Intelligence*, 35, 563–567.
- Strauss, E., Sherman, E.M.S., & Spreen, O. (2006). *A Compendium of Neuropsychological Tests: Administration, norms, and commentary* (3rd ed.). New York, NY: Oxford University Press.
- Wechsler, D. (1997). *Wechsler Adult Intelligence Scale* (3rd ed.). San Antonio, TX: The Psychological Corporation (A Division of Harcourt Assessment Company).
- Zachary, R.A. (1986). *Shipley Institute of Living Scales, revised manual*. Los Angeles, CA: Western Psychological Services.

## APPENDIX

### Formulae for Calculating Estimating IQ From Demographic Variables

*WAIS-III seven-subtest short form (Pilgrim et al., 1999)*

FSIQ estimate produced using the Information, Digit Span, Arithmetic, Similarities, Picture Completion, Block Design, and Digit Symbol subtests.

*Demographics only—Barona (Barona et al., 1984)*

Estimated FSIQ = 54.96 + 0.47 (age) + 1.76 (sex) + 4.71 (race) + 5.02 (education) + 1.89 (occupation) + 0.59 (region).

Standard error of the estimate = 12.14,  
where

Sex: 1 = female; 2 = male

Race: 1 = other; 2 = Black; 3 = White

Occupation: 1 = farm laborers, farm foremen, and laborers (unskilled labor)

2 = operatives/service workers/farmers

and farm managers (semiskilled labor)

3 = not in labor force

4 = craftsmen/foremen (skilled labor)

5 = managerial/official/clerical/sales

6 = professional/technical

Region: 1 = southern

2 = north central

3 = western

4 = northeast

Residence: 1 = rural (<2500)

2 = urban (>2500)

Age:

1 = 16–17	6 = 45–54	11 = 80–84
2 = 18–19	7 = 55–64	12 = 85–89
3 = 20–24	8 = 65–69	13 = 90–94
4 = 25–34	9 = 70–74	14 = 95–99
5 = 35–44	10 = 75–79	15 = >100

  

Education:	
1 = 0–7	4 = 12
2 = 8	5 = 13–15
3 = 9–11	6 = 16+

*Demographics only—Crawford (Crawford & Allan, 1997)*

Predicted FSIQ = 87.14 – (5.21 × occupation) + (1.78 × education) + (0.18 × age),

where

Education credit: 0.05 = for every year spent in school part-time

1 = for every full year of schooling

Occupation: 1 = professional

2 = intermediate

3 = skilled

4 = semiskilled

5 = unskilled

**Demographics and Testing Combined**

OPIE3 for the WAIS-III (Schoenberg et al., 2004a).

OPIE3-4 subtest FSIQ = 35.348 + 0.368 (Vocabulary Raw) + 0.682 (Information Raw) + 0.987 (Matrix Reasoning Raw) + 0.737 (Picture Completion Raw) + 0.175 (age in years) + 0.656 (education) + 0.578 (ethnicity) + 0.341 (region of country).

Standard error of the estimate = 5.68.

OPIE3-2 subtest FSIQ = 45.997 + 0.652 (Vocabulary Raw) + 1.287 (Matrix Reasoning Raw) + 0.157 (age in years) + 1.034 (education) + 0.652 (ethnicity) – 1.015 (gender).

Standard error of the estimate = 6.63.

OPIE3-Vocabulary subtest FSIQ = 57.220 + 0.874 (Vocabulary Raw) + 1.766 (education) + 1.081 (ethnicity) + 0.674 (region of country) – 1.508 (gender).

Standard error of the estimate = 8.35.

OPIE3-Matrix Reasoning FSIQ = 43.678 + 1.943 (Matrix Reasoning Raw) + 0.297 (age) + 3.564 (education) + 1.541 (ethnicity) + 0.543 (region of country) – 1.137 (gender).

Standard error of the estimate = 9.06.

OPIE3-Picture Completion FSIQ = 29.280 + 1.469 (Matrix Reasoning Raw) + 1.242 (Picture Completion Raw) + 0.332 (age) + 3.04 (education) + 1.025 (ethnicity) + 0.557 (region of country) – 1.278 (gender).

Standard error of the estimate = 7.93,

where

Age: age in years

Education coding:

1 = 0–8 years

2 = 9–11 years

3 = 12 years

4 = 13–15 years

5 = 16+ years

Ethnicity coding:

1 = African American

2 = Hispanic

3 = other

4 = Caucasian

Gender coding:

1 = male

2 = female

Region of country coding:

1 = south

2 = north central

3 = northeast

4 = west

**Iowa Test of Basic Skills**

Iowa percentile ranks converted to IQ scores using table 1.1 in Strauss et al. (2006) and then Iowa percentile ranks correlated with WAIS-III FSIQ scores, after Spinks et al. (2006).

**North American Adult Reading Test (Blair & Spreen, 1989)**

Test of 61 irregularly spelled English words.

**SILS-Revised Manual (Zachary, 1986): Los Angeles, CA: Western Psychological Services**

Standard administration and scoring.